USC FBE APPLIED ECONOMICS WORKSHOP
presented by: Mattias Polborn
Friday, Feb. 20, 2015
1:30 pm - 3:00 pm; Room: **HOH-706**

Political Competition in Legislative Elections*

Stefan Krasa† Mattias Polborn‡

February 16, 2015

Abstract

We develop a theory of candidate nomination processes predicated upon the notion that members of the majority party in a legislature collaboratively influence policy. Because of this team aspect, voters care both about their local candidates' positions, and the positions of their parties which are determined by the positions of all their elected representatives in the legislature. We show that candidates may be unable to escape the burden of their party association, and that the primary voters in both parties exploit the median voters' national preferences to nominate the most extreme electable candidates. We also show that gerrymandering affects the equilibrium platforms not only in those districts that become more extreme, but also in those that ideologically do not change.

Keywords: Differentiated candidates, primaries, polarization.

*Both authors gratefully acknowledge financial support from National Science Foundation Grant SES-1261016. Any opinions, findings, and conclusions or recommendations expressed in this paper are those of the authors and do not necessarily reflect the views of the National Science Foundation or any other organization. We are thankful for helpful comments by seminar participants at ITAM, Toulouse, Illinois, Princeton, Western Illinois, Iowa, Michigan, Notre Dame and UC Irvine.

†Department of Economics, University of Illinois, 1407 W. Gregory Dr., Urbana, IL, 61801. E-mail: skrasa@uiuc.edu

‡Department of Economics and Department of Political Science, University of Illinois, 1407 W. Gregory Dr., Urbana, IL, 61801. E-mail: polborn@uiuc.edu.

1 Introduction

In the most basic model of representative democracy, voters elect legislative representatives whose positions reflect the preferences of their respective districts' median voters. These representatives convene in an amorphous assembly — one in which there are no parties, or parties at least do not play an important role —, and national policy is set, in equilibrium, to correspond to the preferences of the median representative in this assembly. Thus, in this basic model, the legislature is composed of representatives who are very moderate relative to the voters who elect them, and actual policy and legislation reflects the most moderate position in this assembly of moderates. Suffice it to say that few observers of Congress believe that reality corresponds closely to these predictions. The central question is why this is the case.

In this paper, we propose a model of electoral competition that can account for a much higher degree of polarization in the legislature, and which is based on two realistic ingredients: First, the majority party in a legislature is an important power center influencing the crafting of policy. Coordination of decision-making and voting according to the majority preferences in the majority party increases the influence of each majority party legislator on the policy outcome (Eguia, 2011a,b). Second, legislative candidates are nominated by policy-motivated primary voters who take both the general election and the legislation process into account when deciding whom to nominate.

The importance of parties is uncontroversial among scholars of legislatures. However, there is surprisingly little analysis of how the fact that each candidate is connected to a party and thus, implicitly, to the positions of candidates of that party from other districts influences the types of candidates who are nominated by their party to run for legislative office and the outcomes of elections in different legislative districts.

If one were to apply the simplest Downsian model naively to Congressional elections – which much of the empirical literature implicitly does – then it generates counterfactual predictions: In each district, both candidates should adopt the preferred position of the district median voter, and so, policy-wise, all voters should be indifferent between the Democratic candidate and his Republican opponent. Republicans in New England or Democrats in rural districts should have a

substantial chance to be elected to Congress if only they match their opponent's policy platform.[1] Furthermore, in this model framework, gerrymandering districts would not help parties, at least not in terms of increasing the party's expected representation in Congress. It is safe to say that both of these predictions are counterfactual, but understanding why that is so is challenging.

In our model, voters care about both their local candidates' positions and the national positions of their respective parties. These national positions are determined by a rather general function that maps the ideal positions of all of party's legislators into a policy, and satisfies some basic intuitive requirements such as efficiency and monotonicity. In this framework, there are spillovers between different districts: The electoral prospects of candidates in a given district are influenced by the ideological positions of their parties' winning candidates elsewhere. The association with a party that is not attuned with a district's ideological leanings may be poisonous for a candidate even if his own policy positions are tailor-made for his district.

Consider, for example, Lincoln Chafee, the U.S. senator from Rhode Island from 1999 to 2006. In spite of being a Republican, Chafee had taken a number of moderate and liberal positions that brought him in line with voters in his state.[2] In the 2006 election, "exit polls gave Senator Lincoln Chafee, a popular moderate Republican from a long-admired political family, a 62 percent approval rating. But before they exited the polls, most voters rejected him, many feeling it was more important to give the Democrats a chance at controlling the Senate. [...] 'I'm caught between the state party, which I'm very comfortable in, and the national party, which I'm not,' said Mr. Chafee."[3] His Democratic challenger Whitehouse "succeeded by attacking the instances in which Chafee supported his party's conservative congressional leadership (whose personalities and policies were very unpopular, state-wide)."[4]

In a review of 2006 campaign ads, factcheck.org summarized: "President Bush was far and away the most frequent supporting actor in Democratic ads [...] The strategy is clear: whether

[1] See Table 1 in Winer et al. (2014) for evidence that a significant share of U.S. Senate elections are non-competitive. In 29.4 percent of U.S. Senate elections between 1922 and 2004 without an incumbent running, the winner received a vote share that was at least 20 percentage larger than the loser's vote share.

[2] For example, Chafee was pro-choice, anti-death-penalty, supported gay marriage and voted against the Iraq war (see `http://en.wikipedia.org/wiki/Lincoln_Chafee`).

[3] "A GOP Breed loses its place in New England", New York Times, November 27, 2006.

[4] See `http://en.wikipedia.org/wiki/Lincoln_Chafee`.

they're referring to a Republican candidate as a 'supporter' of the 'Bush agenda' or as a 'rubber-stamp,' Democrats believe the President's low approval ratings are a stone they can use to sink their opponents [...] Democratic Sen. Hillary Clinton of New York got the most mentions in Republican ads holding forth the supposed horrors of a Democratic-controlled Senate [...] The runner-up is 'San Francisco Liberal Nancy Pelosi,' who is mentioned in at least 6 GOP ads as a reason not to vote for a Democrat who would in turn vote to make her Speaker of the House."[5]

We show that "contamination" – as we call this spillover effect – makes most legislative elections uncompetitive and results in an equilibrium in which party members are able to nominate their ideal candidate, rather than the ideal candidate of the district median voter, and nevertheless win by a healthy margin. The other party either cannot effectively compete because, even if it nominates a candidate at the ideal position of the median district voter, that voter still prefers the more extreme competitor because he is associated with an average party position that is ideologically preferred by the district median voter.

In contrast to the classical one-district spatial model, the ideological composition of districts in our model does not only influence the ideological position of elected candidates, but also the chances of parties to win. Thus, partisan incentives for gerrymandering are much larger in our model. We also show that gerrymandering or, more generally, the intensification of the median ideological preferences in some districts, also affects the political equilibrium in those districts where the median voter preferences remain moderate. Thus, our results imply that testing for the causal effect of gerrymandering on polarization in Congress is more complicated than the existing literature has recognized.

Our paper proceeds as follows. Section 2 reviews the related literature. In Section 3, we provide some stylized facts about statewide executive and legislative elections, and explain why they are hard to explain within the standard model that looks at legislative elections in different districts in isolation. In Section 4, we set up the general model, and the main analysis follows in Sections 5 and 6 . We conclude in Section 7.

[5] See http://www.factcheck.org/elections-2006/our_2006_awards.html

2 Related literature

Ever since the seminal work of Downs (1957), the position choice of candidates and the determinants of policy convergence or divergence are central topics in political economy models of elections. While the classical median voter framework identifies reasons for equilibrium platform convergence, there is a large number of subsequent variations of the spatial model of electoral competition that develop different reasons for policy divergence, such as policy motivation (e.g., Wittman 1983; Calvert 1985; Martinelli 2001; Gul and Pesendorfer 2009); entry deterrence (e.g., Palfrey 1984; Callander 2005); and incomplete information among voters or candidates (e.g. Castanheira 2003; Callander 2008; Bernhardt et al. 2009).

Overwhelmingly, the existing literature looks at isolated elections – usually, two candidates compete against each other, and voters care only about their positions. In the probabilistic voting model (e.g., Hinich 1978; Lindbeck and Weibull 1987; Dixit and Londregan 1995; Banks and Duggan 2005), voters also receive "ideological" payoffs that are independent of the candidates' positions. While, to the best of our knowledge, these authors do not interpret the ideological payoff as capturing the effects of the candidate being affiliated with a party, and therefore implicitly the party's other legislators' policy positions, this is a possible interpretation. However, the "ideology shock" in these models is exogenous, so that the main point of interest in our model – how does the fact that a party's national position matters for voters and is determined as an aggregate of all its representatives' positions, affect both the voters' choice between local candidates and the candidates' equilibrium positions? – cannot be analyzed in these models.

