Identification and Semi-parametric Estimation of Dynamic Games of Electoral Competition*

Holger Sieg
University of Pennsylvania

Chamna Yoon
Baruch College

*We would like to thank Flavio Cunha, John Duggan, Antonio Merlo, George Mailath, Andy Postlewaite, Michael Peress, Richard Rogerson, Petra Todd, Ken Wolpin, and seminar participants at numerous conferences and workshops for comments and discussions. We would also like to thank Jason Dickhaut for excellent research assistance, helping us assembling the data set used in this paper. Sieg acknowledges the support of the National Science Foundation.
Abstract

We provide a new approach to identify and estimate dynamic games of electoral competition. Building on recent theoretical advances in modeling repeated elections, we study dynamic games in which candidates from two parties compete in statewide elections. Candidates differ in their ideological positions and competence. Both dimensions are initially unobserved by voters, but are (partially) revealed over time through observed incumbents’ behavior. Politicians cannot credibly commit to policies prior to an election. Equilibria can be characterized by asymmetric election standards, which depend on the party of the candidate, the competence of the incumbent, and the history of the game. We show that parts of our model are non-parametrically identified and propose a new semi-parametric estimator. Our empirical analysis draws on data from U.S. gubernatorial elections between 1950 and 2012 and evaluates the impact of term limits on economic policies and welfare. Our findings suggest that voters are willing to accept significant trade-offs in ideology to obtain a more competent governor. We find that term limits lead to more variation in economic policies over time and, thus, reduce voters’ welfare.
1 Introduction

Elections serve an important function in modern democracies. They allow voters to express their support for politicians that share their ideological positions and pursue the policies preferred by voters. In addition, elections provide voters the opportunity to remove from office incumbents that cannot competently perform the duties of the office.\(^2\) Recent advances in economic theory suggest modeling these two aspects of elections using dynamic games with asymmetric information.\(^3\) The purpose of this paper is to discuss identification of such models and develop a new semi-parametric estimator for this class of dynamic games. We implement our estimator using data from U.S. gubernatorial election and evaluate the impact of term limits.

Following Duggan (2000), we model competition among two political parties as the outcome of a dynamic game of repeated elections. Candidates differ in their ideological positions as well as their competence. There is a sequence of elections, and incumbents face a binding two-term limit.\(^4\) An election either pits two untested candidates against one another or an incumbent against an untested challenger. Our model maintains the key assumption that a politician cannot credibly commit to a policy prior to an election. An elected politician, therefore, has a tendency to implement his or her preferred policy once in office (Osborne and Slivinski, 1996, Besley and Coate, 1997). Policy moderation only arises due to the desire to be reelected.

Voters cannot observe the ideological position or the competence (valence) of an untested politician who has not served in office. Priors for these politicians depend solely on the

\(^{2}\)Early papers that focus on political control are Barro (1973), Ferejon (1986) and Banks and Sundaram (1998).

\(^{3}\)Our model follows the recent literature on repeated election and combines elements of Duggan (2000), Bernhardt, Dubey, and Hughson (2004), and Bernhardt, Camara, and Squintani (2011). A recent survey of this literature is given by Duggan and Martinelli (2014).

\(^{4}\)There are binding two-term limits in 23 U.S. states.
candidates’ party. Each governor perfectly reveals his or her competence during the first term. Voters also observe the policy implemented by the governor, which serves as a signal regarding the incumbent’s unobserved ideological type. Voters update their beliefs based on the observed policies and observed competence and vote accordingly. Our model allows for two parties with different distributions of ideological positions, but a large common support. Hence, equilibria are not symmetric with respect to both parties. There are fiscal liberals and conservatives in both parties. Not surprisingly, we find in our empirical analysis that Republicans tend to be more fiscally conservative, on average, than Democrats.

We define and characterize a Perfect Bayesian Equilibrium of the dynamic game. We provide conditions so that our model generates unique reelection thresholds for both parties. These election standards imply ideological thresholds which characterize politicians’ strategies. Politicians from each party can be characterized as belonging to one of three groups. Centrists always implement their preferred policies and are reelected to a second term. Extremists also implement their preferred policies in the first period, but are not reelected to a second term. Term limits do not have a direct impact on these types of politicians. The third group of politicians are Moderates. They have incentives to moderate their policies in the first term to win reelection to a second term. The election standards also depend on the competence of the incumbent. Voters are willing to reelect more extreme politicians as long as they are more competent.

The main objective of this paper is to establish non-parametric identification of key parts of the model and develop a new semi-parametric estimator.5 One empirical challenge

---

is that both ideology and competence are latent from the perspective of the econometrician. However, we observe a variety of policy outcomes that are primarily a function of ideology. We consider three such outcomes: expenditures per capita, taxes per capita, and minimum wages. These are outcomes that are expected to be highly correlated with ideology that differentiates fiscal conservatives from fiscal liberals. Ideology is, therefore, implicitly defined as differences in opinion about the preferred size of government. In addition, we collect outcomes that primarily depend on competence. Here, we focus on measures such as workers compensations, state income growth, and state debt borrowing costs. Our proxies for competence are, therefore, the ability of the governor to manage important economic programs, to lower borrowing costs, and to generate economic growth.

We assume that all observed policies are noisy measures of ideology and competence. Recall that Extremists will implement their preferred policy in the first period. Centrists also implement their preferred policy in the first period, but Moderates do not. Both get reelected to a second term, and it is, therefore, hard to distinguish among these types based on first period policies alone. However, Centrists and Moderates implement their preferred ideological policies in the second period because of the binding term limit. As a consequence, we can use first period policies of Extremists and second period policies of Moderates and Centrists to identify the distribution of ideology (and competence). Carneiro, Hansen, and Heckman (2003) show that we need to observe, at least, five different policy outcomes. In addition, they impose the exclusion restriction that two of these outcomes are only functions of one latent factor. Following their approach, we can then treat the distributions of ideology and competence as latent and identify the factor loadings associated with each policy outcome. Following Cunha, Heckman, and Schennach (2010), we can appeal to Kot-

---

Knight and Schiff (2010) estimate a model of social learning in presidential primaries.

The first four measures were used in Besley and Case (1995). Similar additional measures were proposed by Alt, Bueno de Mesquita, and Rose (2011) to disentangle accountability from competence.

larski’s Theorem to non-parametrically identify the underlying distributions of ideology and competence. We implement this procedure separately for each party, thus identifying the distributions of ideological positions for Democratic and Republican candidates as well as the distributions of measurement or implementation error.

We then show that we can use the observed probabilities of extremism in each party (conditional on observed policies) together with the equilibrium properties of the model to identify the election thresholds, the benefits that candidates from each party assign to holding office, and the relative importance of competence in voters’ preferences. We also can predict the degree of policy moderation implied in periods in which term limits are not binding. We construct additional over-identifying restrictions based on the observed first period policy choices by centrists and moderates to test the validity of our model specification. Finally, we show how to identify the underlying distribution of voter preferences based on the observed vote shares of incumbents (conditional on observed policies) that are reelected to a second term. Our proofs of identification are constructive and can be used to design a semi-parametric estimator of our model.

Our data set consists of all gubernatorial elections in the U.S. held between 1950 and 2012. Our empirical findings suggest that the benefits from holding office are significant and large in economic magnitude. As a consequence, the prospects of reelection provide strong incentives for moderate governors to move towards the center of the ideological spectrum during their first term in office. Voters are willing to accept significant trade-offs in ideology to obtain a more competent governor. There are no significant differences in competence

---

8 Differences between ideology and observed outcomes may arise due to the complexities of the legislative process or shocks during the implementation process.

9 Our paper is related to growing literature in econometrics that studies identification and estimation of dynamic games. Some recent methodological papers include Pakes, Ostrovsky, and Berry (2007), Bajari, Benkard, and Levin (2007), Aguirregabiria and Mira (2007), and Pesendorfer and Schmidt-Dengler (2008), Merlo and Tung (2012), and Hu and Shum (2013). As discussed in detail below, our model and our approach towards semi-parametric identification and estimation differs significantly from these papers.
across parties in the U.S. We also find that the distribution of voter preferences is similar to
the distribution of political candidates, providing support for citizen-candidate type models.
If anything, voters tend to be more extreme than candidates. Finally, we evaluate the
impact of term limits by comparing the outcomes to a model without term limits. Term
limits restrict the ability of incumbents to run for reelection and, therefore, eliminate the
incentives for policy moderation in the second term. Term limits also change the incentives
for voters at the ballot box. We find that term limits also lead to more variation in economic
policies over time and, thus, reduce voters’ welfare.

