

ACCOUNTABILITY AND COMPETENCE IN ELECTIONS: THEORY AND EVIDENCE FROM U.S. TERM LIMITS*

James Alt[†]
Ethan Bueno de Mesquita[‡]
Shanna Rose[§]

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 tax growth, spending growth, 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. We find that the two effects are of comparable magnitudes. 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.

* We thank Chris Berry, Tim Besley, John Gasper, Christina Gathmann, Rema Hanna, Ilyana Kuziemko, Rene Lindstadt, Rob van Houweling, and seminar participants at the University of Chicago, University of Chile, Florida State University, UC-Berkeley, New York University, the 2006 Annual Meeting of the Midwest Political Science Association, the 2007 Annual Meeting of the American Political Science Association, and the 2007 NBER Political Economy Meetings for helpful comments.

[†] Frank G. Thomson Professor of Government, Department of Government, Harvard University. jalt@latte.harvard.edu

[‡] Associate Professor, Harris School of Public Policy Studies, The University of Chicago. bdm@uchicago.edu

[§] Assistant Professor, Wagner School of Public Service, New York University. shanna.rose@nyu.edu

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., Rogoff 1990; Zaller 1998; Fearon, 1999; Ashworth 2005; Gordon, Huber, and Landa 2007; Ashworth and Bueno de Mesquita 2008; Gowrisankaran, Mitchell, and Moro 2008;).¹

Empirically distinguishing the accountability and competence effects of elections has proven difficult. In theoretical models, both effects operate in the same direction (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. Existing empirical research has estimated the overall effect of elections on performance (i.e., the combined accountability and competence effects) but has not been able to determine whether each of these effects really exists and, if so, what their relative magnitudes are.

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

¹ 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, Besley 2006, Besley and Smart 2007, Canes-Wrone and Shotts 2007, and Snyder and Ting 2009.

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 tax growth, spending growth, 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.

1. Literature Review

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.

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. In a widely-cited paper on gubernatorial term limits, Besley and Case (1995) find evidence of electoral accountability: in the American states between 1950 and

1986, per capita spending and taxes were higher under term-limited governors.² 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).

One omitted variable that may account for this puzzling result is 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 all of the studies mentioned above, they focus exclusively on accountability, simply including a dummy variable for term-limited governors in their regressions.³

We know of only one study that has attempted to take advantage of differences in term limit lengths to isolate the competence effect of elections. 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, such as Mexico and Turkey, than in countries with two-term limits, such as Portugal and Sri Lanka, consistent with a competence effect. However, studying the effects of term limits across countries is complicated by well-known problems of institutional heterogeneity; the American states have the advantage of sharing a broadly common institutional environment. Thus, in the remainder of the paper we develop a formal model that compares performance under different term-limit regimes and test its predictions using data from the American states.

² In a related study, 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.

³ 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.

2. Accountability, Competence, and Term Limits: Model and Implications

Most formal models of elections that include both moral hazard and adverse selection either restrict attention to a finitely repeated game (e.g., Maskin and Tirole 2004; Ashworth 2005; Persson and Tabellini 2000; Canes-Wrone, Herron, and Shotts 2001; Besley and Smart 2007; Canes-Wrone and Shotts 2007; Snyder and Ting 2009), assume that politicians are limited to two terms within an infinitely repeated game (e.g., Banks and Sundaram 1998), or restrict the set of equilibria in a way that makes the accountability and competence effects difficult to disentangle (e.g., Banks and Sundaram 1993). To model the effects of term limits of different lengths on the accountability and competence effects of elections, we need an infinitely repeated agency model of elections that allows for different term limit rules.

Some existing formal models do allow for such infinite repetition (e.g., Duggan 2000, Bansk and Duggan 2006, Besley 2006, Smart and Sturm 2006).⁴ However, those models focus on fidelity between voter and politician preferences (on the adverse selection side) and between voter policy preferences and the policy implemented (on the moral hazard side). Since we are interested in the effect of elections on economic performance rather than ideological congruence, it is more appropriate to focus on a model in which adverse selection relates to incumbent competence and moral hazard relates to incumbent effort. We provide a model of elections with moral hazard and adverse selection that is simple enough to be tractable while still incorporating the dynamics needed to study the natural experiment created by term limits in our empirical work. The model both rationalizes our empirical strategy of using term limits to isolate the effects we set out to estimate and highlights some complications and nuances that we will have to account for.⁵ The model also allows us to discuss a variety of ways in which term limits affect voter welfare.⁶

⁴ Also see Snyder and Ting (2009) for a model of repeated elections that incorporates the role of interest groups.