Our model belongs to the class of differentiated candidates models (Soubeyran 2009; Krasa and Polborn 2010a,b, 2012, 2014; Camara 2012). In these models, candidates have some fixed "characteristics" and choose "positions" in order to maximize their probability of winning. Voters care about outcomes derived from a combination of characteristics and positions. In contrast to existing differentiated candidates models, voters' preferences over characteristics (i.e., the candidates' party affiliations) are endogenously derived from the positions of Democrats and Republicans in other districts.

Erikson and Romero (1990) and Adams and Merrill (2003) introduce an influential model framework in the political science literature in which voters receive, in addition to the payoff from the elected candidate's position, a "partisan" payoff from the candidate's party affiliation. However, this partisan payoff is exogenous and orthogonal to the policy positions chosen by the candidates. Our contribution to this literature is to provide a microfoundation for these partisan payoffs, to show how they depend on the equilibrium polarization between the parties' candidates in other districts, and how they in turn affect the candidates' equilibrium positions.

The legislative part of our model assumes that parties in Congress have a strong influence on policy outcomes. A significant number of models explain why parties matter. Conditional party government theory (Rohde, 2010; Aldrich, 1995) and endogenous party government theory (Volden and Bergman, 2006; Patty, 2008) argue that party leaders can use incentives and resources to ensure cohesiveness of their party. Procedural cartel theory (Cox and McCubbins, 2005) argues that party leadership can at least enforce voting discipline over procedural issues, and Diermeier and Vlaicu (2011) provide a theory where legislators endogenously choose procedures and institutions that lead to powerful parties. All these models of the importance of parties in Congress take the preference distribution of legislators as exogenously given, while our model provides for an electoral model and thus endogenizes the types of elected legislators.

Since we assume that the nomination decision is made by a policy-motivated party median voter, our model is related to the literature on policy-motivated candidates pioneered by Wittman (1983) and Calvert (1985), who assume that *candidates* are the ones who are policy-motivated and get to choose the platform that they run on. In our model, the effective choice of platform is made by the primary election median voter,[6] but this change does not substantively affect the analysis. This approach is also taken by Coleman (1972) and Owen and Grofman (2006). To our knowledge, no paper in this literature analyzes policy-motivated policy selectors in the type of "linked" elections in different districts that we focus on.

Our results are relevant for the large empirical literature that analyzes how primaries, the ideo-

[6]Implicitly, we assume that either candidates can commit to an ideological position in the primary, or that candidates are citizen-candidates with an ideal position that is common knowledge.

logical composition of districts and especially the partisan gerrymandering of districts affects the ideological positions of representatives in Congress (e.g., Lee et al. 2004; McCarty et al. 2009; Hirano et al. 2010). Most empirical papers in this literature do not include a formal model from which they derive predictions about the "expected" correlations, but rather take the intuition from the isolated election model and simply transfer them to the setting of legislative elections. For example, there is a general expectation in the empirical literature that the positions of district representatives, i.e. U.S. Senators or House members, measured by their DW-Nominate score, "should" reflect the conservativeness of their districts. Our model shows that this transfer of results derived in the isolated-election model to legislative elections is not always justified, and that the candidates' equilibrium positions may correspond to the preferences of the parties' respective primary electorates rather than those of the district median voter.

3 Consistent lopsided elections: A puzzle for the single-district model

In this section, we argue that the influence of the electorate's preference distribution on the parties' performance is substantially larger in legislative elections than in executive ones. This stylized fact is puzzling when viewed through the lens of the simplistic one-district spatial model which does not distinguish between executive and legislative elections. As we show, one can interpret our model as a resolution of this puzzle.

3.1 Some stylized facts

The simplest Downsian model predicts that both candidates in a plurality rule election choose their position at the median voter's ideal point, so that all voters are indifferent between the candidates. A rather liberal or conservative district should not provide a particular advantage – in terms of the probability of winning the district – to Democrats or Republicans. In Section 3.2, we look at somewhat more sophisticated one-district models of candidate competition, but argue that this

intuition is quite robust.

In practice, it is well known that the ideological preferences of voters do affect the electoral chances of the different parties' candidates – we talk of "deep red" (or blue) states, implying that the candidates of the ideologically favored party have a much clearer path to victory than their opposition.

However, we now argue that the voters' ideological preferences have a substantially larger effect in legislative elections than in executive ones. To demonstrate this phenomenon, we consider Gubernatorial and U.S. Senate elections from 1978 to 2012. While both of these types of contests are high-profile, state-wide races, Gubernatorial elections are for executive positions while Senate elections are for legislative ones. Consistent with the empirical literature, we measure the median state ideology by its Partisan Voting Index (PVI), which is calculated as the difference of the state's average Democratic and Republican Party's vote share in the past two U.S. Presidential elections, relative to the nation's average share of the same.[7]

The dependent variable is the difference between the Democrat's and the Republican's vote share of the two party vote in a particular election. In addition to the main independent variables of interest (*PVI* and *PVI*×Senate election), we use incumbency dummies and year fixed effects in order to control for the electoral advantage of incumbents, and for election-cycle national shocks in favor of one party.

Table 1 summarizes the results, with the first column as the baseline case (all years since 1978, all states). For Gubernatorial elections (the omitted category), the PVI coefficient indicates that a one point increase in the Democratic vote share in Presidential elections increases the Democratic gubernatorial candidate's vote share only by about 0.519 points. In contrast, in Senate elections, the same ideological shift increases the Democratic Senate candidate's vote share by $0.519 + 0.645 = 1.164$ points, more than twice the effect in Gubernatorial elections; evidently,

[7]For example, if, in a particular state, Democratic presidential candidates run ahead of Republicans by 7 percent (on average in the last two elections), while nationally, Democratic candidates win by 3 percent (in the same two elections), then the state has a PVI of 7% – 3% = 0.04. Also note that vote shares are calculated relative to the two-party vote, i.e., votes for minor parties are eliminated before the vote share percentages are calculated. See `http://cookpolitical.com/application/writable/uploads/2012_PVI_by_District.pdf` for the PVI based on the 2004 and 2008 Presidential elections.

the difference between executive and legislative elections is substantial and highly significant. The remaining three columns confirm the qualitative robustness of this difference if we restrict to elections after 1990 and if we exclude the political South.[8]

Table 1: Senate and Gubernatorial elections

	All States		Without Confederacy States	
	1978-2012	1990-2012	1978-2012	1990-2012
PVI	0.519***	0.589***	0.529***	0.614***
	(0.111)	(0.124)	(0.117)	(0.132)
PVI× Senate	0.645***	0.596***	0.597***	0.514***
	(0.149)	(0.167)	(0.156)	(0.177)
N	1103	702	871	553
R^2	0.551	0.595	0.571	0.62

*** indicates significance at the 1% level.
Additional explanatory variables used: Election type (Senate or Governor), year dummies, and incumbency status.
Data Source: Congressional Quarterly, `http://library.cqpress.com/elections/`

A coefficient of about 1 for Senate elections is quite remarkable — if Senate candidates were hard-wired at their Presidential party position, irrespective of whether such a position is competitive in their respective state, then this should result in a coefficient of (about) 1. Any degree of willingness of the disadvantaged candidate to adjust his position to better fit the state's voter preferences should reduce the advantage of the opponent, and thus the estimated coefficient. Somehow, only gubernatorial candidates appear (at least to some extent) capable of such a position adjustment, while Senate candidates are not.

3.2 Inconsistency with the simple single-district model

These stylized facts are difficult to reconcile with the standard model of political competition that implicitly assumes that the electoral competition between the two candidates in each district is

[8] At least until the 1990s, there were a lot of conservative Southern Democrats in state politics in the South, so it is useful to check that our results are not just driven by this region of the country.

not influenced by what happens outside the district. Specifically, it is very difficult to set up a one-district model in which a particular party wins almost certainly, and does so with a substantial winning margin.

Without loss of generality, let the median voter be located at zero and the party medians at m_D and m_R. Even if party medians are far apart from each other, parties have to nominate relatively moderate candidates in order to remain competitive. This is obvious for the model without uncertainty where both parties nominate candidates that maximize the median voter's utility, i.e., $x_D = x_R = 0$, and both parties have equal vote shares, even if one party's ideal point is substantially closer to the median voter's ideal position than the opposition's. Since this is true for arbitrary ideal positions of the parties, it implies that, even if party members become more extreme, i.e., m_D moves to the left and m_R to the right, the equilibrium policies remain moderate, and the margin of victory is close to zero (if the distribution of voter types is continuous).

