

AN INFORMATIONAL RATIONALE FOR POLITICAL PARTIES¹

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

October, 2001

¹Forthcoming in *American Journal of Political Science* 46(1): 90-110. We gratefully acknowledge the financial support of National Science Foundation Grant SES-0079035. We thank Tim Feddersen, Scott Desposato, Dhammika Dharmapala, John Roemer, and seminar participants at the University of North Carolina, Stanford, UCLA, and Yale for helpful comments.

Abstract

This paper studies a model of political 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 costs on their members, or less able to screen out certain types of candidates, their platforms move further apart.

1. Introduction

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, helping voters make decisions and helping 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 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. Much effort is currently being devoted to measuring the effects of party on various legislative decisions, such as committee assignments, agenda control, and roll call voting.¹ Moreover, the scholars conducting this research often invoke a 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 probably requires tying the empirical work more closely to theory.

This paper develops a simple model of party brand names that provides tentative answers to some of the questions above. We hope the underlying framework proves 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 most 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 matters. Unions care about

¹See, *e.g.*, Rohde (1991), Kiewiet and McCubbins (1991), Cox and McCubbins (1993), Snyder and Groseclose (2000), Hager and Talbert (2000), and Ansolabehere, *et al.* (2001b).

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 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. Democratic candidates tend to be liberal, and Republicans tend to be conservative. 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, often the most important members. Parties provide many of the resources politicians use to win office, including campaign contributions, lists of likely contributors, and campaign workers. The party’s winning candidates work together to formulate policy in national, state, and local governments. The party also plays 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 policy preferences or ideologies are close to the preferences of the bulk of a party’s members will tend to incur low costs from party membership. On the other hand, candidates with preferences or ideologies that are far from the center of a party will face 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 yields electoral benefits.

Our model brings these two ideas together. Party labels may be valuable to candidates

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

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. In this case, in equilibrium only candidates of a similar ideological stripe will be willing to signal their type by running under a party's banner. Parties can then serve as effective screening devices, as in Spence (1974).

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.³ 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.⁴ Voters learn about various *groups* of candidates. In particular, they learn summary statistics (mean, variance) about the candidates affiliated with the major political parties. We might imagine that voters obtain this relatively crude information from media coverage.

The basic model makes three major predictions. First, parties are effective at aggregating candidates and communicating 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

³Of course, parties might also play a role in mitigating moral hazard problems. For example, they might reduce the degree to which their members cater to narrow special interests.

⁴This is the assumption in Alesina (1988).

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 8 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 model relates 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, numerous papers have analyzed models of elections and policy outcomes when there is moral hazard, adverse selection, or both.⁵ 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 (1992) are more relevant for our work, because they provide a theoretical rationale for political parties. 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.

Second, our work adds to the substantial literature on the role of party platforms in electoral competition. In the Downsian model, 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 party (or candidate) policy preferences, third party entry, uncertainty about voter preferences, reputational costs, valence issues, and decentralized party decision making.⁶ As explained above, our model illustrates another

⁵See Fearon (1999) for an excellent discussion of this work.

⁶See, *e.g.*, Wittman (1983), Palfrey (1984), Calvert (1985), Bernhardt and Ingberman (1985), Londregan

reason non-centrist, divergent 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.⁷ 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 to solve the adverse selection problems posed by elections, we establish a motivation 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 screen 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. In the model below, the amount of party discipline or testing is not a strategic choice variable; but in future work it will be.

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 6-8 consider various extensions of the basic model. Section 9 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.” Data on party switchers, committee assignments, and retirements in Congress motivate this. The second

and Romer (1993), and Snyder (1994).

⁷See, for example, Aldrich (1994), Dion and Huber (1996), and Diermeier and Feddersen (1998). Austen-Smith and Banks (1988) and Baron (1993) are exceptions, but they treat parties as unitary actors.

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 conflicts that are 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 conflicts some members face. Andy Ireland (Fla.) switched to the Republican Party in 1984. Observers attributed his move to “the frustration of being joined in caucus with colleagues few of whom shared his views.”⁸ Richard Shelby (Al.) switched to the Republican Party in 1994. “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’.”⁹ Billy Tauzin (La.) 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.”¹⁰

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.”¹¹ Mike Parker (Miss.), who also switched to the Republicans in 1995, “was clearly uncomfortable with the

⁸ *Almanac of American Politics*, 1985, p. 299.

⁹ *The Commercial Appeal*, November 10, 1994.

¹⁰ *The Washington Post*, August 8, 1995.

¹¹ *Almanac of American Politics*, 1997, p. 1371.

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.”¹² Nathan Deal (Ga.) switched to the Republican Party in 1995 after finding himself leading the opposed by most Democrats and the Democratic leadership on a Clean Water Bill he drafted. “It was an uncomfortable and awkward situation where a bill he supported was opposed by the party leadership.”¹³

The pattern of House committee assignments provides additional evidence that parties impose differential costs and benefits on members. If party leaders influence the allocation of valuable committee slots, then party loyalists may 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.¹⁴

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).¹⁵

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

¹² *Almanac of American Politics*, 1997, p. 815.

¹³ *Roll Call*, April 13, 1995.

¹⁴ See Rohde and Shepsle (1973), Smith and Ray (1983), Cox and McCubbins (1993), and Sinclair (1995).

¹⁵ 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.

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).¹⁶

These findings are not too surprising. Several well-known studies have shown that the parties' incumbents are polarized and only moderately responsive to constituency ideology.¹⁷ More recently, Ansolabehere, *et al.* (2001a) 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. Note also 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. 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.¹⁸

In the basic model below we make the simplifying assumption that voters cannot discriminate within parties at all. In section 8 we relax this assumption, and study an extension where voters sometimes learn the true positions of the candidates running in a race.

¹⁶A recent study finds 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 perception 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.

¹⁷See, *e.g.*, Stone (1980) and Poole and Rosenthal (1997).

¹⁸See, *e.g.*, Erikson (1971), Johannes and McAdams (1981), Wright (1978), Erikson and Wright (1989, 1997), and Bernstein (1989). Challengers' positions also 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).

3. The Basic Model

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 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 $X \equiv [-1, 1]$.

We consider cases with one party and cases with two parties. In both, there is a set of unaffiliated candidates which we denote by U . When there are two parties, we denote them by L and R . Each party chooses a platform, x_L or x_R , on X . 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 sections 6-7 we study different party goals and more decentralized party decision making processes.

Candidates are driven by achieving office, and, if elected, policy.¹⁹ Election winners receive utility $w > 0$ from holding office, while losers receive zero. In each district, there is set of potential candidates, with ideal points distributed uniformly on $[-1, 1]$. These 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 through their party affiliations.

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.²⁰

¹⁹We could assume instead that candidates also care about policy when they lose, but this complicates the analysis and appears to add little in return.

²⁰Aldrich and Bianco (1992) also study a model in which candidates can choose which party to join. However, they focus on purely office-seeking candidates, there is no spatial element in their model, and 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 only one office. Also, the

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 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—forcing party members to attend meetings with fellow partisans, wrap themselves in party symbols, and occasionally vote the party line. 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.²¹

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).²²

Under the first interpretation, the parameter α reflects the magnitude of the costs born by candidates, or the degree of party discipline in a party. Under the second interpretation, α reflects the discriminating power of the screening technology. In what follows, we adopt the language of the first interpretation—candidates choosing whether they wish to join 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

electoral model focuses on the trade-offs candidates face between attracting resources from party activists and appealing to the median voter.

²¹The required activities do *not* affect voters’ utility. 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.

²²We 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.

office as a party member. We assume that $\theta < 1$. As will be clear shortly, this insures that party affiliation will never appeal to all types of candidates.

Voters are also policy driven, but they care only about the policy chosen by the official who wins in the district where they live.²³ Each voter has quadratic utility, and an ideal point in $[-1, 1]$. Denote the median ideal point in a generic district by y , and denote the median of district medians by Y . For most of the analysis 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, but later we relax it somewhat.

The sequence of game play is as follows.

