AN INFORMATIONAL RATIONALE
FOR POLITICAL PARTIES\textsuperscript{1}

James M. Snyder, Jr.
Departments of Political Science and Economics
Massachusetts Institute of Technology

Michael M. Ting
Department of Political Science
University of North Carolina at Chapel Hill

December, 2000

\textsuperscript{1}We thank Tim Fedderson, Scott Desposato, Dhammika Dharmapala, and seminar participants at the University of North Carolina, Stanford, and UCLA for helpful comments.
Abstract

What role do parties play in electoral competition? Why would parties exert discipline on their members? This paper links both questions by modeling parties as informative “brands” to voters. Voters across a large number of constituencies are assumed to be risk averse and incompletely informed about candidate ideal policies, and candidates are unable to commit to a declared policy platform. In this environment, parties can play a critical role by aggregating ideologically similar candidates and signaling their preferences to voters. This signaling is effective because party membership imposes costs, which screen out candidates whose preferences are not sufficiently close to the party’s platform. We find that when party labels are very informative, the parties’ platforms converge. When party labels are less informative, however, platforms diverge, because taking an extreme position allows a party to reduce the variance of its members’ preferences. As parties become less able to impose discipline on their members, or less able to screen out certain types of candidates, their platforms move further apart.
1. Introduction

Much of the current theorizing about political parties in the U.S. emphasizes the importance of parties as producers of political brand names. Proponents claim that these brand names are valuable both to voters and to candidates—they help voters make decisions, and they help candidates win elections. Downs (1957) made this argument in his early work, and recent studies have given it new life (e.g., Kiewiet and McCubbins, 1991; Cox and McCubbins, 1993; Aldrich, 1995).

To our knowledge, however, no one has provided a satisfactory formalization of party labels as brand names. Filling this gap is important for two reasons. First, several theoretical questions remain largely unanswered. Under what circumstances should voters use party labels in deciding how to vote? What kinds of equilibrium policies should result when voters do rely on party labels? Will outcomes be more efficient? Under what circumstances, and on what issues, will parties try to build and maintain differentiated brand names?

Second, the lack of a well-specified family of models with clear predictions hampers empirical research. Currently, a large amount of effort is being devoted to measuring the effects of party on various legislative behaviors, especially committee assignments, agenda control, and roll call voting.\(^1\) Moreover, the scholars conducting this research often invoke the brand name argument to explain their findings. However, none of the research provides a rigorous test of parties-as-brand-names against other possibilities. Further progress in this area may require tying the empirical work more closely to theory.

This paper analyzes a simple model of party brand names that provides tentative answers to some of the questions above. The underlying framework may also prove useful for developing richer models that lead to further insights and help guide empirical research.

We begin by observing two major differences between party labels and interest group endorsements. The first difference is that parties, unlike labor unions, business associations, and other special interest groups, do not have “natural” issue positions. A union endorsement means a candidate is relatively pro-labor, a Chamber of Commerce endorsement means a candidate is relatively pro-business, a Sierra Club endorsement means a candidate is relatively pro-environment, and so on. It is easy to infer the positions of these groups, in large

part because they are *constrained* by their members’ common interests on many non-political (or not-necessarily-political) matters. Unions and businesses bargain over wages and working conditions, businesses use the Chamber of Commerce for advertising and information, and most Sierra Club members share an interest in outdoor recreation. These common interests anchor the groups. Parties do not face such constraints, however, because they are more broad-based organizations that exist only for political purposes. A “Party X” label carries no automatic or natural meaning. Indeed, the spatial theory of electoral competition is predicated on the idea that parties are free to change their positions.²

Of course, in practice most party labels do carry relatively precise meanings. Democrats tend to be liberal, Republicans are conservative, and Libertarians are fiscally conservative but socially liberal. The question is, *how is this sustained as an equilibrium phenomenon?*

The second feature distinguishing parties from interest groups is the amount of interaction they have with candidates. Candidates have only “arms length” relations with most of the interest groups that endorse them. On the other hand, candidates interact intensely and frequently with party officials, party activists, and other candidates running under the party label. They are members of the party, in many cases the most important members. This is a natural state of affairs. Parties provide many of the resources politicians use to win office, including detailed knowledge of constituency characteristics, campaign contributions and lists of other likely contributors, and campaign workers. Each party’s winning candidates work together to formulate policy in national, state, and local governments. Parties also play a key role in determining the career path of candidates who want to be career politicians.

This intense and frequent interaction means that party membership can be costly. More importantly, *the costs will be different for different types of candidates.* Candidates whose personal policy preferences or ideologies are close to the preferences held by the bulk of a party’s members will tend to incur relatively low costs from party membership. On the other hand, candidates with policy preferences or ideologies that are quite far from the center of a party will face much higher costs, and such candidates may decide not to join the party as a result. These differential costs become especially relevant when we consider that one option available to potential candidates is not to run at all. Given their options outside politics, the costs of joining the “wrong” party can be quite large, even if joining the wrong party

²Some parties are probably constrained on some issues, especially parties created by interest groups or social movements. Examples are the Prohibition Party and Right-to-Life Party in the U.S., certain labor and socialist parties in Europe, and the Green parties around the world.
carries electoral benefits.

Our model brings these two ideas together. Party labels may be valuable to candidates and voters because they provide low-cost information about the preferences of groups of candidates across multiple offices. However, the message conveyed by a party label is determined by the set of candidates who run under it. As a result, a party’s label is informative only if the types of candidates who run under it are limited. Parties might be able to restrict access to the label. Alternatively, sorting into parties might occur as a result of candidates’ own choices. We posit that party membership entails costs, and these costs vary across candidates depending on the candidates’ preferences. In equilibrium, only candidates of a similar ideological stripe are willing to signal their type by running under a party’s banner. Parties can then serve as effective screening devices, as in Spence (1974).\(^3\)

Our model also provides a rationale for party discipline. In equilibrium, the differential costs born by different types of candidates are precisely what gives a party label its value. To increase the value of their labels, parties will typically want to impose “discipline” on the candidates who run under their label whenever they can. Alternatively, parties will want to “test” candidates for ideological or policy correctness, and withhold the party label in some cases. If either strategy is successful, the party’s platform will become a useful informational conduit between candidates and voters.

This emphasis on informational asymmetries leads naturally to the consideration of two standard theoretical problems: adverse selection and moral hazard. We focus on the adverse selection problem, and, to keep the analysis simple, assume the moral hazard problem is completely unsolvable.\(^4\) In the model, parties announce platforms prior to a set of simultaneous elections in a large number of districts. Candidates have policy preferences, and these preferences are unobservable to voters. In addition, we assume that candidates cannot commit to a platform, and therefore pursue their ideal policies after gaining office.\(^5\) Voters learn about the distributions of various groups of candidates. In particular, they learn summary statistics (mean, variance) about the groups of candidates affiliated with the major political parties. We assume that voters obtain this relatively crude information costlessly from the

---

\(^3\)Our work is also related to the economics literature on statistical discrimination and group reputations, such as Arrow (1973), Akerlof (1976), and Tirole (1996).

\(^4\)Of course, parties might also play a role in mitigating moral hazard problems. For example, they might try to reduce the amount of time their members spend catering to narrow special interests, or give incentives for their members to place more weight on constituents and less weight on their own policy preferences.

\(^5\)This is the assumption in Alesina (1988).
media. Voters may also be sophisticated in making inferences from the media’s coverage.

The basic model makes three major predictions. First, parties are effective at aggregating candidates and signaling their preferences. Even in a single party system voter welfare is increased by the information party candidates are able to convey relative to unaffiliated candidates. Interestingly, that party’s platform will be responsive to the ideal policies of candidates rather than voters. A two party system may convey even more information. When the parties’ platforms converge, these platforms follow the standard Downsian logic of tracking the ideal policies of median voters, not candidates. Second, in the two party case the parties’ platforms converge when the cost of party membership is high or parties have strong screening technologies. Platforms diverge, however, when the cost of party membership is low or parties have only weak screening technologies. Each party stakes out an extreme position in order to reduce the ideological heterogeneity of its membership, and thereby make its label more meaningful to voters. Third, party membership is endogenous. Candidates do not always affiliate with the nearest party, since they also take the electoral benefits of party membership into account. Even candidates with the same ideal policy may affiliate with different parties, depending on their district’s characteristics.

An extension of the basic model provides an additional set of predictions. The basic model assumes that voters never learn about individual candidates’ preferences. In section 7 we relax this assumption. The model can then account for the fact that the same party label appears to mean different things in different places—why, for example, Massachusetts Republicans are generally more liberal than Georgia Republicans. It also provides one reason parties might want to limit their ability to screen or discipline their members.

Our work is relevant to three segments of the existing literatures on elections and parties: elections and policy outcomes under conditions of asymmetric information, the equilibrium location of party platforms, and the disciplining of elected officeholders and candidates.

With respect to the first issue, a variety of papers have analyzed models of elections and policy outcomes when there is moral hazard, adverse selection, or both. However, these models only consider a single office (although some incorporate repeat play), and they make no real distinction between candidates and parties—each candidate essentially is a party. Alesina and Spear (1988) and Harrington (1992b) are more relevant for our work. In

---

those papers, parties act as long-lived organizations that help short-lived politicians commit to implementing policies that are electorally attractive but different from the politicians’ ideal policies. In Alesina and Spear the parties do this by transferring “income” across members. In Harrington, politicians invest in their party’s reputation in order to increase its probability of winning future elections. They do this because they care about the policies that are implemented after they leave office.

Second, our work adds to the substantial literature on the role of party platforms in electoral competition. Under the standard spatial theory of electoral competition, a party platform represents the policy positions of its candidates. In the Downsian framework, where voters are perfectly informed and parties act as unified teams, equilibrium platforms in a two-party system are both located at the median of a unidimensional policy space. This convergence result has troubled many observers, and has motivated theorists to search for reasons to expect divergence. Among the proposed explanations are the policy preferences of parties or candidates, the threat of entry by third parties, the costs of reputations, the existence of “valence” issues, “dividing up” the set of offices, and decentralized party decision making. As explained above, our model illustrates another reason divergence—or, more precisely, non-centrist positions—can occur: the desire for greater ideological purity.

Finally, our model begins to connect the formal work on the legislative policy-making roles of parties with the formal work on elections. Most models of parties-in-the-legislature take what happens in the electoral arena as exogenous. While the preferences of legislators in these models may be viewed as “induced preferences” that depend on electoral concerns (constituency preferences, campaign contributions), there is no explicit model of the electoral consequences of legislative decisions and actions. Also, in most of these models the parties do not impose “discipline” on their members, but rather they exert control over the agenda. In fact, parties typically have no incentive to discipline their members in these models. By considering the importance of parties in helping solve the adverse selection problems posed by elections, we establish a motivation for party discipline. As noted above, parties need to maintain their brand names, and this requires either discipline or careful screening, or both.

---


We proceed as follows. Section 2 presents empirical evidence that justifies a few key assumptions of the model. Section 3 lays out the assumptions of the model. Section 4 presents an example of equilibrium in the one party case. Section 5 derives the main results of electoral competition in a two party system. Sections 6 and 7 consider various extensions of the basic model. Section 8 concludes and proposes a few extensions for future research.

2. A Few Motivating Facts

This section provides empirical evidence to help justify two of the key assumptions of our model. The first assumption is that there are real costs associated with party membership, and that these costs are higher for a party’s “mavericks” than for its “loyalists.” We provide data on party switchers, committee assignments, and retirements in Congress to motivate this. The second assumption is that voters are much more knowledgeable about the ideological positions of parties than the positions of particular candidates. Data from the National Election Studies strongly support this notion.