Thus, if party members become more extreme then the Downsian model predicts at most a small effect on policy: Party members continue to nominate moderate candidates, and both parties receive approximately one-half of the votes. In contrast, it is a widespread view that the rise of activist and more ideological party members has resulted in more extreme candidates being nominated for office (e.g., Fiorina et al. (2006)). Further, many political commentators and scholars diagnosed a rise in polarization between the two parties. In order to generate such polarization in a standard model with policy-motivation, uncertainty, e.g., about the median voter's location, must increase. In other words, we would need that the quality of political polls deteriorates over time, which is somewhat implausible.

The prediction that both candidates *in executive elections* (i.e., those where the elected candidate can set policy without being tied to their party) will be competitive is borne out in U.S. presidential elections. For example, between 1988 and 2012 the difference between the Republican and Democratic vote share in Presidential elections was between -5.6% and 7.7%, with a median of -0.5%.[9] Furthermore, the results of Table 1 above indicate that Gubernatorial elections,

[9]U.S. presidents are elected in many districts through the electoral college system rather than by a majority of the popular vote. However, to the extent that state ideological leanings are fixed and known, the objective for the parties' primary electorates is essentially to nominate a candidate who can win in the decisive swing state.

even in ideologically skewed states, are at least considerably more competitive than the Presidential election in those same districts. In contrast, as shown above, many legislative elections result in one party receiving a substantially higher vote share than its opposition.

Can we generate lopsided outcomes if one of the candidates has a "valence" advantage? Suppose that the net-valence of the Republican candidate, $v_R - v_D$ is $\varepsilon > 0$. Then, in equilibrium, $x_D = 0$ and $x_R = \sqrt{\varepsilon}$. Given these positions, the median voter at 0 is again indifferent: If he votes for D the utility is 0, if he votes for R the utility is $-x_R^2 + \varepsilon = 0$, but in equilibrium he supports the Republican who wins the election. If voter types are continuously distributed, then the margin of victory is (almost) zero.

In order to generate a vote margin that is bounded away from 0, one would have to assume a valence advantage that is so large that the median voter prefers the favored candidate even if he is located at his party median's ideal point, and the opposition candidate is located at the median voter's ideal point. Usually, valence is interpreted as a small personal preference; for rational voters, most of the utility-relevant payoff from a legislature should come from the laws the legislature enacts, rather than from legislators' valences. Furthermore, it would be hard to understand why one party should be consistently much better than the other party in terms of the quality of candidates that they select, and why that party should necessarily be the one that is ideologically closer to the median voter.

It is easy to show that, if there is some uncertainty about the median, then the higher valence candidate wins with probability close to one, but the margin of victory is close to zero. However, if the uncertainty is not too large, then the winning margin is close to zero. The reason is that in any Downsian model without uncertainty, the median voter is indifferent between the candidates, while with some uncertainty he is close to indifferent, and hence the electorate splits close to 50-50. One of the key insights of this paper is that this is not longer true in a multi-district setting.

4 Model

We consider a polity divided into N districts. In each district $i \in I = \{1, \ldots, N\}$, both candidates who run for district i's seat in the legislature are attached to one of two parties ("Democrats" and "Republicans"), respectively. Let $x_{i,D}$ and $x_{i,R}$ denote the positions of the candidates competing in district i.

Voter utility. The utility of a voter with ideal position θ from a candidate with individual position x and party position $\overline{x}$ is given by

$$(1 - \gamma)v_\theta(\overline{x}) + \gamma v_\theta(x), \tag{1}$$

where $\gamma \in [0, 1]$. Here, the first term is the utility from the legislature's policy, and the second term is the utility from the policy position of a district's representative. Note that $\gamma = 1$ corresponds to the standard case where voters only care about their local candidates' positions, and $\gamma = 0$ means that voters care only about national party positions and not their own representative's position per se. We assume that $v_\theta(x)$ is jointly continuous in θ and x, as well as single-peaked with a bliss point at θ, i.e., $v_\theta(x)$ is strictly increasing in x for $x < \theta$ and strictly decreasing for $x > \theta$.

The position of the median voter in district i is denoted by M_i. From an ex-ante perspective, M_i may be uncertain, in which case we assume that the position of the median voter in district i would be given by a cdf $\Phi_i(M_i)$. For the main part of the paper we assume that there is no uncertainty.

The determination of party positions. The parties' policy positions are determined by the policy positions of the winning candidates of each party, formally described by a policy selection function ξ. Let $K \subset I$ denote the set of winning candidates of a party and $x \in \mathbb{R}^N$ the positions of all candidates (winning and losing). Then the party's policy is given by $\xi_K(x) \in \mathbb{R}$. We assume that only elected representatives can influence the party's policy. Formally, let $\overline{x}, \overline{x}' \in \mathbb{R}^N$ be two policy vectors that are identical for elected candidates, i.e., $\overline{x}_i = \overline{x}'_i$ for all $i \in K \subset I$. Then the party's policy is identical with $\overline{x}$ and $\overline{x}'$, i.e., $\xi_K(\overline{x}) = \xi_K(\overline{x}')$. In Section 5, we will make some very natural assumption about the policy selection functions, and derive some further properties.

Timeline. The game proceeds as follows:

Stage 1 In each district, the local members of each party simultaneously select their candidates, who are then committed to their policies $x_{i,D}, x_{i,R} \in \mathbb{R}$. We assume that the nomination process can be summarized by the preference parameter of a "decisive primary voter," whom we can think of as the median party member in the district, and whose ideal positions are denoted by $m_{i,D}$ and $m_{i,R}$ for Democrats and Republicans, respectively.

Stage 2 In each district i, the median voter M_i is realized, if there is uncertainty. The median observes the candidate positions $x_{i,D}$ and $x_{i,R}$ in his own district, and chooses whom to vote for. Note that, for the other districts, the median voter does not observe the candidates' positions, but he has expectations that are correct in equilibrium.

Stage 3 The elected candidates from all districts form the legislature, which determines the national party positions via the functions ξ_K, and payoffs are realized; no strategic decisions take place in this stage.

Equilibrium concept. In the case without uncertainty about the district median voters' preferred positions, we consider subgame-perfect equilibria in pure strategies. That is, district median voters vote for the local candidate whom they prefer, given their (correct) expectation of the positions of candidates and voting outcomes in other districts, and the decisive party primary voters in each district nominate a candidate who maximizes their respective expected utility, given the equilibrium behavior of general election median voters.

In the case that the preferred position of median voters is uncertain, the equilibrium concept is essentially modified to perfect Bayesian equilibrium. A formal definition is in the Appendix.

5 Properties of the policy selection function

In this section, we define a class of policy selection functions that satisfy some rather weak assumptions. Our results, derived in the next section, hold whenever the policy selection functions

satisfy these conditions. First, we assume that each ξ_K is continuous.

Continuity: ξ_K is continuous on $\mathbb{R}^N$ for every $K \subset I$.

Furthermore, the policy choice function is *consistent* in the following sense: First, if there is only one representative, then this representative determines the party's policy in the legislature. Second, consider a situation in which, if a party's set of elected candidates is K, then $\xi_K(x) = y$. Now suppose that an additional legislator from the same party is elected, whose position is exactly y. Then the party's national position remains at y.

Consistency:
1. If only one member of a party is elected then the party's position is that representative's position, i.e., $\xi_{\{i\}}(x) = x_i$.
2. Let $\xi_K(x) = y$, and $i \notin K$ such that $x_i = y$. Then $\xi_{K\cup\{i\}}(x_{-i}, y) = y$.

Finally, suppose that the set of representatives is constant, but their positions shift to the right from x to $x + \Delta$, where $\Delta > 0$ is a vector of shifts to each representative's position. Then we assume that the party's policy cannot move by more than $\max_{i\in K} \Delta_i$ to the right. We call this property *stability*.

Stability: $\xi_K(x) \leq \xi_K(x + \Delta) \leq \xi_K(x) + \max_{i\in K} \Delta_i$ for all $\Delta \geq 0$.

We show in Lemma 1 that an analogous condition for stability with respect to position shifts to the left holds automatically if the stability condition above is satisfied.

Note that the approach used to describe how a party's policy is determined from the ideal points of its elected representatives is very general. For example, it is possible for specific legislators to have more influence on the aggregate party position than others (i.e., we do not need to assume that ξ_K is symmetric with respect to elected representatives' positions), and it allows for each legislator's influence to depend on the make-up of the caucus in which he serves. In principle, we could allow for the policy selection function to also depend on the party (i.e., to allow for $\xi_K^D \neq \xi_K^R$). All our results would go through with this party-specific policy-selection function, but we prefer to keep the notation a bit simpler.

Also note that it is necessary to define $\xi_\emptyset$, the policy selection function when K is the empty set, in order to analyze situations where one party wins in all districts. The first assumption of stability guarantees reasonable off-equilibrium path beliefs for the case that a member of a party without elected representatives is elected.

There are many examples of policy selection functions that satisfy continuity, consistency and stability. We provide three examples here.

