Parties as Teams and the Nomination of Legislative Candidates∗

Preliminary and Incomplete

Stefan Krasa† Mattias Polborn‡

October 21, 2013

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, a candidate’s party label matters for voters, in addition to his own policy positions: For example, in a liberal district, electing even a liberal Republican may be unattractive for voters because it increases the chance that Republicans obtain the majority in Congress, thereby increasing the power of more conservative Republicans. We show that candidates may be unable to escape the burden of their party association, and that primary voters in both parties are likely to nominate extremist 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.
†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, are 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 electorate at-large, and actual policy and legislation reflects the most moderate point of view 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 build a model that can account for a much higher degree of polarization in the legislature, and which is based on two realistic ingredients: First, parties in the legislature matter for law-making, which creates important spillover effects between the candidates of the same party who run in different districts. Second, legislative candidates are nominated by policy-motivated primary voters who take this legislation process into account when deciding whom to nominate.

The majority party in a legislature is an important power center influencing the crafting of policy. This influence stems from two distinct, but interdependent motivations. First, 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, as in any large organization, specialization is an ubiquitous feature of modern legislatures because no single legislator can be an expert in all policy areas. As a consequence of specialization, individual legislators have considerably more influence on policy in their area of specialization than in other areas where they primarily rely on the expertise of their fellow party members (Shepsle and Weingast, 1987; Gilligan and Krehbiel, 1989).

Clearly, these are far from controversial or novel insights for scholars of legislatures. However, there is surprisingly little analysis of how these organizational features of legislatures influence the types of candidates who are nominated by their party to run for legislative office, and 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 outcomes of elections in different legislative districts.

As mentioned above, if one were to apply the simplest Downsian model naively to Congressional elections, then it generates counterfactual predictions: In each district, both candidates should adopt the pre-
ferred position of the district median voter, and so, policy-wise, all voters should be indifferent between the Democratic candidate and his Republican opponent. Assuming that voters also care about idiosyncratic candidate features such as competence, looks or sympathy, Republicans in New England or Democrats in Utah should have a substantial chance to be elected to Congress if only they match their opponent’s policy platform. Furthermore, in this model framework, gerrymandering districts would not help parties, at least not in the sense that it would increase the party’s expected representation in Congress. It is safe to say that both of these predictions are counterfactual, and the Downsian model applied at the legislative district level has a hard time explaining why this is the case.

In our model, representatives of the majority party influence national policy through two different channels: Each representative of the majority party individually influences policy on some set of policy issues, and other policy issues are determined by an internal compromise within the majority party – which is, of course, also affected by the ideological make-up of the majority party. Voters’ and representatives’ policy payoffs are determined by aggregating over all policy areas. In addition, voters also receive a payoff from their representative’s valence (e.g., quality of constituent service or absence of corruption or other scandals). In the general election for the legislature, voters vote for their preferred candidate, taking into account the two ways in which their local representatives may change the policy outcome: First, the district result may change which party is the majority party in Congress, and second, if they elect a candidate who will be in the majority party, they may affect the ideological composition of the majority party.

In this framework, there are spillovers between different districts: The electoral prospects of candidates in a given district are influenced by the expected ideological position 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, the U.S. senator from Rhode Island from 1999 to 2006, Lincoln Chafee. In spite of being a Republican, Chafee had taken a number of socially moderate and liberal positions that brought him in line with voters in his state,\(^1\) and his average net approval between May 2005 and June 2006 was a healthy 13 percentage points.\(^2\) Wikipedia describes his 2006 reelection campaign as follows: “Despite

---

\(^1\)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](http://en.wikipedia.org/wiki/Lincoln_Chafee)).

\(^2\)See [http://www.surveyusa.com/client/PollTrack.aspx?g=417c7a9d-d8ec-4083-8e80-ca577e486925&x=0,0](http://www.surveyusa.com/client/PollTrack.aspx?g=417c7a9d-d8ec-4083-8e80-ca577e486925&x=0,0). Net approval (i.e., percentage of respondents who “approve” of the Senator’s performance, minus percentage of disapprovers) based on 14 opinion polls between May 2005 and June 2006 when the fall election campaign started. There is no clear time trend during this period.
Chafee’s high approval ratings statewide, [Rhode Island] had tended to lean heavily toward socially liberal or moderate candidates for many decades. As a result, [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).” Whitehouse won the election 54-46.\(^3\)

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 they’re referring to a Republican candidate as a ‘supporter’ of the ‘Bush agenda’ or as a ‘rubberstamp,’ 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.”\(^4\)

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. 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; or the other party could, in principle, compete, but prefers to nominate a losing extremist. The latter case arises if a winning moderate might “taint” the party’s position in the legislature.

Again, Lincoln Chafee provides an instructive illustration of this principle. Before the 2006 general elections, conservative Republicans mounted a primary challenge. Chafee defeated his challenger who had attacked him for not being sufficiently conservative only by a margin of 53 percent to 47 percent, and there is reason to believe that a majority of “real” Republicans would have preferred to replace a relatively popular incumbent Senator with an extremist whose policy positions would have implied a very low likelihood of

\(^3\)Of course, there are analogous examples of Republican candidates running in conservative districts and linking their Democratic opponents to party leaders such as Nancy Pelosi and Barack Obama. For example, in an ad against Ben Nelson, U.S. Senator from Nebraska and Democrat in a very conservative state, his Republican opponent Pete Ricketts tried to tie Nelson to other Senate Democrats: “[...] Hillary Clinton wants higher taxes, Ted Kennedy liberal justices. Now Ben Nelson isn’t a Ted Kennedy. But he is a Democrat who will vote for Democrats in charge. I won’t.” See [http://www3.nationaljournal.com/members/adspotlight/2006/10/1030nesen1.htm](http://www3.nationaljournal.com/members/adspotlight/2006/10/1030nesen1.htm)

prevailing against the Democrat in the general election in Rhode Island.\footnote{In his victory speech, Chafee credited independent voters and disaffiliated Democrats for his victory.} Our model explains why this behavior may be perfectly rational for policy-motivated Republicans: From their point of view, having Chafee as a member of the Republican Senate caucus caused more harm than benefits.

Our model provides a framework in which, in contrast to the classical one-district spatial model, the ideological composition of districts does not only influence the ideological position of elected candidates, but also the chances of parties to win a majority in the legislature. 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.

We then turn to a setting in which the potential for contamination externalities between districts is minimized because we assume that all districts are identical; because there is a symmetric equilibrium in this setting, it is not the case that any candidate in any district is uncompetitive because he is contaminated by his party association. We do this to identify a second effect in legislative elections called “dilution,” which affects the voters’ trade-off between valence and policy positions. Dilution arises from the fact that each representative has only limited influence on policy. From the perspective of a district median voter, valence is therefore more important than socially optimal (as those policies that are influenced by the local representative of course affect everyone in the nation). This implies that general election voters are less sensitive to ideological extremism, and therefore local primary voters can choose more extreme candidates.

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 model, and the main analysis follows in Sections 5, 6 and 7. We conclude in Section 8.

\section{Related literature}

Ever since the seminal work of Downs (1957), the issue of position choice by candidates and what determines whether the equilibrium features policy convergence or divergence is arguably the central question...
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; Bernhardt et al. 2009; Callander 2008).

Overwhelmingly, the existing literature looks at isolated elections – there are (usually) two candidates who are running against each other, and voters care about the positions of the two candidates. In the probabilistic voting model (e.g., Hinich 1978; Lindbeck and Weibull 1987; Dixit and Londregan 1995; Banks and Duggan 2005) and some other models (Adams and Merrill 2003), voters may also receive “partisan” or “ideological” payoffs that are independent of the position that a candidate takes, and one may be inclined to interpret this as also capturing the effects of the candidate being affiliated with a party, and therefore implicitly other legislators and their policy positions. However, the “ideology shocks” in these models are completely exogenous, i.e., they are not derived from the equilibrium positions of candidates in other districts. Thus, the main point of interest in our model – how does the fact that legislative candidates have to work together with other legislators in order to set new policy affect both the equilibrium positions of candidates and the choice of voters between their local candidates? – cannot be analyzed in these models.

In principle, there are only very few elections that perfectly fit the election-in-isolation framework because even executive officials such as state governors or the U.S. President have to work with other actors in order to implement new policy. Notwithstanding this fact, the simple spatial model and its results are usually also applied to 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 more or less reflect the conservativeness of their districts (say, measured by their Partisan Voting Index6). Our model will show that this transfer of results derived in the isolated-election model to legislative elections is not always justified, and that the equilibrium positions of candidates may reflect more the preferences of the parties’ respective primary electorates than those of the district median voter.

Our model is related to the class of differentiated candidates models (Krasa and Polborn 2010a,b, 2012,

---

6The Partisan Voting Index (PVI) is based on the two party Presidential vote in a district relative to two-party vote shares in the nation at-large. See http://cookpolitical.com/application/writable/uploads/2012_PVI_by_District.pdf for the PVI based on the 2004 and 2008 Presidential elections.
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 do not have exogenous preferences over characteristics (here, the candidates’ party affiliations), but rather, these preferences are endogenously derived from how the election of the Democratic and Republican local candidate would affect the national policy.

We assume that the nomination decision is made by a policy-motivated party median voter. For this reason, our model is related to the literature on the position choice of policy-motivated candidates pioneered by Wittman (1983) and Calvert (1985). In most of this literature, it is assumed 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 someone else than the candidate (namely a primary election median voter), 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.

The results of our analysis are relevant for the large empirical literature that analyzes how primaries, the ideological composition of districts and especially the partisan gerrymandering of districts affects the ideological positions of their representatives in Congress (e.g., 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. We will show that this approach is somewhat problematic.

The legislative part of our model assumes that parties in Congress have a meaningful influence on policy outcomes. There is a significant number of models that 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. In contrast, the pivotal politics model (Krehbiel, 1992, 1993, 2010) argues that only the median legislator’s preferences

\[7\text{Implicitly, we assume that either candidates can commit to an ideological position in the primary, or that candidates are citizen-candidates with an ideal position that is common knowledge.} \]
influence the policy outcome. All these models of the importance of parties in Congress take the preference distribution of legislators as exogenously given. This is the main contribution of our model to this literature: It provides for an electoral model and thus endogenizes the types of elected legislators.