⁵ Though see Schwabe (2009) for a related model—with richer action and outcome spaces and farsighted voters—that yields different predictions from ours. In particular, in a compelling class of equilibria, politicians adjust effort, given

The Model

Consider a 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 $a \in \{\underline{a}, \bar{a}\}$. 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). The game is infinitely repeated.

We will say that there is a t -period term limit if a politician can only serve in office for t periods. A politician becomes *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 her private information.

The outcome space is $\{L, H\}$. We interpret H as the politician successfully managing the economy and/or fiscal policy in a given year. The function $f(a, \theta) : \{\underline{a}, \bar{a}\} \times \{\theta_C, \theta_I\} \rightarrow [0, 1]$, maps types and actions into a probability of achieving the outcome H , given by:

$$f(\bar{a}, \theta_H) = 1$$

$$f(\underline{a}, \theta_H) = \gamma$$

$$f(\bar{a}, \theta_L) = 0$$

$$f(\underline{a}, \theta_L) = 0.$$

reputation, such that the voter is indifferent between an incumbent with a good reputation and a randomly selected challenger. Meirowitz (2007) presents a similar result in a model with ideological policy choice and uncertain budget constraints.

⁶ In the conclusion we discuss some positive implications of term limits which are not included in our models,

Only competent types can successfully manage the economy, and effort (or foregone rents) increases the probability of a good outcome for competent types, i.e., $\gamma < 1$. We also assume that γ is large enough to create selection effects in a two-term-limit system. This requires that the voter prefer an incumbent who is competent for certain but 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 so there would be no possibility of selection on competence. Hence, this assumption allows us to model elections with both accountability and competence in the most parsimonious way.

In any given term, there is also some chance that an incumbent transitions from competent to incompetent after choosing her effort level but before the outcome is determined. The probability that a competent type becomes incompetent in her t^{th} term is ε^t . This possibility captures idiosyncratic changes in the strategic environment that may cause a previously effective incumbent to become ineffective for reasons beyond her control.⁷ The fact that the transition probability is decreasing over time reflects the idea that incumbents become more effective over time due to experience. All qualitative results continue to hold if the transition probability is not decreasing with terms in office.

The benefit of reelection is $B > 0$. At the beginning of the game, Nature chooses the cost of high effort, c , and makes it common knowledge. 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 $\bar{c} = \frac{(1-\gamma)B \sum_{i=1}^{\infty} \delta^i \prod_{j=1}^i (1-\varepsilon^j)}{1+(1-\gamma) \sum_{i=1}^{\infty} \delta^i \prod_{j=1}^i (1-\varepsilon^j)}$. The upper bound on the support assures that the costs of effort are

low enough, relative to the benefits of office, that a competent politician would be willing to exert high effort in exchange for a promise of indefinite reelection as long as he remained competent. Without

⁷ We could introduce an analogous transition probability from incompetent to competent. It would not change the analysis but would clutter notation.

loss of generality, we normalize the cost of low effort to zero. The politician's payoff from any period in which she is in office is:

$$B - c(a),$$

where $c(\bar{a}) = c$ and $c(\underline{a}) = 0$. Politicians discount the future by δ .

We assume that voters are myopic, caring only about the next period. Without loss of generality, the payoff to the voter of the outcome H is assumed to be 1 and the payoff to the outcome L is assumed to be 0.

Denote by $\mu(O, a, \mu', t)$ the voter's posterior belief that the incumbent will be a competent type in her t^{th} term, given that the outcome in her $t-1^{\text{th}}$ term was O , the voter believes the politician's action in her $t-1^{\text{th}}$ term was a , and the voter believed the probability the incumbent was competent in her $t-1^{\text{th}}$ term was μ' . Then we have:

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

We focus on pure strategy Perfect Bayesian Equilibria in Markov Strategies (Maskin and Tirole 2001). The restriction to Markov Strategies requires that equilibria be stationary in the payoff-relevant state variables: beliefs about ability and terms remaining.