Before we turn to our analysis, we offer a few more observations regarding the related
literature in political economy. Downs (1957) developed the canonical theoretical model of
a single election in which candidates can commit to policies prior to an election. Alesina
(1988) extends the basic static framework and considers a repeated election model with two
candidates. An alternative approach to the Downsian approach is based on the citizen-
candidate literature, which goes back to Osborne and Slivinski (1996) and Besley and Coate
(1997). These models are based on the notion that candidates cannot commit to policies
prior to an election. Most of the citizen-candidate literature focuses on one-shot elections.
Duggan (2000) introduced repeated elections into a citizen-candidate model with asymmetric
information. This model was extended to account for term limits by Bernhardt, Dubey, and
Hughson (2004). Bernhardt, Camara, and Squintani (2011) consider the trade-off between
competence and ideology in a dynamic model without term limits. These models provide the
basic framework that we estimate in this paper. We allow for asymmetries in the underlying
distributions of ideology of candidates from the two competing parties. As a consequence,
equilibria are not symmetric. Duggan and Fey (2006) consider repeated elections within a
Downsian model and office motivated candidates. Benefits of holding office play a role in
our model as well. Banks and Duggan (2008) consider repeated elections when the policy
space is multi-dimensional. They characterize the set of equilibria in simple voting and
policy strategies. Aragones, Palfrey, and Postelwaite (2007) also consider a repeated election
model with two candidates, but allow for reputation effects which lead to policy moderation in equilibrium. To our knowledge, this is the first paper that has shown how to identify and estimate these types of dynamic games.

A related literature deals with the trade-off between moral hazard (accountability) and adverse selection (competence) in elections.\textsuperscript{10} These models are based on the notation that “good” politicians need to exert effort to convince voters that they are competent. These models also give rise to election standards that are used by voters to provide incentives for politicians.\textsuperscript{11} We focus on the adverse selection problem and abstract from the effort decision treating ideology as a predetermined, but initially unobserved source of heterogeneity.

The seminal empirical paper on term limits is Besley and Case (1995). They consider two different agency models with term limits. The empirical analysis is based on a fixed effect panel data estimator using U.S. data from gubernatorial election from 1950-1986. (Besley and Case (2003) extends the analysis to the mid 1990’s.) They show that term limits affect policy choices as predicted by the model we estimate. As we discuss in detail below, their results are consistent with our approach and our findings. Another prominent empirical paper on term limits is Daniel and Lott (1997) who provide evidence that term limits increase the probability that incumbents lose elections. This finding is also consistent with our model. Stone and Simas (2010) discuss the empirical literature that has analyzed the relationship between valence and ideology. Most of the empirical studies focus on federal elections.

The rest of the paper is organized as follows. Section 2 presents our model which is based on the recent literature on dynamic games of electoral competition. Section 3 discusses identification and estimation. Section 4 introduces our data set. Section 5 presents our

\textsuperscript{10}See, for example, Banks and Sundaram (1998), Person and Tabellini (2000), Maskin and Tirole (2004), Ashworth (2005), Snyder and Ting (2008), and Gowrisankaran, Mitchell, and Moro (2008).

\textsuperscript{11}For a recent discussion of the empirical literature that has tried to distinguish between these two effects, see Alt, Bueno de Mesquita, and Rose (2011).
empirical results. Section 6 discusses the policy implications. We offer conclusions in Section 7.

2 A Dynamic Game of Repeated Elections

Following Duggan (2000), Bernhardt, Dubey, and Hughson (2004), and Bernhardt, Camara, and Squintani (2011), we consider a dynamic game that captures the repeated elections of a governor in a state that has adopted a two-period term limit for the office holder. There is a continuum of infinitely lived voters that differ by their ideological location, $\theta \in \mathbb{R}$. Voter’s utility depends on the policy that a governor implements when in office and the managerial ability of competence of the governor, denoted by $a$. The period utility that voter $\theta$ receives from a governor who implements a policy $x \in \mathbb{R}$ and has competence $a \in A = [A_L, A_H]$ is given by:

$$u(\theta, x, a) = L_\theta(x) + \lambda a$$

where $L_\theta(x)$ is a symmetric, single-peaked loss function. For simplicity we assume that:

$$L_\theta(x) = -|\theta - x|$$

The distribution function of voters’ preferences in the society is given by $F_\theta(\cdot)$. The median voter is located at $\theta = 0$. Voters maximize expected life-time utility. Period utilities are discounted using a common factor, $\beta$.

Politicians receive a payoff both from being in office and from the position that they personally take while in office. A governor with ideology $\rho$ and competence $a$ who locates at $x$ derives period utility

$$v(\rho, a, x) = v(\rho, x) = -|\rho - x| + y$$

where $y > 0$ captures the payoff from holding office. Politicians share a common discount factor, $\gamma$ and maximize life-time utility.\footnote{For expositional reasons, we consider the case in which only the distribution of politician’s ideal points}
We focus on the empirically relevant case where challengers are chosen by opposing parties, denoted by D and R.\textsuperscript{13} We assume that heterogeneity in ideology among politicians of party $j$ is given by a distribution, denoted by $F_{j}^{\rho}(\cdot)$. The two distributions share a common support. We also assume that competence is uncorrelated with a candidate’s ideology and follows a distribution denoted by $F_{j}^{a}(\cdot)$.

The ideological positions of politicians are private information, not observed by other candidates or voters. Voters hold beliefs about ideologies, observe policies taken in office and update beliefs about incumbents. The competence of the challenger is also initially private information. We assume that an elected governor reveals his competence to other candidates and voters during his first term. The competence of an incumbent is, thus, common knowledge.

If there is no incumbent (such as the first period of the game), there is an election between two untried challengers, one from each party. Whenever two untried politician compete against each other in an election, the outcome is determined by a coin toss that elects a politician of party $D$ with probability $p_{D}$. The politician that wins the election then becomes the incumbent.\textsuperscript{14}

Consider the case in which a Republican has been elected to office for a first term in period $t$. The elected governor implements a policy during his first term, denoted by $x_{t}$, which is observed by all voters. Voters update their beliefs about the ideological type and the competence of the incumbent. At the beginning of period $t + 1$, the Republican incumbent with known competence $a$ then faces a challenger from party $D$ in an election. If the incumbent is reelected, he serves a second term in period $t + 1$. Since he faces a binding depend on the party. Extending the model to account for additional sources of heterogeneity by party are straightforward and discussed in the estimation section.

\textsuperscript{13}In contrast to most previous papers we do not assume that party R (D) consists of all candidates with ideology $\rho < 0$ ($\rho > 0$). Our empirical results indicate that this assumption is empirically not valid.

\textsuperscript{14}It is not difficult to endogenize this winning probability by making it a function of the value functions associated with both candidates.
term limit, a second term incumbent implements his preferred policy, $x_{t+1} = \rho$. In period $t + 2$ there is an open election, since the incumbent cannot run for reelection. The game at the beginning of period $t + 2$ is exactly like the game at the beginning of period $t$. We focus on a stage-undominated Perfect Bayesian Equilibrium (PBE).

First term politicians have strategies, $\delta_j(\rho, a)$, that map ideology and competence into first term policies. These strategies are party specific. The history of past positions for an incumbent that has served one period is given by $x_{t-1}$. An incumbent’s strategy is a function $\delta_j(\rho, a, x_{t-1})$ that assigns a policy for each history, politician’s ideology, competence and party, $j \in \{D, R\}$.

A voting strategy for an election with an established incumbent from party $j$ is given by a function $\alpha_j(\theta, x_{t-1}, a)$ that maps the voters type and the observed history and competence into the probability of voting for an incumbent that belongs to party $j$.\footnote{The voting strategy for open elections is a mixed strategy, i.e. a coin toss.} We focus on anonymous sincere voting strategies, i.e. voting strategies that only depend on the incumbent’s personal history and party membership. Voting is sincere if

- $\alpha_R(\theta, x_{t-1}, a) = 1$ if voters prefer the Republican incumbent.
- $\alpha_R(\theta, x_{t-1}, a) = 0$ if voters strictly prefer the Democratic challenger.

Similarly, we can define voting strategies if the incumbent is a Democrat. Voters do not use weakly dominated strategies that hinge on the fact that a voter is not pivotal.