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

In this section, we want to argue that the influence of ideology on the parties’ performance is substantially larger in legislative elections than in executive elections. 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 will 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 the next subsection, 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 influence 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 ideological preferences of voters 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. Both of these types of contests are state-wide races, but evidently, Gubernatorial elections are for executive positions while Senate elections are for legislative positions.

Consistent with the empirical literature in political science, 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 Presidential elections, relative to the nation’s average share of the same.\(^8\)

\(^8\)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
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). In Gubernatorial elections (the omitted category), the coefficient of the PVI variable indicates that a one point increase in the ideological preference of the state for Democrats increases the Democratic gubernatorial candidate’s vote share only by about 0.519 points. In contrast, in Senate elections, the same ideological shift of the state 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, the difference between executive and legislative elections is substantial and highly significant. The remaining 3 columns confirm the qualitative robustness of this difference if we restrict to elections after 1990 and if we exclude the political South.9

Note that 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 appropriate or indeed 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 fit better to the state’s voter preferences should reduce the advantage of the opponent, and thus the estimated coefficient. Somehow, only gubernatorial candidates appear (at least in parts) capable of such a position adjustment, while Senate candidates are not.

### 3.2 Inconsistency with the simple single-district model

We now turn to the interpretation of these stylized facts, and will argue that they are difficult to reconcile with the standard model of political competition that assumes that any election is effectively a competition between the two candidates in each district that is (implicitly) not influenced by what happens outside the district. Specifically, we want to point out the problem that 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.

Suppose, without loss of generality, that the median voter is located at zero and that the party medians 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.

9The reason for excluding the South is that, 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.
Table 1: Senate and Gubernatorial elections

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>PVI</td>
<td>0.519*** (0.111)</td>
<td>0.589*** (0.124)</td>
<td>0.529*** (0.117)</td>
<td>0.614*** (0.132)</td>
</tr>
<tr>
<td>PVI x Senate</td>
<td>0.645*** (0.149)</td>
<td>0.596*** (0.167)</td>
<td>0.597*** (0.156)</td>
<td>0.514*** (0.177)</td>
</tr>
<tr>
<td>N</td>
<td>1103</td>
<td>702</td>
<td>871</td>
<td>553</td>
</tr>
<tr>
<td>$R^2$</td>
<td>0.551</td>
<td>0.595</td>
<td>0.571</td>
<td>0.62</td>
</tr>
</tbody>
</table>

*** indicates significance at the 1% level.

Additional explanatory variables used: Election type (Senate or Governor), year dummies, and incumbency status.


are at $M_D$ and $M_R$. Even if party members are ideologically very polarized, parties will nominate candidates who are “competitive” from an ex-ante perspective in the sense that they have relatively moderate policy positions and a substantial probability of winning. This is obvious for the model without uncertainty where, in equilibrium, both parties will nominate candidates that maximize the median voter’s utility, i.e., $x_D = x_R = 0$ and both parties will 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, this 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 effectively zero (if the distribution of voter types is continuous).

In Proposition 6 in the Appendix, we show that these results also hold when there is uncertainty about candidates’ valence or the ideal position of the median voter. Thus, if party members become more extreme then the Downsian model predict that this has either no or only a small influence on policy: Party members still 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 party members, has resulted in more extreme candidates being nominated for office (e.g., Fiorina et al. (2006)). Further, many political commentators and scholars argue that there is a rise in polarization between the two parties. In order to generate such a polarization in a standard model with policy-motivation, we have to increase uncertainty, e.g., about the median voter’s ideal position. In other words, we would need political polls to have been much more
accurate in the past than they are now, which is probably not plausible.

The prediction that both candidates in executive elections (i.e., those where the elected candidate can meaningfully influence policy without being tied to their party) will be competitive is borne out in the 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%. Furthermore, the results of Table 1 above indicate that Gubernatorial elections 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 the distribution over types is continuous, then the vote margin by which the favored candidate wins is (very close to) 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, the 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.

In Proposition 7 in the Appendix we show that policies will remain moderate and the margin of victory will be close to zero even if there is some uncertainty about the location of the median. However, if the uncertainty is not too large then winning margin will be 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 will be “close” to indifferent, and hence the electorate will split close to 50-50. One of the key insights of this paper is that this is not longer true in a multi-district setting.

10U.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.
4 Model

We consider a polity that is divided into $2n + 1$ districts, ordered according to their ideological preference, so that district $n + 1$ is the median district. Each district $i$ elects one representative to the legislature, who is characterized by a position $x_i$. All candidates are attached to one of two parties, Democrats and Republicans.

Let $x = (x_{D,i}, x_{R,i})_{i=1,...,2n+1}$ be the positions of candidates in all districts, and $X$ the set of all such positions. Let $k_i \in \{D, R\}$ denote the party of the winning candidate in district $i$, and $K = (k_i)_{i=1,...,2n+1}$.

A policy by the legislature can be deterministic, i.e., $\bar{x} \in \mathbb{R}$, or a finite lottery $L = \{(\gamma_j, \bar{x}_j) | j \in J\}$, where $\gamma_j$ is the probability that policy $\bar{x}_j$ is chosen (i.e., $\bar{x}_j \in \mathbb{R}$, $\gamma_j \geq 0$, and $\sum_{j \in J} \gamma_j = 1$). Let $\mathcal{L}$ denote the set of such lotteries.

We now explain how the positions of legislators translate into actual policy. On the most general level, one can think of this relation as a function $\xi : X \times K \rightarrow \mathcal{L}$, mapping candidates’ positions in all districts and the identities of the winning candidates into the national policy that determines all voters’ policy utility.

For example, we can think of how policy is determined by the legislators from the majority party as follows: Members of the majority party are randomly selected to propose a policy, which must be accepted by a majority of the party’s elected representatives; if the proposal is rejected, a new proposer is selected, and bargaining extends until a proposal is accepted by a majority; payoffs are discounted at a rate $\delta < 1$ per round. The resulting policy function $\xi$ will have to important properties:

First, if all elected members of the majority party share the same position, $\hat{x}$, then $\hat{x}$ will be the legislature’s policy. Second, each representative has some influence on the position, but the extent of that influence is limited if all other legislators agree: If one of the elected representative’s policy $x_{P,i}$ differs from $\hat{x}$, and if that party member is the first proposer then the adopted policy will be $\hat{x}$ between $\hat{x}$ and $x_{P,i}$ — the closer $\delta$ is to one the close $\hat{x}$ will be to $\hat{x}$. If one of the other members of the majority party is the first proposer then the policy will be $\hat{x}$. Thus, if all but one member of the majority party have the same policy position then $\xi$ assigns $\hat{x}$ with probability $1/#P(k)$ and $\hat{x}$ with the remaining probability.

Rather than developing our results for a general function $\xi$, we state all our results for the case that $\xi$ selects the policy position of the median legislator in the majority party. Formally, let $P(k) \in \{D, R\}$ denote the majority party; for example, $P(k) = D$ if $\#(i|k_i = D) > \#(i|k_i = R)$, and let $S(x, k) = \{x_{s,i}|k_i = P\}$ denote the set of positions taken by the majority party legislators. Then $\xi = \text{median}(S(x, k) = \text{median}\{x_{s,i}|k_i = P\}$.
However, as we will explain, most of our results hold with minor adjustments for any mapping $\xi$ that satisfies the two properties discussed above, selection of $\hat{x}$ if all majority party legislators agree on $\hat{x}$, and some minimal influence of each majority party legislator.

For the case of $\xi = \text{median} S(x, k) = \text{median}(x_i[k_i = P])$, the utility of a voter with ideal position $\theta$ from district $i$ is

$$u_\theta(x, k, v_i) = -(\text{median } S(x, k) - \theta)^2 + v_i.$$  

(1)

Here, the first term is the utility from the policy, and the last term is the valence of district $i$’s representative. We will consider valence differences in section 7. Otherwise, $v_i = 0$.

The determination of general election candidates is made through a nomination process in which either national parties or local party members select the nominees. We assume that nomination processes can be summarized by the preference parameter of a “decisive voter”. In the case of the U.S. primary system, we take this decisive voter to be the median party member in the district, whose ideal position is denoted $m_{D,i}$ and $m_{R,i}$ for Democrats and Republicans, respectively. In the case of a national nomination system, we assume that the decisive voter in the (national) median of party members, denoted by a capital $M_D$ and $M_R$, respectively.

At the time of the nomination decision, the ideal position of potential candidates are known, but there may be uncertainty about the candidates’ valence levels and/or the ideal position of the district median voters. At the time of the general election, voters observe the realization of uncertainty in their own district, but not in the other districts so that, for example, the party that wins the majority is still uncertain at the time of election. Formally, uncertainty about the state of the world is captured by a collection $(\omega_i)_{i=1,...,2n-1}$ of random variables for each district, where each $\omega_i \in \Omega_i$ determines $v_{i,\ell}(\omega_i)$, the valence of candidate $\ell$ in district $i$, and $M_i(\omega_i)$, the identity of district $i$’s median voter.

In a slight abuse of notation, let $k_i : (M_i(\omega_i), \omega_i, x_{i,D}, x_{i,R}) \rightarrow \{D, R\}$ denote the voting strategy of the district $i$ median voter $M_i(\omega_i)$ when choosing between candidates $x_{i,D}$ and $x_{i,R}$ in state $\omega_i$. Note that the identity of the majority party and its ideological composition are random variables, so when we want to emphasize this, we can write $S(x, k, \omega)$.

**Definition 1** A collection of policies $x \in X$ and of voter strategies $k(x, \omega) = (k_i)_{i=1,...,2n+1}$ is a pure strategy equilibrium if and only if
1. for every district $i$ and for every state $\omega_i$ in that district, the median voter $M_i(\omega_i)$ chooses his optimal candidate: If $w_i(M_i(\omega_i), \omega_i, x_i, D, x_i, R) = P_i$, and $\bar{P}_i$ denotes the other candidate, then

$$E_{\omega_i} u_{M_i(\omega_i)}(x, k_i(\cdot), k_{-i}(\cdot), v_i, P_i(\omega_i)) \geq E_{\omega_i} u_{M_i(\omega_i)}(x, k'_i(\cdot), k_{-i}(\cdot), v_i, P_i(\omega_i))$$ for all $k'_i \in \{D, R\}$ (2)

where the expectation is taken over the realization of uncertainty in the other districts;

2. the candidate choices of the decisive primary voters $m_{i,P}$ are optimal for them, respectively:

$$E_{\omega_i} u_{m_{i,P}}(x, k_i(\cdot), v_{i,k_i}(\cdot), v_i) \geq E_{\omega_i} u_{m_{i,P}}(x, k'_i(\cdot), v_{i,k'_i}(\cdot), v_i)$$ for both parties $P$ and all alternative positions $\bar{x}_{i,P}$ (3)

where the expectation is taken with respect to the ex-ante distribution of $\omega$.