1. *Platform Selection.* Parties choose platforms x_L and x_R simultaneously, observed by all players.
2. *Candidate Nomination Bids.* In each district, all potential candidates simultaneously choose which party label or labels (R , L , or U) they would accept if offered a nomination. Voters do not observe these choices.
3. *Candidate Selection.* In each district, Nature randomly draws a party L candidate from the set of those who would accept the party L nomination (if the set is non-empty). Similarly, Nature randomly selects a party R candidate and an unaffiliated candidate.
4. *Voting.* In each district, winners are decided by plurality rule.

Note that the set of candidates running in each district, and their party affiliations, depends on two decisions: the choices by potential candidates about which parties' nominations they would accept (if any), and Nature's choices about which of the potential candidates actually receive nominations. Each candidate's nomination bid is a three-element set, where for example the bid $\{L, \emptyset, U\}$ indicates a willingness to run as either a party L candidate or unaffiliated, but not as a party R candidate. Let \mathbf{a} denote the vector of choices available to voters in a typical district, where $a_j \in \{L, R, U\}$.²⁴

²³In legislative elections, voters should also care about the electoral outcomes in other districts, as in Austen-Smith (1984). We leave this for future work.

²⁴Allowing more than one unaffiliated candidate to run does not change the analysis. Since voters perceive all unaffiliated candidates to be identical, our assumption that there is at most one of them effectively maximizes the power of unaffiliated candidates in the model.

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, Perfect Bayesian equilibria of the game are characterized by the following four elements:

Party Platforms, $x_L \in X$ and $x_R \in X$.

Candidate Nomination Bids, $d: X^2 \rightarrow \{\emptyset, L\} \times \{\emptyset, R\} \times \{\emptyset, U\}$ for each potential candidate.

Voter Beliefs, $b_j: X^2 \times \{\emptyset, L\} \times \{\emptyset, R\} \times \{\emptyset, U\} \rightarrow \mathcal{B}[-1, 1]$ for each candidate $j = 1, \dots, |\mathbf{a}|$.²⁵

Voting Functions, $v: X^2 \times \{\emptyset, L\} \times \{\emptyset, R\} \times \{\emptyset, U\} \rightarrow \{1, \dots, |\mathbf{a}|\}$.

In words, each party chooses a platform, and all potential candidates declare which party nominations they would accept.²⁶ Based on the set of nomination bids $\{d\}$, Nature randomly selects a vector of actual candidates \mathbf{a} , which contains at most one candidate with each label. Voters use the platforms and affiliations, plus Bayes' Rule, to form beliefs about the intervals in which candidates' ideal points lie. 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.

Given these expectations, Bayes' Rule implies that 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

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

where μ_i and σ_i^2 are the mean and variance of the ideal points of the set of candidates willing to affiliate with party i . Since voters cannot observe whether or not unaffiliated candidates sought party nominations but were rejected by Nature, they are unable to infer

²⁵We can denote voters' updated beliefs as subsets of X (more formally, Borel sets on X) because they are uniform. This happens 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 note 32.

²⁶As will be clear below, restricting the platforms to the interval $[-1, 1]$ is not essential.

anything new about their preferences. To see why, suppose that some types of potential candidates find it profitable to declare interest in running unaffiliated in equilibrium. Since such candidates pay no type-specific costs, *all* potential candidates must then be willing to run unaffiliated. So, in equilibrium voters simply use their priors in evaluating unaffiliated candidates. Although not technically a party, we extend the notation in the obvious way so that $\mu_U = 0$ and $\sigma_U^2 = 1/3$. The median voter's expected utility if an unaffiliated candidate wins is therefore $E[-(y-z)^2|a_j = U] = -y^2 - 1/3$.

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 is affiliated. Thus, if a party L candidate enters out of equilibrium, voters in that district make the same inference about her preferences that voters in other districts would make about a party L candidate who enters in equilibrium.

We make five tie-breaking assumptions to simplify the analysis. First, when a voter is indifferent between a party candidate and an unaffiliated candidate, she votes for the party candidate. Second, when a voter is indifferent between two party candidates even though the party platforms are distinct, she votes for party R 's candidate. Third, potential candidates do not offer affiliation with a party if they expect that party to lose, or if they would receive negative utility upon winning.²⁷ Fourth, potential candidates offer to run unaffiliated even if they expect to lose. Finally, when party platforms are identical, 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.²⁸

Equilibria in this game are partially separating, because the choice of party affiliation may serve as a noisy signal of a candidate's preferences (*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

²⁷This assumption is necessary because there are an infinite number of potential candidates, and the probability of being selected by Nature to be a candidate is therefore zero. By ruling out these weakly dominated strategies, we essentially impose a robustness requirement where potential candidates behave as if the candidate pool were finite.

²⁸While 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.

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 one party. To conserve notation we drop the party subscript.

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.

The Candidate Affiliation Decision. Suppose the party position is x . All potential candidates weakly prefer running unaffiliated to not running, and thus declare interest in doing so. On the other hand, if a potential candidate believes that she will win by joining the party (*i.e.*, declaring interest and being chosen by Nature), then she will join if $w - \alpha(x-z)^2 \geq c$, or $z \in [x - \sqrt{(w-c)/\alpha}, x + \sqrt{(w-c)/\alpha}] = [x-\theta, x+\theta]$. 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,

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

and

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

Notice that x not only affects μ , but it also affects σ . In particular, σ 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 σ above.) Moreover, this is not something particular to the uniform distribution, but a property that holds for a broad class of symmetric, single-peaked, distributions.²⁹

Notice also that when x is not too extreme (the middle case above), party members are

²⁹In 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.

distributed symmetrically around x , so $\mu(x) = x$. On the other hand, when x is extreme, μ is less extreme than x , that is, $|\mu(x)| < |x|$.

The Voting Decision. Voters use Bayes' Rule to determine expectations and variances of the positions of candidates, based on their prior knowledge of the distribution of candidate ideal points and the candidates' nomination bid strategies. Thus, if a candidate belongs to the party, 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 is 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 the affiliated candidate to the unaffiliated candidate? Clearly, the ideal points of party members must belong to a strict subset of $[-1, 1]$, so that some 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

$$-(y-\mu(x))^2 - \sigma^2 > -y^2 - 1/3. \quad (3)$$

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.³⁰ Thus if the party chooses platform $x = 0$, all candidates with ideal points within $[-\theta, \theta]$ join the party if offered the opportunity. These candidates beat unaffiliated candidates in every district because of the informational benefits (*i.e.*, variance reduction) offered by the party.

While the one-party case is 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 center of the distribution of *candidate* ideal points, rather than the distribution of voters' ideal points. Of course, the center of the candidate distribution will often be near the center of the voter distribution, *i.e.*, Y will be near 0. In such cases, we should see moderate platforms even in one-party systems. When it does not hold, however, the model predicts that the party will track candidates rather than voters. 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.

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.

The Voting Decision. As above, voters use Bayes' Rule to determine the expected policy positions of the candidates. If a candidate belongs to 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 is unaffiliated, then voters use their priors and assume that her ideal point is drawn from $U[-1, 1]$.

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

³⁰As 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.

defines the voter who is indifferent between candidates (provided $\mu_i \neq \mu_j$):

$$C_{ij} = \frac{\mu_i + \mu_j}{2} + \frac{\sigma_i^2 - \sigma_j^2}{2(\mu_i - \mu_j)}. \quad (4)$$

Consider first the problem of choosing between an affiliated candidate and an unaffiliated candidate. This analysis is the same as that in section 4, 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 the top case, if θ 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.³¹

The Party Platform Decision. There are two types of equilibria, depending on the value of θ . 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 θ is sufficiently large.

Proposition 1. If $\theta < \sqrt{3}/2$, then the unique equilibrium platforms are $(x_L^*, x_R^*) = (0, 0)$ (*convergent equilibrium*). If $\theta > \sqrt{3}/2$, then the unique equilibrium platforms are $(x_L^*, x_R^*) = (\frac{1}{2} - \theta, \theta - \frac{1}{2})$ (*divergent equilibrium*). ■

In both types of equilibria, each party wins half of the districts. Figures 1 and 2 display the resulting platform locations, party affiliations, and election outcomes as a function of candidate ideal points and district medians. Figure 1 shows the convergent equilibrium, and Figure 2 shows the divergent equilibrium. Note that in the divergent case, the distance between the parties is increasing in θ .