Median: $\xi_K(x)$ is the median position of the elected candidates, i.e. $\xi_K(x) = \text{median}\{x_i | i \in K\}$ (and $\xi_\emptyset(x)$ is arbitrary.)

Weighted Average: $\xi_K(x)$ is a weighted average of the elected candidates' positions, i.e.

$$\xi_K(x) = \frac{\sum_{i \in K} \alpha_i x_i}{\sum_{i \in K} \alpha_i},$$

where $\alpha_i \geq 0$. Again, $\xi_\emptyset(x)$ is arbitrary.

Sequential Dictatorship: For example, the elected candidate in the "lowest" district determines the policy, i.e., if $i = \min K$, then $\xi_K(x) = x_i$ — again $\xi_\emptyset(x)$ is arbitrary.

The first selection function, the median of elected representatives, is, for example, used in Snyder (1994) and Ansolabehere et al. (2012). One can interpret this function as a formalization of what has become known as the "Hastert Rule" or "majority of the majority rule" which stipulates that the Speaker of the House should admit legislative proposals to the floor of the House if and only if they have the support of the majority of representatives in the majority party, and thus provides a powerful position to the median of the majority party.

In contrast to the median, where marginally only changes of the median legislator have an effect on the party position, marginal changes in each legislator's position matter for the resulting policy under the weighted average function (as long as the weight on the legislator is non-zero). This function also provides an easy way to model some legislators (say, those with higher levels of seniority) as more powerful than others.

Taking this idea of more and less powerful legislators to the limit results in the sequential dictatorship function which builds on an exogenous ranking of legislators. Only the most powerful legislator (among those elected) matters for the party's position.[10]

It is also true that any convex combination of these policy selection functions also satisfy the three conditions. The following lemmas list some properties implied by the above conditions. Lemma 1 first shows that consistency and stability imply Pareto efficiency, and that the election of another representative cannot move the party position by more than the difference between the previous position and the new representative's position. Second, stability implies that ξ is weakly monotone. If ξ is differentiable, then all derivatives are non-negative, and they sum up to not more than 1. Finally, stability (which was defined only for policy shifts to the right) implies that an analogous condition is satisfied for policy shifts to the left.

Lemma 1

1. *Suppose that ξ_K satisfies consistency, and stability. Then*

   (a) *$\xi_K(x)$ is Pareto efficient, i.e.,* $\min_{i\in K} x_i \leq \xi_K(x) \leq \max_{i\in K} x_i$.

   (b) *Suppose an additional representative i is elected. Then the new policy is not more extreme than i's policy. That is, let $i \notin K$. If $x_i \geq (\leq)\xi_K(x)$ then $x_i \geq (\leq)\xi_{K\cup\{i\}}(x)$.*

2. *If ξ_k is stable then ξ_k is weakly monotone, i.e., $\xi_K(x) \leq \xi_K(y)$ for all $x \leq y$.*

3. *Suppose that ξ_K is continuously differentiable. Then stability implies: (i) $\frac{\partial \xi_K(x)}{\partial x_i} \geq 0$ for every $i \in K$, and (ii) $\sum_{i\in K} \frac{\partial \xi_K(x)}{\partial x_i} \leq 1$. Conversely, suppose that ξ_K is differentiable and (i) and (ii) hold then ξ_K satisfies stability.*

4. *ξ_K satisfies stability if and only if $\xi_K(x) \geq \xi_K(x-\Delta) \geq \xi_K(x) - \min_{i\in K} \Delta_i$ for all $\Delta \geq 0$.*

All proofs are relegated to the Appendix. Lemma 2 shows that the set of policy selection functions that satisfy the three conditions is closed and convex. For example, a function where the position is

[10]The sequential dictatorship function is a limit in the following sense: Let the weight factor vector be given by $\alpha = (w^{2N+1}, w^{2N}, \ldots, w)$. As $w \to \infty$, the weight of the lowest elected representative relative to all other elected representatives goes to infinity.

a weighted average between the median and the average position also satisfies the three conditions. Lemma 2 thus allows for a rich set of policy selection functions in which there are some particularly influential legislators, but where also the position of the median caucus member matters.

Lemma 2 *The set of all policy selection functions that satisfy continuity, consistency, and stability is closed and convex with respect to the supremum norm.*

6 Equilibrium Analysis

6.1 Convergence and divergence

We now turn to the analysis of the equilibrium of the game. Our first result considers a setting in which all districts are homogeneous (or, at least, each has a median voter with the same policy preferences). In this case, there exists a unique pure strategy equilibrium in which all candidates locate at the preferred position of the median voter.

Proposition 1 *Suppose that the median voters in all districts are located at 0. Furthermore, the decisive Democratic primary voters are located* $m_{i,D} < 0$ *and the decisive Republican primary voters are at* $m_{i,R} > 0$*. Then*

1. *There exists a pure strategy equilibrium in which* $x_{D,i} = x_{R,i} = 0$ *for all districts.*
2. $x_{D,i} = x_{R,i} = 0$ *in all equilibria in which the local primary voters use pure strategies.*

The intuition for this result is straightforward and follows the same centripetal logic underlying the classical Downsian model: If the candidates in all other districts converge, the median voter in any given district has no party preference, and will therefore choose the local candidate who is closest to his ideal position. This, in turn, forces the two local primary voters into a Bertrand-like competition for the support of the median voter.

The result provides a useful baseline against which to interpret the results in the following propositions where, with heterogeneous districts, equilibrium policies diverge. Proposition 1

shows that increased polarization of party primary electorates per se does not lead to policy divergence, as long as general election voters remain moderate.

As a first step towards divergence, Proposition 2 shows that there is no equilibrium in which the candidates' proposed policies are the same in all districts if not all median voters have the same ideal point.

Proposition 2 *Suppose there exist two districts i, j in which the median voters' position differs (i.e., $M_i \neq M_j$) and primary voters are more extreme in both districts than the median voters (i.e., $m_{i,D}, m_{D,j} < M_i, M_j < m_{i,R}, m_{j,R}$). Then there does not exist an equilibrium in which policies are the same across districts.*

It is easy to understand why this has to be the case: If all candidates were to locate at the same position $\bar{x}$, efficiency of ξ implies that both national party positions are $\bar{x}$. Since median voters' ideal positions are different, there must be at least one district median voter who is located strictly to the left or to the right of $\bar{x}$. In this district, both the median voter in the general election and one of the primary voters prefer a candidate located at the median voter's ideal point, which implies that the original strategy profile was not an equilibrium.

Proposition 2 indicates that policy positions differ across districts. However, it could be possible that national policies converge. For a very general setting, this question is hard to analyze, so Proposition 3 analyzes policy divergence in a setting where there are only two types of districts, left-leaning ones with a median voter located at $-\bar{M}$, and right-leaning ones with a median voter located at $\bar{M}$. We show that, in equilibrium, Democrats and Republicans win all left- and right-leaning districts, respectively. Because all players in all districts of the same preference type take the same action in equilibrium, all winning Democratic candidates have the same position, which is therefore the national Democratic position (irrespective of the particular functional form of ξ, and analogously for the Republicans.

Proposition 3 *Suppose that all $m_{D,i} = -\bar{m} < 0$ and $m_{R,i} = \bar{m}$. Further, suppose that the district median voters are located at $M_i = -\bar{M} < 0$ for $i \leq k$ and $M_i = \bar{M}$ for $i > k$. Then in equilibrium*

Democrats win in districts $i \leq k$ with policies $x_{i,D} = \bar{x}_D < M_i$. Republicans win in districts $i > k$ with policies $x_{i,R} = \bar{x}_R > M_i$.

Importantly, Proposition 3 shows that the position of winning Democratic candidates is to the left of the median voter in left-meaning districts, and the position of winning Republican candidates is to the right of the median voter in right-leaning districts. Intuitively, since parties represent the preferences of ideologically different districts, district median voters have a strict preference for one of the parties, and that party therefore has a certain leeway in choosing a candidate that is more extreme than the median voter would like, but is still electable.

The extent of this leeway is the greater, the more voters care about the national party position (i.e., the lower is γ), because of two effects: First, for given national positions, an increase in $1 - \gamma$ implies that the difference between the two parties' national positions weighs more heavily for the district median voter, and he is therefore willing to tolerate a more extreme position from his preferred party's local candidate. Second, the direct effect just described implies that in *all* districts candidates move to more extreme positions, and therefore the national party positions adjust accordingly. Since the median voter is closer to his preferred party's position, with any concave utility function, the median voter's utility difference between his favorite party and the other party increases, and this again implies that he is now willing to tolerate an even more extreme position from his favorite party's candidate. This self-reinforcing reaction may be so strong that it only stops when the primary voters' ideal points are reached.