We will also sometimes use the following simple equilibrium refinement criterion: Suppose that voters in district $i$ believe that, with probability $\gamma$ very close to zero, their representative will determine the national policy (provided that he is in the majority party), while with probability $1 - \gamma$ the policy is determined as described above. In particular,

$$u_{\theta}^{\text{refine}}(x, k, v_i) = \begin{cases} (1 - \gamma)u_\theta(x, k, v_i) + \gamma(x_i, k_i - \theta)^2 & \text{if } k_i \text{ is the majority party;} \\ u_\theta(x, k, v_i) & \text{otherwise.} \end{cases}$$ (4)

Our refinement reduces attention to those equilibria of the game that are limits of equilibria of the $\gamma$ perturbed model, as $\gamma \to 0$.

The main bite of this refinement is that, if the primary voters of the winning party have a choice between several different candidates who would all win the general election, then they choose the one who is closest to their ideal point, even if their local representative does not matter for the location of the national policy in equilibrium (because he does not affect the location of the median). As explained above, in a legislative bargaining model of policy determination within the majority party, each legislator’s ideal position has some chance of mattering as long as the representative is recognized with positive probability, so this is another way of justifying this refinement concept.
5 Externalities between Districts

5.1 Contamination in the case without uncertainty

We start with a standard Downsian model without uncertainty about the median voters’ ideal points, and where all candidates have the same valence, because this very simple framework already allows us to identify contamination, one of the key effects in our model.

As a benchmark, remember that, if there is a single district with a median Democratic primary voter at \(-m\), a Republican primary voter at \(m\) and a general election median voter located at \(M\), then both primary voters in equilibrium select candidates with \(x = M\). The general election median voter is indifferent between the parties, and may thus choose either the Democrat or the Republican with probability 1 in a pure strategy equilibrium.

To gain an intuition for the effects of parties as teams in legislatures, consider the example analyzed in Figure 1: There are five districts, with the median general election voter located at \(M_i\) in district \(i\), where \(M_1 < M_2 < \ldots < M_5\), and the median primary voters again located at \(-m\) and \(m\) in each district. Proposition 1 shows that the results of this simple example remain qualitatively robust for an arbitrary number of electoral districts.

Without major loss of generality, suppose that \(M_3 < 0\) (the case that \(M_3 > 0\) is clearly symmetric). To simplify the exposition, we consider only pure strategy equilibria. We first show that selecting candidates who are located at the median in each district is no longer an equilibrium, due to policy externalities between different districts.

To see this, consider panel (a) of Figure 1 where we have a situation in which the candidates in all districts are located at the ideal positions of their respective median voters, \(x_{D,i} = x_{R,i} = M_i\). Assume that districts 1 and 3 elect their Democratic candidates, and districts 4 and 5 elect their Republican candidates (elected candidates’ labels and positions are indicated by the small box around the party label). Consider the problem of district 2’s median voter. If he elects the Democrat, the policy is equal to his preferred outcome, while if he elects the Republican, the national policy is \(M_4\). Thus, he strictly prefers the Democrat.

However, if district 2’s median voter strictly prefers the Democratic candidate, then the Democratic primary voter could select a more extreme candidate and still be guaranteed to win. Since this makes the Democratic primary voter strictly better off, the scenario depicted in Figure 1 (a) is not an equilibrium.
Intuitively, in the one-district model, the general election median voter is indifferent when both parties choose the same policy, but with multiple districts this is no longer true since the positions of legislators from the other districts generate the externalities just discussed; and, if the median voter in a district is not indifferent, the party of the preferred candidate can do better for themselves if they nominate a more ideologically extreme candidate because that more extreme candidate will still be elected.

Is there an equilibrium in which all candidates locate at the median of the median district (Figure 1 (b))? If the Democrats win a majority in equilibrium (or indeed, if they have any positive probability of winning), then the median voters in districts 1 and 2 would prefer a more left-leaning candidate, as would the Democratic primary voters; so this is not an equilibrium. Similarly, if the Republicans win, then both median voters in districts 4 and 5 and the Republican primary voters would prefer more conservative candidates.

Third, is there an equilibrium in which just one party’s candidates (the Republicans in Figure 1 (c)) are at the median of district 3? If there was, the moderate party would win, but if it wins, both the median voters
of the more extreme districts 4 and 5, and its primary voters, would benefit from nominating more extreme candidates.

We now turn to Figure 1(d) which shows an equilibrium in which all parties nominate their ideal candidates, except for Republicans in district 3; they nominate the median voter’s ideal candidate, but nevertheless lose to an extreme Democrat.

To see that this is an equilibrium, let us start by considering the most competitive district, district 3. While the local Republican candidate is ideal from the median voter’s perspective, his election will move the national policy from $-m$ (the policy if the Democrat wins) to $+m$ (the median policy of the Republican legislators if they win a majority). Since the median voter in district 3 prefers $-m$ to $+m$, he will stick to the Democrats even if they nominate the most extreme candidate, and so the Democratic primary voters can get away with nominating their ideal candidate.

Note that, in this equilibrium, voters split in each district according to which of the extreme positions that the parties offer ($-m$ or $m$) they prefer: All voter types with $\theta < 0$ vote for the Democrats, and all voter types $\theta > 0$ vote for the Republicans. Since there is no district in which the median is at 0, all districts have a winning margin that is bounded away from 0. Also, sufficiently small valence shocks, or uncertainty about the ideal position of the median voter, would therefore not change the qualitative features of the equilibrium, as sufficiently small uncertainty cannot change the outcome anywhere.

The equilibrium depicted in Figure 1(d) is not unique; for example, the moderate Republican who loses in district 3 could, of course, also locate at a more extreme position. Similarly, the extreme Republicans in districts 4 and 5 could locate at a more moderate position as long as it is not too moderate — it cannot be the case that district 3’s general election median voter would prefer the position of the Republicans in districts 4 and 5 to the Democrats’ in districts 1 and 2, because then he would vote for the Republican; and once district 3 votes for Republicans, the Republican primary voters in districts 4 and 5 would much rather nominate more extremist candidates. In principle, it would be possible to refine the away equilibria with more moderate positions for the losing party’s candidates, by assuming (for example) that primary voters select the candidate with the closest ideological position to their own ideal one, as long as this does not affect the expected national policy.

Why do Democrats in districts 4 and 5 (where they do not win) not nominate more moderate candidates? Indeed, such a more moderate candidate would be elected because, as member of the majority party, he
would moderate the implemented policy. However, that is exactly the reason why the Democrats do not want to nominate a more moderate candidate – if he were to be nominated and elected, the implemented policy would be less desirable from the point of view of the decisive Democratic primary voter.

Finally, we should also note that the assumption that the general election median voters in the different districts have different ideal positions is decisive for the equilibrium polarization result. If all median voters are at the same position, i.e., $M_i = M$ for all $i$, then in equilibrium all candidates are located at $M$; of course, identical medians in all districts is a highly non-generic case.

We now state our results for the general case of $2n + 1$ districts.

**Proposition 1** Suppose there are $2n + 1$ districts ordered according to the ideology of their general election median voters (i.e., $M_i < M_{i+1}$ for all $i$), and assume that Democratic and Republican primary voters are located at $-m$ and $m$, respectively. Then:

1. There exist pure strategy equilibria that satisfy the equilibrium refinement. In all equilibria the legislature’s policy is located at the ideal point of the median primary voter of the majority party ($m$ if Republicans win, and $-m$ if Democrats win).

2. There do not exist equilibria in which candidates in each district $i$ located at the district median $M_i$, or where all candidates are located at the median of the median district, $M_{n+1}$.

3. If one party wins with $n + 1$ seats, then the policies of at least $n$ elected candidates of the winning party are located at the party members’ ideal point. The resulting policy of the legislature is $m$ if Republicans win, and $-m$ if Democrats win.

Proposition 1 provides a very stark contrast to the intuition that is based on the single district model. Rather than a legislature that looks like an assembly of district median voters from all over the country, Proposition 1 predicts a legislature in which the majority party is as extreme as the median voter in that party’s primary. General elections lose their disciplining and moderating power because voters (correctly) anticipate that both parties are extreme.

Note also that, in equilibrium, the median voters in most or all districts have a strict preference for the election winner in their district, implying that winning margins are bounded away from zero. Thus, the model can generate safe districts without appealing to a large “party valence” or “partisanship” that is also unrelated to policy.
Moreover, which party wins a majority of seats in the legislature depends on the ideological inclinations of the median voter in the median district. Remember that party median positions are normalized in a way that they are at $-m$ and $m$, i.e., a voter at 0 is indifferent between these positions. If the median voter in the median district prefers the Democratic party position to the Republican one, then, in equilibrium, Democrats will win a majority of seats, and vice versa.

Both the result that the identity of the majority party in Congress depends on the preference of the median voter in the median district between the two extreme party positions, and the result that the equilibrium policy is extremely partisan show that gerrymandering of districts will have a tremendously important effect on the equilibrium policy in this model. It is important to note that research that purports to show empirically that “gerrymandering has no influence on polarization” does not actually do that – rather, it shows some empirical stylized facts that are entirely consistent with the model that we present here: Notably, that the ideological voting position of representatives in Congress is influenced very little by the ideological preferences of district voters (as captured by the district’s PVI), but depends strongly on the party affiliation of the representative.

As mentioned in the introduction, 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. For example, a Democrat has a hard time to be elected in Wyoming, even if he adopted a platform that would actually be preferred by the median voter in his district to the platform of the Republican). The contamination effect in our model produces exactly such a result because voters in a biased district are reluctant to vote for the representative of the party whose representatives from other districts are unpopular. For example, a vote for the Democrat in Wyoming would also strengthen the probability that Democrats from other parts of the country get the chance to enact legislation, and their positions are too liberal to be palatable for Wyoming voters. Furthermore, our model shows that the implied security to win the districts in which the median voter significantly favors one of the parties may induce the favored party’s primary voters to nominate more extreme candidates than optimal for the district’s median voter.