For small values of θ , 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 θ is large, however, the convergent equilibrium cannot exist because a party that locates in the center will attract candidates from such a wide range of the ideological spectrum that its label conveys relatively little information about the preferences of its members. Parties then have an incentive to adopt

³¹Different tie-breaking rules can produce losing party candidates, since potential candidates assess a zero probability of being selected (given that a positive measure of candidates declares interest) and are thus indifferent between joining a party and not. Moreover, one might imagine that parties subsidize such candidates just slightly, so that they will run and “do their part” in maintaining the party's brand nationally. Thus the absence of losing-party candidates is not fundamental to the model.

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, $\theta = \sqrt{3}/2$, both convergent and divergent equilibria exist.³²

There are no non-centrist, convergent equilibria, even when θ is large. This is because of 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).

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 θ 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.

We conclude this section with two results on the distribution of candidates. In the analysis above the distribution of candidate ideal points is fixed, and Proposition 1 characterizes equilibrium platforms as θ varies. We could instead fix θ , and vary the distribution of

³²The results do not depend on the uniform distribution. For example, we can 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. (Because these values have full support on the real line, we also allow platforms to be located anywhere on \mathfrak{R} and allow θ to be any non-negative number.) Then a result similar to Proposition 1 holds: when θ is small then the equilibrium is convergent, with $(x_L^{N*}, x_R^{N*}) = (0, 0)$, and when θ is large enough then the equilibrium is divergent, with $(x_L^{N*}, x_R^{N*}) = (-x^{N*}(\theta), x^{N*}(\theta))$. As in the uniform case, divergence is caused by the lower variance of candidates when platforms are located in the tails of a unimodal distribution. The boundary between the two cases is given by $\theta \approx 1.61$. It is impossible to derive an analytic solution for the equilibrium platforms in the divergent case, but we can calculate the equilibrium for various values of θ . For example, if $\theta = 1.7$ then $x^{N*} = .55$, and if $\theta = 2.0$ then $x^{N*} = 1.10$.

candidate types. Specifically, suppose candidate ideal points are distributed uniformly on $[-M, M]$. Then an increase in M is equivalent to a decrease in θ . In fact, it is straightforward to show that if $\theta < \frac{M\sqrt{3}}{2}$, then the equilibrium platforms converge to 0, and if $\theta > \frac{M\sqrt{3}}{2}$, then the equilibrium platforms diverge to $x_R^* = -x_L^* = \theta - \frac{M}{2}$. Thus, increasing M increases the likelihood that parties converge, and decreases the degree of divergence when the equilibrium remains divergent. The intuition is straightforward. If M is large, then a party must move far away from the center of district medians in order to reduce its variance. If M is too large, then the loss in voter utility from the extreme platform always outweighs the variance reduction, so parties will not find such a move profitable and the equilibrium is convergent.

In the analysis above the distribution of candidate ideal points and the distribution of district medians are centered at the same point ($Y = 0$). While this may be the most natural assumption, certain factors might cause $Y \neq 0$. Gerrymandering and malapportionment (*e.g.*, over-representation of rural areas) are possibilities. Also, the time and money needed to run for political office might limit the set of potential candidates to a select subset of the citizenry.³³ 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 . This result generalizes, partially, to cases where $Y \neq 0$.

Comment 3. Suppose $|Y| \leq \sqrt{(4-\theta^2)/3}$, 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 $|Y| > \sqrt{(4-\theta^2)/3}$, then Y is so extreme that a party located at Y will lose against unaffiliated candidates in some districts. In such cases, $x_L = x_R = Y$ may not be an equilibrium. 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 just to the right of Y .

³³We might expect the variance of candidate ideal points to be greater than the variance of district medians. This is because (i) 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 (ii) 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 extreme, districts. Since we made no assumption about the variance of district medians, the results above clearly hold regardless of this variance.

6. Parties that Maximize Total Net Benefits of Their Members

In this section 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 for each party i is:

$$\begin{aligned} B_i(x_L, x_R) &= S_i(x_L, x_R) \int_{\max\{-1, x_i - \theta\}}^{\min\{x_i + \theta, 1\}} [w - \alpha(z - x_i)^2] dz \\ &= \alpha S_i(x_L, x_R) [\theta^2 - \sigma_i^2(x_i) - (\mu(x_i) - x_i)^2] \\ &= \begin{cases} \alpha S_i(x_L, x_R) [2\theta^2 + \theta(1 + x_i) - (1 + x_i)^2] / 3 & \text{for } x_i \in [-1, \theta - 1] \\ \alpha S_i(x_L, x_R) [2\theta^2 / 3] & \text{for } x_i \in [\theta - 1, 1 - \theta] \\ \alpha S_i(x_L, x_R) [2\theta^2 + \theta(1 - x_i) - (1 - x_i)^2] / 3 & \text{for } x_i \in (1 - \theta, 1] \end{cases} \end{aligned}$$

For large values of θ , the equilibrium platforms under this objective are difficult to characterize. It is also difficult to characterize the maximum value of θ such that equilibrium platforms remain convergent. However, an approximate analog to Proposition 1 is captured by the following result.

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

Figure 3 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 3 have three implications. First, the range of θ 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 ($\frac{\sqrt{3}}{2}$) 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 $\theta \rightarrow 1$. This is the opposite of what happens when parties maximize their share of offices.

7. 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 platform divergence.

Comment 5. The platform pair (x_L^{D*}, x_R^{D*}) is an equilibrium if and only if $x_L^{D*} \in [\theta - 1, 1 - \theta]$ and $x_R^{D*} \in [\theta - 1, 1 - \theta]$. ■

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.

8. Behavior When Voters May Learn More about Candidates

Up to now we have assumed that voters never learn anything about any particular candidates. This precludes the possibility of “personal”, candidate-centered campaigns, which are common for certain offices in the U.S. and some other countries. 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. 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 ψ is the same for all districts, and that whether or not a particular race is hot is independent of what happens in 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

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

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. The model therefore captures the idea that Massachusetts Republicans are more liberal than Georgia Republicans because voters in Massachusetts are more liberal than voters in Georgia. Party candidates are “responsive” to district ideology.

Note finally that in almost all districts, the unaffiliated candidate wins with positive probability *ex ante* because some her type will be closer to the district median than the party candidate. The extended model thus 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 $\alpha \rightarrow \infty$, $\bar{p}(1, \alpha) = \bar{p}(-1, \alpha) \rightarrow \frac{1}{2}$, $\bar{p}(0, \alpha) \rightarrow 1$, and $\bar{p}(\alpha) \rightarrow \frac{2}{3}$. 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 $\bar{p}(.25) \approx .70 > \frac{2}{3}$. Because of the responsive candidate affiliation decisions, the party wins more hot races in the extreme districts when α 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.

9. 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 wish to join different parties in different districts.

Party platforms are also endogenous, and inter-party competition has a large effect on platform positioning. With one party, there exists a continuum of equilibria, in which the party chooses a platform near the center of the distribution of *candidate* ideal points, and party candidates win in every district. When two parties are present, the parties do not locate themselves for maximum candidate coverage. Instead, competitive pressures give them incentives both to locate at the center of the distribution of *district medians*, as well as to sharpen their message by choosing more extreme positions. Thus our model refines the standard Downsian logic: both convergent and divergent equilibria exist, depending on the benefit of holding office relative to the cost of party membership. In all cases, however, parties are able to convey brand name information through their platform choices.

We plan to pursue several extensions of the model in future work, some of which are mentioned above. First, the “hot race” analysis of section 8 must be solved for the two party case. It is expected that such a model will produce similar comparative statics to the one-party case, with more candidate responsiveness to district type the higher is ψ . These may potentially be tested using measures of media coverage across localities. Second, we will explore the mechanisms that parties use to select candidates. One way to do this is to allow parties to choose θ (perhaps through a primary system) and to characterize optimal disciplining or screening schemes. While parties in our basic model clearly have incentives to reduce θ , 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.