Voter beliefs about a Republican incumbent’s ideology for all possible histories are given by the common belief function $P_R(\rho | x_{t-1}, a)$ which is the cumulative probability that a Republican incumbent has ideology less than $\rho$ given the observed history $x_{t-1}$ and competence, $a$.

Consider a time period $t$, which is the second period for an incumbent, i.e. the incumbent
is term-limited. He, therefore, solves the following optimization problem:

$$\max_{x_t} [-|\rho - x_t| + y]$$

(4)

As a consequence, a term-limited governor will implement his preferred policy in the second period, \(x_t = \rho\).

Let \(V^o(\theta)\) denote the expected discounted utility of electing a new governor in an open election. Notice that this value function is time independent since it does not depend on the history of the game.

Let \(V^D(\theta)\) denote the expected discounted utility of electing a new governor from party \(D\). The expected discounted utility if a Republican incumbent is reelected to serve a second term is given by:

$$V^{I,R}(\theta, x_{t-1}, a) = -E\left(|x_t - \theta| \bigg| x_{t-1}, a\right) + \lambda a + \beta E[V^o(\theta)]$$

(5)

If the incumbent is a Republican, sincere voting then implies that:

1. \(\alpha_R(\theta, x_{t-1}, a) = 1\) if \(V^{I,R}(\theta, x_{t-1}, a) \geq V^D(\theta)\)

2. \(\alpha_R(\theta, x_{t-1}, a) = 0\) if \(V^{I,R}(\theta, x_{t-1}, a) < V^D(\theta)\)

A similar condition holds for a Democratic incumbent.

**Definition 1** An equilibrium then consists of two strategy functions for voters (one for each party), two common belief functions, two strategies for untested politicians, and two strategies for incumbents, such that:

- the candidates maximize expected utility given their own ideology, competence, and voters’ strategies,
- the voters vote sincerely given the candidates’ and incumbents’ strategies,
beliefs are consistent with candidates’ and incumbents’ strategies and updated according to Bayes’ Rule.

We are now in a position to characterize the key properties of equilibrium. We have the following result:

**Proposition 1** Voters adopt time invariant election standards for incumbents of each party with competence \(a\) that are given by \([s_j(a), \bar{s}_j(a)]\), \(j \in \{D,R\}\). An incumbent with competence \(a\) belonging to party \(j\) is reelected if and only if the observed policy in the first period in office is within the interval given by \([s_j(a), \bar{s}_j(a)]\).

In the rest of this section, we provide some intuition for the result in Proposition 1 and outline a proof.\(^{16}\)

Note that the resulting equilibrium is not necessarily symmetric, i.e. politicians from different parties face different election standards. Figure 1 illustrates the election standards that arise for each party using our estimated model. We plot the upper and lower election standard as a function of competence. Note that there are significant differences in election standards across parties. Moreover, both parties have a lower and upper threshold reflecting the fact that there are liberals and conservatives in each party. Finally, it is also fairly straight forward to show that election standards get tighter as the benefits from holding office increase.

The election standards then imply ideological thresholds for politicians which implicitly characterize politicians’ strategies. Let us define \(\bar{\rho}_R(a)\) such that:

\[
\gamma y - |\bar{\rho}_R(a) - \bar{s}_R(a)| = 0
\]

(6)

Notice that a Republican politician with ideology \(\bar{\rho}_R(a)\) and competence \(a\) is indifferent between implementing policy \(\bar{s}_R(a)\) and being reelected and implementing policy \(\bar{\rho}_R(a)\) and

\(^{16}\)A formal proof of this result in a symmetric model without term limits is given by Bernhardt, Camara, and Squintani (2011).
not being reelected. Similarly define

$$\gamma y - |\rho_{R}(a) - s_{R}(a)| = 0$$  \hspace{2cm} (7)$$

Notice that a Republican politician with ideology $\rho_{R}(a)$ is indifferent between implementing policy $s_{R}(a)$ and being reelected and implementing policy $\rho_{R}(a)$ and not being reelected.

Then consider a Republican politician that has competence $a$ and has just been elected in period $t$ to serve his first term. The incumbent’s optimization problem implies the following decision rules:

- $\rho < \rho_{\underline{R}}(a)$ then $x_t = \rho$, expecting to lose reelection in $t + 1$.
- $\rho \in (\rho_{\underline{R}}(a), s_{R}(a))$ then $x_t = s_{R}(a)$, expecting to win reelection in $t + 1$.
- $\rho \in (s_{R}(a), \bar{s}_{R}(a))$ then $x_t = \rho$, expecting to win reelection in $t + 1$.
- $\rho \in (\bar{s}_{R}(a), \bar{\rho}_{R}(a))$ then $x_t = \bar{s}_{R}(a)$, expecting to win reelection in $t + 1$.  


• \( \rho > \bar{\rho}_R(a) \) then \( x_t = \rho \), expecting to lose reelection in \( t + 1 \).

This equilibrium can be supported by the following voting beliefs.

• If \( x_t < \underline{\rho}_R(a) \), then \( P_R(\rho|x_t, a) = 0 \) for all \( \rho < x_t \) and \( P_R(\rho|x_t, a) = 1 \) for all \( \rho \geq x_t \). (Left Extremists)

• If \( x_t = \underline{s}_R(a) \) then \( P_R(\rho|x_t, a) = F^\rho_R(\rho|\rho \in [\underline{\rho}_R(a), \underline{s}_R(a)]). \) (Left-leaning Moderates)

• If \( x_t \in (\underline{s}_R(a), \bar{s}_R(a)) \), then \( P_R(\rho|x_t, a) = 0 \) for all \( \rho < x_t \) and \( P_R(\rho|x_t, a) = 1 \) for all \( \rho \geq x_t \). (Centrists)

• If \( x_t = \bar{s}_R(a) \) then \( P_R(\rho|x_t, a) = F^\rho_R(\rho|\rho \in [\bar{s}_R(a), \bar{\rho}_R(a)]). \) (Right-leaning Moderates)

• If \( x_t > \bar{\rho}_R(a) \), then \( P_R(\rho|x_t, a) = 0 \) for all \( \rho < x_t \) and \( P_R(\rho|x_t, a) = 1 \) for all \( \rho \geq x_t \). (Right Extremists)

• If \( x_t \in (\underline{\rho}_R(a), \underline{s}_R(a)) \), then \( P_R(\rho|x_t, a) = 0 \) for all \( \rho < \underline{\rho}_R(a) \) and \( P_R(\rho|x_t, a) = 1 \) for all \( \rho \geq \underline{\rho}_R(a) \). (Beliefs when off-equilibrium deviations occur left center.)

• If \( x_t \in (\bar{s}_R(a), \bar{\rho}_R(a)) \), then \( P_R(\rho|x_t, a) = 0 \) for all \( \rho < \bar{\rho}_R(a) \) and \( P_R(\rho|x_t, a) = 1 \) for all \( \rho \geq \bar{\rho}_R(a) \). (Beliefs when off-equilibrium deviations occur right center.)

Suppose that preferences satisfy a single-crossing property so that the median voter is decisive.\(^{17}\) A Republican incumbent will be reelected if the median voter prefers the incumbent to an untested challenger from the Democratic party. That holds, if and only if:

\[
-\mathbb{E} \left( |\rho| \mid \rho \in [\underline{\rho}_R(a), \underline{s}_R(a)] \right) + \lambda a + \beta V^o(0) \geq V^D(0) \tag{8}
\]

Similarly, the median voter prefers the Republican incumbent at \( x_t = \underline{s}_R(a) \) to the challenger from party \( D \) if and only if

\[
-\mathbb{E} \left( |\rho| \mid \rho \in [\bar{s}_R(a), \bar{\rho}_R(a)] \right) + \lambda a + \beta V^o(0) \geq V^D(0) \tag{9}
\]

\(^{17}\)We will discuss below how to verify this assumption.
Equilibrium also requires that the median voter does not prefer a politician that locates at \( x_t = \rho_R(a) \) to the challenger:

\[
-|\rho_R(a)| + \lambda a + \beta V^o(0) \leq V^D(0)
\]

Similarly, the median voter does not prefer a politician that locates at \( x_t = \bar{\rho}_R(a) \) to the challenger:

\[
-|\bar{\rho}_R(a)| + \lambda a + \beta V^o(0) \leq V^D(0)
\]

Here we will focus on equilibria with maximal sincere beliefs which satisfy:

\[
-E \left( |\rho| \mid \rho \in [\rho_R(a), s_R(a)] \right) + \lambda a + \beta V^o(0) = V^D(0)
\]  

\[
-E \left( |\rho| \mid \rho \in [\bar{s}_R(a), \bar{\rho}_R(a)] \right) + \lambda a + \beta V^o(0) = V^D(0)
\]

Equations (5), (6), (11) and (12) then define election standards and cut-off points.