In contrast to the standard intuition in the Downsian model that suggests that two-party competition will lead both parties to promote very moderate positions in order to appeal to the median voter, positions are very extreme in the equilibrium of our model: Most candidates, and even all candidates who actually win, are at the ideal positions of their partisan supporters (more precisely, the median primary voter). Since one
of these two extremist parties necessarily wins and therefore controls the legislature, we should not expect
that this legislature will necessarily adopt policies with broad popular support, as long as they are unpopular
with their own party base. Furthermore, if only left-wing and right-wing extremists populate the legislature,
they may have problems compromising with each other (if, say, a compromise is necessary because Senate
and House are dominated by different parties).

For example, an October 7 Washington Post opinion poll shows 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 – even those held by a majority of Republican House members – oppose 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 voters who actually vote in Republican primaries, there is an
even larger majority in favor of the shutdown. There are media reports that many Republican representatives
“would like” to end the government shutdown, but are afraid that taking this position publicly (say, in a vote)
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 is a very justified fear: It is not the case that primary voters who refuse to
renominate a “moderate” and replace him with an extremist are irrational – they play their optimal equilib-
rium strategy because even the extremist is very likely to win, even in a relatively moderate district. This
makes the primary threat so credible.

An alternative explanation for non-median policy outcomes is that a strong lobby in favor of the minority
position is able to “buy” the support of legislators. Such a lobbying explanation is most plausible if benefits
on the minority side are highly concentrated, and if the issue itself is a relatively minor issue for those
voters who are in the majority. Arguably, neither case applies to the government shutdown. Moreover,
the U.S. Chamber of Commerce, an traditionally important donor for Republicans has urged their allies in
Congress not to shutdown the government and to increase the debt limit,11 and has promised special financial
contributions to members who come under pressure in primaries because of votes in favor of this position.12

11See http://www.uschamber.com/issues/letters/2013/multi-industry-coalition-letter-regarding-
continuing-resolution-and-debt-limit.
12See http://talkingpointsmemo.com/news/with-traditional-gop-allies-defecting-big-business-leaders-
take-sides-with-obama.
It appears unlikely to us that Republicans decided to shutdown the government in order to attract campaign contributions.

Another theory that sometimes generates non-median equilibrium policies is the probabilistic voting model which shows that minority positions may win in equilibrium if the minority feels sufficiently more intensely about the issue than the majority. This also does not appear a plausible interpretation for the government shutdown – 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 percent “disapprove somewhat.” Thus, intensity about the issue appears higher among those who disapprove.

6 The effects of gerrymandering

We now turn to an analysis of the effects of “gerrymandering” (i.e., changes in the ideal positions of district median voters) on equilibrium policy divergence between parties. This is an important issue because many observers suspect that polarization in Congress may be caused by gerrymandering of districts, that is, the creation of districts that are internally ideologically homogeneous (either very conservative or very liberal) and in which representatives are much more afraid of a primary challenge from within their own party than of the opposition candidate in the general election.

On the other hand, there is an empirical literature that argues that gerrymandering has no significant effect on polarization between the two parties in Congress. Essentially, this literature argues that polarization and gerrymandering increase at the same time, but that there is no causal relationship between the two. In this view, the principal driver of polarization between parties is “sorting” in the sense that conservative districts are increasingly represented by Republicans.

McCarty et al. (2009) use the DW-Nominate score of representatives as their dependent variable, and analyze how it correlates with district characteristics (e.g., the district’s PVI) and the party affiliation of the district representative. They argue that “for a given set of constituency characteristics, a Republican representative compiles an increasingly more conservative record than a Democrat does. Gerrymandering cannot account for this form of polarization” (p. 667).

Implicitly, this argument assumes that the structure of voter preferences in the different districts (which
is influenced by gerrymandering) does not affect the equilibrium positions of candidates in any but the most direct way: For example, if a district is gerrymandered to be more conservative, then positions of candidates in that district will be more conservative, but there are no spill-over effects on the positions of candidates in other districts that remain moderate.

In the following, we want to analyze whether this is a logically justified argument. To do this, we extend the framework analyzed in the last section to allow for uncertainty about the position of the median voters in the different districts. Some uncertainty about voter preferences is realistic and therefore an important robustness feature.

We start with the benchmark case of full inter-district homogeneity, i.e., district median voters in all districts are independent draws from the same distribution.

**Proposition 2** Suppose that, in all districts, the district median voter is an iid draw from a distribution \( \Phi \) that is symmetric around zero. Then, in the symmetric equilibrium, the policies are independent of the number of districts and given by \( \bar{x} = \frac{m}{\sqrt{2\phi(0)m}} \).

Proposition 2 provides the most competitive benchmark case as all districts are symmetric. Interestingly, the equilibrium candidate positions in this equilibrium are independent of the number of districts. Thus, in particular, they are the same as in the case that there is only one district.

We now turn to the case that there are liberal and conservative-leaning districts. We let the position of the median voter be distributed symmetrically around \( M_{n+1} \) in the median district \( n+1 \).

**Proposition 3** Suppose the districts medians for all districts \( i \neq n + 1 \) are deterministic, and for all deterministic districts \( i, j \neq n + 1 \) we have \( M_i < M_j \) for all \( i < j \). Further, let \( M_n < 0 \) and \( M_{n+2} > 0 \).

Suppose further that the median in district \( n + 1 \) is stochastic: The cdf of \( M_{n+1} \) is symmetric around its median \( M \) and is given by \( \Phi_M(x) = \Phi_0(x - M) \), where \( \Phi_0 \) is the cdf of a symmetric distribution around 0, with a density \( \phi_0 \) that is continuously differentiable, and satisfies \( \phi_0(x) > 0 \) for all \( x \).

1. Suppose that \( n \geq 3 \). Then, there exists a pure strategy equilibrium with the following properties

   (a) The optimal policies in districts \( i \neq n + 1 \) are \( x_{D,i} = -m, x_{R,i} = m \).
(b) The optimal policies in district \( n + 1 \) are given by

\[
\begin{align*}
    x_D &= -\frac{\Phi_0(-M)m}{\Phi_0(-M) + \phi_0(-M)m}, \\
    x_R &= \frac{(1 - \Phi_0(-M))m}{(1 - \Phi_0(-M)) + \phi_0(-M)m}.
\end{align*}
\]

(5)

(c) If \( n \geq 4 \) then in any pure strategy equilibrium that satisfies the refinement and in which the Democrat wins in districts \( i < n \) and the Republican in districts \( i > n + 1 \), positions in district \( n + 1 \) are given by (5) and \( x_{D,i} = -m \) for \( n - 1 \) districts \( i \leq n \), and \( x_{R,i} = m \) for \( n - 1 \) districts \( i > n + 1 \).

2. If \( M = 0 \) then polarization is the same as in the case where all districts are identical, i.e., \( x_D - x_R = 2m/(1 + 2\phi_0(0)m) \).

3. There exists \( \bar{m} < \infty \) such that for all \( M \neq 0 \) polarization is larger than in the identical district case when \( m > \bar{m} \) and smaller if \( m < \bar{m} \), i.e., \( x_D - x_R > 2m/(1 + 2\phi_0(0)m) \) if \( m > \bar{m} \) and \( x_D - x_R < 2m/(1 + 2\phi_0(0)m) \) if \( m < \bar{m} \).

Proposition 3 shows that polarization in the other districts affects equilibrium positions in the most competitive district. Specifically, if \( M \neq 0 \) and party ideal position are relatively close (\( m < \bar{m} \)), then the fact that competition for control of the legislature is now focused on the median district induces parties to compete more vigorously in the median district (i.e., \( x_D - x_R \) is smaller than in the case all homogeneous districts).

In contrast, if party ideal positions are relatively far apart from each other (\( m > \bar{m} \)), then the parties diverge in the median district. Then reason is that the median voter in the median district becomes less responsive to the local candidates when party ideal positions (and thus, the parties’ candidates in the other districts) are farther apart from each other. This decreased sensitivity of the general election median voter induces the parties to go for more extreme candidates even in the median district.

In summary, our model shows, for elections to a legislature, the subtle dependency of optimal decisions in one district (both by general election voters, and by the party primary voters) on (the expectation about) decisions made in other districts. Thus, it is not possible to argue that the effects of gerrymandering a particular district should necessarily be contained to just the gerrymandered district.
7 Dilution: The Externality generated by Valence

Up to this point we have analyzed the externalities generated by differences between districts. If all districts are indeed identical and there are no valence differences between candidates, then Proposition 2 above shows that the equilibria of multi-district setting are the same as in the one district case. However, if candidates have stochastic valence differences, an effect that we refer to as “dilution” arises, which is not present in the single district Downsian model.

To see this, consider one particular district, say district 1, of the $2n + 1$ identical districts. Assuming that there is a symmetric equilibrium so that each party wins with probability $1/2$ in each district, the probability that the election in district 1 decides who wins a majority in the legislature converges to zero as $2n + 1$ goes to infinity, so the effect that the local election has on national policy goes to zero.

The key questions is how “dilution” affects the nomination of candidates. It diminishes the benefit of primary voters to nominate an extreme candidate, but also the cost of this action: When nominating a candidate, primary voters will take into account that voters are less influenced by policy differences, and hence nominating a more extreme candidate results in fewer votes lost compared to the single district case.

The following Proposition 4 analyzes the interaction of these two effects and shows that, for any symmetric distribution of net valence shocks, the disciplining force of the general election completely vanishes in the limit. As the number of districts grows large, candidate positions become very extreme: The equilibrium candidate platforms converge to the Democratic and Republican primary voters’ ideal points, respectively.

**Proposition 4** Suppose that the median voter in the general election is located at zero in all of the $2n + 1$ identical districts, and that the valence shocks are iid draws from a continuous distribution that is symmetric around 0. Then the political equilibrium platforms $x_{D,i}$ and $x_{R,i}$ in each district converge to the ideal points of the primary voters ($-m$ for the Democrats, and $+m$ for the Republicans), as the number of districts grows large.