The first piece of evidence supporting the idea of costly party membership is that very few politicians switch parties during their career. This is true even in the U.S., where political career ladders do not appear to be dominated by hierarchical party structures and politicians need not be “party men” to acquire positions of power. For example, only about .3% of all congressmen serving since 1900 have ever switched between the major parties during their congressional service. Moreover, the switches that do occur can probably be explained in terms of “cognitive dissonance” of the sort assumed consistent with our assumptions about the costs of party membership. For example, almost all of the recent switchers in Congress and state governments are conservative southern Democrats who found themselves increasingly alienated inside the Democratic party as the party moved to the left.

A few examples of party switchers give a sense of the cognitive dissonance members sometimes feel. Andy Ireland (Fla.) switched to the Republican Party in 1984. The *Almanac of American Politics* (1985, p. 299) attributes his move in part to “the frustration of being joined in caucus with colleagues few of whom shared his views,” and to the fact that “the Democrats treated him less well; President Reagan and the technicians at the Republican National Committee welcomed him warmly.” Richard Shelby, who switched to the Republican Party in 1994, was probably the most conservative Democratic Senator at the time. “Shelby said he had mistakenly believed there was room for a conservative Southerner in
the Democratic party, ‘but I can tell you there’s not, there’s not room’.” (*The Commercial Appeal*, November 10, 1994) Billy Tauzin (La.), a founder of the Blue Dog Coalition of conservative Democrats and a supporter of the Republican’s “Contract with America” in 1994, also switched to the Republicans in 1995. He explained: “There is no room for conservatives in the Democratic Party. It is determined to be a party by and for liberals,” and “I will not change my votes or my ideas or ideals. I will simply be with people who appreciate me. I will have a chance to be inside the room for a change, as decisions are being made.” (*The Washington Post*, August 8, 1995)

Some switchers were more specific about the costs of opposing the party leadership. Greg Laughlin (Tex.), who switched to the Republican Party in 1995, explained: “I tried to be part of the Democratic team, but I was miserable on some of the votes I cast. Since I’ve been a Republican, I haven’t cast one hard vote. I’m really comfortable where I am.” (*Almanac of American Politics*, 1997, p. 1371) Mike Parker (Miss.), who also switched to the Republicans in 1995, “was clearly uncomfortable with the Democratic Party some time before his switch.” He said: “I think people understand the mendacity I had been through with the Democratic Party... things they had done trying to censure me, slap me around.” (*Almanac of American Politics*, 1997, p. 815) Nathan Deal (Ga.) switched to the Republican Party in 1995 after finding himself leading the opposition forces on a Clean Water Bill. He drafted the bill considered by the committee but was then opposed by the committee’s other Democrats and the Democratic leadership. “It was an uncomfortable and awkward situation where a bill he supported was opposed by the party leadership.” (*Roll Call*, April 13, 1995)

Additional evidence that parties impose differential costs (benefits) on members is revealed in the pattern of House committee assignments. To the degree that party leaders influence the allocation of valuable committee slots, we may expect party loyalists to have better opportunities than mavericks. Conventional wisdom has long supported this notion. As one anonymous House member reported in Shepsle (1978, p. 145), “We are elected by the party as you know. You have to be acceptable to get on Ways and Means.” Numerous studies have found more systematic evidence that party loyalty helps House members obtain desirable committee assignments (*e.g.*, Rohde and Shepsle, 1973; Smith and Ray, 1983; Cox and McCubbins, 1993; Sinclair, 1995).

---

9 There is little evidence that any of these members switched for electoral purposes. Ansolabehere, *et al.* (1999) show that party switchers generally lost votes after their switch.
The retirement patterns of House members provide a final piece of evidence on the incidence of party membership costs. For example, Hibbing (1982) finds that between 1959 and 1978, House members with lower party unity scores were more likely to retire than other members, and Kiewiet and Zeng (1993) find that conservative Democrats had higher retirement rates than others over most of the postwar period. None of these studies employ a direct measure of ideological distance from the party mean (or median), so we conducted our own analysis. Specifically, we ran a probit analysis of retirements as a function of ideological distance from the party mean, age, and a number of other standard variables. The estimates are shown in Table 1. As the results show, party extremists, defined as those whose roll-call voting scores are greater than one standard deviation from the party mean, are more than 25% more likely to retire in any given congressional term than non-extremists (see the first and last rows of the table).

The second assumption is that voters learn little about the policy preferences of individual candidates, but do learn a lot about the parties. In fact, survey data from the National Election Studies suggest that party labels convey most of what voters know about candidate ideologies. Table 2 presents the results of several regressions. The observations are NES respondents, and the dependent variables are respondents’ placements of their representatives on a 7-point liberal/conservative ideological scale. The main independent variables are members’ “true” ideologies as measured by W-Nominate scores, and members’ party affiliations. Focusing on the bottom half of the table, we see the following: (i) overall, voters appear to be able to distinguish between liberals and conservatives (see column 1, which shows that members’ W-Nominate scores predict voter placements effectively); (ii) this appears to be driven mainly by inter-party differences (see column 2, which shows that a party dummy does just as well as W-Nominate scores); and (iii) voters do not appear to distinguish very well between liberals and conservatives within a party (see columns 3 and 4).

These findings should not be too surprising. Several well-known studies have shown that party candidates and party activists are polarized, and even incumbents are only moderately polarized.

---

10 We use the well-known W-Nominate scores, based on roll-call voting, to measure “ideology” (see, e.g., Poole and Rosenthal (1997). Using distance from the party median produces nearly identical results.

11 Two recent studies have produced results with a similar flavor. Franklin (1991) uses the 1988 NES Senate Election Study to study voter placements of senators on a liberal/conservative scale, and finds that voter perceptions of the location of a senator’s party is a better predictor of voter placements than a senator’s ACU score. However, this study does not control for actual party affiliation independently of roll call voting score, as we do in Table 2. Alvarez (1997) applies Franklin’s specification to House incumbents using the 1993 NES Pilot Study and finds similar results.
“responsive” to constituency ideology (e.g., McCloskey, et al., 1960; Stone, 1980; Poole and Rosenthal, 1997). More recently, Ansolabehere, et al. (1999b) study the 1996 House races and find that there are two “pools” of candidates, one Republican and one Democratic, that barely overlap. Erikson and Wright (1997) find similar patterns for the 1994 elections. Another point worth emphasizing is that the results in Table 2 only cover incumbents, since only incumbents have voting records in Congress. Voters have even less information about non-incumbents, so they probably rely even more on party cues when evaluating them (e.g., Mann and Wolfinger, 1980). Also, non-incumbents exhibit even less responsiveness than incumbents—for example, moderate or conservative districts attract Democratic candidates that are only slightly more moderate or conservative than the Democrats running in very liberal districts (Ansolabehere, et al., 1999b).

Finally, although most studies find that congressional incumbents’ issue positions or roll-call voting records have some effect on election outcomes, these effects are typically small, and much smaller than the effects of partisanship.12 Challengers’ positions generally appear to have little or no effect on the vote. Nor do candidates’ positions appear to matter in open-seat contests—instead, partisanship dominates (e.g., Erikson and Wright, 1989, 1997).

In the basic model below we make the extreme assumption that voters cannot discriminate within parties at all. This simplifies the analysis and allows us to focus on certain elements. The assumption can be relaxed, however. In section 7 we study an extension where voters sometimes learn the true positions of the candidates running in a race.

3. The Basic Model

3.1. Environment and Players

There are three kinds of players: parties, candidates, and voters. There is a continuum of voters, divided into a continuum of constituencies or districts. Each district elects one

12See, for example, Erikson (1971), Johannes and McAdams (1981), Wright (1978), Erikson and Wright (1989, 1993, 1997), and Bernstein (1989). In the district-level regressions with incumbent vote-share as the dependent variable and district partisanship and incumbent position as the independent variables (plus year dummies, and sometimes a measure of challenger quality), the beta coefficient on district partisanship is usually 3-4 times as large as that on the incumbent’s position. In many cases, the estimated effect of the incumbent’s position is not significantly different from zero for the members of one or the other party (e.g., Erikson, 1971). In all cases, the estimates imply that even large changes in an incumbent’s roll-call voting record have modest effects—changing a Republican incumbent’s ADA score from 0 to 50 would typically increase his or her vote share by 4-5 percentage points. Of course, these results are only suggestive, because district ideology is never directly measured.
official by plurality rule from a set of competing candidates. The winning candidate takes office, and then implements policy. The space of feasible policies is the interval \([-1, 1]\).

We consider cases with one party and cases with two parties. When there are two parties, we denote them by \(L\) and \(R\). Each party chooses a platform, \(x_L\) or \(x_R\). In some equilibria the platforms diverge—we always choose the equilibrium with \(x_L < x_R\), so we can associate \(L\) with “left” and \(R\) with “right.” Given a pair of platforms, let \(S_i(x_L, x_R)\) be the share of offices won by party \(i\)’s candidates, \(i \in \{L, R\}\). In the basic model we assume the goal of each party is to maximize its share of the offices. In section 6 we study different party goals and also a more decentralized party decision making process.

Candidates are driven by achieving office, and, if elected, policy.\(^{13}\) Election winners receive utility \(w > 0\) from holding office, while losers receive zero. Candidates’ ideal points are independent draws from a uniform distribution on \([-1, 1]\), and are private information. Candidates can neither communicate their ideal points to voters directly, nor commit to a specific platform upon election. They may therefore only communicate their preferences with their affiliation choices. We denote these choices by \(a\). Candidates may run unaffiliated \((a = U)\), or join a party \((a = L\) or \(a = R)\).

Joining a party is costly. This cost is larger the greater is the distance between a candidate’s ideal point and the platform of the party she joins. Thus, a candidate with an ideal point \(z\) who affiliates with party \(i\) and wins office receives utility

\[
w - \alpha (x_i - z)^2 - c,
\]

where \(\alpha > 0\) and \(c \in [0, w]\). If the candidate loses, she receives \(-c\). Unaffiliated candidates pay no costs, so \(c\) reflects the costs of running “credible” campaigns as party members.\(^{14}\)

The costs of party membership may reflect a variety of factors. First, candidates may simply dislike associating with people who hold policy or ideological views that are far from theirs. Campaigning and holding office under a party’s label may entail frequent meetings

\(^{13}\)So, candidates have lexicographic preferences. We could assume instead that candidates also care about policy when they lose, but this complicates the analysis and appears to add little in return.

\(^{14}\)Aldrich and Bianco (1992) also study a model in which candidates can choose which party to join. They focus on purely office-seeking candidates, however, and there is no spatial element in their model. Also, they treat voters’ strategies as exogenous. Aldrich’s work on party activists is a bit closer in spirit, but again there are large differences. Aldrich (1983) focuses on the decisions by activists to join parties, and has no model of elections. Aldrich and McGinnis (1989) has both activists and elections, but there is just one district. Also, the electoral model focuses on the trade-offs candidates face between choosing positions that attract resources from party activists (which are then used to generate turnout), and choosing positions that are attractive to the median voter.
with other party members and party supporters, making speeches and engaging in debates at these meetings, and so on. Alternatively, candidates may dislike being associated with views that are far from theirs, because they believe it reflects badly on them. The costs may also reflect discipline imposed by party leaders. Even forcing party members to attend meetings with fellow partisans, wrap themselves in party symbols, and occasionally vote the party line may be enough. Candidates whose ideological views are opposed to the views held by most others in the party may find that the cognitive dissonance is too high of a price to pay, especially when weighed against their options outside politics. On the other hand, candidates whose preferences are close to the platform will feel relatively unencumbered by such duties.\textsuperscript{15} The quotations in section 2 suggest that these costs are real.\textsuperscript{16}

A second interpretation of our formulation is that parties have imperfect “screening devices.” As will be clear shortly, each party would like to limit the types of candidates who join it. A party might therefore give candidates a test of “ideological correctness”, and only nominate candidates who pass the test. Our assumptions can be interpreted as a specification of the available testing technology. All candidates with $z$ close enough to $x_i$ are able to pass the test, while those with $z$ too far away cannot (i.e., they will be smoked-out).\textsuperscript{17}

Under the first interpretation, the parameter $\alpha$ reflects the magnitude of the costs born by candidates, or the degree of party discipline in a party. Under the second interpretation, $\alpha$ reflects the discriminating power of the screening technology. In what follows, we adopt the language of the first interpretation—candidates choosing whether or not to affiliate but forced to pay a cost—although we believe both interpretations have merit.