Similarly, we can derive election standards for Democratic incumbents denoted by \( s_D(a) \) and \( \bar{s}_D(a) \), as well as cut-off points \( \rho_D(a) \) and \( \bar{\rho}_D(a) \).

Finally, we need to verify that the median voter is in fact decisive. The value function of voter \( \theta \) for electing an untried Democratic challenger is given by the following expression:

\[
V^D(\theta) = \int_A \int_{-\infty}^{\rho_D(a)} -|\rho - \theta| + \lambda a + \beta V^R(\theta) dF^\rho_D(\rho) dF^a_D(a)
+ \int_A \int_{s_D(a)}^{\rho_D(a)} -|s_D(a) - \theta| + \lambda a + \beta(-|\rho - \theta| + \lambda a) + \beta^2 V^o(\theta) dF^\rho_D(\rho) dF^a_D(a)
+ \int_A \int_{\bar{s}_D(a)}^{\rho_D(a)} (1 + \beta)(-|\rho - \theta| + \lambda a) + \beta^2 V^o(\theta) dF^\rho_D(\rho) dF^a_D(a)
+ \int_A \int_{\bar{s}_D(a)}^{\rho_D(a)} -|s_D(a) - \theta| + \lambda a + \beta(-|\rho - \theta| + \lambda a) + \beta^2 V^o(\theta) dF^\rho_D(\rho) dF^a_D(a)
+ \int_A \int_{\rho_D(a)}^{\infty} -|\rho - \theta| + \lambda a + \beta V^R(\theta) dF^\rho_D(\rho) dF^a_D(a)
+ \int_A \int_{\rho_D(a)}^{\infty} -|\rho - \theta| + \lambda a + \beta V^R(\theta) dF^\rho_D(\rho) dF^a_D(a)
\]

A similar equation holds for \( V^R(\theta) \). Finally, we have:

\[
V^o(\theta) = p_D V^D(\theta) + (1 - p_D)V^R(\theta)
\]
where $p_D$ is the probability that an untried Democrat wins an open election.\footnote{Note that we implicitly assume that an extremist runs for a the second term, but loses the election and is replaced by candidates from the opposing party. Alternatively, we could assume that extremists do not run in the second period and the election is an open election.}

To finish the argument, we need to verify that the value functions are single-peaked in $\theta$. While we do not have a general proof for this result, we can numerically verify these conditions for each specification that we consider in estimation. To illustrate these issues, we plot the key value functions based on our parameter estimates.

First consider the case in which an incumbent is pitted against an untested challenger. Figure 2 plots the value functions associated with a Republican incumbent and an untested challenger from the Democratic party. We consider two interesting cases. The upper panel of Figure 2 plots $V^{I,R}(\theta, \text{lower}, a)$ which refers to the case when the last period policies was given by $x_{t-1} = s_R(a)$. This is the case of a moderate conservative Republican. The lower panel of Figure 2 plots $V^{I,R}(\theta, \text{upper}, a)$, which refers to the case when the last period policy was given by $x_{t-1} = \bar{s}_R(a)$. The incumbent is thus a moderate liberal Republican. By construction the two value functions intersect in both panels of Figure 2 at zero which is the location of the median voter. The median voter is, therefore, indifferent between reelecting the incumbent or electing the challenger. The two plotted value functions only intersect once and, thus, satisfy a single-crossing property. The set of voters who prefers the Republican incumbent is a connected set. As a consequence, the median voter at zero is decisive.

## 3 Identification and Estimation

### 3.1 Identifying the Distributions of Ideology and Competence

The model is based on the assumption that voters observe the policy implemented by the governor, $x$, as well as the competence, $a$. However, we as econometricians do not observe $a$
Figure 2: Value Functions: $V^{I,R}$ and $V^{D}$
or \( x \). Instead we observe a variety of policy outcomes which are imperfect measures of \( x \) and \( a \). We consider three such outcomes in the empirical analysis (expenditures per capita, taxes per capita, and minimum wages) that are expected to be highly correlated with ideology that differentiates fiscal conservatives from fiscal liberals. Ideology is, therefore, implicitly defined as differences in opinion about the preferred size of government. In addition, we observe additional outcomes that not just reflect ideology, but depend on competence. Here we focus on measures such as workers compensations, state income growth, and state debt borrowing cost. We, therefore, implicitly define competence as the managerial ability of the governor to run important economic programs, to lower borrowing costs, and to generate economic growth.

To formalize these ideas we follow Carneiro, Hansen, and Heckman (2003) (CHH) and interpret the observed policies are noisy measures of two latent factors \( x \) and \( a \).\(^{19}\) We specify and estimate a separate system of measurement equations for each party, \( j \in \{D, R\} \). CHH also show that we need, at least, five different measurements to identify a model with two latent factors.\(^{20}\) Moreover, we need to impose an exclusion restriction, that basically implies that a subset of outcomes are only a function of ideology.

For expositional simplicity, suppose we only have five outcomes and assume that the

\[^{19}\text{Differences between observed policies and policies favored by governors may arise due to the complexities of the legislative process or shocks during the implementation process.}\]

\[^{20}\text{More generally, let } L \text{ be the number of measurement equations and } K \text{ be the number of factors. CHH show that identification requires that } \]

\[ L \geq 2K + 1 \tag{16} \]
measurement system is given by

\begin{align*}
    z_1 &= x + \epsilon_1 \\
    z_2 &= \mu_{21} x + \epsilon_2 \\
    z_3 &= \mu_{31} x + a + \epsilon_3 \\
    z_4 &= \mu_{41} x + \mu_{42} a + \epsilon_4 \\
    z_5 &= \mu_{51} x + \mu_{52} a + \epsilon_5
\end{align*}

(17)

where $x$ is ideology and $a$ is competence. Note that the normalizations in the first and third equation are necessary to impose a scaling on the latent factors. The exclusion restrictions require, that we have, at least, two outcomes that are primarily driven by ideology and do not depend on competence. The level of taxes and expenditures are good candidates. Ideology is, therefore, defined as differences in option about the optimal size of government.\textsuperscript{21}

Notice that governors always reveal their true preferences in their last term in office. Thus, we exclude the first period policy choices of governors who won reelection. Using the subpopulation that consists of policies enacted by one term governors and second term policies of reelected governors, we first identify $\mu_{ji}$ using the observed covariances. Following CHH we obtain identification of the factor loadings as follows. Let $j=1,2$ and $l=1,2,3,4,5$, $(j \neq l)$, we then have:

\begin{align*}
    \text{cov}(z_j, z_l) &= \mu_{j1} \mu_{l1} \sigma_x^2 \tag{18}
\end{align*}

In particular

\begin{align*}
    \text{cov}(z_1, z_l) &= \mu_{l1} \sigma_x^2 \tag{19} \\
    \text{cov}(z_2, z_l) &= \mu_{l1} \mu_{21} \sigma_x^2
\end{align*}

Hence we have:

\begin{align*}
    \mu_{21} = \frac{\text{cov}(z_2, z_3)}{\text{cov}(z_1, z_3)} = \frac{\text{cov}(z_2, z_4)}{\text{cov}(z_1, z_4)} \tag{20}
\end{align*}

\textsuperscript{21}In addition, we also use minimum wages in the empirical analysis as a measure of ideology.
and $\text{cov}(z_1, z_2) = \mu_{21}\sigma^2_x$ identifies $\sigma^2_x$. $\text{cov}(z_1, z_3) = \mu_{31}\sigma^2_x$ identifies $\mu_{31}$. $\text{cov}(z_1, z_4) = \mu_{41}\sigma^2_x$ identifies $\mu_{41}$. $\text{cov}(z_1, z_5) = \mu_{51}\sigma^2_x$ identifies $\mu_{51}$.