Intuitively, a model in which policy becomes less and less important relative to valence is isomorphic to a model in which the importance of policy remains constant, but the variance of the valence shocks goes to infinity. In such a model, the probability that starting from a symmetric situation, any small change in platform affects the winner of the election goes to zero so that parties have a strong incentive to nominate more extreme candidates.
It is important to understand that the extremism induced by dilution is based on an externality, but it is an externality between the general election median voters, not one between different local party median voters. As far as policy is concerned, the incentives of, say, the Democratic primary voters in different districts are perfectly aligned with each other as they have the same preferences, and all of them get to consume the same national policy.

In contrast, there is an important externality between the general election median voters in different districts: If the election in some district $i$ pits an extreme, but high valence candidate against a moderate, but lower valence candidate, then all median voters in other districts prefer the moderate candidate because district $i$’s representative affects the national policy that influences their utility, while valence is district-specific. In contrast, the median voter of district $i$ will likely be better off with the higher valence extremist. Because parties take the equilibrium response of general election voters (rather than their socially optimal response) into account when choosing their nominees, parties nominate relatively extreme candidates.

We next show that if candidates were nominated by the parties, then the candidates nominated in both parties are moderates with a policy platform close to the median voter’s ideal point of zero. Specifically, as in the case of a single district there is moderation, and if the standard deviation of the valence shock goes to zero, then both platforms converge to the general election median voter’s ideal point.

**Proposition 5** Suppose the same assumption as in Theorem 4 are satisfied, but in each party, all candidates are chosen by a central organization whose preference coincides with that of the party’s median voter $M_\ell$, $\ell \in \{D, R\}$. Then the following holds in the symmetric equilibrium as $n \to \infty$.

1. The divergence of party platforms from the median voter’s ideal point, $|\bar{x}_\ell|$ is bounded from above by $0.5 \sqrt{\pi/\phi(0)}$. Hence the polarization of candidates is limited.

2. Platforms $\bar{x}_D$ and $\bar{x}_R$ converge to zero as the standard deviation of the valence shock goes to zero.

**8 Discussion**

In this paper, we develop a theory of candidate nomination processes predicated upon the notion that members of the majority party in Congress collaboratively influence policy. This assumption is both empirically reasonable and substantively important for the results: Much of the existing literature implicitly assumes
that voters evaluate their local candidates based on the positions that they themselves take (but not on the party label under which they run). As we have shown, such a model cannot explain why some districts are “safe” for one party, and it implies a considerably more moderate nomination behavior than the one that arises as equilibrium in the American primary system in our model.

In our model, a candidate’s association with candidates of the same party that run in other districts both dilutes and contaminates the effect of the candidate’s own position. We have shown that this generates an incentive for voters to focus less on the candidates’ own position positions when deciding whom to vote for, and an incentive for parties to nominate more extreme candidates than in elections in which a candidate has autonomous policy influence (such as elections for executive leadership positions).

If empirical analysis of the effects of gerrymandering is based on applying such a naive Downsian model, results may appear counterintuitive and lead to incorrect inferences about the importance of gerrymandering. For example, in many of our examples, the ideal position of the district median voter does not affect the equilibrium position of candidates at the margin, but the total effect of gerrymandering on polarization in Congress may nevertheless be substantial (and actually larger than in the naive Downsian 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 (McCarty et al. 2009). 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.

While we assume a particular channel through which this influence works — namely that all majority party legislators have influence because of informational problems and the resulting specialization in different policy areas among legislators — our results would clearly be qualitatively stable if we were to assume a different causal mechanism, as long as every legislator of the majority party has some influence on policy, and legislators in the majority party have more influence on policy than those in the minority.
9 Appendix

9.1 Proofs

Lemma 1 Let \( x(\omega), k(\omega), \) and \( x'(\omega), k'(\omega) \) be a collection of policies and winning districts in each state \( \omega \), where \( \omega \) occurs with probability \( \mu(\{\omega\}) \) and \( \mu'(\{\omega\}) \), respectively. Then unless the functions \( \theta \mapsto E\left[u^\text{refine}_\theta(x, k, 0)\right]\) and \( \theta \mapsto E'\left[u^\text{refine}_\theta(x', k', 0)\right]\) are identical, they intersect at at most one point.

Proof. Note that there exists a collection of non-negative weights \( \beta_i, i \in I \) and \( \beta'_j, j \in J \) such that utility can be written as \( -\sum_{i \in I} \beta_i (\theta - x_i)^2 \) and \( -\sum_{j \in J} \beta'_j (\theta - x'_j)^2 \), where \( \sum_{i \in I} \beta_i = \sum_{j \in J} \beta'_j = 1 \). The derivative of both expressions with respect to \( \theta \) are therefore \( -2\theta + 2\sum_{i \in I} \beta_i x_i \) and \( -2\theta + 2\sum_{j \in J} \beta'_j x'_j \). Thus, the derivatives are either identical (if \( \sum_{i \in I} \beta_i x_i = \sum_{j \in J} \beta'_j x'_j \)) or one is strictly larger than the other for all \( \theta \). In the first case the curves are either identical, or they never intersect. In the second case, they can intersect at most once. ■

Proof of Proposition 1. Without loss of generality assume that \( M_{n+1} < 0 \). We first show existence. In the equilibrium, all districts in \( i \leq n + 1 \) vote for Democrats, and all districts \( i > n + 1 \) for Republicans. Policies are given by \( x_{D,i} = -m \) and \( x_{R,i} = m \) for all \( i = 1, \ldots, 2n + 1 \). In the proposed equilibrium Democrats win a majority and the implemented policy is \( -m \).

First, note that median voters have no incentive to switch to the other party. If the median voter in a district \( i \leq n + 1 \) switches to the Republican, then Republicans win a majority in the legislature and policy \( m \) will be implemented. Since \( M_i \leq M_{n+1} < 0 \) for all \( i \leq n + 1 \) this deviations makes median voters worse off. Further, median voters in districts \( i > n + 1 \) cannot improve by switching to the Democrats, since the implemented policy remains unchanged.

Next, we shows that it is not optimal for parties to nominated different candidates. First, in districts \( i \leq n + 1 \) median party voters receive their ideal candidate, and hence it is not optimal to deviate. A deviation of a Republican candidate in district \( i \) to \( M_i \) does not lead to election of the Republican. In particular, the utility from voting for the Democrat is \( -(M_i + m)^2 \), while the payoff from the Republican is \( -(1-\gamma)(M_i - m)^2 \). Since \( M_i < 0 \) the payoff from the Democrat exceeds the payoff from the Republican for small \( \gamma \).

Now consider districts \( i > n + 1 \). The Democrats already receive the ideal policy and have therefore no incentive to deviate. Similarly, the Republicans are already elected in this districts. Given that deviations in
one district cannot be observed in another district, a deviation by a Republican in a district \( i > n + 1 \) cannot win a majority in the legislature.

Now consider without loss of generality a pure strategy equilibrium in which Democrats win a majority in the legislature. Let \( S \) be the set of all district in which the Democrats win. Note that \( x_{D,i} \leq M_i \) for all \( i \in S \). Else, if \( x_{D,i} > M_i \) then by moving the policy to \( M_i \) the median is either unaffected, or also moves closer to \( M_i \). Even if the median policy is unaffected, \( M_i \) receives the \( \gamma \) weighted utility from policy \( x_{D,i} \) is therefore strictly better off.

Now consider the case where \( |S| = n + 1 \), i.e., Democrats have a one-vote majority. Suppose that \( x_{D,i} > -m \). The median Democratic party member would clearly be strictly better off if \( x_{D,i} \) is moved to the left, because the legislature’s policy moves to the left, or because of the \( \gamma \) weighted utility received from the district’s representative. Hence, to exclude such a deviation, the median voter \( M_i \) must be indifferent between the Democrat and Republican. This, however, implies that \( x_{R,i} = M_i \), else if \( x_{R,i} > M_i \) the Republican would be elected in the district by moving close to \( M_i \), and win a majority in the legislature.

Now suppose there are two districts \( i, j \in S \), \( i < j \) with \( x_{D,i}, x_{D,j} > -m \). Then \( x_{R,i} = M_i \) and \( x_{R,j} = M_j \) and the median voters in districts \( i \) and \( j \) are indifferent between the candidates. Let \( \bar{x}_D \) be the policy of Democratic legislature if both districts vote for the Democrat. If district \( k = i, j \) deviates and votes for the Republican the resulting policies are \( \bar{x}_R^k \), \( k = i, j \), where \( \bar{x}_R^i \geq \bar{x}_D \) since \( x_{R,i} > x_{R,j} \). Further, \( M_j \) strictly prefers \( \bar{x}_R^i \) to \( \bar{x}_R^j \), where the strict preference is generated by the \( \gamma \)-weighted utility of policies \( x_{R,i} \) versus \( x_{R,j} \). Median voter \( M_i \) is indifferent between \( \bar{x}_R^i \) and \( \bar{x}_D \). Thus, Lemma 1 implies that \( M_j \) strictly prefers \( \bar{x}_R^i \) to \( x_D \) (note that not all voters \( \theta \) can be indifferent between the parties for all \( \gamma \)).

Thus, \( M_j \) strictly prefers \( \bar{x}_R^i \) to \( \bar{x}_D \), a contradiction. Hence \( x_{D,i} > -m \) can hold in at most one district \( i \in S \).

Now suppose that \( |S| > n + 1 \). Again let \( \bar{x}_D \) be the implemented policy, and consider a district \( i \in S \). If \( x_{D,i} > \bar{x}_D \) then the Democrats would be better of losing the district, i.e., they could choose \( x_{D,i} = -m \) in which case the Democrats in district \( i \) would be strictly better whether or not they would get elected. Thus, the most moderate equilibrium is one which all districts go for the Democrat, and \( x_{D,i} = M_n \) for all \( i \geq n \).