For convenience, we define $\theta = \sqrt{(w-c)/\alpha}$ as a statistic for the relative benefit of holding office as a party member. For reasons that will be clear shortly, we assume that $\theta < 1$.

Voters are also policy driven, but they care only about the policy chosen by the official who wins in the district where they live.\textsuperscript{18} Each voter has quadratic utility, and an ideal

\textsuperscript{15}The required activities do not affect voters’ utility.

\textsuperscript{16}Parties might also impose discipline by taking actions that affect $w$. For example, suppose the parties learn each winning candidate’s $z$, but only after a candidate has held office for a while. The parties can then influence $w$ for the rest of the candidate’s career. Moreover, they can ensure higher values of $w$ for candidates with $z$ closer to $x_i$—e.g., by promoting candidates who are more faithful to the party’s platform to positions of power and prestige, by using party resources to help candidates get reelected, and so on.

\textsuperscript{17}We could generalize this to allow a more continuous technology—e.g., by specifying a smooth function for the probability of passing the test, and making the natural assumption that candidates with $z$’s closer to $x_i$ have higher probabilities of passing—but the qualitative results would be the same as in our model.

\textsuperscript{18}In future work we will study cases where voters also care about the electoral outcomes in other districts, as in Austen-Smith (1984, 1986), Snyder (1994), and Ansolabehere and Snyder (1997). That is of course the
point in $[-1, 1]$. Denote the median ideal point in a generic district by $y$, and denote the median of district medians by $Y$. We assume $Y = 0$, so the distributions of district medians and candidate ideal points have the same median. This is the simplest, and possibly the most sensible assumption. In section 6.1 we consider cases where $Y \neq 0$.

3.2. Sequence

Figure 1 illustrates the sequence of gameplay. The moves are as follows.

(1) **Platform Selection.** Parties choose positions $x_L$ and $x_R$ simultaneously, observed by all.

(2) **Candidate Generation.** Nature randomly draws an infinite sequence of potential candidates i.i.d. from $U[-1, 1]$.

(3) **Candidate Affiliation.** Party $L$ offers affiliation to the first candidate, who may either affiliate or not run. Likewise, party $R$ offers affiliation to the second candidate. Each party continues to offer affiliation to subsequent candidates in turn until someone affiliates. The remaining candidates then choose whether to run unaffiliated or not run. Voters observe the set of candidates and their affiliations, but do not observe the process.\(^{19}\)

(4) **Voting.** Winners are decided by plurality rule.

Voters are imperfectly informed about the ideal points of candidates, but can update their beliefs based on candidates’ party affiliations. If party $i$’s candidate wins in a district, $i \in \{L, R\}$, then the expected utility of the median voter in the district is

$$E[-(y-z)^2|a=i] = -(y - \mu_i)^2 - \sigma_i^2,$$

where $\mu_i$ and $\sigma_i^2$ are the mean and variance of ideal points of party $i$’s candidates.\(^{20}\)

Since voters cannot observe whether unaffiliated candidates declined a party affiliation that was offered (or failed a test), they are unable to infer anything new about the preferences

\(^{19}\)The inability of parties to observe the ideal points of potential candidates makes the selection process equivalent to one of randomizing among willing candidates. A variety of selection mechanisms produce the same results—the key assumption is that voters do not observe whether candidates turn down party labels.

\(^{20}\)Expected utility takes this simple form because the utility function is quadratic. Also, we do not need to assume that voters are making sophisticated probability calculations. Rather, voters might learn the mean and variance of the ideal points of affiliated and unaffiliated candidates to the voters from the news media.
of unaffiliated candidates. So, if a candidate is unaffiliated, then voters simply use their priors. The median voter’s expected utility if an unaffiliated candidate wins is therefore

\[ E[-(y-z)^2|a = U] = -y^2 - 1/3. \]

3.3. Equilibrium

Because steps (2)-(4) form a proper subgame, we may initially focus without loss of generality on a single district, expanding the analysis to all districts when we consider the platform location decision. Within a single district, Bayesian equilibria of the game are characterized by the following four elements:

- **Party platforms**, \( x_L \in [-1, 1] \) and \( x_R \in [-1, 1] \).
- **Candidate affiliations**, \( a_j : \{x_L, x_R\} \rightarrow \{L, R, U\} \) for each candidate \( j \).
- **Beliefs**, \( b_j : \{x_L, x_R, a_j\} \rightarrow [-1, 1] \) about each candidate \( j \).
- **Voting functions**, \( v : \{x_L, x_R, a\} \rightarrow \{1, ..., |a|\} \).

In words, each party must choose a platform, and candidates who are given the opportunity to run must choose their affiliations (subject to the constraint of one candidate per party).\(^{21}\) Voters use these decisions, plus Bayes’ Rule, to form beliefs about the intervals in which candidates’ ideal points lie.\(^{22}\) These beliefs must be consistent with the candidates’ affiliation decisions. Since the candidate selection rules are random, we conserve on notation by eliminating references to candidates who are unwilling or unable to run. Finally, voters vote on the basis of their updated beliefs. In each constituency, all players anticipate that the median voter’s most preferred candidate will win the election. This outcome can be supported by fully strategic voting decisions. In fact, it is the unique outcome associated with strong Nash equilibrium voting strategies.

In any Bayesian game, the model must be closed by specifying out-of-equilibrium beliefs. We assume simply that voters believe that the preferences of any candidates who enter out of equilibrium are consistent with those of the group with which the candidate affiliated. Thus, if a party \( L \) candidate enters out of equilibrium, voters in that district make the same

---

\(^{21}\) As will be clear below, restricting the platforms to the interval \([-1, 1]\) is not essential.

\(^{22}\) We denote voters’ updated beliefs simply by an interval because they are uniform. Voters’ beliefs are uniform because the distribution of candidate ideal points is uniform and candidates are selected randomly. More complicated distributions yield more complicated beliefs, but qualitatively similar outcomes. We show some calculations using the normal distribution in section 6.
inference about her preferences that voters in other districts would make about a party $L$ candidate who enters in equilibrium.

We make four tie-breaking assumptions to simplify the analysis. These assumptions are made for technical reasons, and are not critical for the results. First, since voters cannot distinguish between unaffiliated candidates, we assume they vote for the first unaffiliated candidate to enter the election to all others. This implies that the actions of all other unaffiliated candidates are irrelevant, and allows us to restrict attention to elections with at most one unaffiliated candidate.\footnote{Since voters perceive all such candidates to be identical, this assumption effectively maximizes the power of unaffiliated candidates in the model.} Second, when a voter is indifferent between a party candidate and an unaffiliated candidate, she votes for the party candidate. Third, when a voter is indifferent between two party candidates even though the party platforms are distinct, she votes for party $R$’s candidate. Fourth, when party platforms are identical (and all voters should therefore be indifferent between the parties’ candidates), we assume that voters vote in a deterministic fashion such that each party wins exactly half of the districts in which the party candidates are preferred to unaffiliated candidates.\footnote{While a bit artificial, this last assumption simplifies the analysis considerably. If voters randomize in their decisions (perhaps the more natural assumption), and party candidates win with probability 1/2, then the set of candidates willing to join each party jumps discontinuously at the point where the parties converge. Our assumption avoids this discontinuity.}

Equilibria in this game are partially separating, because the choice of party affiliation may serve as a noisy signal of a candidate’s preferences (\emph{i.e.}, their type). Separating equilibria cannot exist because the action space of the candidates is far smaller than their type space. Except in a trivial case where voters do not update their beliefs, pooling equilibria do not exist either, because party affiliation appeals to some types of candidates more than others.

4. Elections in a One Party System

To illustrate some basic characteristics of the model, this section develops the model with just one party. The setup described in the previous section is therefore modified in obvious ways to include only party $L$. Also, to conserve notation we drop party subscripts.

Although backward induction suggests that we should begin by analyzing the voting decision, the candidate affiliation decision is more interesting so we discuss it first.

\textit{The Candidate Affiliation Decision.} Suppose the party position is $x$. If a candidate believes that she will win an election by joining the party, then she will join if \( w - \alpha(x-z)^2 \geq c, \)
or \( z \in \left[ x - \sqrt{(w-c)/\alpha}, x + \sqrt{(w-c)/\alpha} \right] = [x-\theta, x+\theta] \). If \( z \) is too far from \( x \), then the cost of party membership is too high. Given these affiliation decisions, together with the assumption that the overall distribution of candidate ideal points is \( U[-1,1] \), the mean and variance of the party’s candidates’ ideal points are, respectively,

\[
\mu(x) = \begin{cases} 
(x+\theta-1)/2 & \text{for } x \in [-1, \theta-1) \\
 x & \text{for } x \in [\theta-1, 1-\theta] \\
(x-\theta+1)/2 & \text{for } x \in (1-\theta, 1] 
\end{cases}
\]

and

\[
\sigma^2(x) = \begin{cases} 
(1+x+\theta)^2/12 & \text{for } x \in [-1, \theta-1) \\
 \theta^2/3 & \text{for } x \in [\theta-1, 1-\theta] \\
(1-x+\theta)^2/12 & \text{for } x \in (1-\theta, 1] 
\end{cases}
\]

Notice that \( x \) not only affects \( \mu \), but it also affects \( \sigma \). In particular, \( \sigma \) is smaller when the party adopts relatively extreme positions, and larger when the party adopts relatively centrist positions. (This is easily seen by examining the equation for \( \sigma \) above.) Moreover, this is not something particular to the uniform distribution, but a more general property of symmetric, single-peaked, distributions.\(^{25}\)

Notice also that when \( x \) is not too extreme (the middle case above), party members are distributed symmetrically around \( x \), so \( \mu(x) = x \). On the other hand, when \( x \) is extreme, \( \mu \) is less extreme than \( x \), that is, \(|\mu(x)| < |x|\).

The Voting Decision. Voters use Bayes’ Rule to determine the expected policy positions of the candidates, based on their prior knowledge of the distribution of candidate ideal points and the candidates’ affiliation strategies.

Suppose voters’ beliefs are consistent with the affiliation decisions above. That is, if a candidate affiliates herself with the party (\( i.e., a = L \)), then voters infer that her ideal point is distributed uniformly on the interval of ideal points identified with the party. In this case, the interval is \([\max\{-1, x-\theta\}, \min\{x+\theta, 1\}] \). If a candidate remains unaffiliated, then no information is conveyed about her policy preferences, so voters assume that her ideal point is distributed uniformly on \([-1, 1] \).