One-Term Limit

In the absence of any chance of reelection, effort imposes costs without offering any benefits for the politician. As such, a politician who can only serve one term will choose \underline{a} in that term. The following result is immediate and stated without proof.

Proposition 1. *In the unique MPE of the game with a one term limit, the incumbent chooses \underline{a} regardless of type.*

Two-Term Limit

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. The voter forms his beliefs by observing outcomes. 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. These intuitions are formalized in the following result.⁸

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

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

⁸ The proofs of this and all subsequent results appear in the appendix.

This result implies that first-term outcomes are better for the voter under two-term limits than under one-term limits due to the 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. This is because, although no term-limited incumbent has an incentive to exert high effort, there is a competence effect. Only competent types survive the first election and, moreover, due to experience, they are less likely to transition away from competence.

No Term Limits

The absence of term limits enhances the incentive effect even further, since it provides the incumbent with a longer time horizon. In particular, unlike in the case of a two-term limit, the voter is able to provide second- (and later-) term politicians with electoral incentives. Moreover, selection effects continue to operate. Thus, we have the following result:

Proposition 3 *There is an infinite, increasing sequence $\{\tilde{c}_1, \tilde{c}_2, \dots\}$ satisfying $F(\tilde{c}_t) = 1$ for all t , such that in any equilibrium of the game with no term limits:*

- *A competent incumbent in her t^{th} term chooses \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 t^{th} -term incumbent is competent is $1 - \varepsilon^t$.

As we will see in greater detail below, this result demonstrates that there is an equilibrium in which the conditions for sustaining effort in the first period are weaker (i.e., incumbents are willing to exert effort for a wider range of costs) under no term limits than under two-term limits. When politicians can serve

more terms in office, the benefit of winning reelection includes not only the immediate payoff, but also the possibility of future terms, so second-term (and later) incumbents are also willing to exert effort.

Empirical Implications: Comparisons across Systems

In this section we derive empirical implications from the model in order to show how the length of term limits can be used to distinguish and thus potentially estimate our quantities of interest, the accountability and competence effects of elections. The first predictions relate to differences between one- and two-term-limit systems.

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.*
- 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.*

In 4.1 the difference is entirely due to the fact that, lacking electoral incentives in the one-term-limit system, first-term incumbents are less likely to exert effort. In 4.2 the difference is entirely due to the fact that second-term incumbents have higher expected competence than first-term incumbents. Thus, in our terms, differences in performance between term-limited and eligible incumbents in their first term are attributable solely to an accountability effect, while differences in performance between first- and second-term lame ducks are attributable solely to a competence effect. These comparisons are summarized in Table 1.

[Table 1 about here]

The second set of predictions relate to differences between two-term-limit and no-term-limit systems.

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.*
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.*

In 5.1 the difference is entirely due to the fact that, in the two-term-limit system, second-term incumbents are less likely to exert effort (an accountability effect). In 5.2, second-term incumbents are of higher expected competence than first-term incumbents in either system (a competence effect). Moreover, second-term incumbents in a no-term-limit system are more likely to be willing to exert effort than first-term incumbents in a one-term-limit system (an accountability effect). Thus, in our terms, a comparison of second-term eligible incumbents to first-term eligible incumbents estimates a combination of competence and accountability. These comparisons are summarized 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.

Probability of Reelection

In addition, the model has implications for the first-term probability of reelection under different term-limit rules. In particular, reelection in the first term is more likely, the less stringent the

term-limit rule. The intuition is as follows. In both two-term-limit and no-term-limit systems, if first-term competent politicians exert a high level of effort in equilibrium, then competent types are always reelected and incompetent types are never reelected. If first-term competent politicians do not exert high effort, then competent types are reelected with probability γ (i.e., if they happen to get the good outcome) and incompetent types are never reelected. Thus, reelection of competent types is more likely when first-term candidates exert high effort. Since first-term candidates are more likely to exert high effort with no term limits than with two-term limits, the absence of term limits increases the probability of reelection.