To illustrate Proposition 3, suppose that voter utility is given by $-(1-\gamma)(\theta - \bar{x})^2 - \gamma(\theta - x_i)^2$. In equilibrium, the non-favored candidate locates at the median voter's ideal point because a deviation from this position would allow the favored party to nominate an even more extreme candidate. The median voters in left-leaning and right-leaning districts are indifferent between both candidates if

$$-(-\bar{M} - \bar{x}_D)^2 = -(1-\gamma)(-\bar{M} - \bar{x}_R)^2, \text{ and } -(\bar{M} - \bar{x}_R)^2 = -(1-\gamma)(\bar{M} - \bar{x}_D)^2. \quad (2)$$

Note that if $\bar{x}_R = -\bar{x}_D$ then the first equation in (2) holds if and only if the second equation holds. Solving the second equation for $\bar{x}_R$, and using the result from Proposition 3 that $\bar{x}_R \geq \bar{M}$ gives

$$\bar{x}_R = \frac{2(1+\sqrt{1-\gamma})-\gamma}{\gamma}\bar{M}.^{11} \tag{3}$$

Note that the fraction in (3) takes the value 1 for $\gamma = 1$, and is strictly decreasing in γ, so is greater than 1 for all $\gamma < 1$. Thus, as the national party position becomes more important for voters relative to the position of their local candidate, equilibrium policies diverge.

Equation (3) provides the most extreme position that Republicans can adopt in conservative districts and still be elected. As long as this position satisfies $\bar{x}_R \leq \bar{m}$, this is the position that is adopted by Republicans in equilibrium. In this case, the median voter in these right-leaning districts is indifferent between voting for the Democrat and the Republican, but, in a pure strategy equilibrium, he votes for the Republican.[12] The election thus ends with the Republican just barely beating the Democrat in right-leaning districts, and vice versa in left-leaning districts.

If, instead, $\bar{x}_R \geq \bar{m}$, then the Republican primary voter can simply nominate a candidate who proposes the primary voter's ideal point $\bar{m}$. In this case, the district median voter strictly prefers to vote for the Republican candidate (since he is more moderate than $\bar{x}_R$, which would make the median voter indifferent) to voting for the Democratic candidate. In this case, the Republican candidate wins a strict supermajority in the right-leaning districts. The same applies, analogously, in left-leaning districts.

In summary, in equilibrium we have

$$\bar{x}_R = -\bar{x}_D = \min\left\{\frac{2(1+\sqrt{1-\gamma})-\gamma}{\gamma}\bar{M}, \bar{m}\right\} \tag{4}$$

Equation (4) is useful for discussing the effects of a radicalization of general election voters and of primary voters on the equilibrium level of policy divergence. If we start from a situation in which

[11] The second root of the second equation has $\bar{x}_R \leq \bar{M}$ and therefore does not correspond to an equilibrium.

[12] If the median voter were to mix if he is indifferent, then it would be optimal for Republicans to nominate a slightly more moderate candidate than the value given by (3).

the median voter is indifferent between the candidates (i.e., the first term in (4)), then policies become more extreme if the district median voters become more extreme. This could happen as a consequence of more successful gerrymandering which packs more like-minded people into the same district. Note that, because the fraction in (4) is greater than 1 for all $\gamma < 1$, any change in equilibrium positions is larger than the initial shift in the median voters' preferences. Furthermore, changes in the primary voters' extremism $\bar{m}$ have no effect on the candidates' equilibrium position in this case.

In contrast, if the initial polarization level is sufficiently high such that, in equilibrium, Democrats win left-leaning districts and Republican win right-leaning districts *with a strict supermajority*, then this indicates that the favored party's primary voter is essentially unconstrained and can pick his favorite candidate. In this situation, changes in the position of the primary voters move the winning candidate's position one-to-one, while changes of the district median voter's preferred position have no effect on the candidates' positions.

Instead of measuring policy divergence as the distance between the two national party positions, we could also focus on the difference between the local candidates' positions, $x_{R,i} - x_{D,i}$. Remember that the losing candidate is always located at the district median voter's ideal position. In the case that the median voter is indifferent between the candidates (i.e., the first term in (4)), an increase in $\bar{M}$ increases local policy divergence because the favored party's position reacts more strongly than the opposition party's position. If, instead, the favored party's primary voter is unconstrained, then an increase in $\bar{M}$ only affects the opposition candidate's position, and so local policy divergence decreases. An increase in the party primary voter's extremism has either no effect (in case that the primary voter is constrained), or leads to increased local policy divergence (in the second case).

We now discuss some empirical implications of our results so far.

Do primaries cause polarization? This prediction of the model is consistent with the empirical result of Hirano et al. (2010) who analyze the introduction of direct primaries for Congressional elections in seven U.S. states between 1938 and 1976. This reform replaces a nominating body

composed of party bosses – who are conceivably office-motivated and thus inclined to moderate – with one that might be more policy motivated. However, this reform does not significantly change the extremism of representatives (measured by the deviation of a representative's DW-Nominate score from the average score of his party). Similarly, the level of participation in primaries, another plausible proxy for the primary electorate's extremism, has no significant effect on representatives' extremism either.

Note that the institutional reform analyzed by Hirano et al. (2010) takes place in an era in which Congressional districts are very non-gerrymandered and where the two national parties' positions on many issues are close to each other — in fact, in 1950, the American Political Science Association, in a committee report entitled "Toward a more responsible two-party system," bemoaned that parties did not stand for clearly distinctive political points of view. In such a non-polarized environment, Proposition 1 tells us that a reform resulting in more extreme primary voters does not have a large effect on the candidates' equilibrium positions. We will see below that this changes when district preferences are very different. There is a complementarity between party extremism and district heterogeneity.

Extremism in the legislature. The standard intuition for the moderating effects of political competition is based on a naive application of the single district model. Propositions 2 and 3 show that this standard intuition does not carry over to our model of political competition in legislative elections. Rather than an assembly of district median voters from all over the country, our model predicts a legislature in which representatives from almost all districts are more extreme than their median voter. Extremism is not caused by extremism of primary voters per se (see Proposition 1), but differences between the locations of the median voters in different districts enable the primary voters of the favored party to nominate candidates who are closer to their own ideal points and win nevertheless.

In our model, we should not expect that the legislature necessarily adopts policies with broad popular support, as long as they are unpopular with their own party base. For example, an October 7, 2013 Washington Post opinion poll showed that registered voters disapproved of the Republican

party shutting down the government by 71 to 26. Thus, it is very likely that the median voters in most districts – including those held by Republican House members – opposed shutting down the government, but among voters who identify as Republicans, there was a 52-45 majority in favor of the shutdown, and it is likely that, among those who actually vote in Republican primaries, there was an even larger majority in favor of the shutdown.

According to media reports, many Republican representatives "would have liked" to end the government shutdown much sooner, but were afraid that taking this position publicly would put them at risk in their district primary. For example, former House Speaker Dennis Hastert said in an October 7, 2013 interview with NPR: "It used to be they're looking over their shoulders to see who their general [election] opponent is. Now they're looking over their shoulders to see who their primary opponent is."

Our model shows that this fear is justified: Primary voters who refuse to renominate a "moderate" and replace him with a more extremist candidate are not irrational – because even the extremist can win. This makes the primary threat so credible.[13]

6.2 A closer look at inter-district spillovers and gerrymandering

In Proposition 3 we consider a situation in which all districts are either left or right-leaning. If there are also some centrist districts, to what extent are they affected by what happens in more extreme districts?

To answer this question, we assume that, if a majority of legislator have the same policy position x, then the party's position is also x. An example of such a policy selection function ξ is the function that selects the median Democratic and Republican representative's position as the

[13] Alternative explanations for non-median policy outcomes include lobbying and differential preference intensity. Arguably, neither of these alternative explanations is plausible for the government shutdown.

The lobbying explanation (a strong lobby in favor of the minority position is able to "buy" the support of legislators) requires that benefits on the minority side are highly concentrated, while the issue is a relatively minor issue for most voters. The intensity explanation requires that minority voters care more about the issue, but according to the same Washington Post poll cited above, 12 percent of registered voters "strongly approve" of the shutdown, 14 percent "approved somewhat," while 53 percent "strongly disapprove" and 18 percent "disapprove somewhat." Thus, intensity about the issue appears higher among those who disapprove.

respective party position.

Consider first a situation in which there are many centrist and few extremist districts. In this case, a natural conjecture is that the full convergence equilibrium of Proposition 1 is qualitatively robust in the centrist districts: The median legislator in both parties comes from a centrist district, and the position in these districts is the national party position. Thus, party positions do not differ, so that no party has an advantage, even in the "extreme" districts. While it is possible that all left-leaning districts elect Democrats and all right-leaning districts elect Republicans, this does not have to be the case, and even if it is, then the fact that there is no difference between national party positions means that the winning candidates in the extreme districts cannot be more extreme than those districts' median voters.