When does a voter prefer an affiliated candidate to an unaffiliated candidate? Clearly, the ideal points of party members must belong to a strict subset of \([-1, 1] \), so that some

\(^{25}\)In an interesting paper, Enelow and Hinich (1981) assume that the degree of uncertainty voters have about a candidate’s position is a function of the position a candidate adopts. One case they study is where voters are more certain about candidates who adopt extreme positions than they are about candidates who adopt moderate positions. Our model shows how this assumption can be generated endogenously.
differentiation from non-affiliated candidates is possible. This is ensured by the assumption that \( \theta < 1 \). A voter with ideal point \( y \) will vote for the party candidate if

\[
-(y-\mu(x))^2 - \sigma^2 > -y^2 - 1/3.
\]  

(3)

Combining this with equations (1)-(2) above, we can easily calculate the set of voters and districts that prefer affiliated candidates. For \( \mu(x) \neq 0 \), we can rewrite (3) in terms of the “cut-point,” \( C \), which identifies the voter who is indifferent between affiliated and unaffiliated candidates. This in turn determines the set of districts that elect affiliated candidates.

Comment 1. Let \( C \) be defined as follows:

\[
C = \begin{cases} 
\frac{x+\theta}{3} & \text{for } x \in [-1, \theta-1) \\
\frac{x}{2} + \frac{\theta^2-1}{6x} & \text{for } x \in [\theta-1, 1-\theta], x \neq 0 \\
\frac{x-\theta}{3} & \text{for } x \in (1-\theta, 1]
\end{cases}
\]

If \( x < 0 \), then a district with median at \( y \) elects an affiliated candidate iff \( y \leq C \); if \( x > 0 \), then a district with median at \( y \) elects an affiliated candidate iff \( y \geq C \); and if \( x = 0 \), then all districts elect affiliated candidates.

Proof. Proofs of all comments and propositions are in the Appendix.

The Party Platform Decision. Examining Comment 1, it is clear that if \( x \) is located near the extremities of the policy space, then a non-affiliated candidate must win some districts. On the other hand, for values of \( x \) near 0 the party’s candidate will win all districts.\(^{26}\)

Figure 2 illustrates an equilibrium outcome of the one party election game. The party chooses platform \( x = 0 \), and all candidates whose ideal points lie within \([−\theta, \theta]\) join the party if offered the opportunity. These candidates will win the election in every district because of the informational benefits (i.e., variance reduction) offered by the party.

While the one-party case is extremely simple, it has two noteworthy features. First, voters benefit from the information provided by candidates’ affiliation choices even when there is only one party. Second, to exploit its informational advantage and maximize its share of offices, the party must choose a platform near the distribution of candidate ideal

\(^{26}\)As a result, restricting the analysis to equilibria in which voters always vote for the first unaffiliated candidate who enters is innocuous—the party candidate would always defeat any set of unaffiliated candidates.
points, rather than the distribution of voters’ ideal points. Of course, the center of the candidate distribution will typically be near the center of the voter distribution, i.e., $Y$ will be near 0. When this holds, there will be incentives to choose moderate platforms even in one-party systems. When it does not, however, the model predicts that the party will track the candidate distribution rather than the voter distribution. The reason is simple. Locating at the center of the candidate space makes the party’s candidates preferable to unaffiliated candidates in all districts, because the distribution of ideal points among the party’s candidates will have the same mean but a lower variance. And, the party need not worry about capturing the median voter because it does not have to worry about competition from another party.

5. Electoral Competition in a Two Party System

We now turn to the two party model. While the logic of aggregating similarly-minded candidates remains from the previous section, the presence of competition can dramatically affect party incentives. Here we characterize the partially separating equilibria. In the next section we examine the implications of altering or relaxing a few of the basic assumptions.

The Voting Decision. As above, voters use Bayes’ Rule to determine the expected policy positions of the candidates. If a candidate affiliates with a party, then voters infer with certainty that her ideal point lies in the interval of ideal points identified with that party. For example, voters infer that candidates affiliated with party $L$ have ideal points distributed uniformly on the interval $[\max\{-1, x_L-\theta\}, \min\{x_L+\theta, 1\}]$. If a candidate stays unaffiliated, then voters use their priors and assume that her ideal point is drawn from $U[-1,1]$. Although not technically a platform, we extend the notation in the obvious way to non-affiliated candidates, so that $x_U = 0$, $\mu_U = 0$, and $\sigma_U^2 = 1/3$.

Suppose two candidates are running in a district, one from party $i \in \{L, R, U\}$, and one from party $j$ ($j \neq i$). A voter with ideal point $y$ prefers the candidate from party $i$ if $-(y-\mu_i)^2 - \sigma_i^2 > -(y-\mu_j)^2 - \sigma_j^2$. Simple manipulation produces a cut-point, $C_{ij}$, which defines the voter who is indifferent between candidates (provided $\mu_i \neq \mu_j$):

$$C_{ij} = \frac{\mu_i + \mu_j}{2} + \frac{\sigma_i^2 - \sigma_j^2}{2(\mu_i - \mu_j)}.$$

Consider first the problem of choosing between an affiliated candidate and an unaffiliated candidate. This analysis is the same as that in section 3 above, except that the notation in
Comment 1 should be changed slightly to reflect the existence of two parties. In particular, $C$ should be changed to $C_{iU}$, and $x$ to $x_i$. Then Comment 1 describes the set of districts that will choose party $i$’s candidate over an unaffiliated candidate.

If two affiliated candidates run, then, by equation (4), a voter whose ideal point lies mid-way between the party means will prefer the party with the lower variance. Because extreme parties have a lower variance than moderate ones, this expression gives rise to six cases. Comment 2 shows the cut-point, $C_{LR}$, for each of these cases. This cut-point defines the set of districts that will elect $L$ or $R$ candidates, subject to these candidates’ ability to defeat unaffiliated candidates.

**Comment 2. Choosing Between Two Affiliated Candidates.** Suppose $x_L < x_R$, and let $C_{LR}$ be defined as follows:

$$C_{LR} = \begin{cases} 
\frac{x_L+x_R+2\theta-1}{3} & \text{for } x_L \in [-1, \theta-1) \text{ and } x_R \in [-1, \theta-1) \\
\frac{3x_R^2-x_L^2-2x_L\theta+x_L+\theta-1}{3(2x_R-x_L-\theta+1)} & \text{for } x_L \in [-1, \theta-1) \text{ and } x_R \in [\theta-1, 1-\theta] \\
\frac{(x_L+x_R)(x_R-x_L-2\theta+1)}{3(x_R-x_L-2\theta+2)} & \text{for } x_L \in [-1, \theta-1) \text{ and } x_R \in (1-\theta, 1] \\
\frac{x_L+x_R}{2} & \text{for } x_L \in [\theta-1, 1-\theta] \text{ and } x_R \in [\theta-1, 1-\theta] \\
\frac{x_R^2-3x_L^2-2x_R\theta+x_R-\theta+1}{3(x_R-2x_L-\theta+1)} & \text{for } x_L \in [\theta-1, 1-\theta] \text{ and } x_R \in (1-\theta, 1] \\
\frac{x_L+x_R-2\theta+1}{3} & \text{for } x_L \in (1-\theta, 1] \text{ and } x_R \in (1-\theta, 1] 
\end{cases}$$

A district with median at $y$ prefers party $L$’s candidate to $R$’s candidate iff $y < C_{LR}$. ■

Most of these cut-points respond to party platform locations in an intuitive manner. Near the edges of policy space, however, the trade-off between mean and variance in candidate positions becomes particularly acute. For example, in case (1) (top case), if $\theta$ is sufficiently large then $C_{LR} > x_R$. Thus, even voters located to the right of $x_R$ might prefer party $L$ because its more extreme position provides greater variance reduction.

**The Candidate Affiliation Decision.** As before, if joining party $i$ guarantees that a candidate will win office, then she will join the party if her ideal point $z$ satisfies $w-\alpha(x_i-z)^2 \geq c$, or $z \in [x_i-\theta, x_i+\theta]$. That is, the candidate joins if $z$ is close enough to $x_i$. If $z$ is too far from $x_i$ then the candidate would rather not hold office—the cost of party membership is
too high. This constraint justifies the voter beliefs discussed in the voting decision problem. Joining a party is weakly dominated unless a candidate can win with positive probability by affiliating. Thus with the tie-breaking rule adopted above (party $R$ wins ties), equilibria exist in which only one party candidate runs for any given office. However, the party expecting to lose with certainty may attract a candidate in equilibrium if $c = 0$.

The Party Platform Decision. There are two types of equilibria, depending on the value of $\theta$. The first type features convergence to the global median, 0. The second type features divergent platforms, in which the parties locate on opposite sides of 0. These equilibria are only possible when $\theta$ is sufficiently large.

Proposition 1. If $\theta < \sqrt{3}/2$, then the unique equilibrium is $(x^*_L, x^*_R) = (0, 0)$ (convergent equilibrium). If $\theta > \sqrt{3}/2$, then the unique equilibrium is $(x^*_L, x^*_R) = (\frac{1}{2} - \theta, \theta - \frac{1}{2})$ (divergent equilibrium).

In both types of equilibria, each party wins half of the districts. Figures 3a and 3b display the resulting platform locations, party affiliations, and election outcomes as a function of candidate ideal points and district medians. Figure 3a shows the convergent equilibrium, and Figure 3b shows the divergent equilibrium. Note that in the divergent case, the distance between the parties is increasing in $\theta$.

For small values of $\theta$, each party’s label conveys a large amount of information about its candidates, regardless of how centrist or extremist the party is. As a result, the logic behind traditional median convergence holds. When $\theta$ is large, however, the convergent equilibrium is unstable because a party that locates in the center will attract candidates from such a wide range of the ideological spectrum that the party label conveys relatively little information about the preferences of its members. Parties then have an incentive to adopt more extreme positions, in order to reduce the variance in their members’ policy preferences. This reduction more than compensates for the loss of votes due to an unattractive average position. At the “threshold” value, where $\theta = \sqrt{3}/2$, both convergent and divergent equilibria exist.

There are no “non-centrist” convergent equilibria, even when $\theta$ is large. This is because of

---

27If $c = 0$, then joining a party is costless if a candidate loses, and weakly dominated equilibria also exist where a candidate lying outside the interval occupied by the losing party enters as that party’s candidate. Such equilibria may be eliminated by a “trembling” argument. If instead of losing the candidate wins with probability $\epsilon > 0$, then she does strictly better by not joining a party.
the threat posed by unaffiliated candidates. For example, suppose \( \theta > 1/2 \) and both parties converged to some \( x \in (1 - \theta, \theta) \). Then districts with medians near \(-1\) will prefer unaffiliated candidates. Since the parties divide the set of districts not won by unaffiliated candidates evenly between them, each party will win less than half of the districts. If one party changes its position to \(-x\), however, then no unaffiliated candidates will win, and each party will win exactly half of the races. Unaffiliated candidates play a role similar to the “third-party waiting in the wings” in Palfrey (1984). If unaffiliated candidates are not allowed, then non-centrist convergent equilibria do exist (in addition to the divergent equilibrium).

Finally, we can make a few statements about voter welfare. First, when the parties converge in equilibrium, voter welfare is clearly the same with two parties as it is with one. So, all voters are better off with two parties than with no parties. Second, when the parties diverge in equilibrium, a majority of voters in each district—and, therefore, a majority of voters overall—are better off with two parties than with one party. However, some voters are worse off. Specifically, voters with ideal policies greater than \( \frac{2\theta^2}{3} - \frac{1}{2} \) who reside in districts with medians \( y < 0 \), and voters with ideal policies less than \( \frac{1}{2} - \frac{2\theta^2}{3} \) who reside in districts with medians \( y > 0 \), prefer one party to two. Thus as \( \theta \) increases, more voters prefer two parties to one party. Third, a straightforward calculation shows that when the parties diverge in equilibrium, all voters are better off with two parties than with no parties.