Proposition 6. *The probability of reelection for first-term incumbents is higher with no term limits than with two-term limits.*

3. 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.⁹ The sample period is 1950-2000 unless otherwise noted. We adopt a specification similar to those of Besley and Case (1995), 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 1995, 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. In order to integrate our study with this literature, we use the log of per capita spending and taxes as dependent variables.

⁹ 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.

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.¹⁰ 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 3.¹¹

[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.”

¹⁰ 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.

¹¹ 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.

We also include a variety of economic and political controls. We follow Besley and Case (1995, 2003) in controlling 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 add controls relating to the party composition of government (Alt and Lowry 2000): Democratic governor, Democratic house, Democratic senate, divided government, and 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, which determines the number of terms a successful candidate could serve and thus the potential value of office, could 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).

Finally, we include state fixed effects, year fixed effects, and state-specific linear time trends to address any remaining unobserved heterogeneity.

4. 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 uses Proposition 4 to estimate these effects by running the following regression on the sample of states that changed from one- to two-term limits:

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

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.

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 the growth of per capita spending and taxes is 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 incomes. 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. Intriguingly, 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.

[Table 4 about here]

Robustness

As noted above, a possible source of heterogeneity in this 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. We control for the incumbent's prior political experience 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.

Any analysis of the policy effects of institutions raises inevitable concerns about institutional endogeneity. In the case of the switch from one- to two-term limits, 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 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. Again, importantly for what follows, in no case do we find reason to believe that the accountability and competence effects are of different magnitudes.

Resolving an Empirical Puzzle

These results help to resolve an important puzzle in the empirical literature. As described in the literature review, Besley and Case (1995) 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 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.

5. 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 sample period. This allows us to avoid problems with the specification used by Besley and Case (1995; 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.¹²

We run the following regression:

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

The omitted category is second-term governors who are eligible for reelection.¹³ 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. Since performance is expected

¹² Moreover, many of the 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 our estimates of the effects of term limits. Maryland and Ohio adopted two-term limits in the 1950s. In order to include these states, our sample begins in 1960. We find similar results for the sample beginning in 1950 and excluding these two states.

¹³ It also includes a handful of governors in their third term or higher. All results are robust to excluding these governors.

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 the growth of 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 the growth of per capita taxes is 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 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 both our theoretical story and our data, 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 the growth of per capita taxes is 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 growth or borrowing costs.

Probability of Reelection

Our final test relates to Proposition 6, which predicts that the first-term probability of reelection is higher under no term limits than under two-term limits. A difference in means test reveals that the first-term probability of reelection is 0.66 in states without term limits compared to only 0.63 in states with two-term limits, as predicted, but this difference falls short of statistical significance. When we add control variables for the governor's party affiliation, the governor's years of experience, economic growth and per capita tax growth (both of which were shown to drive the reelection of incumbents in Table 5), and year effects, the result remains qualitatively unchanged.

However, among those states that adopted two-term limits during the sample period, the first-term probability of winning reelection was 0.75 before and 0.64 after the adoption of two-term limits. This difference is statistically significant at the ten percent level, and is robust to the inclusion of the aforementioned controls and year effects to account for changes in the probability of reelection over time. Of course, this estimate is potentially affected by the same endogeneity issues that we highlighted in the section on one- versus two-term limits. Nonetheless, the estimate suggests that proponents of term limits—who often cite reducing the incumbency advantage as a potential benefit—could well be right: term limits do appear to increase turnover, not just by removing lame ducks but also by reducing the probability of reelection among eligible incumbents.

6. Conclusions

This paper attempts to shed light on the role of elections in representative democracy. We develop 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” (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 separate accountability from competence in a common set of elections. While both of these effects have long existed in theory, this paper is, to our knowledge, 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 and spending grow more rapidly, 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 tax and spending growth 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 provided further confirmation of 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 gain 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 experience effect. However, since high 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, for which we also provided some evidence. 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 2007) of why voters and interest groups take the positions they do on term limits.

Appendix

Proof of Proposition 2.

Since incompetent types can never achieve the outcome H , in any equilibrium they choose low effort in every period. Further, in the second term, the incumbent will choose low effort even if competent, since there is no prospect of reelection.