In contrast, consider a situation in which there are now "many" extreme districts and only "few" moderate districts. Specifically, assume that the median voters in extreme districts are located at $M_i = -\bar{M}_e < 0$ when $i \leq k$ and $M_i = \bar{M}_e > 0$ when $i > N - k$ (so, there are k left-leaning and k right-leaning districts). In the "centrist" districts, the district median voter is located at $M_i = \bar{M}_c$.

In equilibrium, the representatives from the extreme districts all have the same ideal policy, and since they outnumber the representatives from moderate districts, our assumption on ξ implies that party policies are completely determined by their respective position. For extreme districts, the strategic situation is the same as in Proposition 3, so that equilibrium platforms in the extreme districts $\bar{x}_D$ and $\bar{x}_R$ are given by (4).

If $\bar{M}_c = 0$, i.e., the median voters in moderate districts are located equidistantly from the national Democratic and Republican positions, then the equilibrium policies in these moderate districts n are $x_{n,D} = x_{n,R} = 0$. Now suppose instead that median voter in the centrist district is slightly conservative, i.e., $\bar{M}_c > 0$, so that the district median voter prefers the Republican national position. Thus, in equilibrium the local Democrat locates at $\bar{M}_c$ while the Republican can choose a position to the right of $\bar{M}_c$ that satisfies

$$-(1-\gamma)(\bar{M}_c - \bar{x}_R)^2 - \gamma(\bar{M}_c - x_{n,R})^2 = -(1-\gamma)(\bar{M}_c - \bar{x}_D)^2. \tag{5}$$

Using the fact that, in equilibrium, $\bar{x}_D = -\bar{x}_R$ it follows that

$$x_{n,R} = \bar{M}_c + 2\sqrt{\bar{x}_R \bar{M}_c}\sqrt{\frac{1-\gamma}{\gamma}} \tag{6}$$

Polarization in moderate districts, i.e., the difference between the local candidates positions $x_{n,R} - x_{n,D}$ is equal to the second term in (6). This is increasing in the extent to which the moderate district median voters prefer one party over the other (i.e., $\bar{M}_c$), and also increases as the party positions determined by the extreme districts become more extreme (i.e., $\bar{x}_R \uparrow$).

These spill-over results have important empirical implications.

Implications for the effects of gerrymandering. In our model, positions in a particular district do not just depend on the ideal positions of local (primary and general election) voters, but are also affected by what happens in other districts. Thus, gerrymandering can affect equilibrium positions not just in those districts where a gerrymander changes the make-up of the electorate, but also in all other districts. Specifically, a radicalization of the electorate in gerrymandered districts will also lead to increased polarization in those districts that were not gerrymandered.

McCarty et al. (2009) argue that, while Congress has become more polarized in a time during which electoral districts became more heterogeneous due to gerrymandering, this is merely a temporal coincidence. Their core argument is as follows: There are some districts in which no gerrymandering can occur (e.g., in small states that have only one or two House districts). Local polarization also increases in these districts, which they argue cannot be caused by "gerrymandering." They then difference out the extent of polarization in these districts and argue that only the remaining excess polarization in these other districts can be considered as caused by gerrymandering. Our model highlights the logical flaw in this argument: In the language of randomized controlled trials, gerrymandering is not a treatment that only affects the treatment group.

Formally, consider a treatment that directly changes the extremism of representatives from

liberal and conservative districts. Differentiating (6), we have

$$\frac{\partial x_{n,R}}{\partial \bar{x}_R} = \sqrt{\frac{\bar{M}_c}{\bar{x}_R}} \sqrt{\frac{1-\gamma}{\gamma}} \tag{7}$$

This is certainly positive – polarization spills over to "untreated" districts – and may even be larger than 1, meaning that the "treatment effect" is larger in "untreated" districts than in "treated" ones. Clearly, attempting to difference out the treatment effect is particularly problematic in this case.

Vote pattern. The standard spatial model in which voters look at their district's candidate positions in isolation cannot explain why there are "safe" districts that are essentially guaranteed to be won by one party's candidate, such as rural districts for the Republicans or inner cities for Democrats. In contrast, in our model, median voters in most districts have a strict preference for one of the two parties.

In equilibrium, the winning margin in an interval of districts close to the national median is close to zero, and independent of the specific district's exact ideological preferences. In contrast, in the more extreme districts where parties are unconstrained and can nominate their respective ideal candidates, the favored party will have a higher vote margin. This is exactly the pattern that Winer et al. (2014) find empirically for U.S. Senate elections between 1922 and 2004: For a range of moderate states (i.e., a range of states with a PVI sufficiently close to zero), the estimated marginal effect of a state's PVI on the vote difference between Democrats and Republicans is close to zero, while that same marginal effect is much larger for states that are outside this range.

7 Conclusion

Much of the existing literature on electoral competition in legislative elections implicitly assumes that voters evaluate their local candidates based on their positions, but not on the party label under which they run. Such a model implies that both parties nominate candidates who are very close to the preferences of the respective district median voters. Therefore, even in districts with rather ex-

treme preferences, both parties' candidates should be competitive, and the position of Democratic and Republican Congressmen elected from similar districts should be very similar. It is safe to say that these predictions are not borne out in reality, and to understand why this is the case is of first-order importance for our understanding of the American democratic system.

In this paper, we have developed a theory of candidate nomination processes predicated upon the notion that majority party legislators collaboratively influence policy. This assumption is appears reasonable and yields fundamentally different results.

In our model, a candidate's association with candidates of the same party that run in other districts generates an incentive for voters to focus less on the candidates' own position positions when deciding whom to vote for — local candidates are "contaminated" by their party association. This leads to less competitive local elections, providing the ideologically favored party with the leeway to nominate more extreme candidates who are nevertheless elected. As a consequence, the equilibrium of our model can explain how electoral competition can beget a very polarized legislature.

Our analysis has two additional important empirical implications. First, it can explain why a district's ideological preferences have a smaller partisan effect in elections in which a candidate has a more autonomous policy influence, such as elections for executive leadership positions than in legislative elections. Of course, in reality, even executive leader positions are not entirely autonomous, so there will be some contamination in executive elections as well, but we would expect this effect to be smaller than in legislative elections, and this expectation is borne out in our empirical analysis of Senate and Gubernatorial elections in Section 3.

Second, much of the existing empirical analysis of the effects of gerrymandering on polarization in Congress is implicitly based on applying a naive model in which voters care only about the local candidates' positions. Such a model may lead to incorrect inferences about the importance of gerrymandering. For example, the ideal position of the district median voter often does not affect the equilibrium position of candidates at the margin in our model, but the total effect of gerrymandering on polarization in Congress may nevertheless be substantial, and actually may be much larger than in the naive model. Thus, one cannot infer that gerrymandering does not matter

for polarization in Congress from showing that there is no marginal effect of changes in district medians on ideological positions of legislators, and that the difference in voting records of Republicans and Democrats representing the same or very similar districts has increased. In general, an implication of our model for empirical work is that legislator behavior in different districts is intricately connected rather than independent, and this implies that one needs to be very careful with claims that difference-in-difference approaches can identify causation.

8 Appendix

8.1 Proofs for the case of no uncertainty

Proof of Lemma 1. *Item 1:* Follows immediately from stability. In particular, let $y \geq x$ and define $\Delta = y - x$. Then stability implies $\xi_K(y) = \xi_K(x + \Delta) \geq \xi_K(x) + \min_{i \in K} \Delta_i = \xi_K(x)$.

Item 2: Let $z = \xi_K(x)$ and let $x' = (x_{-i}, z)$. Then consistency implies that $\xi_{K\cup\{i\}}(x') = \xi_K(x)$. Now suppose that $x_i \geq \xi_K(x)$. Let $y \in \mathbb{R}^N$ such that $y_j = 0$ for all $j \neq i$ and $y_i = x_i - z$. Since $z = \xi_K(x)$ it follows that $y_i \geq 0$. Thus, stability implies that $\xi_{K\cup\{i\}}(x') \leq \xi_{K\cup\{i\}}(x'+y) \leq \xi_{K\cup\{i\}}(x)+y_i = z + x_i - z = x_i$. Thus, the statement of item 2 follows.

Item 3: Let $\delta > 0$ and let e_i be the i'th unit vector, i.e., all entries $j \neq i$ of e_i are zero, and entry i is 1. Then stability implies

$$\frac{\xi_K(x + \delta e_i) - \xi_K(x)}{\delta} \geq 0.$$

Taking the limit for $\delta \to 0$ implies that $\frac{\partial \xi_K(x)}{\partial x_i} \geq 0$.