6. Variations and Extensions of the Basic Model

6.1. Variances and Means of District Medians and Candidates Differ

In the model analyzed above, the distribution of candidate ideal points and the distribution of district medians are centered at the same point. While this is perhaps the natural assumption, we might expect the variance of candidate ideal points to be greater than the variance of district medians. This is because (1) district medians are summary statistics, and will therefore tend to have a lower variance than the underlying distribution of individual ideal points, while candidates are individuals; and (2) survey data indicate that candidates tend to have more extreme preferences than voters. One factor that may work in the opposite direction is gerrymandering, since lines are often drawn to produce a disproportionate number of “safe,” and therefore “extremist,” districts. In the analysis above we made no assumption about variances, so the results clearly hold even if the variances are different.

Another possibility is that the mean of the candidate ideal points differs from the median of district medians, that is, \( Y \neq 0 \). This might be produced by gerrymandering (e.g., over-
representation of rural areas), or because the time and money needed to run for political office limits the set of potential candidates to a select subset of the citizenry. In section 4 we showed that when there is only one party, then it will locate at or near the mean of the distribution of candidate ideal points, 0. When there are two parties, however, the parties locate at the median of the district median ideal points, $Y$. To some degree, this result generalizes for $Y \neq 0$.

Comment 3. Suppose $|Y| \leq \sqrt{\frac{4-\theta^2}{3}}$, and $\theta \leq \frac{\sqrt{3}(2Y^2+Y+1)}{2(Y+1)}$. Then the unique equilibrium party platforms are $x^*_L = x^*_R = Y$.  ■

If $|Y| > \sqrt{\frac{4-\theta^2}{3}}$, then $Y$ is so extreme that party candidates who locate at $Y$ sometimes lose against unaffiliated candidates. For example, if $x_L = x_R = Y > 0$ implies that $C_{LU} = C_{RU} > -1$, then each party wins strictly less than half of the offices. However, one or the other party can win almost exactly half of the offices by choosing a platform near $Y$. So, $x_L = x_R = Y$ is not an equilibrium. Further work is necessary to characterize the equilibria in these cases.

6.2. Normally Distributed Candidate Ideal Points and District Medians

This subsection shows that the qualitative features of our main results do not depend on the uniform distribution. We maintain all the assumptions of the basic model, but let candidate ideal points and district medians both be normally distributed with mean 0 and variance 1. Then a result similar to Proposition 1 holds: when $\theta$ is small then the equilibrium is convergent, with $(x^N_L, x^N_R) = (0,0)$, and when $\theta$ is large enough then the equilibrium is divergent, with $(x^N_L, x^N_R) = (-x^N(\theta), x^N(\theta))$. The boundary between the two cases is given by $\theta \approx 1.61$. We cannot derive an analytic solution for the equilibrium in the divergent case, but Table 3 below provides calculations for various values of $\theta$. 21
Table 3
Equilibrium When Candidate Ideal Points and District Medians are Both $N(0, 1)$

<table>
<thead>
<tr>
<th>$\theta$</th>
<th>$x^{N*}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.60</td>
<td>0.000</td>
</tr>
<tr>
<td>1.62</td>
<td>0.134</td>
</tr>
<tr>
<td>1.65</td>
<td>0.356</td>
</tr>
<tr>
<td>1.70</td>
<td>0.548</td>
</tr>
<tr>
<td>1.80</td>
<td>0.790</td>
</tr>
<tr>
<td>1.90</td>
<td>0.963</td>
</tr>
<tr>
<td>2.00</td>
<td>1.105</td>
</tr>
</tbody>
</table>

6.3. Parties that Maximize Total Net Benefits of Their Members

In this subsection we examine the impact of different party goals. Specifically, suppose that instead of maximizing the share of offices won, each party maximizes the aggregate expected net benefits, or “surplus,” of its members. This is an especially compelling goal if transfers are possible inside parties. To the extent that something approximating Lindahl pricing is possible, there would be approximately unanimous agreement inside the party in support of the surplus-maximizing platform.

For simplicity we return to the case where the distributions of candidates and district medians is uniform on $[-1, 1]$ and assume $c = 0$. Recalling that $S_i(x_L, x_R)$ is the share of offices won by party $i \in \{L, R\}$, the new objective function each party $i$ is:

$$B_i(x_L, x_R) = S_i(x_L, x_R) \int_{\max(-1, x_i - \theta)}^{\min(x_i + \theta, 1)} [w - \alpha(z - x_i)^2 \, dz]$$

$$= \alpha S_i(x_L, x_R) [\theta^2 - \sigma_i^2(x_i) - (\mu(x_i) - x_i)^2]$$

$$= \begin{cases} 
\alpha S_i(x_L, x_R) [2\theta^2 + \theta(1+x_i) - (1+x_i)^2]/3 & \text{for } x_i \in [-1, \theta-1) \\
\alpha S_i(x_L, x_R) [2\theta^2/3] & \text{for } x_i \in [\theta-1, 1-\theta] \\
\alpha S_i(x_L, x_R) [2\theta^2 + \theta(1-x_i) - (1-x_i)^2]/3 & \text{for } x_i \in (1-\theta, 1] 
\end{cases}$$

The equilibrium platforms under this objective are difficult to characterize, but intuitive. The following result is an analog to Proposition 1.

Comment 4. If $\theta \leq 3\sqrt{15}/16$, then the unique equilibrium is $(x_L^{B*}, x_R^{B*}) = (0, 0)$. If $\theta \geq \sqrt{3}/2$, then the unique equilibrium $(x_L^{B*}, x_R^{B*}) = (-x^{B*}, x^{B*})$, where $x^{B*} \in (\theta - \frac{1}{2}, 1 - \frac{\theta}{2})$. ■

Figure 4 presents numerical calculations of the equilibrium. For comparison, the figure
also shows the equilibrium when parties are concerned only with maximizing the share of offices they win. Comment 4 and Figure 4 have three implications. First, the range of $\theta$ that supports convergence in equilibrium under surplus maximization is smaller than that under office-share maximization. Specifically, the threshold value for a convergent equilibrium drops from about .87 ($\sqrt{3}/2$) to about .77. Second, the amount of divergence is always greater under surplus maximization. Intuitively, this is because surplus-maximizing parties care explicitly about the variance of the distribution of their members’ ideal points. Parties gain more from choosing relatively extreme positions, because these positions tend to attract candidates who are closer to the platform and thus value membership highly. Finally, when parties maximize surplus and the equilibrium is divergent, the amount of divergence decreases as $\theta \to 1$. This is the opposite of what happens when parties maximize their share of offices.

6.4. “Democratic” Parties

In the analysis above, party platforms are chosen by a pair of actors called “parties.” The implicit assumption is that these actors are party leaders of some sort, with dictatorial powers to choose their party’s platform. In this section we assume instead that parties are more democratic institutions, along the lines of Austen-Smith (1984) and Snyder (1994), and use a collective choice process to choose their platforms. Specifically, we assume that each party chooses its platform by simple majority-rule, and all of the party’s candidates have one vote each. We define an equilibrium as a pair of platforms that “reproduce themselves.” That is, $(x^*_L, x^*_R)$ is an equilibrium if, given the candidate affiliation decisions produced by $x_L$ and $x_R$, a majority of the candidates who join party $L$ prefer $x_L$ to any other platform, and a majority of the candidates who join party $R$ prefer $x_R$ to any other platform.

There is now a range of possible equilibria, but also a limit on the extremity of party platforms.

*Comment 5.* The platform pair $(x^*_L, x^*_R)$ is an equilibrium if and only if $x^*_L \in [\theta - 1, 1 - \theta]$ and $x^*_R \in [\theta - 1, 1 - \theta]$. □

This result rules out divergent equilibria with platforms that are as extreme as those in Proposition 1 and Comment 4. At the same time, it introduces a constraint on the centrist tendencies of party platforms. Given a sufficiently moderate platform, any additional movement toward the center will alienate a majority of the party’s existing membership,
and will be opposed. Also, if one party happens to have a platform closer to 0 (caused, for example, by a recent shift in voter preferences) then it will be the majority party.

7. Behavior When Voters May Learn More about Candidates

Up to now we have assumed that voters never learn anything about any particular candidates. Some readers may therefore have concluded that the analysis is largely irrelevant to the study of U.S. federal elections. Where are the “personal”, candidate-centered campaigns, so prominent in U.S. presidential, congressional, and gubernatorial elections? In this section we introduce the possibility that voters learn about individual candidates, as well as parties. We focus on a particular case, which serves to illustrate the basic logic of the situation. We will conduct a more complete analysis in future work.

We return to the case of one party. Also, we assume that in each district there is at most one unaffiliated candidate. Finally, we assume there is an exogenous probability $\psi > 0$ that a race becomes “hot,” and when a race is hot voters learn the ideal points of all candidates running in the race. This captures in a crude but tractable way the idea that in a hot race there is a disproportionate amount of media attention, a high amount of campaign spending on both sides, and a relatively high level of voter information about the contending candidates. We assume $\psi$ is the same for all districts, and that whether or not a particular race is hot is independent of what is happening in all other races.

The following result characterizes the probabilities that party candidates win in hot races.

**Comment 6.** Let $p(z, y)$ be the probability that an affiliated candidate with an ideal point at $z$ wins a hot race in a district with median at $y$. Then

$$p(z, y) = \begin{cases} 
(1+z)/2 & \text{for } (z, y) \text{ such that } z \leq 2y - 1 \\
1 - y + z & \text{for } (z, y) \text{ such that } 2y - 1 < z < y \\
1 + y - z & \text{for } (z, y) \text{ such that } y \leq z < 2y + 1 \\
(1-z)/2 & \text{for } (z, y) \text{ such that } z \geq 2y + 1 
\end{cases}$$

Assume $c > 0$, so party candidates pay a strictly positive cost of running for office. Also, suppose affiliated candidates always win in “cold” races, where voters do not learn the candidates’ positions. (This holds provided $x$ is close enough to 0.) Then a candidate with ideal point $z$ is willing to join the party if and only if

$$[1-\psi+\psi p(z, y)][w-\alpha(x-z)^2] \geq c.$$
Assume $c < w(1 - \psi/2)$. This guarantees that when $x = 0$, candidates with $z = x$ are willing to join the party in all districts (i.e., the party can always at least get a “party hack” to run). Let $Z(x, y)$ be the set of $z$’s willing to join the party in a district with median $y$. Let $\mu(x, y)$ and $\sigma^2(x, y)$ be the mean and variance, respectively, of the $z$’s of those willing to join.

The next comment describes certain characteristics the party will have in equilibrium, given candidates’ optimal affiliation decisions.

Comment 7. Suppose $x = 0$. For all districts with $y > 0$, $\mu(x, y) > 0$ and $\sigma^2(x, y) < \theta^2/3$; and for all districts with $y < 0$, $\mu(x, y) < 0$ and $\sigma^2(x, y) < \theta^2/3$. ■

Thus, in equilibrium the party label will mean different things in different districts. In left-leaning districts the party will tend to attract more candidates from the left than from the right, and the mean position will be to the left of zero. Similarly, in right-leaning districts the party will tend to attract more candidates from the right, and its mean position will be to the right of zero. The model therefore captures the idea that Massachusetts Republicans are more liberal than Georgia Republicans because voters in Massachusetts voters are more liberal than voters in Georgia. Party candidates are “responsive” to district ideology.

Finally, the extended model also provides an intuition about why parties might want to limit their ability to screen or discipline their members. Suppose the party wants to maximize the expected share of offices won by its candidates. If the party imposes too much discipline, then the ideological range of its candidates shrinks to a tiny interval around its platform. But this limits the extent to which the party’s candidates will be responsive to district ideology, and therefore reduces the number of hot races the party wins.