Perhaps most ambitiously, the model provides a promising basis for examining other 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 labels are unimportant in highly centralized systems, because there are few

important elected offices. Races will usually 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 preference of a larger number of officials.

Our model is highly stylized, and is best viewed mainly as a theoretical framework on which to build. Nonetheless, the model makes some predictions that may be testable, possibly through cross-sectional analyses of U.S. states or a set of countries, or time series analyses of particular countries or states. One clear prediction is that if the power of parties' screening or disciplining technologies increases, then parties should become more homogeneous and platforms should converge. A similar effect should be observed if the pool of candidate types becomes more heterogeneous. Measuring the heterogeneity of the pool of potential candidates, however, or parties' abilities to screen or discipline members, is no easy task. One possibility is that primary elections reduce the power to screen and discipline candidates. If so, the widespread adoption of primaries in the U.S. around the turn of the century should have led the parties to become more heterogeneous, and caused their platforms to diverge. Also, assuming the candidate pool reflects the population at large, we have the somewhat paradoxical result that party platforms should exhibit more divergence in homogeneous countries than in heterogeneous countries.

Another prediction of the model is that in moving from a one-party regime to a divergent two-party system, the original party should become more ideologically extreme (either to the left or the right) and more homogeneous, and voter information should increase. This might be tested by studying state Democratic parties in the south during the past thirty years. The recent Mexican experience may soon provide yet another case to study. Finally, it is likely that the two-party version of the hot-race model of section 8 will lead to the prediction that intra-party "responsiveness" to constituency preferences increases as the probability of a hot race increases. If so, this could be tested by looking across offices—for example, the probability of a hot race is probably higher for U.S. Senate and gubernatorial races than it is for other offices. We hope that extensions and refinements of the basic model will lead to even more interesting hypotheses and tests.

APPENDIX

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 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 $C = \frac{x}{2} + \frac{\theta^2/3-1/3}{2x} = \frac{x}{2} + \frac{\theta^2-1}{6x}$. 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 an unaffiliated one. Third, if $x \in (1 - \theta, 1]$, then $C = \frac{x-\theta+1}{2} + \frac{(-x+\theta+1)^2/12-1/3}{2(x-\theta+1)/2} = \frac{x-\theta}{3}$. 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 $y < C_{LR} = \frac{\mu_L + \mu_R}{2} + \frac{\sigma_L^2 - \sigma_R^2}{2(\mu_L - \mu_R)}$. There are six cases, depending on the locations of x_L and x_R .

First, if $x_L \in [-1, \theta - 1)$ and $x_R \in [-1, \theta - 1)$, then $\mu_L = \frac{x_L + \theta - 1}{2}$, $\mu_R = \frac{x_R + \theta - 1}{2}$, $\sigma_L^2 = \frac{(x_L + \theta + 1)^2}{12}$, and $\sigma_R^2 = \frac{(x_R + \theta + 1)^2}{12}$. Substituting into (4) yields: $C_{LR} = \frac{x_L + x_R + 2\theta - 1}{3}$. Second, if $x_L \in [-1, \theta - 1)$ and $x_R \in [\theta - 1, 1 - \theta]$, then $\mu_L = \frac{x_L + \theta - 1}{2}$, $\mu_R = x_R$, $\sigma_L^2 = \frac{(x_L + \theta + 1)^2}{12}$, and $\sigma_R^2 = \frac{\theta^2}{3}$. Substituting into (4) yields: $C_{LR} = \frac{3x_R^2 - x_L^2 - 2x_L\theta + x_L + \theta - 1}{3(2x_R - x_L - \theta + 1)}$. Third, if $x_L \in [-1, \theta - 1)$ and $x_R \in (1 - \theta, 1]$, then $\mu_L = \frac{x_L + \theta - 1}{2}$, $\mu_R = \frac{x_R - \theta + 1}{2}$, $\sigma_L^2 = \frac{(x_L + \theta + 1)^2}{12}$, and $\sigma_R^2 = \frac{(1 - x_R + \theta)^2}{12}$. Substituting into (4), $C_{LR} = \frac{(x_L + x_R)(x_R - x_L - 2\theta + 1)}{3(x_R - x_L - 2\theta + 2)}$. 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 $\sigma_L^2 = \sigma_R^2 = \frac{\theta^2}{3}$ and substituting into (4) yields: $C_{LR} = \frac{x_L + x_R}{2}$. Fifth, if $x_L \in [\theta - 1, 1 - \theta]$ and $x_R \in (1 - \theta, 1]$, then by symmetry with the second case, $C_{LR} = \frac{x_R^2 - 3x_L^2 - 2x_R\theta + x_R - \theta + 1}{3(x_R - 2x_L - \theta + 1)}$. Finally, if $x_L \in (1 - \theta, 1]$ and $x_R \in (1 - \theta, 1]$, then by symmetry with the first case, $C_{LR} = \frac{x_L + x_R - 2\theta + 1}{3}$. ■

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, iff $\frac{x_L^2 - x_L + 1}{1 - 2x_L} > \theta$. 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 there is no equilibrium 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 $(x_L, x_R) = (\frac{1}{2} - \theta, \theta - \frac{1}{2})$. Only two orderings of C_{LU} , C_{RU} , and C_{LR} are possible: $C_{LU} < C_{LR} < C_{RU}$, and $C_{RU} < C_{LR} < C_{LU}$. If $C_{LU} < C_{LR} < C_{RU}$, then S_R is determined by C_{RU} . But C_{LR} is continuous in x_R , and $\frac{\partial C_{RU}}{\partial x_R} < 0$, 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 $C_{LU} < C_{LR} < C_{RU}$ 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 $C_{RU} \leq C_{LR} \leq C_{LU}$ 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 $S_L = 1 - S_R$. 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, $S_L(Y, Y) = S_R(Y, Y) = \frac{1}{2}$.

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 $\sigma_L^2 < \sigma_R^2$. Substituting $x_R = Y$ into the equation for C_{LR} yields $C_{LR} = \frac{3Y^2 - x_L^2 - 2x_L\theta + x_L + \theta - 1}{3(2Y - x_L - \theta + 1)}$. The denominator is positive, so $C_{LR} < 0$ iff the numerator is negative, that is, iff $\frac{-3Y^2 + x_L^2 - x_L + 1}{1 - 2x_L} > \theta$. Differentiating, the minimum value of the left-hand side over $[-1, 0)$ is $\frac{\sqrt{3}(2Y^2 + Y + 1)}{2(1 + Y)} = \theta^S$, which occurs at $x_L = \frac{1}{2}[1 - \sqrt{3} + 3Y - \sqrt{3}Y]$. 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 $S_L(x_L, \hat{x}_R(x_L)) < \frac{1}{2}$. But, by a symmetric argument to that above, $S_L(Y, \hat{x}_R(x_L)) \geq \frac{1}{2}$ (*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. ■

Proof of Comment 4. Convergent Case. Suppose $x_R = 0$. If $x_L \in [\theta - 1, 1 - \theta]$, then $B_L(x_L, x_R) = [2\theta^2/3]S_L(x_L, x_R)$. So, on the interval $[\theta - 1, 1 - \theta]$, party L maximizes net benefits iff it maximizes its share of the offices. By the proof the convergent case of Proposition 1, party L 's best response on this interval is 0. Next, on the interval $[-1, \theta - 1)$, the maximum

per-member surplus attainable by party L is $\frac{3\theta^2}{4}$ (by choosing $x_L = \frac{\theta}{2} - 1$), while the surplus at $x_L = 0$ is $\frac{2\theta^2}{3}$. Thus, there exists no best response on $[-1, \theta - 1]$ if $B(x_L, 0) < \frac{\theta^2}{3}$, or equivalently $C_{LR} < -\frac{1}{9}$. For reasons that will be clear in the uniqueness portion of the proof we solve for the sufficient condition $C_{LR} \leq -\frac{1}{8}$. By case (2) of Comment 2 and some manipulation, this obtains if $8x_L^2 + 16x_L\theta - 5x_L - 5\theta + 5 \geq 0$. The left-hand side of this expression is convex and minimized at $x_L = \frac{5}{16} - \theta$. Substituting and solving for θ , we see that $C_{LR} \leq -\frac{1}{8}$ if $\theta \leq \frac{3\sqrt{15}}{16}$. By symmetry, $(0, 0)$ is an equilibrium.