Let $e_K = \sum_{i \in K} e_i$. Then stability and the fact that policy is only determined by representatives $i \in I$ implies that

$$\xi_K(x + \delta e_K) - \xi_K(x) = \xi_K(x + \delta e) - \xi_K(x) \leq \delta.$$

Let $K_1 \subset K_2 \subset \ldots \subset K_m = K$ such that $|K_i| = i$. Define $e_{K_0} = 0$. Then

$$\sum_{i=1}^{m} \Big(\xi_K(x + \delta e_{K_{i-1}} + \delta e_i) - \xi_K(x + \delta e_{K_{i-1}})\Big) = \xi_K(x + \delta e_K) - \xi_K(x) \leq \delta.$$

Dividing both sides by δ and taking the limit for $\delta \to 0$ immediately implies that $\sum_{i \in K} \frac{\partial \xi_K(x)}{\partial x_i} \leq 1$.

To prove the converse let $\Delta \in \mathbb{R}^n$. Because ξ_k is differentiable there exists y such that

$$\xi_K(x + \Delta) = \xi_K(x) + \sum_{i=1}^{N} \frac{\partial \xi_K(z)}{\partial x_i} \Delta_i.$$

Since only elected representatives $i \in K$ determine the policy we get $\frac{\partial \xi_K(z)}{\partial x_i} = 0$ for $i \notin K$. Using the

fact that all derivatives are non-negative and add up to 1 we therefore get

$$\xi_K(x+\Delta) = \xi_K(x) + \sum_{i\in K} \frac{\partial \xi_K(z)}{\partial x_i}\Delta_i \leq \xi_K(x) + \max_{i\in K}\Delta_i \sum_{i\in K} \frac{\partial \xi_K(z)}{\partial x_i} \leq \xi_K(x) + \max_{i\in K}\Delta_i.$$

Similarly, it follows that

$$\xi_K(x+\Delta) = \xi_K(x) + \sum_{i\in K} \frac{\partial \xi_K(z)}{\partial x_i}\Delta_i \geq \xi_K(x) + \min_{i\in K}\Delta_i \sum_{i\in K} \frac{\partial \xi_K(z)}{\partial x_i} \geq \xi_K(x) + \min_{i\in K}\Delta_i.$$

Item 4: We proceed by way of induction over the number of representatives. If $|K| = 1$ consistency implies $\xi_K(x) = x$ and hence Pareto efficiency. Now suppose we have shown the result for a party of size $n-1$. Let K be a party of size n, and denote by x_i, $i = 1, \ldots, n$ the positions of the n elected members. Let $k = \arg\min\{x_i | i \leq n\}$. Let $x = (x_i)_{i\leq n}$ and $x_{-k} = (x_1, \ldots, x_{k-1}, x_{k+1}, x_n)$. induction hypothesis implies that $\xi_{K\setminus\{k\}}(x_{-k}) \geq \min\{x_i | i \neq k\}$. Thus, $\xi_{K\setminus\{k\}}(x_{-k})(x_{-k}) \geq x_k$. Item 2 therefore implies $\xi_K(x) \geq x_k$. Hence $\xi_K(x) \geq \min_i x_i$. The argument that $\xi_n(x) \leq \max x_i$ is analogous.

■

Proof of Lemma 2. It is immediate that ξ_K satisfies continuity and stability. It remains to prove consistency. If $|K| = 1$ then consistency implies $\xi_K^n(x) = x$ for every $n \in \mathbb{N}$. Taking the limit for $n \to \infty$ implies $\xi_K(x) = x$.

Now suppose that $|K| > 1$. Consistency implies $\xi_{K^n\cup\{i\}}(x_{-i}, \xi_K^n(x)) = \xi_K^n(x)$. Uniform converge implies that in the limit $\xi_{K\cup\{i\}}(x_{-i}, \xi_K(x)) = \xi_K(x)$. Thus, consistency is satisfied for ξ_K. ■

Proof of Proposition 1. It is immediate that $x_{D,i} = x_{R,i} = 0$ is an equilibrium, since efficiency of ξ implies that the party positions are also at the median. Thus, any district primary voter who deviates to a candidate not located at zero loses in the general election.

Let $\bar{x}_D$ and $\bar{x}_R$ denote the parties' policies in equilibrium.[14] First, suppose by way of contradiction that $\bar{x}_R < 0$. Consider a district i in which the Republican wins with policy $x_{R,i}$.

[14]Note that if one party wins all the races, then stability implies that if a member of the other party deviates, the party's policy is the same as his policy. All arguments in the proof below remain valid for this special case.

The median voter must be indifferent between the two candidates, else the Republican primary voter can move $x_{R,i}$ marginally to the right. By continuity of ξ this would also only move $\bar{x}_R$ by a small amount, and hence the median voter in district i would still prefer the Republican (note that this deviation is not observable in districts $i \neq j$ and therefore does not change who wins or loses in these districts). By monotonicity of ξ, $\bar{x}_R$ will be moved to the right, and hence the Republican primary voter is strictly better off from this deviation, a contradiction.

If $x_{R,i} < 0$, then both the median voter and the median Republican voter are made strictly better off by moving $x_{R,i}$ marginally to the right. By monotonicity and continuity, this also moves $\bar{x}_R$ marginally to the right, i.e., toward the median voter. Hence to have an equilibrium $x_{R,i} \geq 0$ for all districts in which the Republican wins. The efficiency property of ξ therefore implies that $\bar{x}_R \geq 0$. Similarly, it follows that $\bar{x}_D \leq 0$.

Assume, by way of contradiction, that $\bar{x}_D \leq 0 \leq \bar{x}_R$ with $\bar{x}_D \neq \bar{x}_R$. There are five distinct possibilities:

Case 1: $\bar{x}_R = -\bar{x}_D > 0$. Lemma 1 implies that there exists a district i where $x_{D,i} \leq \bar{x}_D$ and the Democrat wins (so it must be true that, in equilibrium, $x_{R,i} \geq \bar{x}_R$). But then, the Republican primary voter could deviate to $x_{R,i} = \bar{x}_R - \varepsilon$, which would win and increase his utility.

Case 2: $\bar{x}_R = 0$ and $\bar{x}_D < 0$. Again, consider a district i where $x_{D,i} \leq \bar{x}_D$ and the Democrat wins. The Republican primary voter could deviate to $x_{R,i} = 0$, which would win and increase his utility.

Case 3: $\bar{x}_R > 0$ and $\bar{x}_D < 0$, with $|\bar{x}_D| > \bar{x}_R$. Again, consider a district i where $x_{D,i} \leq \bar{x}_D$ and the Democrat wins. The Republican primary voter could deviate to $x_{R,i} = \bar{x}_R$, which would win and increase his utility.

The remaining two cases — $\bar{x}_D = 0$ and $\bar{x}_R > 0$, and $\bar{x}_R > 0$ and $\bar{x}_D < 0$, with $|\bar{x}_D| < \bar{x}_R$ — are analogous to cases 2 and 3. Thus, $\bar{x}_D = \bar{x}_R = 0$.

Finally, suppose that $x_{D,i} \neq x_{R,i}$ in a district i. If the median voter in district i strictly preferred the winner, say the Republican candidate, then we must have $x_{R,i} = m_{i,R}$ (otherwise, the Republican primary voter could increase $x_{i,R}$, and the Republican would still win). Consider the effect of a

deviation of the Democratic primary voter to $\tilde{x}_{D,i} = 0$, so that, by stability, the Democratic party position remains at 0. Clearly, the median voter now prefers the Democratic candidate, and the Democratic primary voter is strictly better off as well. ■

Proof of Proposition 2. Suppose by way of contradiction that there exists an equilibrium in which all policies are located at some point $\bar{x}$. Lemma 1 therefore implies that party positions $\bar{x}_D = \bar{x}_R = \bar{x}$. Since $M_i \neq M_j$ we have either $M_i \neq \bar{x}$ or $M_j \neq \bar{x}$.

Without loss of generality suppose that $M_i < \bar{x}$. Then policy $x_{i,D} = M_i$ would make both the median voter and the median Democrat strictly better off — the latter follows because by assumption $m_{i,D} < M_i$. Since such a deviation is profitable, the original policies cannot be an equilibrium. ■

Proof of Proposition 3. If the elected representatives all have the same positions, $\bar{x}_D$, and $\bar{x}_R$, respectively, then Lemma 1 implies that these are also the party positions. If $\bar{x}_D > -\bar{m}$ in a district $i < k$, then the median voter must be indifferent between the two candidates, else, the Democrat could move further to the left and still get elected. This also implies that $x_{i,R} = M_i = -\bar{M}_i$ in these districts, else the Republican could win by moving marginally to the left.