An example illustrates the logic. Let $\bar{p}(y, \alpha)$ be the average probability that a party candidate wins a hot race in a district with median $y$, given the candidates’ optimal affiliation choices, and let $\bar{p}(\alpha)$ be the average of $\bar{p}(y, \alpha)$ over $y$. Suppose there are 3 types of districts, with medians at $-1$, 0, and 1, and an equal number of each type. Also, suppose $w = 1$, $c = .5$, and $\psi = .5$. Then as $\alpha \to \infty$, $\bar{p}(1, \alpha) = \bar{p}(-1, \alpha) \to 1/2$, $\bar{p}(0, \alpha) \to 1$, and $\bar{p}(\alpha) \to 2/3$. However, if the party could choose $\alpha = .25$, then $Z(0, -1) \approx [-1, .27]$, $Z(0, 1) \approx [-.27, 1]$, and $Z(0, 0) \approx [-.53, .53]$. Thus, $\bar{p}(1, .25) = \bar{p}(-1, .25) \approx .68$, $\bar{p}(0, .25) = .74$, and $\bar{p}(.25) \approx .70 > 2/3$. Because of the responsive candidate affiliation decisions, the party wins more hot races in the extreme districts when $\alpha$ is small. Of course, the party loses more hot races in moderate districts, but as the example shows the gain can easily offset the loss.
8. Discussion

We have constructed a simple model of informative party brand names. Importantly, candidates’ party affiliation choices—and thus the information conveyed by party labels—are endogenous. The affiliation decision is vital to candidates, since they cannot communicate with voters in any other way. Candidates do not simply join the party with the nearest platform, but rather take into account the position of their district’s median voter and the electoral consequences of their affiliation decision. As a result, candidates with identical preferences may affiliate with different parties in different districts.

Party platforms are also endogenous, and consistent with the Downsian tradition we observe a strong centripetal force. Our model introduces a refinement to the standard logic, however. Since affiliation is fundamentally about signaling preferences, parties also have incentives to sharpen their message by choosing more extreme positions. Combined with the usual incentives for median convergence, this produces both convergent and divergent equilibria, depending on the benefit of holding office relative to the costs of party membership.

Inter-party competition has a large affect on platform positioning. With one party, there exists a continuum of equilibria in which party candidates win in every district. In all of these, the party chooses a platform near the center of the distribution of candidate ideal points. When two parties are present, the parties do not locate themselves for maximum candidate coverage. Instead, competitive pressures force them either to take positions at the center of the distribution of district medians, or to take extreme positions. In all cases, however, parties are able to convey brand name information through their platform choices.

We conclude with four extensions not discussed in the main text above. First, the “hot race” analysis of Section 7 must be extended to multiple parties to provide a fuller account of party responsiveness to electoral preferences. Second, we must further explore the mechanisms parties use to select candidates. One way to do this is to allow parties to choose $\theta$ (perhaps through a primary system) and to characterize optimal disciplining or screening schemes. While parties in our basic model clearly have incentives to reduce $\theta$, the discussion of hot races suggests that “big tents” may be necessary for maintaining candidate responsiveness. Parties might also be able to prevent potentially friendly candidates from running unaffiliated, in addition to being able to screen out undesirable candidates. This would tend to cause voters’ assessments of unaffiliated candidates to become less noisy, and might thereby cause the opposing party to lose some districts to unaffiliateds.
Third, and perhaps most ambitiously, the model provides a promising platform for examining alternative political systems. The number of parties may be expanded or made endogenous, and the impact of different electoral rules may also be examined. The comparative perspective may also shed light on the circumstances under which brands will matter. We may expect, for instance, that in highly centralized systems, labels are unimportant because there are few important elected offices. Races will therefore be very likely to be “hot,” and the incentives to screen or discipline party members will be weak. In decentralized political systems, however, brand names may be more in demand by voters, and these brands will also tend to be more stable because they represent the “average” of a larger number of officials.

Finally, we note that our general theoretical approach is applicable to a variety of settings in which actors wish to signal their spatial position. One promising application may be the examination of interest group endorsements of candidates or products. As with party affiliation, an endorsement may reduce the variance of voters’ or consumers’ assessments. Informational intermediaries therefore have a much wider substantive scope than party competition alone, and fleshing out their equilibrium properties remains an important challenge.
Proof of Comment 1. To find the cut-point between an affiliated candidate and an unaffiliated candidate, substitute equations (1) and (2) into (3). There are three sub cases. First, if $x \in [-1, \theta - 1)$, then we have $C = \frac{x + \theta - 1}{2} + \frac{(x + \theta + 1)^2/12 - 1/3}{2(x + \theta - 1)/2} = \frac{x + \theta}{3}$. Since $x < 0$, the median voter in a district prefers the affiliated candidate if $y \leq C$.

Second, if $x \in [\theta - 1, 1 - \theta]$, then equation (3) becomes $C = \frac{x + \theta - 1}{2} + \frac{(x + \theta + 1)^2/12 - 1/3}{2(x + \theta - 1)/2} = \frac{x + \theta}{3}$. Clearly, if $x < 0$ then the median voter in a district prefers the affiliated candidate if $y \leq C$, and if $x > 0$ then the median voter in a district prefers the affiliated candidate if $y \geq C$. Note also that if $x \in [1 - \sqrt{(4 - \theta)^2}/3, 1 + \sqrt{(4 - \theta)^2}/3]$, then all districts prefer an affiliated candidate over an unaffiliated one.

Third, if $x \in (1 - \theta, 1]$, then (3) becomes $C = \frac{x + \theta - 1}{2} + \frac{(x + \theta + 1)^2/12 - 1/3}{2(x + \theta - 1)/2} = \frac{x + \theta}{3}$. Since $x > 0$, the median voter in a district prefers the affiliated candidate if $y \geq C$. ■

Proof of Comment 2. Suppose $x_L < x_R$. Also, suppose $\mu_L < \mu_R$ (we check that this holds below). Then, by equation (4), a district with median voter $y$ prefers party $L$’s candidate if

$$y < C_{LR} = \frac{\mu_L + \mu_R}{2} + \frac{\sigma_L^2 - \sigma_R^2}{2(\mu_L - \mu_R)}. \quad (A.1)$$

There are six cases, depending on the locations of $x_L$ and $x_R$.

First, if $x_L$ and $x_R$ are both in the interval $[-1, \theta - 1)$, then both candidates are on the far left but $L$ has a lower variance. Substituting $\mu_L = \frac{x_L + \theta - 1}{2}$, $\mu_R = \frac{x_R + \theta - 1}{2}$, $\sigma_L^2 = \frac{(x_L + \theta + 1)^2}{12}$, and $\sigma_R^2 = \frac{(x_R + \theta + 1)^2}{3}$ into (A.1) yields: $y < C_{LR} = \frac{x_L + x_R + 2\theta - 1}{3}$.

Second, if $x_L \in [-1, \theta - 1)$ and $x_R \in [\theta - 1, 1 - \theta]$, then party $L$’s candidate is on the far left and has a lower variance. Substituting $\mu_L = \frac{x_L + \theta - 1}{2}$, $\mu_R = x_R$, $\sigma_L^2 = \frac{(x_L + \theta + 1)^2}{12}$, and $\sigma_R^2 = \frac{\theta^2}{3}$ into (A.1) yields: $y < C_{LR} = \frac{3x_L^2 - x_R^2 - 2x_L \theta x_L + \theta - 1}{3(2x_R - x_L - \theta + 1)}$.

Third, if $x_L \in [-1, \theta - 1)$ and $x_R \in (1 - \theta, 1]$, then both parties’ candidates are extreme, but on opposite sides of 0. Substituting $\mu_L = \frac{x_L + \theta - 1}{2}$, $\mu_R = x_R - \theta + 12$, $\sigma_L^2 = \frac{(x_L + \theta + 1)^2}{12}$, and $\sigma_R^2 = \frac{(x_R - \theta + 1)^2}{12}$ into (A.1) yields: $y < C_{LR} = \frac{(x_L + x_R)(x_R - x_L - 2\theta + 1)}{3(x_R - x_L - 2\theta + 2)}$.

Fourth, if $x_L$ and $x_R$ are both in the interval $[\theta - 1, 1 - \theta]$, then $\mu_L = x_L$, $\mu_R = x_R$, and $\sigma_L^2 = \sigma_R^2 = \frac{\theta^2}{3}$. Inspection of (A.1) clearly reveals that a voter located at $y$ chooses the candidate whose party mean is closer. Substituting into (A.1) yields: $y < C_{LR} = \frac{x_L + x_R}{2}$.
Fifth, if \( x_L \in [\theta - 1, 1 - \theta] \) and \( x_R \in (1 - \theta, 1] \), then party \( R \)'s candidate is on the far right and has a lower variance. Substituting \( \mu_L = x_L \), \( \mu_R = \frac{x_R - \theta + 1}{2} \), \( \sigma_L^2 = \frac{\sigma^2}{3} \), and \( \sigma_R^2 = \frac{(1-x_R+\theta)^2}{12} \) into (A.1) yields: \( y < C_{LR} = \frac{x_R^2 - 3x_L^2 - 2x_L\theta x_R + \theta + 1}{3(x_R - 2x_L - \theta + 1)} \).

Finally, if \( x_L \) and \( x_R \) are both in the interval \((1 - \theta, 1] \), then both parties are on the extreme right. Substituting \( \mu_L = \frac{x_L - \theta + 1}{2} \), \( \mu_R = \frac{x_R - \theta + 1}{2} \), \( \sigma_L^2 = \frac{(1-x_L+\theta)^2}{12} \), and \( \sigma_R^2 = \frac{(1-x_R+\theta)^2}{12} \) into (A.1) yields \( y < C_{LR} = \frac{x_L + x_R - 3\theta + 3}{3} \).

**Proof of Proposition 1.**

**Convergent Case.** First, we show that if \( \theta \leq \frac{\sqrt{3}}{2} \) then the unique equilibrium is \( x_L^* = x_R^* = 0 \). Suppose \( x_L = x_R = 0 \). By case (2) of Comment 1, all voters prefer both party \( L \) and party \( R \) candidates to unaffiliated candidates, so no unaffiliated candidates win. Also, by the tie-breaking rule the parties divide the districts evenly. So, \( S_L(0,0) = S_R(0,0) = \frac{1}{2} \).

Now consider whether party \( L \) could profitably deviate to \( x_L < 0 \). Given such a deviation, \( L \)'s candidates can only win districts with \( y < C_{LR} \). If \( C_{LR} < 0 \), then \( S_L(x_L,0) < \frac{1}{2} \). So, if \( C_{LR} < 0 \) for all \( x_L < 0 \), then no such profitable deviations exist.

There are two cases. First, if \( x_L \in [\theta - 1, 0) \), then case (4) of Comment 2 applies, so \( C_{LR} = \frac{\mu_L + \mu_R}{2} < 0 \). Second, if \( x_L \in [-1, \theta - 1) \), then case (2) of Comment 2 applies, and \( \sigma_L < \sigma_R \) (so, party \( L \) might gain from its lower variance). Substituting \( x_R = 0 \) into the equation for \( C_{LR} \) yields \( C_{LR} = -\frac{x_L^2 - 2x_L\theta x_L + \theta + 1}{3(-x_L - \theta + 1)} \). The denominator is positive, so \( C_{LR} < 0 \) iff the numerator is negative, that is, \( \frac{x_L^2 - x_L + 1}{x_L - \theta} > \theta \). Differentiating, the minimum value of the left-hand side over the interval \([-1, 0)\) is \( \frac{\sqrt{3}}{2} \), which occurs at \( x_L = \frac{1-\sqrt{3}}{2} \). So, if \( \theta < \frac{\sqrt{3}}{2} \), then \( C_{LR} < 0 \) for all \( x_L \in [-1, \theta - 1) \).