We then proceed to the next equations and note that for $j=3,4$ and $l=3,4,5$ ($j \neq l$) we have:

$$\text{cov}(z_j, z_l) = \mu_{j1}\mu_{l1}\sigma^2_x + \mu_{j2}\mu_{l2}\sigma^2_a$$

In particular:

$$\text{cov}(z_3, z_5) - \mu_{31}\mu_{51}\sigma^2_x = \mu_{52}\sigma^2_a$$

$$\text{cov}(z_4, z_5) - \mu_{41}\mu_{51}\sigma^2_x = \mu_{42}\mu_{52}\sigma^2_a$$

Hence we have:

$$\mu_{42} = \frac{\text{cov}(z_4, z_5) - \mu_{41}\mu_{51}\sigma^2_x}{\text{cov}(z_3, z_5) - \mu_{31}\mu_{51}\sigma^2_x}$$

and $\text{cov}(z_3, z_4) - \mu_{31}\mu_{41}\sigma^2_x = \mu_{42}\sigma^2_a$ identifies $\sigma^2_a$. $\text{cov}(z_3, z_5) - \mu_{31}\mu_{51}\sigma^2_x = \mu_{52}\sigma^2_a$ identifies $\mu_{52}$.

Following Cunha, Heckman, and Schennach (2010) (CHS) we obtain non-parametric identification of the distributions of $x$ and $a$ by repeatedly applying Kotlarski’s Theorem using a convenient transformation of the measures. In particular, define:

$$\bar{z}_1 = z_1 = x + \epsilon_1 = x + \bar{\epsilon}_1$$

$$\bar{z}_2 = \frac{z_2}{\mu_{21}} = x + \frac{\epsilon_2}{\mu_{21}} = x + \bar{\epsilon}_2$$

Kotlarski’s Theorem then implies that the characteristic functions of $x = \rho$ and $\bar{\epsilon}_i$ are given by:

$$\varphi_\rho(t) = \exp \left( \int^t \frac{\varphi^1_n(0, u)}{\varphi_n(0, u)} du \right)$$

$$\varphi_{\bar{\epsilon}_1}(t) = \frac{\varphi_n(t, 0)}{\varphi_\rho(t)}$$

$$\varphi_{\bar{\epsilon}_2}(t) = \frac{\varphi_n(0, t)}{\varphi_\rho(t)}$$
where $\varphi_n$ is the joint characteristic function of $\tilde{z}_j^1$ and $\tilde{z}_j^2$ for the restricted sample, and $\varphi_n'(0,u)$ denotes the derivative of this function with respect to its first argument. We can then use the standard inversion formula to estimate the densities based on the characteristic functions:

$$f_\rho(x) = \frac{1}{2\pi} \int_{-T}^{T} \exp(-itx)\varphi_\rho(t)dt$$  \hspace{1cm} (29)

$$f_{\bar{\epsilon}_i}(x) = \frac{1}{2\pi} \int_{-T}^{T} \exp(-itx)\varphi_{\bar{\epsilon}_i}(t)dt \hspace{1cm} i = 1, 2$$ \hspace{1cm} (30)

where $T$ is a smoothing parameter. Next define:

$$\tilde{z}_3 = z_3 - \mu_{31}z_1 = a + \epsilon_3 + \mu_{31}\epsilon_1 = a + \bar{\epsilon}_3$$ \hspace{1cm} (31)

$$\tilde{z}_4 = \frac{1}{\mu_{42}} z_4 - \frac{\mu_{41}}{\mu_{42}\mu_{21}}z_2 = a + \frac{1}{\mu_{42}}\epsilon_4 + \frac{\mu_{41}}{\mu_{42}\mu_{21}}\epsilon_2 = a + \bar{\epsilon}_4$$ \hspace{1cm} (32)

Applying Kotlarski’s Theorem on the two transformed measurements above yields the distribution of $a$.

Finally note that we can apply this methodology for each party separately, thus identifying the distributions of ideology and competence for each party.

### 3.2 Identifying the Benefits of Holding Office and the Value of Competence

Next we discuss how to identify the benefits of holding office for governors, $y$, and the benefits associated with electing more able politicians, $\lambda$.\textsuperscript{22} Fix the level of competence at level $a$ and suppose, initially, that the econometrician observes $a$. Equilibrium requires that the

\textsuperscript{22}We can estimate $p_D$ as the share of Democratic winners in open elections. We normalize the discount factor of politicians to be equal to one ($\gamma = 1$) since $\gamma$ and $y$ are not separately identified.
following conditions hold for each level of $a$:

\[
\begin{align*}
-E \left( |\rho| \mid \rho \in [s_D(a) - y, s_D(a)] \right) + \lambda a + \beta V^o(0) &= V^R(0) \\
-E \left( |\rho| \mid \rho \in [\bar{s}_D(a), \bar{s}_D(a) + y] \right) + \lambda a + \beta V^o(0) &= V^R(0) \\
-E \left( |\rho| \mid \rho \in [s_R(a) - y, s_R(a)] \right) + \lambda a + \beta V^o(0) &= V^D(0) \\
-E \left( |\rho| \mid \rho \in [\bar{s}_R(a), \bar{s}_R(a) + y] \right) + \lambda a + \beta V^o(0) &= V^D(0)
\end{align*}
\]

Given $(y, \lambda)$, we find that the election standards $s_j(a)$ and $\bar{s}_j(a)$, $j = R < D$ are uniquely determined by these four equations.\(^{23}\) In addition, the fraction of extremists in each party $j$ conditional on $a$ satisfies:

\[
Pr\{\text{Lose Reelection}|a, j\} - \left( F^p_{\bar{s}}(\bar{s}_j(a) - y) + 1 - F^p_{\bar{s}}(\bar{s}_j(a) + y) \right) = 0
\]

(34)

Note that the first term of the equation above is identified if we observe $a$. The second term is a known function of the two parameters $\lambda$ and $y$. Equation (34), therefore, generates moment conditions for each value of $a$ and each party $j$. $(y, \lambda)$ are identified if these moment conditions satisfy a standard rank condition.

Intuitively speaking, an increase in $y$ (holding $\lambda$ constant) implies a higher willingness to accept the election standards for all incumbents and, as a consequence, leads to a decrease in the probability of electoral defeat for all candidates, independent of their ability. An increase in $\lambda$ (holding $y$ constant) implies a decrease in probability of defeat for high ability politicians and increase in the probability of defeat for low ability politicians since the elections standards become steeper functions of ability. As a consequence, $\lambda$ is identified by the variation of $Pr\{\text{Lose Reelection}|a, j\}$ in $a$, $y$ is identified from the unconditional probability of defeat.

In practice $a$ is not directly observed by the econometrician, instead we observe a vector of measurements $z$ that are correlated with $a$. As a consequence, attention focuses on the

\(^{23}\)Given that this system of equations is non-linear, uniqueness is not guaranteed. We find that it seems to hold for all parametric version of our model that we have explored.
following moment condition:

\[
E[Pr\{\text{Lose Reelection}|a, j]\vert z] - \int Pr\{\text{Lose Reelection}|a, j\} f(a\vert z) \, dz = 0 \quad (35)
\]

Note that the first term in equation (35) is identified and can be consistently estimated for each value of \( z \) based on the observed data. The second term is a known function of \( y, \lambda \). Finally, note that \( f(a\vert z) \) is non-parametrically identified as discussed in Section 3.1. Using a similar argument as above, we conclude that the moment conditions in equation (35) together with the equilibrium restrictions in (33) identify the two parameters of interest.

Finally, we can exploit the variation of the election probabilities by party to identify an extended version of the model, in which the benefits of holding office, \( y_j \), are party-specific.

### 3.3 Identifying the Distribution of Voters’ Ideal Points

The distribution of voters’ ideological ideal points can be identified from the observed vote shares of incumbents that are reelected to a second term. Again, we first consider the case in which \( a \) and \( x_{t-1} \) are observed without error by the econometrician. Consider a fiscally conservative Republican \((x_{t-1} \leq 0)\), and recall that the conditional value function associated with reelecting the incumbent of voter \( \theta \) is given by:

\[
V^{I,R}(\theta, x_{t-1}, a) = -E[|x_t - \theta| \vert x_{t-1}, a] + \lambda a + \beta E[V^o(\theta)] \quad (36)
\]

Let \( V^D(\theta) \) denote the conditional value function of electing an untested Democratic candidate. Given sincere voting, the set of individuals that vote for the conservative Republican incumbent is defined as:

\[
I_R(x_{t-1}, a) = \{ \theta \, \big| \, V^{I,R}(\theta, x_{t-1}, a) \geq V^D(\theta) \} \quad (37)
\]

If preferences satisfy a single-crossing property, this is a connected set.\(^{24}\) Hence, there exists

\(^{24}\)We numerically check this condition during each step of the estimation algorithm and find that it typically holds for our specifications.
a unique cut-off point, denoted by $\theta_R(x_{t-1}, a)$ such that all voters with $\theta \leq \theta_R(x_{t-1}, a)$ will vote for the incumbent while all votes with $\theta > \theta_R(x_{t-1}, a)$ will vote for the challenger.