We now show uniqueness. Suppose that there exists an equilibrium in which $x_L < 0$. If $x_L \geq \theta - 1$, then an identical argument to the uniqueness proof of the convergent case of Proposition 1 holds. If $x_L < \theta - 1$, then by the above argument $B_R(x_L, 0) \geq \alpha \frac{9}{16} \frac{2\theta^2}{3} = \frac{3\alpha\theta^2}{8}$. There are three cases. First, suppose that R 's best response satisfies $\hat{x}_R^B(x_L) \in (1 - \theta, 1]$. Since the maximum per-member surplus attainable is $\frac{3\theta^2}{4}$, party R must attain $S_R(x_L, \hat{x}_R^B(x_L)) \geq \frac{1}{2}$ to prefer $\hat{x}_R^B(x_L)$ to 0. Thus, $S_L(x_L, \hat{x}_R^B(x_L)) \leq \frac{1}{2}$, and by the above argument L could do strictly better by locating at 0. Second, if $\hat{x}_R^B(x_L) \in [\theta - 1, 1 - \theta]$, then since R maximizes its seat share, $S_R(x_L, \hat{x}_R^B(x_L)) \geq \frac{9}{16}$ and hence $B_L(x_L, \hat{x}_R^B(x_L)) \leq \alpha \frac{7}{16} \frac{3\theta^2}{4} = \frac{21\alpha\theta^2}{64}$. But by choosing $x_L = 0$, party L could receive $B_L(0, \hat{x}_R^B(x_L)) \geq \alpha \frac{1}{2} \frac{2\theta^2}{3} = \frac{\alpha\theta^2}{3} > \frac{21\alpha\theta^2}{64}$. Finally, if $\hat{x}_R^B(x_L) \in [-1, \theta - 1]$ then $S_i(x_L, \hat{x}_R^B(x_L)) < \frac{1}{2}$ for some party i , and by the above argument that party would prefer $x_i = 0$. Thus $x_L < 0$ cannot be a part of an equilibrium, a contradiction. The other cases hold by symmetry.

Divergent Case. We break the proof up into a series of claims. Suppose $\theta \geq \frac{\sqrt{3}}{2}$.

Claim 1. Any equilibrium must satisfy $x_L^{B*} \in [-1, \theta - 1)$, $x_R^{B*} \in (1 - \theta, 1]$. ■

First, we address whether both platforms may lie in the interval $[\theta - 1, 1 - \theta]$. Recall that party i maximizes net benefits on this interval iff it maximizes seat share ($S_i(\cdot)$). If $\theta > \frac{\sqrt{3}}{2}$, then by Proposition 1 there is no equilibrium with x_L and x_R both in $[\theta - 1, 1 - \theta]$ when parties maximize $S_i(\cdot)$, so there is also no such equilibrium when parties maximize $B_i(\cdot)$. If $\theta = \frac{\sqrt{3}}{2}$, then by the proof of Proposition 1 the only mutual best response platforms both in $[\theta - 1, 1 - \theta]$ are $(0, 0)$, where $S_L(0, 0) = S_R(0, 0) = \frac{1}{2}$. But party R achieves a higher per candidate expected surplus at $x_R = \theta - \frac{1}{2} > 1 - \theta$, while $S_R(0, \theta - \frac{1}{2}) = \frac{1}{2}$. Thus $(0, 0)$ is not an equilibrium.

Second, if $x_L^{B*} \in (1 - \theta, 1]$ and $x_R^{B*} \in (1 - \theta, 1]$, then party L can achieve the same

surplus by locating at $-x_L^{B*} \notin (1-\theta, 1]$. By case (6) of Comment 2 and case (3) of Comment 1, $S_L(-x_L^{B*}, x_R^{B*}) > S_L(x_L^{B*}, x_R^{B*})$ if $\frac{1}{2}[\min\{C_{LR}, C_{LU}\} + 1] > \frac{1}{2}[C_{LR} - C_{LU}]$, or $\min\{\frac{-x_L^{B*} + x_R^{B*} + 2\theta - 1}{3}, \frac{-x_L^{B*} + \theta}{3}\} + 1 > \frac{x_L^{B*} + x_R^{B*} - 2\theta + 1}{3} - \frac{x_L^{B*} - \theta}{3}$. This condition obtains trivially. Third, the result for the interval $[-1, \theta - 1]$ follows by symmetry.

Fourth, suppose $x_L^{B*} \in [\theta - 1, 1 - \theta]$ and $x_R^{B*} \in (1 - \theta, 1]$. If $S_L(x_L^{B*}, x_R^{B*}) \leq \frac{1}{2}$, then $B_L(\max\{-x_R^{B*}, -\theta\}, x_R^{B*}) > B_L(x_L^{B*}, x_R^{B*})$. If $S_L(x_L^{B*}, x_R^{B*}) > \frac{1}{2}$, then since 0 was not chosen by party R , $B_R(x_L^{B*}, x_R^{B*}) \geq B_L(x_L^{B*}, x_R^{B*})$, with the inequality strict if $x_L^{B*} \neq 0$. But since candidate expected surplus is equal at $-x_R^{B*}$ and x_R^{B*} , $B_L(-x_R^{B*}, x_R^{B*}) \geq B_R(x_L^{B*}, x_R^{B*})$, with the inequality strict if $x_L^{B*} = 0$. Thus $B_L(-x_R^{B*}, x_R^{B*}) > B_L(x_L^{B*}, x_R^{B*})$ and x_L^{B*} cannot be a best response to x_R^{B*} . Finally, the case $x_R^{B*} \in [-1, \theta - 1]$, $x_L^{B*} \in [\theta - 1, 1 - \theta]$ holds by symmetry.

Claim 2. If $x_R \in (1 - \theta, 1]$, then the unique best response on $[-1, \theta - 1]$ satisfies $\hat{x}_L^B(x_R) \in (\frac{\theta}{2} - 1, \hat{x}_L(x_R))$, where $\hat{x}_L(x_R)$ is party L 's best response to x_R when it maximizes its share of the offices (from the proof of Proposition 1). ■

We first reduce the range of x_L where $\hat{x}_L^B(x_R)$ must lie. Differentiating $B_L(x_L, x_R)$, we obtain $\frac{\partial B_L}{\partial x_L} = \frac{\alpha}{3}[\frac{\partial S_L}{\partial x_L}(2\theta^2 + \theta(1 + x_L) - (1 + x_L)^2) + S_L(x_L, x_R)(\theta - 2(1 + x_L))]$. Consider each term in this expression. Clearly, $2\theta^2 + \theta(1 + x_L) - (1 + x_L)^2 > 0$, and $S_L(x_L, x_R) > 0$. If $x_L > (=)(<) \frac{\theta}{2} - 1$, then $\theta - 2(1 + x_L) < (=)(>) 0$. Finally, because the distribution of districts is assumed to be uniform, $S_L(x_L, x_R) = \frac{1 + C_{LR}}{2}$. Substituting from case (3) of Comment 2 and differentiating $S_L(x_L, x_R)$ yields $\frac{\partial S_L}{\partial x_L} = \frac{1}{6}[\frac{2x_L + 2\theta - 1}{x_L - x_R + 2\theta - 2} - \frac{(x_L + x_R)(x_L - x_R + 2\theta - 1)}{(x_L - x_R + 2\theta - 2)^2}]$. Differentiating again yields $\frac{\partial^2 S_L}{\partial x_L^2} = \frac{1}{3}[\frac{1}{x_L - x_R + 2\theta - 2} - \frac{2x_L + 2\theta - 1}{(x_L - x_R + 2\theta - 2)^2} + \frac{(x_L + x_R)(x_L - x_R + 2\theta - 1)}{(x_L - x_R + 2\theta - 2)^3}]$. Reducing, $\frac{\partial^2 S_L}{\partial x_L^2} < 0$ if $(x_L + x_R + 1)(x_L - x_R + 2\theta - 2) < (x_L + x_R)(x_L - x_R + 2\theta - 1)$, or $x_R > \theta - 1$. Thus $S_L(x_L, x_R)$ is concave if $x_R \in (1 - \theta, 1]$. We therefore also know that $\hat{x}_L(x_R) = x_R - 2\theta + 2 - \sqrt{2x_R - 2\theta + 2}$, and $\frac{\partial S_L}{\partial x_L} > 0$ iff $x_L < \hat{x}_L(x_R)$.