Thus, if \( \theta \leq \frac{\sqrt{3}}{2} \), then no profitable deviations exist for party \( L \). A symmetric argument holds for \( R \), so \( x_L^* = x_R^* = 0 \) is an equilibrium.

Next, we establish uniqueness. Suppose \( \theta \leq \frac{\sqrt{3}}{2} \), consider any \( x_L < 0 \), and let \( \hat{x}_R(x_L) \) be a best response by party \( R \). As shown above, if \( x_R = 0 \) then \( C_{LR} < 0 \), so \( S_R(x_L,0) > \frac{1}{2} \). So, \( S_R(x_L, \hat{x}_R(x_L)) > \frac{1}{2} \). This implies that \( S_L(x_L, \hat{x}_R(x_L)) < \frac{1}{2} \). But, by a symmetric argument to that above, \( S_L(0, \hat{x}_R(x_L)) \geq \frac{1}{2} \) (i.e., by choosing \( x_L = 0 \) party \( L \) can insure that it wins at least half of the districts). So, \( x_L \) is not a best response by party \( L \) to \( \hat{x}_R(x_L) \). So, \( x_L < 0 \) cannot be part of an equilibrium. The other cases follow by symmetry.

**Divergent case.** We now show that if \( \theta > \frac{\sqrt{3}}{2} \) then the unique equilibrium with \( x_L \leq x_R \) is \((x_L^*, x_R^*) = (\frac{1}{2} - \theta, \theta - \frac{1}{2}) \). (There is also a symmetric equilibrium with \( x_L^* \) and \( x_R^* \) reversed.)
First, we show that \( (x_L^*, x_R^*) = (\frac{1}{2} - \theta, \theta - \frac{1}{2}) \) is an equilibrium. Suppose \( x_L = \frac{1}{2} - \theta \). Since 
\[ \theta > \frac{\sqrt{3}}{2} > \frac{3}{4}, \frac{1}{2} - \theta < \theta - 1. \]
So, case (1) of Comment 1 applies to \( x_L \), and \( C_{LU} > 0 \). Also, if \( x_R = \theta - \frac{1}{2} \) then case (3) of Comment 1 applies to \( x_R \), and \( C_{RU} < 0 \). We now consider all possible choices of \( x_R \). Suppose \( x_R \in (1 - \theta, 1) \). Case (3) of Comment 2 applies, so 
\[ C_{LR} = (x_R - \theta + \frac{1}{2})^2/D_1, \]
where \( D_1 > 0 \). If \( x_R = \theta - \frac{1}{2} \), then \( C_{LR} = 0 \) and \( C_{RU} < 0 \), so 
\[ S_R(x_L, x_R) = \frac{1}{2}. \]
If \( x_R \neq \theta - \frac{1}{2} \), then \( C_{LR} > 0 \), so \( S_R(x_L, x_R) < \frac{1}{2} \). So, \( R \)'s best response in the interval \((1 - \theta, 1)\) is 
\[ x_R = \theta - \frac{1}{2}, \]
which produces 
\[ S_L(x_L, x_R) = S_R(x_L, x_R) = \frac{1}{2}. \]
Next, consider \( x_R \in [\theta - 1, 1 - \theta] \). Case (2) of Comment 2 applies, so 
\[ C_{LR} = (3x_R^2 + \theta^2 - \frac{3}{4})/D_2, \]
where \( D_2 > 0 \). Since \( \theta > \frac{\sqrt{3}}{2} \), \( C_{LR} > 0 \), so \( S_R(x_L, x_R) < \frac{1}{2} \). So, party \( R \) strictly prefers \( \theta - \frac{1}{2} \) to any \( x_R \in [\theta - 1, 1 - \theta] \). Finally, consider \( x_R \in [-1, \theta - 1) \). Case (1) of Comment 2 applies, so 
\[ C_{LR} = (x_R - \theta + \frac{1}{2})/3. \]
If \( x_R > \theta - \frac{1}{2} = x_L \), then \( C_{LR} > 0 \), so \( S_R(x_L, x_R) < \frac{1}{2} \). If \( x_R < \theta - \frac{1}{2} = x_L \), then \( C_{LR} < 0 \), and again \( S_R(x_L, x_R) < \frac{1}{2} \). If \( x_R = \theta - \frac{1}{2} = x_L \), then by the tie-breaking assumption party \( R \) and party \( L \) each win exactly half of the districts with medians \( y \in [-1, C_{LU}] \). But \( C_{LU} < 1 \), so again \( S_R(x_L, x_R) < \frac{1}{2} \). So, party \( R \)'s unique best response to \( x_L = \frac{1}{2} - \theta \) is \( x_R = \theta - \frac{1}{2} \). A symmetric argument shows that \( x_L = \frac{1}{2} - \theta \) is party \( L \)'s unique best response to \( x_R = \theta - \frac{1}{2} \). So, \( (x_L^*, x_R^*) = (\frac{1}{2} - \theta, \theta - \frac{1}{2}) \) is an equilibrium.

We now show uniqueness. First, we show that no equilibrium is possible where \( x_L \) and \( x_R \) are both in \([\theta - 1, 1 - \theta]\). Case (4) of Comment 2 applies, so 
\[ C_{LR} = \frac{x_L + x_R}{2}. \]
Also, if \( x_i = 0 \) then party \( i \) never loses to an unaffiliated candidate. Suppose \( x_L \neq 0 \). If party \( R \) chooses \( x_R = 0 \), then 
\[ S_R(x_L, x_R) > \frac{1}{2}, \]
so \( S_L(x_L, x_R) < \frac{1}{2} \). But \( S_L(0, 0) = \frac{1}{2} \), so \( x_L \) is not a best response to \( x_R = 0 \). So, the only possible equilibrium is \( x_L = x_R = 0 \). But, from the convergent case above, \( x_L = x_R = 0 \) is not an equilibrium when \( \theta > \frac{\sqrt{3}}{2} \).

Next, we show that no equilibrium is possible where \( x_L \) and \( x_R \) are both in \([-1, \theta - 1)\). By symmetry, this also implies there is no equilibrium where \( x_L \) and \( x_R \) are both in \((1 - \theta, 1)\). Suppose \(-1 \leq x_L < x_R < \theta - 1 \). Then case (1) of Comment 1 and case (1) of Comment 2 apply. So, 
\[ C_{LR} < C_{LU} < C_{RU}, \]
and party \( L \) wins all districts with medians \( y < C_{LR} \). But \( \frac{\partial C_{LR}}{\partial x_L} > 0 \), so \( \frac{\partial S_L}{\partial x_L} > 0 \), so \( x_L \) is not a best response to \( x_R \). Next, suppose \(-1 \leq x_L = x_R < \theta - 1 \). By the tie-breaking assumption, party \( L \) and party \( R \) each win exactly half of the districts with medians \( y \in [-1, C_{LU}] \). By case (1) of Comment 1, \( C_{LU} < 1 \), so 
\[ S_L(x_L, x_L) = S_R(x_L, x_L) < \frac{1}{2}. \]
If party \( R \) switches to \( x_R' = -x_L \), then its candidates win all districts with \( y \in [-C_{LU}, 1] \). If \( C_{LU} > 0 \), then 
\[ S_R(x_L, -x_L) = \frac{1}{2} > S_R(x_L, x_L), \]
so \( x_R \) is not a best-response to \( x_L \). If \( C_{LU} < 0 \), then 
\[ S_R(x_L, -x_L) = 2S_R(x_L, x_L), \]
so again \( x_R \) is not a
best-response to $x_L$.

Last, we show that the only equilibrium where $x_L \in [-1, \theta-1)$ and $x_R \in (1-\theta, 1]$ is the point $(x_L, x_R) = (\frac{1}{2} - \theta, \theta - \frac{1}{2})$. Only two orderings of $C_L$, $C_R$, and $C_{LR}$ are possible: $C_L < C_{LR} < C_R$, and $C_R < C_L < C_{LL}$. If $C_L < C_{LR} < C_R$, then $S_R$ is determined by $C_R$. But $C_{LR}$ is continuous in $x_R$, and $\frac{\partial C_{LR}}{\partial x_R} < 0$, so there exists $x_R' < x_R$ such that $S_R(x_L, x_R') > S_R(x_L, x_R)$. Thus, if $x_R$ implies that $C_L < C_{LR} < C_R$ holds, then $x_R$ is not a best response to $x_L$. By a symmetric argument, $x_L$ is not a best response to $x_R$, either. If $C_R < C_{LR} < C_L$ holds, then no unaffiliated candidates win, so $S_R > \frac{1}{2}$ iff $C_{LR} < 0$, $S_R = \frac{1}{2}$ iff $C_{LR} = 0$, and $S_R = 1 - S_R$. As shown above, if $x_L = \frac{1}{2} - \theta$ then $C_{LR} = (x_R - \theta + \frac{1}{2})^2/D_1$, where $D_1 > 0$. So, if $x_R \neq \theta - \frac{1}{2}$, then party $L$ can insure that $S_L > 0$ by choosing $x_L = \frac{1}{2} - \theta$. So, if $x^*_L$ is a best response to $x_R$, then $S_L(x^*_L, x_R) > \frac{1}{2}$ and $S_R(x^*_L, x_R) < \frac{1}{2}$. But then $x_R$ is not a best response to $x^*_L$, since party $R$ can insure $S_R = \frac{1}{2}$ by choosing $x_R' = -x^*_L$. So, there is no equilibrium in which $x_R \neq \theta - \frac{1}{2}$. A symmetric argument shows that there is no equilibrium in which $x_L \neq \frac{1}{2} - \theta$.

**Proof of Comment 3.**

Suppose $x_L = x_R = Y$. It is easily shown that $[1 - \sqrt{(4-\theta^2)/3}, -1 + \sqrt{(4-\theta^2)/3}] \subset [\theta - 1, 1 - \theta]$, so $x_L$ and $x_R$ are both in $[\theta - 1, 1 - \theta]$. Also, by case (2) of Comment 1, if party $i$ chooses $x_i \in [1 - \sqrt{(4-\theta^2)/3}, -1 + \sqrt{(4-\theta^2)/3}]$, then all voters prefer that party’s candidates to unaffiliated candidates. So, $S_L(Y,Y) = S_L(Y,Y) = \frac{1}{2}$.

Now consider whether party $L$ could profitably deviate to $x_L < Y$. Given such a deviation, $L$’s candidates can only win districts with $y < C_{LR}$. If $C_{LR} < Y$, then $S_L(x_L, Y) < \frac{1}{2}$. So, if $C_{LR} < Y$ for all $x_L < Y$, then no such profitable deviations exist.

There are two cases. First, if $x_L \in [\theta - 1, 0)$, then case (4) of Comment 2 applies, so $C_{LR} = \frac{\mu_L + \mu_R}{2} < 0$. Second, if $x_L \in [-1, \theta - 1)$, then case (2) of Comment 2 applies, and $\sigma_L < \sigma_R$ (so, party $L$ might gain from its lower variance). Substituting $x_R = Y$ into the equation for $C_{LR}$ yields $C_{LR} = \frac{3Y^2 + 2x_L^2 - 2x_L \theta + \theta^2}{3(2Y - x_L - \theta + 1)}$. The denominator is positive, so $C_{LR} < 0$ iff the numerator is negative, that is, iff \(-3Y^2 + 2x_L^2 - 2x_L \theta + \theta^2 < 0\). Differentiating, the minimum value of the left-hand side over the interval $[-1, 0)$ is $\frac{\sqrt{5}(2Y^2 - Y + 1)}{2(1 + Y)} = \theta^S$, which occurs at $x_L = \frac{1}{2}[1 - \sqrt{5} + 3Y - \sqrt{3Y}]$. So, if $\theta < \theta^S$, then $C_{LR} < 0$ for all $x_L \in [-1, \theta - 1)$.