Note that both results imply there do not exist equilibria in which all candidates are located at the district medians, \( M_i \) or at the median of the median district, \( M_{n+1} \).
Proof of Proposition 2. Let \( p_k \) be the probability that \( k \) districts out of \( n \) vote for a Republican. Consider a particular district and suppose that the Republican in the district offers policy \( y \) while in all remaining districts Democrats are located at \(-\bar{\theta}\) and Republicans at \(\bar{\theta}\). We again use the refinement that policy \( y \) is implemented with probability \( \gamma \) where \( \gamma \rightarrow 0 \). Then if the median voter, now located at \( M \) votes for a Democrat, the payoff is given by

\[
-\left( \sum_{k=0}^{n} p_k \right) (M + \bar{\theta})^2 - \left( \sum_{k=n+1}^{2n} p_k \right) (M - \bar{\theta})^2
\]

If the median voter elects the Republican then

\[
-\left( \sum_{k=0}^{n} p_k \right) (M + \bar{\theta})^2 - \left( \sum_{k=n}^{2n} p_k \right) ((1-\gamma)(M - \bar{\theta})^2 + \gamma (M - y)^2).
\]

The median voter prefers the Democrat to the Republican if the payoff in (6) is higher than the payoff in (7). Thus,

\[
p_n (M + \bar{\theta})^2 + \left( \sum_{k=n+1}^{2n} p_k \right) (M - \bar{\theta})^2 \leq \left( \sum_{k=n}^{2n} p_k \right) ((1-\gamma)(M - \bar{\theta})^2 + \gamma (M - y)^2),
\]

which simplifies to

\[
4\theta p_n M + \gamma \left( \sum_{k=n}^{2n} p_k \right) (M - \bar{\theta})^2 \leq \gamma \left( \sum_{k=n}^{2n} p_k \right) (M - y)^2
\]

Thus, the Democrat is preferred by a median voter \( M \) if \( M \leq M(\bar{\theta}, y) \), where

the probability that the Democrat is elected is given by

\[
M(\bar{\theta}, y) = \frac{1}{2} \frac{\gamma \left( \sum_{k=n}^{2n} p_k \right) (\bar{\theta}^2 - y^2)}{\gamma \left( \sum_{k=n}^{2n} p_k \right) (\bar{\theta} - y) - 2 p_n \bar{\theta}}.
\]

The Republican primary voter therefore maximizes

\[
-\Phi(M(\bar{\theta}, y)) \left( \sum_{k=0}^{n} p_k (m + \bar{\theta})^2 + \sum_{k=n+1}^{2n} p_k (m - \bar{\theta})^2 \right)
\]

\[
- (1 - \Phi(M(\bar{\theta}, y))) \left( \sum_{k=0}^{n-1} p_k (m + \bar{\theta})^2 + \sum_{k=n}^{2n} p_k ((1-\gamma)(m - \bar{\theta})^2 + \gamma (m - y)^2) \right).
\]
The first order condition evaluated at \( y = \bar{\theta} \) is

\[
-\gamma \frac{\sum_{k=n}^{2n} p_k}{2p_n} \phi(0) \left( \frac{p_n(m + \bar{\theta})^2 - p_n(m - \bar{\theta})^2}{2} \right) - (1 - \Phi(0)) \left( \sum_{k=n}^{2n} p_k \right) 2\gamma(\bar{\theta} - m) = 0
\]

Note that the solution of (10) is independent of \( \gamma \). Further, we can simplify the equation to get

\[
2\phi(0)\bar{\theta}m = 2(1 - \Phi(0))(m - \bar{\theta}).
\]

Since \( \Phi(0) = 1/2 \) we get

\[
\bar{\theta} = \frac{m}{1 + 2\phi(0)m},
\]

which is independent of \( n \), and is the same solution as in the single-district case. Hence, there is no policy dilution. Also note that the policy \( \bar{\theta} \) is bounded above by \( 1/(2\phi(0)) \) even has \( m_R \) becomes arbitrarily large.

\[ \blacksquare \]

**Proof of Proposition 3.** We first show that in any pure strategy equilibrium that satisfies the refinement, \( x_{D,i} = -m \) for all except possible one district \( i \leq n \).

Suppose there exists \( i < j \leq n \) with \( x_{D,i} > -m \). As in the proof of Proposition 1 we can argue that \( x_{R,i} = M_i \) and \( x_{R,j} = M_j \). In particular, the median voter in the districts must be indifferent between the candidates, else the Democrat could move the position further to the left. For the same reason the Republicans must be located at the median, else by moving towards the median they would get elected.

Let \( \bar{x}(\omega) \), \( \omega = \omega_D, \omega_R \) be the lottery over policies if districts \( i \) and \( j \) vote for the Democrat. If, instead, \( M_i \) votes for the Republican, then the resulting policy vector is \( \bar{x}'(\omega) \), \( \omega = \omega_D, \omega_R \). If, instead, district \( j \) elects the Republican, then the resulting lottery \( \bar{x}'(\omega) \), \( \omega = \omega_D, \omega_R \) simply replaces \( x_{i,D} \) by \( x_{j,D} \). By assumption \( M_i \) is indifferent between them. Thus, Lemma 1 implies that \( M_j \) strictly prefers lottery \( \bar{x}'(\omega) \) to \( \bar{x}(\omega) \). However, in addition he prefers lottery \( \bar{x}'(\omega) \) to \( \bar{x}(\omega) \). As a consequence \( M_j \) is not indifferent between the candidates, a contradiction. Finally, note that if \( n - 1 \) districts select the \( -m \) then \( -m \) is the median if \( n - 1 \geq 3 \).

The argument that \( x_{R,i} = m \) for \( n - 1 \) districts \( i > n + 1 \) follows similarly.

Now suppose districts 1 to \( n \) elect a Democrat at \( -m \), and districts \( i > n + 1 \) a Republican a \( m \). Then the
cutoff primary voter in district $n + 1$ who is indifferent by the candidates satisfies

$$-(1 - \gamma)(M + m)^2 - \gamma(M - x_D)^2 = -(1 - \gamma)(M - m)^2 - \gamma(M - x_R)^2.$$ 

Thus,

$$M(x_D, x_R) = \frac{1}{2 \gamma(x_R - x_D) + 2(1 - \gamma)m}.$$ 

The Democratic primary voter solves

$$\max_{x_D} -\Phi(M(x_D, x_R)) \gamma(-m - x_D)^2 - (1 - \Phi(M(x_D, x_R))) (1 - \gamma)4m^2 + \gamma(-m - x_R)^2.$$ 

The first order condition is

$$-\phi(M) \frac{\partial M(x_D, x_R)}{\partial x_D} \left(\gamma \left((m + x_D)^2 - (m - x_R)^2\right) - (1 - \gamma)4m^2\right) - 2\gamma \Phi(M)(m + x_D) = 0.$$ 

Note that

$$\frac{\partial}{\partial x_D} M(x_D, x_R) \bigg|_{\gamma=0} = -\frac{1}{2} \frac{x_D}{m}.$$ 

It is also immediate that $M(x_D, x_R) = 0$ for $\gamma = 0$. Thus, dividing both sides of (12) by $\gamma$ and then setting $\gamma = 0$ yields

$$\phi(0)(-x_D)m = \Phi(0)(m + x_D).$$ 

The Republican primary voter solves

$$\max_{x_R} -\Phi(M(x_D, x_R)) \left((1 - \gamma)4m^2 + \gamma(m - x_D)^2\right) - (1 - \Phi(M(x_D, x_R))) \gamma(m - x_R)^2.$$ 

The first order condition is

$$-\phi(M) \frac{\partial M(x_D, x_R)}{\partial x_R} \left((1 - \gamma)4m^2 + \gamma((m - x_D)^2 - (m - x_R)^2)\right) + 2\gamma (1 - \Phi(M))(m - x_R) = 0.$$ 

It follows that

$$\frac{\partial}{\partial x_R} M(x_D, x_R) \bigg|_{\gamma=0} = -\frac{1}{2} \frac{x_R}{m}.$$
Again, dividing by $\gamma$, setting $\gamma = 0$ and using the fact that $M = 0$ when $\gamma = 0$, yields

$$\phi(0)x_Rm = (1 - \Phi(0)) (m - x_R). \tag{14}$$

This implies (5).

It is easy to verify the second order. Now suppose a median voter in district $i < n + 1$ deviates and elects a Republican. Then since $n \geq 3$ the resulting median policy is $m$ independent of who wins in district $n + 1$. Since $M_i < 0$, the median voter is worse off. Similarly, $M_i > 0$ for $i > n + 1$ implies that median voters in these districts are worse off switching to the Democrats. Finally, note that median Democratic party members in districts $i < n$ receive their ideal policy $m$, and similarly, Republicans receive their ideal policy $m$ in districts $i > n + 1$. Hence, it is not optimal to deviate, and we have therefore an equilibrium.

If $M_3 = 0$ then $\Phi(0) = 0.5$, hence the distance between the policies is the same as in Proposition 2, i.e., as in the case where all districts are symmetric.

Using the fact that $F$ is symmetric and hence $f'(0) = 0$ and $F(0) = 0.5$, it is easy to verify that

$$\frac{\partial(x_R - x_D)}{\partial M_3} \bigg|_{M=0} = 0.$$  

Hence, if $\frac{\partial(x_R - x_D)}{\partial M_3} \bigg|_{M=0} < 0$, $M = 0$ is a local maximum, and polarization, i.e., the distance between the policies is smaller in a neighborhood of $M = 0$. The reverse is true if $\frac{\partial(x_R - x_D)}{\partial M_3} \bigg|_{M=0} > 0$.

Again, using the fact that $f'(0) = 0$ and $F(0) = 0.5$ it follows that

$$\frac{\partial^2(x_R - x_D)}{\partial M_3^2} \bigg|_{M=0} = -\frac{4m^2 \left(8f(0)^3 + f''(0) + 2mf''(0)f(0)\right)}{(1 + 2mf(0))^3}.$$  

Thus, the second derivative is positive, and polarization increase, if and only if $8f(0)^3 + f''(0) + 2mf''(0)f(0) < 0$, i.e., if $m > \bar{m}$, where

$$\bar{m} = -\left(\frac{4f(0)^2}{f''(0)} + \frac{1}{2f(0)}\right).$$

If $\bar{m} > 0$ then for $m < \bar{m}$ polarization decreases. \hfill \blacksquare

**Proof of Proposition 4.** In the general election the median voter selects the candidate that maximizes his utility. In the proposed symmetric equilibrium, each candidate wins with probability 0.5 in each district.
Further, $x_{D,i} = -\bar{x}$ and $x_{R,i} = \bar{x}$ for all districts $i$. Now consider a particular district $i$, and suppose that the Republican in district $i$ deviates to policy $y$. Let $p_k = \binom{2n}{k}0.5^{2n}$, be the probability that $k$ of the remaining districts go to Republicans. Then the payoff of the median voter in district $i$ (who is located at 0) from voting Republican is

$$-\left(\sum_{k=0}^{n-1} p_k\right)\bar{x}^2 - \sum_{k=n}^{2n} p_k \left((1 - \gamma)\bar{x}^2 + \gamma y^2\right) + \epsilon,$$

while the payoff from voting for the Democrat is $-\bar{x}^2$. Thus, the valence shock at which the median voter is indifferent between the candidates is given by

$$\epsilon(y) = \gamma (y^2 - \bar{x}^2) \sum_{k=n}^{2n} p_k.$$

As a consequence, the probability that the Democratic candidate is elected in district $i$ is given by $\Phi(\epsilon(y))$.