Putting these results together and noting that $\hat{x}_L(x_R) > \frac{\theta}{2} - 1$, we see that $\frac{\partial B_L}{\partial x_L} > 0$ if $x_L \leq \frac{\theta}{2} - 1$. If $x_L \geq \hat{x}_L(x_R)$, then $\frac{\partial B_L}{\partial x_L} < 0$. By the continuity of $B_L(\cdot)$, we conclude that $\hat{x}_L^B(x_R) \in (\frac{\theta}{2} - 1, \hat{x}_L(x_R))$. Taking the second derivative of $B_L(x_L, x_R)$, we obtain $\frac{\partial^2 B_L}{\partial x_L^2} = \frac{\alpha}{3}[\frac{\partial^2 S_L}{\partial x_L^2}(2\theta^2 + \theta(1 + x_L) - (1 + x_L)^2) + 2\frac{\partial S_L}{\partial x_L}(\theta - 2(1 + x_L)) - 2S_L(x_L, x_R)]$.

Accumulating the previous results, $\frac{\partial^2 B_L}{\partial x_L^2} < 0$ on $(\frac{\theta}{2} - 1, \hat{x}_L(x_R))$. Thus $B_L(\cdot)$ is pseudo-concave and the first-order condition is sufficient for characterizing party L 's best response over $[-1, \theta - 1]$.

Claim 3. There exists a unique platform pair $(\tilde{x}_L, \tilde{x}_R)$ such that \tilde{x}_L and \tilde{x}_R are mutual best responses on $[-1, \theta-1)$ and $(1-\theta, 1]$, respectively, and $\tilde{x}_L = -\tilde{x}_R$. ■

We consider the class of symmetric platforms $(x_L, -x_L)$ where $x_L \in [-1, \theta-1)$ and find a best response pair. Differentiating $B_L(x_L, x_R)$ and substituting $x_R = -x_L$, we obtain $\frac{\partial B_L}{\partial x_L} = \frac{\alpha}{6} \left[\frac{2x_L+2\theta-1}{3(2x_L+2\theta-2)} (2\theta^2 + \theta(1+x_L) - (1+x_L)^2) + \theta - 2(1+x_L) \right]$.

Simple manipulation shows that $\frac{\partial B_L}{\partial x_L} > 0$ if $x_L \in [-1, \frac{\theta}{2}-1)$ and $\frac{\partial B_L}{\partial x_L} < 0$ if $x_L > \frac{1}{2}-\theta$. Thus $x_L \in [\frac{\theta}{2}-1, \frac{1}{2}-\theta]$ in any symmetric best response platform. To characterize the symmetric best responses in this region, substitute $x_R = -x_L$ into $\frac{\partial^2 B_L}{\partial x_L^2}$ to obtain $\frac{\partial^2 B_L}{\partial x_L^2} = \frac{\alpha}{9} \left[-\frac{2\theta^2 + \theta(1+x_L) - (1+x_L)^2}{(2x_L+2\theta-2)^2} + \frac{2x_L+2\theta-1}{2x_L+2\theta-2} (\theta - 2(1+x_L)) - 6S_L(x_L, x_R) \right]$.

Again, simple manipulation shows that $\frac{\partial^2 B_L}{\partial x_L^2} < 0$ if $x_L \in [\frac{\theta}{2}-1, \frac{1}{2}-\theta]$. Thus the first order condition characterizes a unique best response platform \tilde{x}_L on $[-1, \theta-1)$. By symmetry, $\tilde{x}_R = -\tilde{x}_L$ is party R 's best response on $(\theta-1, 1]$.

Claim 4. There does not exist a platform pair (x'_L, x'_R) such that x'_L and x'_R are mutual best responses on $[-1, \theta-1)$ and $(1-\theta, 1]$, respectively, and $x'_L \neq -x'_R$. ■

Suppose that such a platform pair exists. We first establish that the ‘‘mirror’’ platforms $(-x'_R, -x'_L)$ must also be mutual best responses. For any (x_L, x_R) , the symmetry of C_{LR} and $S_i(\cdot)$ implies that $S_L(-x_R, -x_L) = S_R(x_L, x_R)$ and $\frac{\partial S_L}{\partial x_L}(-x_R, -x_L) = -\frac{\partial S_R}{\partial x_R}(x_L, x_R)$. Thus, $B_L(-x_R, -x_L) = B_R(x_L, x_R)$, and $\frac{\partial B_L}{\partial x_L}(-x_R, \cdot) = -\frac{\partial B_R}{\partial x_R}(\cdot, x_R)$. Since the first-order conditions are sufficient for characterizing best responses, if (x'_L, x'_R) are mutual best responses on $[-1, \theta-1)$ and $(1-\theta, 1]$, respectively, then $(-x'_R, -x'_L)$ must also be.

To show that such asymmetric platforms cannot be in equilibrium, we examine the behavior of $\frac{\partial B_L}{\partial x_L}$ over the set of platforms. Differentiating with respect to x_R , $\frac{\partial^2 B_L}{\partial x_L \partial x_R} = \frac{\alpha}{3} \left[\frac{\partial^2 S_L}{\partial x_L \partial x_R} (2\theta^2 + \theta(1+x_L) - (1+x_L)^2) + \frac{\partial S_L}{\partial x_R} (\theta - 2(1+x_L)) \right]$.

We evaluate $\frac{\partial^2 B_L}{\partial x_L \partial x_R}$ for the set of pairs (x_L, x_R) where $x_L = \tilde{x}_L$. Differentiating, $\frac{\partial^2 S_L}{\partial x_L \partial x_R} < 0$ for $x_R \leq \tilde{x}_R$. Further, $2\theta^2 + \theta(1+x_L) - (1+x_L)^2 > 0$, $\frac{\partial S_L}{\partial x_R} \geq 0$ for $x_R \leq \tilde{x}_R$, and $\theta - 2(1+x_L) < 0$ for $x_L = \tilde{x}_L$. Simplifying, $\frac{\partial^2 B_L}{\partial x_L \partial x_R}(\tilde{x}_L, \cdot) < 0$ for $x_R \leq \tilde{x}_R$. Since \tilde{x}_L and \tilde{x}_R are best responses in their respective regions, $\frac{\partial B_L}{\partial x_L}(\tilde{x}_L, \tilde{x}_R) = 0$. Thus, for any $x_R < \tilde{x}_R$, $\frac{\partial B_L}{\partial x_L}(\tilde{x}_L, x_R) > 0$.

By the proof of Claim 3, if $x_R > (<) \tilde{x}_R$, $\frac{\partial B_L}{\partial x_L}(-x_R, x_R) > (<) 0$. Since the first-order conditions are sufficient for characterizing $\hat{x}_L^B(x_R)$, for any $x_R < \tilde{x}_R$, $\frac{\partial B_L}{\partial x_L}(\tilde{x}_L, x_R) > 0$, and

hence $\hat{x}_L^B(x_R) \in (\tilde{x}_L, -x_R)$. And, for any $x_R > \tilde{x}_R$, $\hat{x}_L^B(x_R) > -x_R$. Summarizing, party L 's best response platform on $[-1, \theta - 1)$ must lie in the following regions:

$$\hat{x}_L^B(x_R) \in \begin{cases} \{x_L, x_R \mid x_L > \tilde{x}_L, x_L < -x_R\} & \text{for } x_R < \tilde{x}_R \\ \{x_L, x_R \mid x_L > -x_R\} & \text{for } x_R > \tilde{x}_R \end{cases}$$

Thus, if (x'_L, x'_R) lies within the above regions, then $(-x'_R, -x'_L)$ does not, a contradiction. Thus, no asymmetric mutual best responses in the stated regions exist.