Figure 3: Vote Shares

Figure 3 illustrates the basic mechanism. We plot the value functions associated with a Democratic challenger and two different Republican incumbents. The set $I_R(x_{t-1}, a)$ is implicitly characterized by the intersection of the incumbent’s and challenger’s value functions. The observed vote share, denoted by $v_t$, then satisfies:

$$F_\theta(\theta_R(x_{t-1}, a)) = v_t$$

(38)

Given that we have identified the parameters of politicians’ preferences and the distribution of politician types in the first two stages, $\theta_R(x_{t-1}, a)$ is known up to the discount factor $\beta$.\textsuperscript{25} We thus conclude that the distribution of voter’s ideal points is point identified at all values that correspond to observed vote shares of incumbents. Values of the distribution for points.

\textsuperscript{25}We set the annual discount factor of voters equal to 0.95.
that are not associated with an incumbent’s vote share can only be bounded or interpolated using semi-nonparametric approximations as discussed in detail below.

In practice, we do not observe the policy, \( x_{t-1} \), and the competence, \( a \). Instead we observe a vector of noisy measures, \( z_{t-1} \) that are correlated with both variables. The expected vote share of the incumbent conditional on observing \( z_{t-1} \) is, therefore, given by:

\[
E[v_t|z_{t-1}] = \int \int F_\theta(\theta_R(x_{t-1}, a)) g_R(x_{t-1}, a|z_{t-1}) dx_{t-1} da
\]  

(39)

where the conditional density \( g_R(x_{t-1}, a|z_{t-1}) \) can be derived from the measurement model in equation (17) and is thus non-parametrically identified. We can use a flexible parametrization for \( F_\theta(\cdot) \) and estimate its parameters by minimizing the squared differences between the observed and the predicted vote shares in equation (39). Moreover, we can use any policy \( z \) to implement this procedure, which then give rise to many potential moment conditions.

\subsection{Semi-parametric Estimation}

The proofs of identification are constructive and can be used to define a moments estimator that is based four different types of orthogonality conditions. First, we use co-variance restrictions to estimate the factor loading coefficients. We can use Kotlarski’s Theorem to estimate the distribution of ideology. To ease the burden of computing equilibria, we approximate the density of \( \rho \) and \( a \) obtained from Kotlarski’s Theorem using semi-nonparametric methods. Following Gallant and Nychka (1987), we define SNP densities as:

\[
f(x) = \left[ \sum_{j=0}^{K} \alpha_j (x - \mu)^j \right]^2 \exp \left[ -\frac{(x - \mu)^2}{\gamma^2} \right], \quad \int f(x) dx = 1
\]

(40)

The second set of moment restrictions arise form the fraction of extremists that are not reelected in the first period conditional on the observed vector of measurements, \( z \). As discussed in the Section 3.2, these moment restrictions together with equilibrium conditions
that define the upper and lower election thresholds can be used to estimate the benefits from holding office and the coefficient on competence.

Third, we can construct orthogonality conditions based on the observed policies of candidates that win reelection. Term limits give rise to policy moderation in the first period relative to the second period. The degree of observed policy moderation depends on the functions and parameters of our model. We solve the model and simulate policy outcomes. We match the moments of the distribution of observed policies for each party for two-term politicians.

Finally, we construct moment conditions by matching the predicted vote shares with the observed vote shares of reelected incumbents as discussed in Section 3.3. By combining orthogonality conditions from these different sources, we obtain an over-identified method of simulated moments estimator. We use bootstrap methods to estimate standard errors for the parameters and functions of interest.

4 Data

Our data set is based on all gubernatorial elections between 1950 and 2012 in the U.S. In constructing our data, we closely follow Besley and Case (1995) (BC) to guarantee that our findings are comparable to their study. We have also replicated and extended their findings using our sample.\footnote{Since our findings are almost identical to the previous set papers, we do not report them in this paper. This analysis is available upon request from the authors.}

The Book of the States provides detail information about gubernatorial term limits. Appendix A summarizes term limits by state during our sample period.\footnote{Note that different states have adopted term limits at different points of time. In addition to the 23 states that constitute our main sample, there are two states limiting governors to serve 8 years out of 12 years in office.} Data on vote
shares, party affiliation, and incumbency status of candidates in gubernatorial elections are based on a web site called www.ourcampaigns.com. Table 1 summarizes the election data.

Table 1: Winners’ Vote Shares, Party Affiliation, and Incumbency Status

<table>
<thead>
<tr>
<th>Variable</th>
<th>Obs</th>
<th>Winner’s Vote Share</th>
<th>Democratic Party</th>
</tr>
</thead>
<tbody>
<tr>
<td>Total</td>
<td>255</td>
<td>0.58</td>
<td>0.53</td>
</tr>
<tr>
<td>Incumbent</td>
<td>103</td>
<td>0.61</td>
<td>0.52</td>
</tr>
<tr>
<td>Challenger against Incumbent</td>
<td>39</td>
<td>0.54</td>
<td>0.62</td>
</tr>
<tr>
<td>Challenger in open election</td>
<td>113</td>
<td>0.57</td>
<td>0.50</td>
</tr>
</tbody>
</table>

Note: The vote share is calculated based on votes cast for the two candidates.

We use the same policy outcome measures as BC. Taxes can be obtained from the state government tax data collected by U.S. Census. We focus on total general sales tax, individual income tax, and corporate net income tax, which account for the vast majority of state tax receipts. Total general expenditures are also obtained from the U.S. Census. The Monthly Labor Review and the Report on the Minimum Wage Commission provide detailed data on minimum wages for each state in the U.S. Finally, data on worker’s compensations can be obtained from the Analysis of Worker’s Compensation Law and the Book of States. Workers compensation is measured as the maximum weekly benefits for temporary total disability. Temporary total disability benefits are paid during the period an employee is unable to work due to the effects of the work-related injury, subject to the waiting period, if applicable.

In addition, we collect a variety of other outcomes that are more closely related to the managerial competence of a governor. Following Alt, Bueno de Mesquita, and Rose (2011), we focus on economic growth and borrowing cost. State total income is from Bureau of Economic Analysis (BEA). Total interest expenditure on debt and total debt are from U.S. Government Census (the same as tax and expenditure).

Table 2 provides some descriptive statistics of our policy outcome measures. All taxes,
Table 2: State Policy and Economic Variables, 1950-2012

<table>
<thead>
<tr>
<th></th>
<th>Mean</th>
<th>Std Dev</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sales tax</td>
<td>367</td>
<td>171</td>
</tr>
<tr>
<td>Income tax</td>
<td>189</td>
<td>186</td>
</tr>
<tr>
<td>Corporate tax</td>
<td>44</td>
<td>36</td>
</tr>
<tr>
<td>State spending</td>
<td>1326</td>
<td>704</td>
</tr>
<tr>
<td>Minimum wage</td>
<td>2.1</td>
<td>1.4</td>
</tr>
<tr>
<td>Maximum weekly benefits</td>
<td>239</td>
<td>108</td>
</tr>
<tr>
<td>Economic growth rate (%)</td>
<td>7.1</td>
<td>4.2</td>
</tr>
<tr>
<td>Borrowing costs (%)</td>
<td>4.9</td>
<td>1.6</td>
</tr>
<tr>
<td>State income</td>
<td>11969</td>
<td>4285</td>
</tr>
<tr>
<td>Population (millions)</td>
<td>4.8</td>
<td>5.1</td>
</tr>
<tr>
<td>Fraction of old (%)</td>
<td>11.1</td>
<td>2.4</td>
</tr>
<tr>
<td>Fraction of young (%)</td>
<td>21.5</td>
<td>3.7</td>
</tr>
<tr>
<td>Governor cannot stand for reelection</td>
<td>0.29</td>
<td>0.46</td>
</tr>
<tr>
<td>Party of governor (=1 if Democrat)</td>
<td>0.55</td>
<td>0.5</td>
</tr>
</tbody>
</table>
income, and expenditure are per capita in 1982 dollars. There exists a fair bit of heterogeneity in policy outcomes across states and time. To account for heterogeneity among states as well as business cycle effects, we regress all policy outcomes on a full set of state and time dummy variables as well as state income, state population, fraction of old, and fraction of young. We then use the time and state adjusted policies when we implement the estimator of our model. This approach is fully consistent with our model if we assume that voters and politicians only care about deviations of policy from state means and time trends.