Thus, if $\theta \leq \theta^S$, then no profitable deviations exist for party $L$. A symmetric argument holds for $R$, so $x_L^* = x_R^* = Y$ is an equilibrium.
Next, we establish uniqueness. Suppose \( \theta \leq \theta^S \), consider any \( x_L < Y \), and let \( \hat{x}_R(x_L) \) be a best response by party \( R \). As shown above, if \( x_R = Y \) then \( C_{LR} < Y \), so \( S_R(x_L, Y) > \frac{1}{2} \). So, \( S_R(x_L, \hat{x}_R(x_L)) > \frac{1}{2} \). This implies that \( S_L(x_L, \hat{x}_R(x_L)) < \frac{1}{2} \) (i.e., by choosing \( x'_L = Y \) party \( L \) can insure that it wins at least half of the districts). So, \( x_L \) is not a best response by party \( L \) to \( \hat{x}_R(x_L) \). So, \( x_L < Y \) cannot be part of an equilibrium. The other cases follow by symmetry. \( \blacksquare \)

**Proof of Comment 4.** Omitted in the interest of space. Available by request from authors.

**Proof of Comment 5.** Suppose \( x_L \in [\theta, 1-\theta] \). Then the ideal points of party \( L \)'s members (candidates) are distributed uniformly on the interval \( [x_L-\theta, x_L+\theta] \), and the median of the members' ideal points is \( x_L \). So, a majority of the party's members prefer \( x_L \) to any proposal \( x'_L \neq x_L \). So, \( x_L \) can be part of an equilibrium. Next, suppose \( x_L \in [-1, \theta-1) \). Then the ideal points of party \( L \)'s members are distributed uniformly on the interval \( [-1, x_L+\theta] \), and the median of the members' ideal points is greater than \( x_L \). So, there exist proposals \( x'_L > x_L \) such that a majority of the party members prefer \( x'_L \) to \( x_L \). So, \( x_L \) cannot be part of an equilibrium. An analogous argument holds for party \( R \). \( \blacksquare \)

**Proof of Comment 6.** Consider an affiliated candidate with ideal point \( z \) running in a hot race in a district with median \( y \). The affiliated candidate wins iff \( |z-y| < |z'-y| \), where \( z' \) is the unaffiliated candidate's ideal point. We work through the four cases. First, if \( z \leq 2y-1 \) (which can hold only if \( y > 0 \)), then the affiliated candidate wins iff \( z' \leq z \), so \( p(z, y) = (1+z)/2 \). Second, if \( z \in (2y-1, y) \), then the affiliated candidate wins iff \( z' \leq z \) or \( z' \geq 2y-z \), so \( p(z, y) = 1-y+z \). Third, if \( z \in [y, 2y+1) \), then the affiliated candidate wins iff \( z' \geq z \) or \( z' \leq 2y-z \), so \( p(z, y) = 1+y-z \). Finally, if \( z \geq 2y+1 \) (which can hold only if \( y < 0 \)), then the affiliated candidate wins iff \( z' \geq z \), so \( p(z, y) = (1-z)/2 \). \( \blacksquare \)

**Proof of Comment 7.** We first prove the following Lemma.

**Lemma.** If \( y > 0 \), then \( p(z, y) > p(-z, y) \) for all \( z > 0 \). If \( y < 0 \), then \( p(z, y) < p(-z, y) \) for all \( z > 0 \). \( \blacksquare \)

**Proof.** Suppose \( y > 0 \). If \( z > 0 \), then \( (z, y) \) is in case (i), (ii), or (iii) of Comment 6. Suppose \( (z, y) \) is in case (i), so \( p(z, y) = (1+z)/2 \). Then \(-z < 0 < z \leq 2y-1 \), so \((-z, y)\) is also in case
(i). So, \( p(z, y) - p(-z, y) = (1+z)/2 - (1-z)/2 = z > 0 \). Next, suppose \((z, y)\) is in case (ii), so \( p(z, y) = 1-y+z \). Since \(-z < 0 < y\), \((-z, y)\) is in case (i) or case (ii). Suppose \((-z, y)\) is in case (ii). Then \( p(-z, y) = 1-y-z \), so \( p(z, y) - p(-z, y) = 2z > 0 \). Next, suppose \((-z, y)\) is in case (i). Then \( p(-z, y) = (1-z)/2 \), so \( p(z, y) - p(-z, y) = z + (z-2y+1)/2 \). This is positive, since \( z > 0 \) and \( z > 2y-1 \). Finally, suppose \((z, y)\) is in case (iii), so \( p(z, y) = 1+y-z \). Since \(-z < 0 < y\), \((-z, y)\) is in case (i) or case (ii). Suppose \((-z, y)\) is in case (ii). Then \( p(-z, y) = 1-y+z \), so \( p(z, y) - p(-z, y) = 2y > 0 \). Next, suppose \((-z, y)\) is in case (i). Then \( p(-z, y) = (1-z)/2 \), so \( p(z, y) - p(-z, y) = y + (1-z)/2 \). This is again positive, since \( y > 0 \) and \( z < y \leq 1 \). A symmetric argument holds for \( y < 0 \). □

We are now ready to prove Comment 7. Suppose \( x = 0 \) and consider \( y > 0 \). Suppose affiliated candidates win all cold races (we check that this holds below). Let \( \underline{z}(y) = \min\{z|z \in Z(0, y)\} \) and \( \overline{z}(y) = \max\{z|z \in Z(0, y)\} \). The assumption \( c < w(1-\psi)/2 \) implies that candidates with \( z = 0 \) strictly prefer to join the party, so \( \underline{z}(y) < 0 \). Clearly, \( p(z, y) > p(\underline{z}(y), y) \) for all \( z \in (\underline{z}(y), 0] \). Also, \( \alpha(z)^2 < \alpha(\underline{z}(y))^2 \) for all such \( z \). So, \( [\underline{z}(y), 0] \subset Z(0, y) \). By Lemma 1, \( p(-z, y) > p(z, y) \) for each \( z \in [\underline{z}(y), 0] \). Also, \( \alpha(-z)^2 = \alpha z^2 \). So, \( [0, -\underline{z}(y)] \subset Z(0, y) \). Lemma 1 also implies that \( p(-\overline{z}(y), y) > p(z(y), y) \). So, there exists \( \overline{z} > -\underline{z}(y) \) such that \([-\overline{z}(y), \overline{z}] \subset Z(0, y) \). It is straightforward (although a bit tedious) to show that for each \( y \) the equation defining the ideal point of a candidate who is indifferent between joining the party and not joining has exactly two real roots on the interval \([-1, 1]\). So, \( Z(0, y) \) is a closed interval for all \( y \). So, the ideal points of affiliated candidates are distributed uniformly over \( Z(0, y) = [\underline{z}(y), \overline{z}(y)] \), where \( \overline{z}(y) > -\underline{z}(y) \).

Clearly, \( \mu(x, y) = [\overline{z}(y) + \underline{z}(y)]/2 > 0 \). Also, \( |z| < \theta \) for all \( z \in Z(0, y) \), so \( \sigma^2(y) < \theta^2/3 \). Thus, \( |\mu(y) - y| < |y| \) and \( \sigma^2(y) < 1/3 \). So, for all \( y > 0 \), in cold races a majority of the voters prefers an affiliated candidate to an unaffiliated candidate. A symmetric argument holds for \( y < 0 \). □
REFERENCES


<table>
<thead>
<tr>
<th>Predicting Retirements from the House, 1938-1998</th>
</tr>
</thead>
<tbody>
<tr>
<td>Distance from Party Mean</td>
</tr>
<tr>
<td>Terms Served</td>
</tr>
<tr>
<td>Age</td>
</tr>
<tr>
<td>(Age)$^2$</td>
</tr>
<tr>
<td>Vote Share</td>
</tr>
<tr>
<td>(Vote Share)$^2$</td>
</tr>
<tr>
<td>Member of Majority</td>
</tr>
<tr>
<td>Congress</td>
</tr>
<tr>
<td>(Congress)$^2$</td>
</tr>
<tr>
<td>constant</td>
</tr>
<tr>
<td>Observations</td>
</tr>
<tr>
<td>Psuedo-R$^2$</td>
</tr>
<tr>
<td>Log likelihood</td>
</tr>
<tr>
<td>% Δ in Retirement Prob.</td>
</tr>
</tbody>
</table>

Probit estimates.
Standard errors in parentheses.
Sample consists of all House Democrats and Republicans who were first elected prior to 1988.
% Δ in Retirement Prob. = \((\text{retirement probability of a member with ideology one standard deviation away from the party mean}) - (\text{retirement probability of a member with ideology at the party mean})\), with all other variables held at their means.
Table 2
Predicting Voter Placements of House Representatives
Dep. Var. = Placement of House Member on 7-Point Lib./Con. Scale

<table>
<thead>
<tr>
<th></th>
<th>All Members</th>
<th>All Members</th>
<th>Democratic Members</th>
<th>Republican Members</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>All Voters</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>W-Nominate Score</td>
<td>1.42 (.04)</td>
<td>---</td>
<td>1.04 (.14)</td>
<td>.67 (.15)</td>
</tr>
<tr>
<td>Party Dummy</td>
<td>---</td>
<td>1.41 (.04)</td>
<td>---</td>
<td>---</td>
</tr>
<tr>
<td>R² within year</td>
<td>.24 (.04)</td>
<td>.23</td>
<td>.02</td>
<td>.01</td>
</tr>
<tr>
<td>R² overall</td>
<td>.24 (.04)</td>
<td>.23</td>
<td>.03</td>
<td>.01</td>
</tr>
<tr>
<td># Obs.</td>
<td>3633</td>
<td>3633</td>
<td>1947</td>
<td>1686</td>
</tr>
<tr>
<td><strong>“Informed” Voters</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>W-Nominate Score</td>
<td>1.90 (.04)</td>
<td>---</td>
<td>1.17 (.14)</td>
<td>.73 (.15)</td>
</tr>
<tr>
<td>Party Dummy</td>
<td>---</td>
<td>1.97 (.04)</td>
<td>---</td>
<td>---</td>
</tr>
<tr>
<td>R² within year</td>
<td>.42 (.04)</td>
<td>.43</td>
<td>.04</td>
<td>.02</td>
</tr>
<tr>
<td>R² overall</td>
<td>.42 (.04)</td>
<td>.43</td>
<td>.04</td>
<td>.02</td>
</tr>
<tr>
<td># Obs.</td>
<td>2708</td>
<td>2746</td>
<td>1389</td>
<td>1319</td>
</tr>
</tbody>
</table>

OLS estimates.
Standard errors in parentheses.
Year dummies included in all regressions.
“Informed” voters are those who reported voting and who also assigned the Democratic party a more liberal position than the Republican party.
Figure 1: Game Sequence

Parties:  
Platform Selection
Nature:  
Candidate Generation
Candidates:  
Candidate Affiliation
Voters:  
Voting

Figure 2: Platform and Affiliation Choices in a Single-Party Electorate

Figure 3a: Platform and Affiliation Choices in a Two Party Electorate
Figure 3b: Platform and Affiliation Choices in a Two Party Electorate

Candidate Ideal Points

\[ \theta - \frac{1}{2} \]

\[ 0 \]

\[ 1/2 - \theta \]

Candidates Join Party L and Win

Candidates Join Party R and Win

District Medians

Figure 4: Surplus- and Seat-Maximizing Equilibria

\[
\begin{align*}
\text{Party } R \\
\text{Platform}
\end{align*}
\]

\[
\begin{align*}
\text{Surplus} & \quad \text{Seats}
\end{align*}
\]

\[
\begin{align*}
\theta
\end{align*}
\]