Now, consider the voter's behavior when faced with a first-term incumbent.

Lemma 1. *In any pure strategy 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 of Lemma 1.

Suppose the voter expects high effort from competent types in the first term. Then sequential rationality requires that the voter reelect the incumbent when his beliefs are μ if

$$\mu(O, \bar{a}, \mu_0, 1)\gamma \geq \mu_0(1 - \varepsilon).$$

Following the outcome H , beliefs are $\mu(H, \bar{a}, \mu_0, 2) = 1 - \varepsilon^2$. The voter will reelect if $(1 - \varepsilon^2)\gamma \geq \mu_0(1 - \varepsilon) \Leftrightarrow (1 + \varepsilon)\gamma \geq \mu_0$, which we have by the fact that $\gamma > \mu_0$.

Following the outcome L the voter's beliefs are that the incumbent's probability of being competent is 0. Thus, sequential rationality implies that the voter will not reelect.

Now suppose the voter expects low effort from competent types in the first term. Then, sequential rationality requires that 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 $\mu(L, \underline{a}, \mu_0(1 - \varepsilon), 2) = \frac{(1 - \varepsilon^2)\mu_0(1 - \varepsilon)(1 - \gamma)}{(1 - \gamma)\mu_0(1 - \varepsilon) + 1 - \mu_0(1 - \varepsilon)}$, which algebra shows is less than $\mu_0(1 - \varepsilon)$ so the voter will not reelect. ■

Now consider first-term effort by competent types.

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

According to Lemma 1, the voter will reelect if and only if the politician succeeds. This implies that the expected payoff to the competent politician of choosing \bar{a} is

$$B - c + (1 - \varepsilon)\delta B$$

The payoff to the competent politician of choosing \underline{a} is

$$B + (1 - \varepsilon)\gamma\delta B.$$

The politician will choose high effort in the first period 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 directly above establishes that, in this case, 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 $c = (1 - \gamma)(1 - \varepsilon)\delta B$.

Here, the first-term incumbent is indifferent and so any action is consistent with equilibrium. However,

whatever the incumbent does in this case, the proof above shows that in any equilibrium play is as described in the proposition whenever $c \neq (1-\gamma)(1-\varepsilon)\delta B$. ■

Proof of Proposition 3.

Consider the following assessment:

- Competent incumbents choose \bar{a} .
- Incompetent incumbents choose \underline{a} .
- Voters reelect if and only if their posterior beliefs are greater than $\mu_0(1-\varepsilon)$.
- Voter's beliefs after observing the outcome H in an incumbent's t^{th} term are that the incumbent is competent with probability $1-\varepsilon^t$. Voter's beliefs after observing the outcome L in an incumbent's t^{th} term is that the incumbent is competent with probability 0.

First, consider consistency of beliefs with actions. Given the politicians' strategies, competent types succeed and incompetent types fail. Competent types then transition to incompetence with probability $1-\varepsilon^t$. This, then, implies the voter's beliefs.

Second consider whether the incompetent type's strategy is sequentially rational. Since incompetent types always fail, regardless of effort, the voter always has posterior beliefs that incompetent types are incompetent with probability 1. This implies that the voter never reelects an incompetent type. Thus, sequential rationality requires that the incompetent type choose \underline{a} , as called for in the assessment.

Third consider whether the competent type's strategy is sequentially rational. If a competent type in her t^{th} term plays the strategy called for, her expected payoff is:

$$B - c + (1 - \varepsilon^t) \delta \left(B - c + (1 - \varepsilon^{t+1}) \delta \left(B - c + (1 - \varepsilon^{t+2}) \delta \dots \right) \right)$$

A one shot deviation to \underline{a} yields an expected payoff of:

$$B + (1 - \varepsilon^t) \gamma \delta \left(B - c + (1 - \varepsilon^{t+1}) \delta \left(B - c + (1 - \varepsilon^{t+2}) \delta \right) \dots \right)$$