Before we turn to the estimation results, it is useful to point out that there are two fairly simple tests of our model that do not require us to implement our estimator. The first test only requires us to compute second moments of observed policies. One way to measure policy moderation is to analyze the differences in the standard deviation of policies adopted in the first and second periods restricting attention to a subsample of policies that were enacted by two-term governors. Broadly speaking, our model implies that the observed standard deviation of policies of successful incumbents should be larger in the second period than in the first period.

Table 3: A Policy Moderation Test

<table>
<thead>
<tr>
<th></th>
<th>std deviation</th>
<th>std deviation</th>
<th>One sided Test</th>
<th>p-value</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>1st period</td>
<td>2nd period</td>
<td></td>
<td></td>
</tr>
<tr>
<td>expenditures</td>
<td>94.98</td>
<td>118.20</td>
<td></td>
<td>0.0000</td>
</tr>
<tr>
<td>taxes</td>
<td>57.02</td>
<td>58.31</td>
<td></td>
<td>0.3393</td>
</tr>
<tr>
<td>minimum wage</td>
<td>0.46</td>
<td>0.51</td>
<td></td>
<td>0.0230</td>
</tr>
</tbody>
</table>

Table 3 reports the empirical results for the three outcomes studied that are primarily driven by ideology. We find that the standard deviation of first period policies is smaller than the standard deviation of second period policies for all three outcome measures that are strongly correlated with ideology. Using conventional levels of significance, the difference
is significantly different from zero in two out of three cases. We have also conducted the same analysis for each party. Our qualitative findings are similar once we condition on party membership. The main difference is that we find more pronounced differences in the standard deviations for Republicans than Democrats, especially for tax and minimum wage policies.

Second, our model also suggests that the effect of policy moderation depends on which side of the median voter a politician finds himself. Both types of moderates need to move towards the center to win reelection. A fiscally conservative politician will adopt higher taxes and expenditures in the first period than in the second period. A fiscally liberal politician will adopt lower taxes and expenditures in the first period than in the second period. Note that this is true for both parties. The degree of policy moderation is, however, party specific.

Table 4: Modified Besley-Case Regressions

<table>
<thead>
<tr>
<th>Variables</th>
<th>expenditure</th>
<th></th>
<th>tax</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>ideology</td>
<td>liberal</td>
<td>conservative</td>
<td>liberal</td>
</tr>
<tr>
<td>Democratic incumbent</td>
<td></td>
<td>-39.94</td>
<td>32.75</td>
<td>-3.34</td>
</tr>
<tr>
<td>1st term</td>
<td></td>
<td>(11.37)</td>
<td>(15.05)</td>
<td>(6.06)</td>
</tr>
<tr>
<td>Republican incumbent</td>
<td></td>
<td>-20.10</td>
<td>58.78</td>
<td>-13.20</td>
</tr>
<tr>
<td>1st term</td>
<td></td>
<td>(14.61)</td>
<td>(13.01)</td>
<td>(6.43)</td>
</tr>
<tr>
<td>Governor’s party is Democratic</td>
<td></td>
<td>3.28</td>
<td>-3.78</td>
<td>-16.03</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(13.10)</td>
<td>(14.05)</td>
<td>(6.22)</td>
</tr>
<tr>
<td>Constant</td>
<td></td>
<td>72.37</td>
<td>-82.42</td>
<td>49.90</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(10.29)</td>
<td>(9.13)</td>
<td>(4.52)</td>
</tr>
</tbody>
</table>

This insight then suggests a modified version of the Besley & Case regression. First, we restrict the sample to politicians that win reelection. Second, we split the sample not only based on party affiliation, but also based on an indicator of ideology, which uses the second
period tax or expenditure policies to classify governors as liberal or conservative. Table 4 reports the results of these regression exercises. We find that the modified Besley & Case regressions are supportive of our modeling strategy. Conservatives adopt higher tax and spending policies in the first period while liberals do exactly the opposite regardless of their party affiliation. Our findings thus suggests that the ideology of the candidate may be more important than party membership in explaining outcomes.

5 Empirical Results

We implement our semi-parametric estimator using the six outcome measures discussed above. Table 5 reports the parameter estimates and estimated standard errors.\(^{28}\) Standard errors are computed using a bootstrap algorithm.

We find that the factor loadings are similar for both parties. A formal test suggests that we cannot reject the null hypothesis that the coefficients are not party specific. Imposing equality of the factor loadings is desirable so that ideology is measured on the same scale for both parties. Recall that we normalize the ideology coefficient of taxes and the coefficient of competence in the income growth equation to be equal to one. Given these normalizations, the signs of the all estimated coefficients are correct. Ideology has a large impact on the first three outcomes: taxes, expenditures, and minimum wages. Fiscal conservatives prefer lower taxes, lower expenditures, and lower minimum wages than fiscal liberals. Competence primarily drives the last three outcomes: income growth, workers compensation, and cost of borrowing. Ideology has negligible impacts on thee outcomes. Higher competence leads to higher economic growth, higher workers compensation, and lower costs of financing debt. We view these results as providing strong positive evidence for our exclusion restrictions.

There are also some differences in how strongly Republicans and Democrats value holding

\(^{28}\)Our estimates for the probability that a Democrat will win an open election, denoted by \(p_D\), is 0.5044 which indicates that open election are very competitive.
Table 5: Estimates

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Estimate</th>
<th>S.E.</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Factor Loadings: Ideology</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Taxes</td>
<td>$\mu_{D11}$</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td>$\mu_{R11}$</td>
<td>1</td>
</tr>
<tr>
<td>Expenditures</td>
<td>$\mu_{D21}$</td>
<td>1.6395</td>
</tr>
<tr>
<td></td>
<td>$\mu_{R21}$</td>
<td>1.6466</td>
</tr>
<tr>
<td>Minimum wages</td>
<td>$\mu_{D31}$</td>
<td>0.0045</td>
</tr>
<tr>
<td></td>
<td>$\mu_{R31}$</td>
<td>0.0043</td>
</tr>
<tr>
<td>Economic growth</td>
<td>$\mu_{D41}$</td>
<td>0.0009</td>
</tr>
<tr>
<td></td>
<td>$\mu_{R41}$</td>
<td>0.0010</td>
</tr>
<tr>
<td>Workers comp</td>
<td>$\mu_{D51}$</td>
<td>0.2016</td>
</tr>
<tr>
<td></td>
<td>$\mu_{R51}$</td>
<td>0.2355</td>
</tr>
<tr>
<td>Debt costs</td>
<td>$\mu_{D61}$</td>
<td>0.0013</td>
</tr>
<tr>
<td></td>
<td>$\mu_{R61}$</td>
<td>0.0016</td>
</tr>
<tr>
<td><strong>Factor Loadings: Competence</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Economic growth</td>
<td>$\mu_{D42}$</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td>$\mu_{R42}$</td>
<td>1</td>
</tr>
<tr>
<td>Workers comp</td>
<td>$\mu_{D52}$</td>
<td>0.9517</td>
</tr>
<tr>
<td></td>
<td>$\mu_{R52}$</td>
<td>0.8787</td>
</tr>
<tr>
<td>Debt costs</td>
<td>$\mu_{D62}$</td>
<td>-0.0727</td>
</tr>
<tr>
<td></td>
<td>$\mu_{R62}$</td>
<td>-0.0763</td>
</tr>
<tr>
<td>Benefits of holding office</td>
<td>$y_D$</td>
<td>0.5254</td>
</tr>
<tr>
<td></td>
<td>$y_R$</td>
<td>0.7348</td>
</tr>
<tr>
<td>Benefits associated with Competence</td>
<td>$\lambda$</td>
<td>0.2033</td>
</tr>
</tbody>
</table>
office with Republicans having stronger preferences for being in office. Our findings suggest that the benefits from holding office are significant and large in economic magnitude. As a consequence, the prospects of reelection provide strong incentives for moderate governors to move towards the center of the ideological spectrum during their first term in office. Based on the model estimates we can then predict the fraction of moderate and centrist politicians for each party. We find that 35.2 (21.9) percent of all Republicans (Democrats) are moderates while 31.6 (38.1) percent are centrists. These estimates suggest that a large fraction of candidates engage in policy moderation during the first term.