Claim 5. $\hat{x}_L^B(\tilde{x}_R) \in [-1, \theta - 1)$. ■

Since L 's surplus for locating in $[\theta - 1, 1 - \theta]$ is lower than her surplus for locating in $[-1, \theta - 1)$, a sufficient condition for there to be no best response in $[\theta - 1, 1 - \theta]$ is that $C_{LR} < 0$. From case (5) of Comment 2, $C_{LR} < 0$ if $3x_L^2 - x_R^2 + 2x_R\theta - x_R + \theta - 1 > 0$. Substituting $\tilde{x}_R > \theta - \frac{1}{2}$, $\tilde{x}_R < 1 - \frac{\theta}{2}$, and $\theta \geq \frac{\sqrt{3}}{2}$, this inequality holds. So, $\hat{x}_L^B(\tilde{x}_R) \notin [\theta - 1, 1 - \theta]$.

By the proof of Claim 1, if $x_L \in (1 - \theta, 1]$, then party L could achieve the identical surplus and higher seat share (and thus higher expected net benefits) by choosing $-x_L$. But $-x_L \in [-1, \theta - 1)$, which by Claim 2 is weakly inferior to \tilde{x}_L . So, $\hat{x}_L^B(\tilde{x}_R) \notin (1 - \theta, 1]$.

Since $\hat{x}_L^B(\tilde{x}_R) \in [-1, \theta - 1)$ and \tilde{x}_L is party L 's best response in $[-1, \theta - 1)$, we conclude that $x_L^{B*} = \tilde{x}_L$. By symmetry, $x_R^{B*} = \tilde{x}_R$. ■

Proof of Comment 5. Suppose $x_L \in [\theta - 1, 1 - \theta]$. 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 . ■

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 $p(z, y) = (1 - z)/2$. ■

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 $z > 0$. ■

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 now 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 $\bar{z}(y) = \max\{z | z \in Z(0, y)\}$. 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 the Lemma, $p(-z, y) > p(z, y)$ for each $z \in [\underline{z}(y), 0]$. Also, $\alpha(-z)^2 = \alpha z^2$. So, $[0, -\underline{z}(y)] \subset Z(0, y)$. The Lemma also implies that $p(-\underline{z}(y), y) > p(\underline{z}(y), y)$. So, there exists $\bar{z} > -\underline{z}(y)$ such that $[-\underline{z}(y), \bar{z}] \subset Z(0, y)$. It is straightforward 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 in $[-1, 1]$. So for all y , $Z(0, y) = [\underline{z}(y), \bar{z}(y)]$, where $\bar{z}(y) > -\underline{z}(y)$, and the ideal points of affiliated candidates are distributed uniformly over $Z(0, y)$.

Clearly, $\mu(x, y) = \frac{\bar{z}(y) + \underline{z}(y)}{2} > 0$. Also, $|z| < \theta$ for all $z \in Z(0, y)$, so $\sigma^2(y) < \frac{\theta^2}{3}$. Thus,

$|\mu(y) - y| < |y|$ and $\sigma^2(y) < \frac{1}{3}$. So, for all $y > 0$, in cold races a majority of the voters prefers the affiliated candidate. A symmetric argument holds for $y < 0$. ■

REFERENCES

- Aldrich, John H. 1983. "A Downsian Spatial Model with Party Activism." *American Political Science Review* 77:974-990.
- Aldrich, John H. 1994. "A Model of a Legislature with Two Parties and a Committee System." *Legislative Studies Quarterly* 19:313-340.
- Aldrich, John H. 1995. *Why Parties?: The Origin and Transformation of Political Parties in America*. Chicago: University of Chicago Press.
- Aldrich, John H., and William T. Bianco. 1992. "A Game-Theoretic Model of Party Affiliation of Candidates and Office Holders." *Mathematical and Computer Modeling* 16:103-116.
- Aldrich, John H., and Michael D. McGinnis. 1983. "A Model of Party Constraints on Optimal Candidate Positions." *Mathematical and Computer Modeling* 12:437-450.
- Alesina, Alberto. 1988. "Credibility and Convergence in a Two-Party System with Rational Voters." *American Economic Review* 78:796-805.
- Alesina, Alberto, and Stephen E. Spear. 1988. "An Overlapping Generations Model of Electoral Competition." *Journal of Public Economics* 37:359-379.
- Ansolabehere, Stephen D., James M. Snyder, Jr., and Charles Stewart, III. 2001a. "Candidate Positioning in U.S. House Elections." *American Journal of Political Science* 45:136-159.
- Ansolabehere, Stephen D., James M. Snyder, Jr., and Charles Stewart, III. 2001b. "The Effects of Party and Preferences on Congressional Roll Call Voting." *Legislative Studies Quarterly*, forthcoming.
- Austen-Smith, David. 1984. "Two-Party Competition with Many Constituencies." *Mathematical Social Sciences* 7:177-198.
- Austen-Smith, David, and Jeffrey S. Banks. 1988. "Elections, Coalitions, and Legislative Outcomes." *American Political Science Review* 82:405-422.
- Baron, David P. 1993. "Government Formation and Endogenous Parties." *American Political Science Review* 87:34-47.
- Bernhardt, Daniel M., and Daniel E. Ingberman. 1985. "Candidate Reputations and the 'Incumbency Effect'." *Journal of Public Economics* 27:47-67.
- Bernstein, Robert A. 1989. *Elections, Representation, and Congressional Voting Behavior: The Myth of Representation*. Englewood Cliffs, NJ: Prentice Hall.
- Calvert, Randall. 1985. "Robustness of the Multidimensional Voting Model: Candidate Motivations, Uncertainty, and Convergence." *American Journal of Political Science* 39:69-95.
- Cox, Gary W., and Matthew D. McCubbins. 1993. *Legislative Leviathan*. Berkeley: University of California Press.

- Diermeier, Daniel, and Timothy J. Feddersen. 1998. "Cohesion in Legislatures and the Vote of Confidence Procedure." *American Political Science Review* 92:611-621.
- Dion, Douglas, and John Huber. 1996. "Party Leadership and Procedural Control in Legislatures." *Journal of Politics* 58:25-53.
- Downs, Anthony. 1957. *An Economic Theory of Democracy*. New York: Harper and Row.
- Enelow, James M., and Melvin J. Hinich. 1981. "A New Approach to Voter Uncertainty in the Downsian Spatial Model." *American Journal of Political Science* 25:483-493.
- Erikson, Robert S. 1971. "The Electoral Impact of Congressional Roll Call Voting." *American Political Science Review* 65:1018-1032.
- Erikson, Robert S., and Gerald C. Wright, Jr. 1989. "Voters, Candidates, and Issues in Congressional Elections." In *Congress Reconsidered*, 4th ed., L.C. Dodd and B.I. Oppenheimer (eds.). Washington: Congressional Quarterly Press.
- Erikson, Robert S., and Gerald C. Wright, Jr. 1997. "Voters, Candidates, and Issues in Congressional Elections." In *Congress Reconsidered*, 6th ed., L.C. Dodd and B.I. Oppenheimer (eds.). Washington: Congressional Quarterly Press.
- Fearon, James D. 1999. "Electoral Accountability and the Control of Politicians: Selecting Good Types Versus Sanctioning Poor Performance." In *Democracy, Accountability, and Representation*, B. Manin, A. Przeworski, and S. Stokes (eds.). Cambridge: Cambridge University Press.
- Franklin, Charles H. 1991. "Eschewing Obfuscation? Campaigns and the Perception of U.S. Senate Incumbents." *American Political Science Review* 85:1193-1214.
- Hager, Gregory L., and Jeffery C. Talbert. 2000. "Looking for the Party Label: Party Influences on Voting in the U.S. House." *Legislative Studies Quarterly* 25:75-99.
- Harrington, Joseph E., Jr. 1992. "The Role of Party Reputation in the Formation of Policy." *Journal of Public Economics* 49:107-121.
- Hibbing, John R. 1982. "Voluntary Retirement from the U.S. House of Representatives: Who Quits?" *American Journal of Political Science* 26:467-484.
- Johannes, John R., and John C. McAdams. 1981. "The Congressional Incumbency Effect: Is It Casework, Policy Compatibility, or Something Else? An Examination of the 1978 Election." *American Journal of Political Science* 25:512-542.
- Kiewiet, D. Roderick, and Mathew D. McCubbins. 1991. *The Logic of Delegation*. Chicago: University of Chicago Press.
- Kiewiet, D. Roderick, and Langche Zeng. 1993. "An Analysis of Congressional Career Decisions, 1947-1986." *American Political Science Review* 87:928-941.
- Londregan, John, and Thomas Romer. 1993. "Polarization, Incumbency, and the Personal Vote." In *Political Economy: Institutions, Information, Competition, and Representation*, W.A. Barnett, M.J. Hinich, and N.J. Schofield (eds.). Cambridge: Cambridge University Press.