From the one-shot deviation principle, there is no profitable deviation if:

$$\begin{aligned} & B(1 - \gamma) \left((1 - \varepsilon^t) \delta + (1 - \varepsilon^t)(1 - \varepsilon^{t+1}) \delta^2 + (1 - \varepsilon^t)(1 - \varepsilon^{t+1})(1 - \varepsilon^{t+2}) \delta^3 + \dots \right) \\ & \geq c + c(1 - \gamma) \left((1 - \varepsilon^t) \delta + (1 - \varepsilon^t)(1 - \varepsilon^{t+1}) \delta^2 + (1 - \varepsilon^t)(1 - \varepsilon^{t+1})(1 - \varepsilon^{t+2}) \delta^3 + \dots \right) \end{aligned}$$

which can be rewritten

$$c \leq \frac{B(1 - \gamma) \sum_{i=1}^{\infty} \delta^i \prod_{j=1}^i (1 - \varepsilon^{t+j})}{1 + (1 - \gamma) \sum_{i=1}^{\infty} \delta^i \prod_{j=1}^i (1 - \varepsilon^{t+j})} \equiv \tilde{c}_t.$$

Now, it is clear that \tilde{c}_t is minimized for $t = I$, so the incumbent least likely to exert high effort is a first-term incumbent. Substituting in $t = I$ we find that $\tilde{c}_1 = \bar{c}$, so $F(\tilde{c}_t) = 1$ for all t .

Finally consider whether the voter's strategy is sequentially rational. Given the incumbents' strategies, the voter's expected payoff from reelection is simply the probability he assigns to the incumbent being competent. His payoff from not reelecting is $\mu_0(1 - \varepsilon)$. This establishes that this assessment is an equilibrium.

To see that there are no other equilibria, first note that the argument above establishes that it cannot be part of an equilibrium for incompetent types to exert effort. Thus, the only remaining 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 above assessment. But, then, the argument above establishes that sequential rationality requires competent types to exert effort. ■

Proof of Proposition 4.

1. Proposition 1 states that first term incumbents in a one-term-limit system will choose \underline{a} . Thus, expected performance is $\mu_0(1-\varepsilon)\gamma$. Proposition 2 states that first-term incumbents in a one-term-limit system will choose \bar{a} if the costs of effort are less than \hat{c} . Thus, expected performance is $\mu_0(1-\varepsilon)(F(\hat{c})+(1-F(\hat{c}))\gamma) > \mu_0(1-\varepsilon)\gamma$.

Remark. Since the probability that first-term incumbents are the competent type is $\mu_0(1-\varepsilon)$ in both systems, it is clear, comparing these two expected performances that performance is higher in the two-term-limit system and that the difference $(\mu_0(1-\varepsilon)F(\hat{c})(1-\gamma))$ is due to the fact that in a two-term-limit system the first-term incumbent may choose high effort.

2. Proposition 2 states that second-term incumbents in the two-term-limit system will choose \underline{a} , and are the competent type with probability $1-\varepsilon^2$. Thus, expected performance is $(1-\varepsilon^2)\gamma$. Comparing this to first-term incumbents in the one-term-limit system (from the previous point), it is clear that second-term lame ducks perform better and that the difference $((1-\varepsilon^2)-(1-\varepsilon)\mu_0)\gamma$ is entirely due to the fact that second-term lame ducks are more likely to be the competent type. ■

Proof of Proposition 5.

1. As shown in the proof of Proposition 4, the expected performance of second-term incumbents in a two-term-limit system is $(1-\varepsilon^2)\gamma$. According to Proposition 3, second-term incumbents in a no-term-limit system are competent when reelected and choose high effort. Thus, the

expected performance of such incumbents is $(1 - \varepsilon^2)$. The difference in expected performance is $(1 - \varepsilon^2)(1 - \gamma)$, which is clearly due entirely to the fact that in the no-term-limit system, the second-term incumbent may exert high effort.

2. As shown above, expected performance of second-term incumbents in the no-term-limit system is $(1 - \varepsilon^2)$. The expected performance of first-term incumbents in the no-term-limit system is $(1 - \varepsilon)\mu_0 < (1 - \varepsilon^2)$. The difference is due entirely to differences in expected competence.