Now consider the choice of median Republican voter $m_R$. Let $\psi(t) = t\phi(t)$, and $\Psi(t) = \int_{-\infty}^{t} \psi(x) \, dx$. Note that, since $\epsilon$ has mean zero and is symmetrically distributed, $-\Psi(t) = \int_{t}^{\infty} \psi(x) \, dx$. The Republican’s expected valence given that $\epsilon \geq t$ (i.e., if he is elected) is therefore given by $-\Psi(t)/(1 - \Phi(t))$.

The utility of the median Republican primary voter $m_R$ from choosing a candidate with position $y$ is given by the probability that the Democrat wins times his policy utility from the Democrat, plus the probability that the Republican wins times the sum of his policy utility and the valence from the Republican.

$$-\Phi(\epsilon(y)) \left(\sum_{k=0}^{n} p_k (m + \bar{x})^2 + \sum_{k=n+1}^{2n} p_k (m - \bar{x})^2\right)$$

$$- \left(1 - \Phi(\epsilon(y))\right) \left(\sum_{k=0}^{n-1} p_k (m + \bar{x})^2 + \gamma \sum_{k=n}^{2n} p_k (m - y)^2 + (1 - \gamma) \sum_{k=n}^{2n} p_k (m - \bar{x})^2\right)$$

$$- \Psi(\epsilon(y)).$$

In a symmetric equilibrium equilibrium, (15) is maximized when $y = \bar{x}$. Thus, taking the derivative of (15) with respect to $y$, substituting $\bar{x}$ for $y$ yields

$$-\phi(0) \epsilon'(\bar{x}) p_n \left((m + \bar{x})^2 - (m - \bar{x})^2\right) + 2(1 - \Phi(0))\gamma \sum_{k=n}^{2n} p_k (m - \bar{x}) - \epsilon'(\bar{x})\psi(0) = 0.$$
In the symmetric equilibrium $\Phi(0) = 0.5$. Further, $\psi(0) = 0$. Thus, (16) simplifies to
\begin{equation}
8p_n\phi(0)m\bar{x}^2 = m - \bar{x}.
\end{equation}

Note that $\lim_{n \to \infty} p_n = 0$. Hence, (17) implies that as $n$ becomes large, $\bar{x}$ converges to $m$, i.e., both the Republican and Democrat primary choose extreme candidates when $n$ is large.

**Proof of Proposition 5.** We consider symmetric equilibria, where all Republicans locate at a point $\bar{x}$ and all Democrats at $-\bar{x}$. Let $p_k = \binom{2n}{k}q^k(1 - q)^{2n-k}$ be the probability that $k$ of the remaining $2n$ districts go for Republicans. Suppose that the Republicans move all candidates to position $y$. Then if the median voter in district $i$ selects a Democrat, the resulting utility is $-\sum_{k=0}^{n} p_k \bar{x}^2 - \sum_{k=n+1}^{2n} p_k y^2$. If the median voter selects the Republican, then the median voter’s utility is $-\sum_{k=0}^{n-1} p_k \bar{x}^2 - \sum_{k=n+1}^{2n} p_k y^2 + \epsilon$. Thus, the median voter selects the Democrat if $\epsilon \leq p_n(y^2 - \bar{x}^2)$. Therefore, the probability that a district elects the Republican is
\begin{equation}
q = 1 - \Phi\left(p_n(y^2 - \bar{x}^2)\right).
\end{equation}

Let $\hat{p}_k = \binom{2n+1}{k}q^k(1 - q)^{2n+1-k}$ be the probability that $k$ of the $2n + 1$ districts elect a Republican. Then the Republican party chooses $y$ to maximize expected utility given by
\begin{equation}
-\sum_{k=0}^{n} \hat{p}_k(-\bar{x} - m)^2 - \sum_{k=n+1}^{2n+1} \hat{p}_k(y - m)^2 - \Psi\left(p_n(y^2 - \bar{x}^2)\right).
\end{equation}

Taking the derivative of (19) with respect to $y$ at $y = \bar{x}$ yields
\begin{equation}
-\sum_{k=0}^{n} \frac{\partial \hat{p}_k}{\partial y}(-\bar{x} - m)^2 - \sum_{k=n+1}^{2n+1} \frac{\partial \hat{p}_k}{\partial y}(\bar{x} - m)^2 - 2 \sum_{k=n+1}^{2n+1} p_k(\bar{x} - m) - \psi(0)\frac{\partial p_n(y^2 - \bar{x}^2)}{\partial y} = 0.
\end{equation}

Taking the derivative of (18) with respect to $y$ at $\bar{x}$ yields
\begin{equation}
\frac{\partial q}{\partial y} = -2\phi(0)\bar{x}p_n.
\end{equation}

33
Further,
\[
\frac{\partial \hat{p}_n}{\partial y} = \left(\frac{2n+1}{k}\right) \left(\frac{kq}{k-1} (1-q)^{2n+1-k} - (2n+1-k)q^k (1-q)^{2n-k}\right) \frac{\partial q}{\partial y}
\]
\[
= -\left(\frac{k}{q} \frac{2n+1-k}{1-q}\right) 2\phi(0) \bar{x} \hat{p}_n
\]
\[
= (2n-2k+1)\phi(0) \bar{x} \hat{p}_n,
\]
where the last equality follows since \( q = 0.5 \) in the symmetric equilibrium. Note that \( \psi(0) = 0 \). Thus, (20) and (22) imply
\[
-\phi(0) \bar{x} p_n \left(\sum_{k=0}^{n} (2n-2k+1) \hat{p}_k (-\bar{x} - m)^2 + \sum_{k=n+1}^{2n+1} (2n-2k+1) \hat{p}_k (\bar{x} - m)^2\right) - 2 \sum_{k=n+1}^{2n+1} \hat{p}_k (\bar{x} - m) = 0.
\]
Since \( q = 0.5 \) it follows that
\[
-\phi(0) \bar{x} p_n \left(\sum_{k=0}^{n} (2n-2k+1) \hat{p}_k (-\bar{x} - m)^2 + \sum_{k=n+1}^{2n+1} (2n-2k+1) \hat{p}_k (\bar{x} - m)^2\right) - 2 \sum_{k=n+1}^{2n+1} \hat{p}_k (\bar{x} - m) = 0.
\]
and hence
\[
-4\phi(0) \bar{x}^2 m p_n \left(\sum_{k=0}^{n} (2n-2k+1) \hat{p}_k \right) - 2 \sum_{k=n+1}^{2n+1} \hat{p}_k (\bar{x} - m) = 0. \tag{23}
\]
Note that
\[
\sum_{k=0}^{n} k \hat{p}_k = \sum_{k=0}^{n} k \frac{(2n+1)!}{k! (2n+1-k)!} 2^{-(2n+1)} = (2n+1) \sum_{k=0}^{n} \frac{(2n)!}{(k-1)! (2n-(k-1))!} 2^{-(2n+1)}
\]
\[
= \frac{2n+1}{2} \sum_{k=0}^{n-1} \binom{2n}{k} 2^{-2n} = \frac{2n+1}{4} (1 - p_n).
\]
Thus,
\[
\sum_{k=0}^{n} (2n-2k+1) \hat{p}_k = p_n \left( n + \frac{1}{2} \right).
\]
Using Stirling’s formula\(^{13}\) we get
\[
p_n \sqrt{n} \sim \frac{\sqrt{2\pi(2n)^{2n+\frac{1}{2}}} e^{-2n}}{n^{2n+1} e^{-2n}} 2^{-2n} \sqrt{n} = \frac{1}{\sqrt{\pi}}
\]
\(^{13}\)Stirling’s formula states that \( n! \sim \sqrt{2\pi n^{n+0.5}} e^{-n} \).
As a consequence,
\[ \lim_{n \to \infty} \sum_{k=0}^{n} (2n - 2k + 1) \hat{p}_k = \lim_{n \to \infty} p_n \left( n + \frac{1}{2} \right) = \frac{1}{\pi}. \]
This and the fact that \( \sum_{k=n+1}^{2n+1} \hat{p}_k = 0.5 \) implies that taking the limit for \( n \to \infty \) in (23) yields
\[ -\frac{4}{\pi} \phi(0) \bar{x}^2 m - (\bar{x} - m) = 0. \]  
(24)

As a consequence,
\[ \bar{x} = -\frac{\pi + \sqrt{\pi^2 + 16\pi \phi(0)m^2}}{8\phi(0)m}. \]
(25)

Clearly, \( 0 < \bar{x} < m \). Further, if \( m \) converges to zero, then \( \bar{x} \) converges to zero. The same is true if \( \phi(0) \) becomes arbitrarily large. If \( m \) becomes arbitrarily large, then \( \bar{x} = 0.5 \sqrt{\pi/\phi(0)}. \]

9.2 Results for the one district model

Proposition 6 Consider the distributed policy model. Suppose there is only one district, that \( \Phi \) is symmetric, and that the hazard rate satisfies
\[ \lim_{x \to \infty} \frac{\phi(x)}{(1 - \Phi(x))} > 0. \] Then

1. Both parties nominate candidates who are located at different sides of the general election median voter, but are more moderate than the decisive voter in their respective party primaries: \( M_D < x_D < 0 < x_R < M_R \).

2. If \( M_D \to -\infty \) or \( M_R \to \infty \) (i.e., the parties’ medians diverge), the equilibrium candidate positions remain bounded, i.e., there exists \( L > 0 \) such that \( |x_D|, |x_R| < L \) for all \( M_D, M_R \).

3. Suppose that the valence shock is \( k \varepsilon \) and that \( \phi(\cdot) > 0 \) in a neighborhood of zero. Then, as \( k \to \infty \), each candidate’s winning probability converges to \( 1/2 \).