- Palfrey, Thomas R. 1984. "Spatial Equilibrium with Entry." *Review of Economic Studies* 51:139-156.
- Poole, Keith T., and Howard Rosenthal. 1997. *Congress: A Political-Economic History of Roll Call Voting*. New York: Oxford University Press.
- Rohde, David W. 1991. *Parties and Leaders in the Postreform House*. Chicago: University of Chicago Press.
- Rohde, David W., and Kenneth Shepsle. 1973. "Democratic Committee Assignments in the House of Representatives: Strategic Aspects of a Social Choice Process." *American Political Science Review* 67:889-905.
- Shepsle, Kenneth. 1978. *The Giant Jigsaw Puzzle*. Chicago: University of Chicago Press.
- Sinclair, Barbara. 1995. *Legislators, Leaders, and Lawmaking: the House of Representatives in the Post-Reform Era*. Baltimore: Johns Hopkins University Press.
- Smith, Steven S., and Bruce A. Ray. 1983. "The Impact of Congressional Reform: House Democratic Committee Assignments." *Congress and the Presidency* 10:219-240.
- Snyder, James M., Jr. 1994. "Safe Seats, Marginal Seats, and Party Platforms: The Logic of Party Platform Differentiation." *Economics and Politics* 6:201-213.
- Snyder, James M., Jr., and Tim Groseclose. 2000. "Estimating Party Influence in Congressional Roll Call Voting." *American Journal of Political Science* 44:193-211.
- Spence, Michael. 1974. *Market Signalling*. Cambridge, MA: Harvard University Press.
- Stone, Walter J. 1980. "The Dynamics of Constituency: Electoral Control in the House." *American Politics Quarterly* 8:399-424.
- Wittman, Donald. 1983. "Candidate Motivation: A Synthesis of Alternatives." *American Political Science Review* 77:142-157.
- Wright, Gerald C. 1978. "Candidates' Policy Positions and Voting in U.S. Congressional Elections." *Legislative Studies Quarterly* 3:445-464.

| Table 1 | |
|---|------------------------------|
| Predicting Retirements from the House, 1938-1998 | |
| Distance from Party Mean | 0.43 (0.13) |
| Terms Served | 0.03 (0.01) |
| Age | -0.05 (0.15) |
| (Age) ² | 0.03 (0.01) |
| Vote Share | -7.79 (1.37) |
| (Vote Share) ² | 4.81 (0.90) |
| Member of Majority | -0.13 (0.04) |
| Congress | -1.20 (0.54) |
| (Congress) ² | 0.07 (0.03) |
| constant | 5.38 (2.48) |
| Observations | 12649 |
| Pseudo-R ² | 0.085 |
| Log likelihood | -2220.9 |
| % Δ in Retirement Prob. | 26.9 |

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. = (retirement probability of a member with ideology one standard deviation away from the party mean)/(retirement probability of a member with ideology at the party mean) - 1, with all other variables held at their means.

| Table 2 | | | | |
|--|---------------|---------------|--------------------|--------------------|
| Predicting Voter Placements of House Representatives | | | | |
| Pooled NES Sample, 1980, 1982, 1986, 1990, 1994, 1996, 1998 | | | | |
| Dep. Var. = Placement of House Member on 7-Point Lib./Con. Scale | | | | |
| | All Members | All Members | Democratic Members | Republican Members |
| All Voters | | | | |
| W-Nominate Score | 1.38 (.04) | — | 1.04 (.12) | .76 (.13) |
| Party Dummy | — | 1.37 (.04) | — | — |
| R ² within year | .23 | .22 | .03 | .02 |
| R ² overall | .23 | .22 | .03 | .01 |
| # Obs. | 4388 | 4388 | 2338 | 2050 |
| “Informed” Voters | | | | |
| W-Nominate Score | 1.85 (.04) | — | 1.16 (.13) | .83 (.13) |
| Party Dummy | — | 1.90 (.04) | — | — |
| R ² within year | .40 | .40 | .05 | .03 |
| R ² overall | .40 | .41 | .05 | .02 |
| # Obs. | 3232 | 3232 | 1649 | 1583 |

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: Platform and Affiliation Choices in a Two Party Electorate

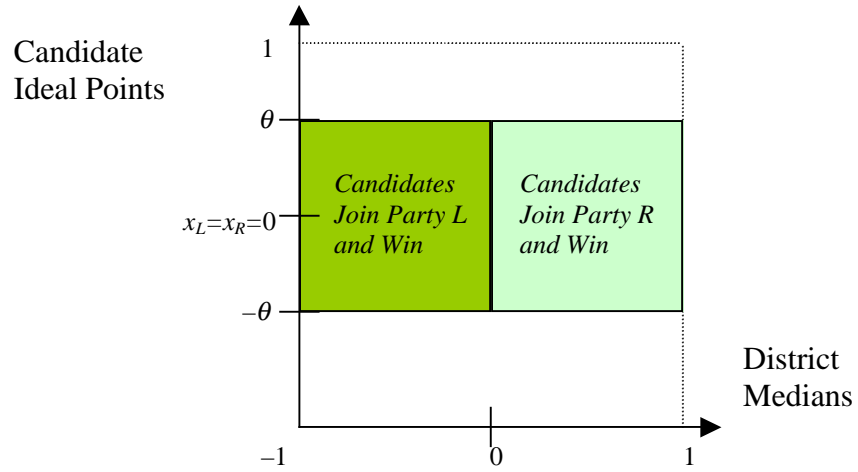


Figure 2: Platform and Affiliation Choices in a Two Party Electorate

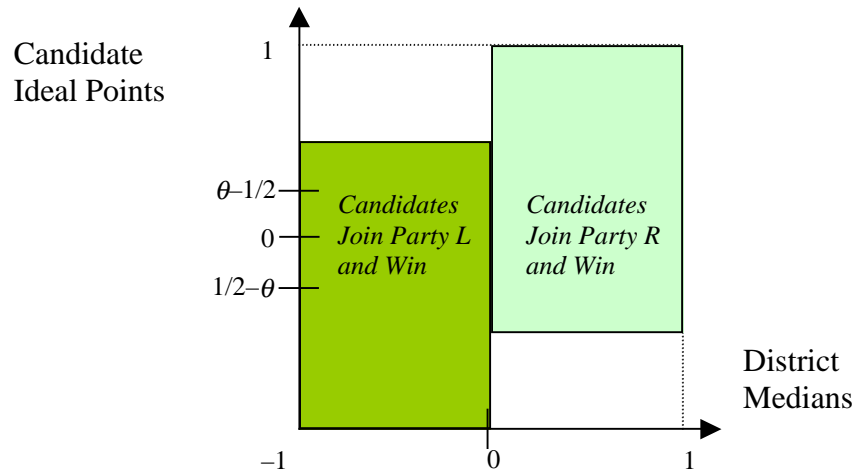


Figure 3: Surplus- and Seat-Maximizing Equilibria