The expected performance of first-term incumbents in a two-term limit system is

$\mu_0(1 - \varepsilon)(F(\hat{c}) + (1 - F(\hat{c}))\gamma)$. To see that there is an accountability effect as well as the

competence effect, it suffices to show that $\hat{c} < \bar{c}$. This requires

$$\begin{aligned}
(1 - \gamma)(1 - \varepsilon)\delta B &< \frac{B(1 - \gamma)\sum_{i=1}^{\infty}\delta^i\prod_{j=1}^i(1 - \varepsilon^j)}{1 + (1 - \gamma)\sum_{i=1}^{\infty}\delta^i\prod_{j=1}^i(1 - \varepsilon^j)} \Leftrightarrow \\
(1 - \varepsilon)\delta + (1 - \varepsilon)\delta(1 - \gamma)\sum_{i=1}^{\infty}\delta^i\prod_{j=1}^i(1 - \varepsilon^j) &< \sum_{i=1}^{\infty}\delta^i\prod_{j=1}^i(1 - \varepsilon^j) \Leftrightarrow \\
(1 - \varepsilon)\delta(1 - \gamma)\sum_{i=1}^{\infty}\delta^i\prod_{j=1}^i(1 - \varepsilon^j) &< \sum_{i=2}^{\infty}\delta^i\prod_{j=2}^i(1 - \varepsilon^j) \Leftrightarrow \\
-(1 - \varepsilon)\delta\gamma\sum_{i=1}^{\infty}\delta^i\prod_{j=1}^i(1 - \varepsilon^j) &< \sum_{i=1}^{\infty}\delta^{i+1}(\varepsilon - \varepsilon^{i+1})\prod_{j=1}^i(1 - \varepsilon^j).
\end{aligned}$$

Now the left-hand side is clearly negative while the right-hand side is clearly positive, establishing the result. ■

Proof of Proposition 6. Under no or two-term limits, the probability of reelection for a first term incumbent is $\mu_0(1-\varepsilon)$ if she chooses high effort and $\mu_0(1-\varepsilon)\gamma$ if she chooses low effort. Thus, the probability of reelection for first-term incumbents under no term limits is $\mu_0(1-\varepsilon)$ whereas under two term limits it is $\mu_0(1-\varepsilon)(F(\hat{c})+(1-F(\hat{c}))\gamma) < \mu_0(1-\varepsilon)$. ■

References

- Alt, James E. and Robert C. Lowry. 2000. "A Dynamic Model of State Budget Outcomes Under Divided Partisan Government." *Journal of Politics* 62: 1035-69.
- Ashworth, Scott. 2005. "Reputational Dynamics and Political Careers." *Journal of Law, Economics and Organization* 21: 441-446.
- Ashworth, Scott and Ethan Bueno de Mesquita. 2006. "Delivering the Goods: Legislative Particularism in Different Electoral and Institutional Settings." *Journal of Politics* 68(1):168-179.
- Ashworth, Scott and Ethan Bueno de Mesquita. 2008. "Electoral Selection, Strategic Challenger Entry, and the Incumbency Advantage." *Journal of Politics* 70(4):1006-1025.
- Banks, Jeffrey S. and John Duggan. 2006. "A Dynamic Model of Democratic Elections in Multidimensional Policy Spaces." University of Rochester typescript.
- Banks, Jeffrey S., and Rangarajan K. Sundaram. 1998. "Optimal Retention in Agency Problems." *Journal of Economic Theory* 82(2): 293-323.
- Banks, Jeffrey S., and Rangarajan K. Sundaram. 1993. "Moral Hazard and Adverse Selection in a Model of Repeated Elections," in Barnett, William A., Melvin J. Hinich, and Norman J. Schofield, eds. *Political Economy: Institutions, Information, Competition, and Representation*. Cambridge: Cambridge University Press.
- 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. 1995. "Does Electoral Accountability Affect Economic Policy Choices? Evidence from Gubernatorial Term Limits." *Quarterly Journal of Economics* 110 (3): 769-798.
- 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): 755-773.
- 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(May):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* 45:532-550.
- 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: 105-136.
- Duggan, John. 2000. "Repeated Elections with Asymmetric Information." *Economics and Politics* 12(2):109-135.
- Ferejohn, John. 1986. "Incumbent Performance and Electoral Control." *Public Choice* 50: 5-26.
- Ferraz, Claudio and Frederico Finan. 2009. "Electoral Accountability and Corruption: Evidence from the Audits of Local Governments." NBER Working Paper 14937.
- Figlio, David N. 1995. "The Effect of Retirement on Political Shirking: Evidence from Congressional Voting." *Public Finance Review* 23(2): 226-241.
- Gordon, Sanford C., Gregory A. Huber, and Dimitri Landa. 2007. "Challenger Entry and Voter Learning." *American Political Science Review* 101(2):303-320.