Voters of all ideological types strongly value competence as indicated by the economically large value of $\lambda$. We, therefore, conclude that there exists a significant trade-off between ideology and competence.

Figure 4 plots SNP and Kotlarski estimates of the distribution of ideological positions and competence by party. There are clearly significant differences in the ideological positions of Republican and Democratic candidates for governor. As expected, the mean of the Democratic distribution is significantly larger than the mean of the Republican distribution, implying that, on average, Democrats prefer higher taxes and expenditures than Republicans. However, there is also much overlap in the relevant support of both distributions. Candidates with negative values can be viewed as “fiscal conservatives,” while candidates with positive values are ” fiscal liberals.” In addition, there is much heterogeneity in competence. However, we do not find any significant differences among parties.

Finally, we implement the third stage of our estimation strategy by matching the predicted voted shares with the observed vote shares of incumbents using a NLLS estimator. Figure 5 plots the estimated densities of ideological positions of voters and candidates. Figure 5 shows that both distributions are similar. Citizen-candidate models assume that these two distributions are identical. Our results provide some support for this assumption. If anything, political candidates tend to be less radical than voters. This finding is plausible since the distribution of politicians is restricted to potential candidates. These candidates...
Figure 4: The Distribution of Ideological Positions and competence by Party
must be viable and are thus screened carefully by parties. Moreover, they typically have to survive an internal primary process to win the party nomination. This process may help to eliminate candidates that are considered to be too extreme in their ideological positions.

6 Policy Implications

We study the impact of of term limits on electoral outcomes and economic policies. The natural benchmark is a model without term limits. We, therefore, solve our model with and without term limits and compare the predictions. To ease the comparison between the two models we simulate a large number of draws from the underlying distribution of politician types. We then compare the two models ignoring the outcomes in the model without term limits that occur after the second period. Table 6 summarizes the impact of eliminating term limits on electoral outcomes.
Table 6: Policy Implications

<table>
<thead>
<tr>
<th></th>
<th>2 Term Limit</th>
<th>No Term Limit</th>
</tr>
</thead>
<tbody>
<tr>
<td>Welfare</td>
<td>-10.30</td>
<td>-9.72</td>
</tr>
<tr>
<td>Competence</td>
<td>-0.025</td>
<td>-0.023</td>
</tr>
<tr>
<td>Fraction Dem. Centrists</td>
<td>0.381</td>
<td>0.362</td>
</tr>
<tr>
<td>Fraction Dem. Moderates</td>
<td>0.219</td>
<td>0.259</td>
</tr>
<tr>
<td>Fraction Rep. Centrists</td>
<td>0.315</td>
<td>0.358</td>
</tr>
<tr>
<td>Fraction Rep. Moderates</td>
<td>0.352</td>
<td>0.334</td>
</tr>
</tbody>
</table>

Policy: Mean

<table>
<thead>
<tr>
<th></th>
<th>2 Term Limit</th>
<th>No Term Limit</th>
</tr>
</thead>
<tbody>
<tr>
<td>Tax</td>
<td>-2.70</td>
<td>-2.79</td>
</tr>
<tr>
<td>Expenditure</td>
<td>-0.034</td>
<td>-0.133</td>
</tr>
<tr>
<td>Minimum Wage</td>
<td>-0.012</td>
<td>-0.013</td>
</tr>
<tr>
<td>Compensation</td>
<td>-0.662</td>
<td>-0.050</td>
</tr>
<tr>
<td>Income Growth</td>
<td>0.014</td>
<td>-0.056</td>
</tr>
<tr>
<td>Borrowing Cost</td>
<td>-0.028</td>
<td>-0.043</td>
</tr>
</tbody>
</table>

Policy: Standard Deviation

<table>
<thead>
<tr>
<th></th>
<th>2 Term Limit</th>
<th>No Term Limit</th>
</tr>
</thead>
<tbody>
<tr>
<td>Tax</td>
<td>78.30</td>
<td>71.22</td>
</tr>
<tr>
<td>Expenditure</td>
<td>127.26</td>
<td>115.80</td>
</tr>
<tr>
<td>Minimum Wage</td>
<td>0.513</td>
<td>0.492</td>
</tr>
<tr>
<td>Compensation</td>
<td>83.25</td>
<td>83.40</td>
</tr>
<tr>
<td>Income Growth</td>
<td>2.88</td>
<td>2.88</td>
</tr>
<tr>
<td>Borrowing Cost</td>
<td>1.06</td>
<td>1.06</td>
</tr>
</tbody>
</table>
We find that election standards are similar across party in a model without term limits. This convergence arises because moderates do not deviate from the election standards in a model without term limits. Thus, differences in the attractiveness of holding office across parties are not important in determining election standards. Table 6 suggests that the fraction of Democratic centrists is reduced by almost 2 percentage points while the fraction of moderates increase by almost 4 percentage points. For Republicans, our model predicts an increase in the fraction of centrists that is partially off-set by a decrease in the fraction of moderates.

Eliminating term limits also has an impact on economic policies, levels of competence, overall welfare. Table 6 also reports policy outcomes and welfare under both regimes. Not surprisingly, we find that the average differences in mean policies is small and not economically significant. We also find that the average level of competence of governors is higher in a model without term limit.

The differences in welfare are largely driven by a reduction in policy variance. Term limits have a significant impact on the variance of policies. Comparing the standard deviations of the economic policies, that are primarily driven by ideology, we find large differences between both models. The standard deviation is almost 10 percent smaller in a model without term limits. Policies that are largely a function of ability are not strongly affected by term limits. This decrease in the variance of policies then implies an increase in overall welfare since voters are risk-averse.

7 Conclusions

Building on recent theoretical advances in modeling repeated elections, we have shown how to identify and estimate dynamic models of electoral competition. Candidates are drawn from different distributions of ideological positions and competence, which captures one of the key
trade-offs faced by voters. The empirical analysis draws on data from U.S. gubernatorial elections between 1950 and 2012. Our analysis provides a new empirical evaluation of the impact of term limits on political competition and policy outcomes in the U.S. Our findings suggest that voters are willing to accept significant trade-offs in ideology to obtain a more competent governor. There are significant benefits associated with holding office. As a consequence, the prospects of reelection provide strong incentives for moderate governors to move towards the center of the ideological spectrum during the first term in office. We also find evidence that the distribution of voter ideal points is similar to the distribution of political candidates providing support for citizen-candidate type models. Moreover, term limits also imply larger variation in observed economic policies such as taxes, expenditures, and minimum wages. Since voters are risk-averse, economic welfare is lower in a model with term limits. We view the methods developed in this paper for estimating dynamic games of electoral competition as promising for future research.
References


choices? Evidence from gubernatorial term limits,” The Quarterly Journal of Economics,
110(3), 769–798.

BESLEY, T., AND A. CASE (2003): “Political institutions and policy choices: evidence from
the United States,” Journal of Economic Literature, pp. 7–73.


Effects with an Application to the Returns to Schooling and Measurement of the

COATE, S., AND M. CONLIN (2004): “A Group Rule-Utilitarian Approach to Voter Turnout:

Economics, 92, 582–96.


from California’s state legislative races.,” Public Choice, 90, 165–84.


A Online Appendix: States with Term Limits

Table 7: Term Limitations by State in 2013

<table>
<thead>
<tr>
<th>State law</th>
<th>State</th>
</tr>
</thead>
<tbody>
<tr>
<td>States with no term limits</td>
<td>CT, IA, ID, IL, MA, MN, ND, NH, NY, TX, UT, VT, WA, WI</td>
</tr>
<tr>
<td>States limiting governors to 1 term in office</td>
<td>VA</td>
</tr>
<tr>
<td>States limiting governors to 8 out of 12 years in office</td>
<td>IN(1973), OR(1987)&lt;sup&gt;c&lt;/sup&gt;</td>
</tr>
<tr>
<td>States limiting governors to 8 out of 16 years in office</td>
<td>MT(1993), WY(1993)</td>
</tr>
</tbody>
</table>

The source is the book of the states. Note that parenthesis shows the year of change if it was after 1950. We have considered 48 constitutional states. a) NC adopted 2 lifetime term limit from 1977 to 1992. b) NM adopted 2 consecutive term limit prior to 1971 and adopted 1 term limit from 1971 to 1990. c) OR adopts 2 consecutive term limit before 1987. OK adopts 2 consecutive term limit from 1967 to 2010.