4. Suppose that the valence shock is \( \varepsilon/k \). Let \( \phi \) be single-peaked, i.e. \( \phi'(x) > 0 \) for all \( x < 0 \) and \( \phi'(x) < 0 \) for \( x > 0 \). Then as \( k \to \infty \) each candidate’s winning probability converges to \( 1/2 \).

Proof of Proposition 6. Let \( x_D \) and \( x_R \) denote the candidates’ policy positions. Then the median voter votes for the Democrat if \( -x_D^2 \geq -x_R^2 + \varepsilon \), and hence the probability that the Democrat wins is \( \Phi(x_R^2 - x_D^2) \).
Thus, in the primary election, the Democratic party median \(M_D\) maximizes his expected utility,

\[
\max_{x_D} -\Phi(x_R^2 - x_D^2)(x_D - M_D)^2 - (1 - \Phi(x_R^2 - x_D^2))(x_R - M_D)^2 - \Psi(x_R^2 - x_D^2).
\] (26)

Similarly, the Republican party median \(M_R\) solves

\[
\max_{x_R} -\Phi(x_R^2 - x_D^2)(x_D - M_R)^2 - (1 - \Phi(x_R^2 - x_D^2))(x_R - M_R)^2 - \Psi(x_R^2 - x_D^2).
\] (27)

The first order conditions are therefore

\[
2x_D\phi(x_R^2 - x_D^2)\left((x_D - M_D)^2 - (x_D - M_D)\right) - 2\Phi(x_R^2 - x_D^2)(x_D - M_D) + 2x_D(x_R^2 - x_D^2)\phi(x_R^2 - x_D^2) = 0
\] (28)

\[
-2x_R\phi(x_R^2 - x_D^2)\left((x_D - M_R)^2 - (x_R - M_R)\right) - 2(1 - \Phi(x_R^2 - x_D^2))(x_R - M_R) - 2x_R(x_R^2 - x_D^2)\phi(x_R^2 - x_D^2) = 0.
\] (29)

It is useful to rewrite the first-order conditions as

\[
2x_DM_D(x_R - x_D)\phi(x_R^2 - x_D^2) - \Phi(x_R^2 - x_D^2)(x_D - M_D) = 0
\] (30)

\[-2x_RM_R(x_R - x_D)\phi(x_R^2 - x_D^2) - (1 - \Phi(x_R^2 - x_D^2))(x_R - M_R) = 0.
\] (31)

Since the first term of (30) is positive, it follows immediately that \(x_D > M_D\), i.e., the Democratic primary median prefers to nominate a candidate who is closer to the general election median than he himself. Similarly, \(x_R < M_R\). Furthermore, it also follows immediately that \(x_D = x_R = 0\) is not a solution.

Next, suppose that \(M_R \to \infty\), while \(M_D\) remains bounded. We show that \(x_R\) remains bounded. Dividing both sides of (31) by \(M_R\phi(\cdot)\) we get

\[-2x_R(x_R - x_D) + \frac{1 - \Phi(x_R^2 - x_D^2)}{\phi(x_R^2 - x_D^2)} \left(1 - \frac{x_R}{M_R}\right) = 0
\] (32)

Note that \(x_D^2 \leq M_D^2\). Therefore, the assumption that \(\lim_{x \to \infty} \phi(x)/(1 - \Phi(x)) > 0\) and the fact that \(\left|\frac{x_R}{M_R}\right| < 1\) implies that the second summand in (32) remains bounded as \(M_R \to \infty\). Hence, (32) \(x_R\) must remain bounded.

The argument for \(M_D \to \infty\) and \(M_R\) bounded is similar. In particular, dividing both sides of (30) by
Let $M_D \phi(\cdot)$ we get
\[ 2x_D(x_R - x_D) + \frac{\Phi(x_R^2 - x_D^2)}{\phi(x_R^2 - x_D^2)} \left(1 - \frac{x_D}{M_D}\right) = 0, \]
which by symmetry of the distribution $\Phi$ implies
\[ 2x_D(x_R - x_D) + \frac{1 - \Phi(x_D^2 - x_R^2)}{\phi(x_D^2 - x_R^2)} \left(1 - \frac{x_D}{M_D}\right) = 0. \tag{33} \]
Again, the summand in (33) remains bounded, and as a consequence $x_D$ must remain bounded as $M_D \to \infty$.

Now, consider the case where both $M_D$ and $M_R$ go to infinity. Suppose that $\limsup_{n \to \infty} x_R^2 - x_D^2 \geq 0$. Then (32) again implies that the second summand is bounded. Hence $x_R(x_R - x_D)$ remains bounded, i.e., both $x_R$ and $x_D$ are bounded. If $\limsup_{n \to \infty} x_R^2 - x_D^2 < 0$, the result follows from (33).

Next, note that changing the valence from $\epsilon$ to $a\epsilon$ for some constant $a > 0$ is equivalent to keeping the valence shock the same and changing the ideal points to $\sqrt{a}M_D$ and $\sqrt{a}M_R$, and strategies to $\sqrt{a}x_D$ and $\sqrt{a}x_R$. This is the case because $(y - x_D)^2 \geq (y - x_R) + \epsilon$ if and only if $(\sqrt{a}y - \sqrt{a}x_D)^2 \geq (\sqrt{a}y - \sqrt{a}x_R) + a\epsilon$.

First, consider the case where the valence shock is $k\epsilon$. This is equivalent to a model in which the ideal points are $M_D/\sqrt{k}$ and $M_R/\sqrt{k}$ and the valence shock is $\epsilon$. If $k \to \infty$ this means that that in (30) and (31) the first summands converge to zero. Since $\phi(x) > 0$ for $x$ near zero, it follows that $x_D = x_R = 0$. Thus, each candidate wins with probability $1/2$.

Finally, suppose that the valence shock is $\epsilon/k$. This is equivalent to a model in which the ideal points are $M_D \sqrt{k}$ and $M_R \sqrt{k}$ and the valence shock is $\epsilon$. If $k \to \infty$ this means that the ideal points go to infinity. As shown above, $x_D$ and $x_R$ remain bounded. Taking the limit for $k \to \infty$ in (32) and (33) therefore imply
\[ 2x_D(x_R - x_D) + \frac{1 - \Phi(x_D^2 - x_R^2)}{\phi(x_D^2 - x_R^2)} = 0, \quad \text{and} \quad 2x_R(x_R - x_D) + \frac{1 - \Phi(x_R^2 - x_D^2)}{\phi(x_R^2 - x_D^2)} = 0. \tag{34} \]

Subtracting the second equation in (34) from the first and diving the result by 2 yields
\[ x_D^2 - x_R^2 = \frac{\Phi(x_R^2 - x_D^2) - \frac{1}{2}}{\phi(x_R^2 - x_D^2)}, \tag{35} \]
Let $h(x) = x\phi(x) - \left(\Phi(x) - \frac{1}{2}\right)$. Then $h(0) = 0$. We now show that 0 is the only solution of the equation $h(x) = 0$.

Note that $h'(x) = \phi'(x)x$. Thus, the assumption that $\phi'(x) > 0$ for $x < 0$, and $\phi'(x) < 0$ for $x > 0$ implies
that \( h'(x) < 0 \) for all \( x \neq 0 \). As a consequence, \( x = 0 \) is the unique solution of \( h(x) = 0 \). Thus, (35) implies \( x_R^2 - x_D^2 = 0 \). As a consequence, the winning probability is 0.5. ■

**Proposition 7** Suppose there is uncertainty about the median given by a family of distributions \( \Phi_n \) on \([-a_n, a_n]\), where \( \lim_{n \to \infty} a_n = 0 \). Suppose that \( \phi'_n \) is uniformly bounded, i.e., there exists \( M > 0 \) such that \( |\phi'_n(x)| < M \) for all \( x \in [-a_n, a_n] \) and for all \( n \). Suppose that the Republican has a net valence advantage \( \varepsilon > 0 \), with \( \varepsilon < 4M^2_R \). Then for all sufficiently large \( n \) the following is an equilibrium. \( x_D = -a_m, \) \( x_R = -a_n + \sqrt{\varepsilon} \) and candidate R wins with probability 1.

**Proof of Proposition 7.** Note that

\[
M(x_D, x_R) = \frac{1}{2} \left( x_R + x_D - \frac{\varepsilon}{x_R - x_D} \right). \tag{36}
\]

The Democratic primary voter therefore solves

\[
\max_{x_D} -\Phi(M(x_D, x_R))(x_D - M_D)^2 - (1 - \Phi(M(x_D, x_R)))(x_R - M_D)^2, \tag{37}
\]

while the Republican primary voters solves

\[
\max_{x_R} -\Phi(M(x_D, x_R))(x_D - M_R)^2 - (1 - \Phi(M(x_D, x_R)))(x_R - M_R)^2. \tag{38}
\]

First, note if \( x_R = -a_n + \sqrt{\varepsilon} \) then the Democrat cannot attract any voters by changing his strategy. In particular, \(-M - x_D)^2 \leq (M + a_n - \sqrt{\varepsilon})^2 + \varepsilon \) for all \( M \geq a_n \). Hence, the Democrat cannot improve by deviating.

Now consider the Republican. The derivative of (38) with respect to \( x_R \) for which \(-a < M(x_D, theta_R) < a_n \) is given by

\[
-\frac{(x_R - x_D)^2 + \varepsilon}{2(x_R - x_D)^2} \phi(M(x_D, x_R)) \left( 2M_R(x_R - x_D) + x_D^2 - x_R^2 \right) - 2 \Phi(M(x_D, x_R))(x_R - M_R) = 0. \tag{39}
\]

Note that \( 2M_R(x_R - x_D) + x_D^2 - x_R^2 \) is bounded away from zero for sufficiently large \( n \). Given that \( \phi'_n \) is uniformly bounded it follows that \( \phi(M) \to \infty \) for all \( x_R \) for which \(-a < M(x_D, theta_R) < a_n \). Thus, the first summand of (39) goes to \(-\infty \) as \( n \to \infty \). The second summand remains bounded. Hence, it is optimal to choose \( x_R \) such that \( M(x_D, x_R) = -a_n \). Hence, \( x_R = -a_n + \sqrt{\varepsilon} \). ■
References


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