- 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* 11:1-17.
- 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.
- Holmström, Bengt. 1999. "Managerial Incentive Problems: A Dynamic Perspective." *Review of Economic Studies* 66: 169–82.
- 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: 73-90.
- King, Gary, Michael Tomz, and Jason Wittenberg, "Making the Most of Statistical Analyses: Improving Interpretation and Presentation" *American Journal of Political Science* 44, no. 2 (April 2000): 347-61.
- 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.
- List, John A., and Daniel M. Sturm. 2006. "How Elections Matter: Theory and Evidence from Environmental Policy." *Quarterly Journal of Economics* 121(4): 1249-1281.
- Lott, John R. and Stephen G. Bronars. 1993. "Time Series Evidence on Shirking in the U.S. House of Representatives." *Public Choice* 76: 125-149.
- 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 (March): 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):1034-1054.
- McArthur, John, and Steven V. Marks. 1988. "Constituent Interest vs. Legislator Ideology: The Role of Political Opportunity Cost." *Economic Inquiry* 26(3): 461-470.
- 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): 347-381.
- Persson, Torsten, and Guido Tabellini. 2000. *Political Economics: Explaining Economic Policy*. Cambridge: MIT Press.
- Schwabe, Rainer. 2009. "Reputation and Accountability in Repeated Elections." Princeton University typescript.
- 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): 419-444.
- Snyder, James M., and Michael M. Ting. 2009. "Interest Groups and the Electoral Control of Politicians." *Journal of Public Economics* (forthcoming).
- Tien, Charles. 2001. "Representatives, Voluntary Retirement, and Shirking in the Last Term." *Public Choice* 106:117-130.

- 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: 444-456.
- 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.

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 $c \leq \hat{c}$	$\mu_0(1-\varepsilon)$	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: $\gamma\left((1-\varepsilon^2)-(1-\varepsilon)\mu_0\right)$

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: $(1-\varepsilon^2)(1-\gamma)$
	Second Term Eligible	\bar{a}	$1-\varepsilon^2$	
First Term Eligible	No Term Limits	\bar{a}	$\mu_0(1-\varepsilon)$	Competence Effect $(1-\varepsilon^2)-\mu_0(1-\varepsilon)$
	2 Term Limit	\bar{a} only if $c \leq \hat{c}$	$\mu_0(1-\varepsilon)$	Competence and Accountability $(1-\varepsilon^2)-\mu_0(1-\varepsilon)(F(\hat{c})+(1-F(\hat{c}))\gamma)$

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 (Accountability)	-0.048** (0.012)	-0.065** (0.015)	-0.039** (0.014)	-0.039** (0.018)	-5.81** (2.18)	-14.04** (3.45)	0.66** (0.27)	0.82** (0.33)
Second-term lame duck (Competence)	-0.041** (0.012)	-0.050** (0.015)	-0.030** (0.015)	-0.029* (0.018)	-6.75** (2.47)	-14.54** (3.44)	0.45 (0.29)	0.54* (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.

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 (Accountability)	0.018 (0.016)	0.016 (0.016)	0.032* (0.017)	0.030* (0.017)	12.21** (3.21)	12.76** (3.20)	-0.28 (0.40)	-0.22 (0.41)
First-term eligible (Competence/accountability)	0.027** (0.011)	0.022** (0.011)	0.045** (0.011)	0.038** (0.012)	10.89** (2.62)	11.89** (2.75)	-0.21 (0.31)	-0.01 (0.32)
First-term eligible non-winner		0.017 (0.012)		0.025* (0.014)		-2.98 (2.94)		-0.68** (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.