If the median voter is not indifferent then the Democratic primary voter will select a candidate at his ideal point $-\bar{m}$. Thus,

$$v_{-\bar{M}}(\bar{x}_D) \geq (1-\gamma)v_{-\bar{M}}(\bar{x}_R) + \gamma v_{-\bar{M}}(-\bar{M}), \tag{8}$$

for $i \leq k$, where $x_D = -\bar{m}$ if the inequality in (8) is strict. Similarly,

$$v_{\bar{M}}(\bar{x}_R) \geq (1-\gamma)v_{\bar{M}}(\bar{x}_D) + \gamma v_{\bar{M}}(\bar{M}), \tag{9}$$

for $i > k$, where $\bar{x}_R = \bar{m}$ if the inequality is strict.

To show that there exists $\bar{x}_D$ and $\bar{x}_R$ that satisfy (8) and (9), consider the function

$$g_D(x, x_R) = v_{-\bar{M}}(x) - \left((1-\gamma)v_{-\bar{M}}(\bar{x}_R) + \gamma v_{-\bar{M}}(-\bar{M})\right),$$

Since $v_{-\bar{M}}(x)$ is maximized at $x = -\bar{M}$ it follows that $g_D(-\bar{M}, x_R) > 0$ for $x_R \geq \bar{M}$. By assumption $v_{-\bar{M}}(x)$ is increasing in x for $x \leq -\bar{M}$. Thus, $g_D(x, x_R)$ is increasing in x on $[-\bar{m}, -\bar{M}]$. Thus, there exists a decreasing function $h\colon [\bar{M}, \bar{m}]$ such that $g_D(h(x_R), x_R) \geq 0$, where the strict inequality holds if and only if $h(x_R) = -\bar{m}$.

Now define

$$g_R(x) = v_{\bar{M}}(x) - \left((1-\gamma)v_{\bar{M}}(h(x)) + \gamma v_{\bar{M}}(\bar{M})\right)$$

Note that $h(\bar{M}) \leq -\bar{M}$. Thus, $g_R(\bar{M}) > 0$. If $g_R(\bar{m}) \geq 0$ then $\bar{x}_R = \bar{m}$ and $\bar{x}_D = h(\bar{x}_R)$ satisfy (8) and (9). Else continuity implies that there exists $\bar{x}_R$ such that $g_R(\bar{x}_R) = 0$. Then $\bar{x}_R$ and $\bar{x}_D = h(x_R)$ solve the two conditions. In both cases $\bar{x}_R < -\bar{M}$ and $\bar{x}_R > \bar{M}$, i.e., candidate positions are more extreme than those of the median voters.

Finally, it remains to prove that in all equilibria winning candidates use the same position.

Suppose by way of contradiction that there exist two districts $i \neq j \leq k$ in which Democrats win with policies $x_{D,i} < x_{D,j}$. Since the median voter is at the same location, it cannot be the case that the median voters in both districts are indifferent between the Democrat and the Republican. Otherwise, $x_{D,i} < x_{D,j}$ implies that in one of the districts $x_{i,R} \neq M_i$. Thus, by moving marginally to the left the Republican would get elected in district *i*, a contradiction.

A similar contradiction follows if the median voter in one district is indifferent and strictly prefers the Democrat in the other district. Finally, if the median voters in districts *i* and *j* both strictly prefer the Democrat, then we must have $x_{D,i} = x_{D,j} = -\bar{m}$, a contradiction.

Similar argument show that it is not possible for the Democrat wins in a district $i > k$. The argument for the Republican is similar. ■

References

Adams, J. and S. Merrill (2003). Voter turnout and candidate strategies in american elections. *Journal of Politics 65*, 161–189.

Aldrich, J. H. (1995). *Why Parties?: The Origin and Transformation of Party Politics in America*. Cambridge Univ Press.

Ansolabehere, S., W. Leblanc, and J. M. Snyder Jr (2012). When parties are not teams: party positions in single-member district and proportional representation systems. *Economic Theory 49*(3), 521–547.

Banks, J. S. and J. Duggan (2005). Probabilistic voting in the spatial model of elections: The theory of office-motivated candidates. In *Social Choice and Strategic Decisions*, pp. 15–56. Springer: Berlin, Heidelberg.

Bernhardt, D., J. Duggan, and F. Squintani (2009). Private polling in elections and voter welfare. *Journal of Economic Theory 144*(5), 2021–2056.

Callander, S. (2005). Electoral competition in heterogeneous districts. *Journal of Political Economy 113*, 1116–1145.

Callander, S. (2008). Political motivations. *Review of Economic Studies 75*, 671–697.

Calvert, R. L. (1985). Robustness of the multidimensional voting model: Candidate motivations, uncertainty, and convergence. *American Journal of Political Science 29*, 69–95.

Camara, O. (2012). Economic policies of heterogeneous politicians. mimeo, USC.

Castanheira, M. (2003). Why vote for losers? *Journal of the European Economic Association 1*(5), 1207–1238.

Coleman, J. S. (1972). The positions of political parties in elections. In R. Niemi (Ed.), *Probability models of collective decision making*, pp. 332–57. Charles E. Merrill Columbus, OH.

Cox, G. W. and M. D. McCubbins (2005). *Setting the agenda: Responsible party government in the US House of Representatives*. Cambridge University Press.

Diermeier, D. and R. Vlaicu (2011). Parties, coalitions, and the internal organization of legislatures. *American Political Science Review 105*(2), 359–380.

Dixit, A. and J. Londregan (1995). Redistributive politics and economic efficiency. *American Political Science Review 89*(4), 856–866.

Downs, A. (1957). *An economic theory of democracy*. Harper.

Eguia, J. X. (2011a). Endogenous parties in an assembly. *American Journal of Political Science 55*, 16–26.

Eguia, J. X. (2011b). Voting blocs, party discipline and party formation. *Games and Economic Behavior 73*, 111–135.

Erikson, R. and D. Romero (1990). Candidate equilibrium and the behavioral model of the vote. *American Political Science Review 84*, 1103–1126.

Fiorina, M., S. Abrams, and J. Pope (2006). *Culture war? The myth of a polarized America*. Pearson Longman.

Gul, F. and W. Pesendorfer (2009). Partisan politics and aggregation failure with ignorant voters. *Journal of Economic Theory 144*, 146–174.

Hinich, M. (1978). The mean versus the median in spatial voting games. In P. Ordeshook (Ed.), *Game Theory and Political Science New York*. NYU Press: New York.

Hirano, S., J. M. Snyder Jr, S. Ansolabehere, and J. M. Hansen (2010). Primary elections and partisan polarization in the us congress. *Quarterly Journal of Political Science 5*(2), 169–191.

Krasa, S. and M. Polborn (2010a). The binary policy model. *Journal of Economic Theory 145*, 661–688.

Krasa, S. and M. Polborn (2010b). Competition between specialized candidates. *American Political Science Review 104*(4), 745–765.

Krasa, S. and M. Polborn (2012). Political competition between differentiated candidates. *Games and Economic Behavior 76*, 249–271.

Krasa, S. and M. Polborn (2014). Social ideology and taxes in a differentiated candidates framework. *American Economic Review 104*, 308–22.

Lee, D. S., E. Moretti, and M. J. Butler (2004). Do voters affect or elect policies? Evidence from the US House. *The Quarterly Journal of Economics 119*(3), 807–859.

Lindbeck, A. and J. Weibull (1987). Balanced-budget redistribution as the outcome of political competition. *Public Choice 52*, 273–297.

Martinelli, C. (2001). Elections with privately informed parties and voters. *Public Choice 108*(1), 147–167.

McCarty, N., K. T. Poole, and H. Rosenthal (2009). Does gerrymandering cause polarization? *American Journal of Political Science 53*(3), 666–680.

Owen, G. and B. Grofman (2006). Two-stage electoral competition in two-party contests: persistent divergence of party positions. *Social Choice and Welfare 26*(3), 547–569.

Palfrey, T. R. (1984). Spatial equilibrium with entry. *Review of Economic Studies 51*, 139–156.

Patty, J. W. (2008). Equilibrium party government. *American Journal of Political Science 52*(3), 636–655.

Rohde, D. W. (2010). *Parties and leaders in the postreform House*. University of Chicago Press.

Snyder, J. M. (1994). Safe seats, marginal seats and party platforms: The logic of platform differentiation. *Economics & Politics 6*(3), 201–213.

Soubeyran, R. (2009). Does a disadvantaged candidate choose an extremist position? *Annales d'Economie et de Statistique 93-94*, 328–348.

Volden, C. and E. Bergman (2006). How strong should our party be? party member preferences over party cohesion. *Legislative Studies Quarterly 31*(1), 71–104.

Winer, S., L. Kenny, and B. Grofman (2014). Explaining Variation in Competitiveness of U.S. Senate Elections, 1922-2004. *Public Choice*. forthcoming.

Wittman, D. (1983). Candidate motivation: A synthesis of alternative theories. *American Political Science Review 77*, 142–157.