

When parties are not teams: party positions in single-member district and proportional representation systems

Stephen Ansolabehere · William Leblanc · James M. Snyder Jr.

Received: 6 August 2010 / Accepted: 2 March 2011 / Published online: 23 March 2011
© Springer-Verlag 2011

Abstract Theoretical analyses of party positions commonly assume that parties act as teams to maximize their legislative representation. This assumption runs counter to another line of theorizing in which individual legislators maximize their own chances of winning reelection. To resolve this tension, the paper presents a model of party platform choice that relaxes only the assumption that parties are teams in the classical two-party spatial model. Platforms are chosen by majority rule among all legislators within a party. Politicians seek to win their own seats in the legislature, but they must run under a common party label. In both single-member district and proportional representation systems, equilibrium platforms are shown to diverge substantially, with one party located near the 25th percentile of the voter distribution and the other near the 75th percentile, rather than converge to the median. The model also yields predictions concerning short-term economic shocks, incumbency advantages, and gerrymandering.

Keywords Political economy · Elections · Political parties

JEL Classification I235: Political Economy

We wish to thank Kristen Kanthak, Jiyeon Kim, Rebecca Morton, and participants at workshops at Harvard University and Princeton University for their comments.

S. Ansolabehere · J. M. Snyder Jr. (✉)
Department of Government, Harvard University,
Cambridge, MA, USA
e-mail: jsnyder@gov.harvard.edu

W. Leblanc
Department of Political Science, Massachusetts Institute of Technology,
Needham, MA, USA

1 Introduction

This paper presents a model of ideological voters choosing between two competing parties. These parties are made up of election-oriented, self-interested incumbents running for reelection in districts. Furthermore, the model assumes that voters, possibly due to informational constraints, can only observe one universal party position for each party. That is, members of each party are forced to run under one common “platform.” However, the candidates of the party do not have identical goals—each candidate represents a particular district and wishes to win her particular district. The model assumes that the party position that voters use to decide which party to support is determined by current incumbents in the party. This can either be interpreted literally, as incumbents voting based on majority rule on a platform, or as an approximation of voters perceiving the aggregate behavior of partisans during the last legislative session and thereby observing the implicit party positions. (For ease of argument, the model uses the literal interpretation.) Finally, the model assumes that incumbents are not certain about voter choices. This is represented by a valence term in voter utilities that makes all voters more likely to support one party over the other, independent of ideology. These and other supporting axioms are formalized in section three.

The basic argument is as follows. In this setup, there exists a cut-point such that all voters left of the cut-point support the left party, and all voters right of the cut-point support the right party. We will show that each incumbent has induced ordinal preferences that correspond to the preferences of the median voter of the district she represents. As long as each party won at least some districts in the last election, the median incumbent in the left (right) party will desire a platform left (right) of the preferences of the median voter in the median district. Since one platform is left of center, and the other platform is right of center, the platforms are diverged.

While the basic argument above for divergence is fairly simple, working through the model in detail yields other, more subtle, testable predictions. These predictions are discussed in sections four and five. The degree of divergence remains large even as the uncertainty about valence becomes small (unlike [Calvert \(1985\)](#) and similar models). In section four, we will show that if the distribution of districts is symmetrical, there is a sense in which platforms at the first and third quartiles of district medians are stable. We will show the ordering of parties (left and right) endures over time. We will show that parties that do well in the past tend to moderate and therefore do better than they would have otherwise in the future. Perhaps, most surprisingly of all, we will show that if the distribution of districts is asymmetrical (as might be caused by majority–minority districting), a party can be permanently electorally disadvantaged, even with ideological voters and flexible party platforms.

In section six, we will contrast the districted case above with a party list, proportional representation system. Surprisingly, most of the conclusions remain the same, with the distribution of voters replacing the role of the distribution of district medians. Of course, if the distribution of districts does not equal the distribution of voters, there may be significant policy implications. For example, party positions may be further apart, and election results may be less sensitive to economic shocks.

2 Previous literature

The seminal analytical model of electoral competition among parties, due to [Hotelling \(1929\)](#) and [Downs \(1957\)](#), begins with the assumption that parties are teams. All candidates and legislators within a party unite around a common goal of winning control of government and present a common ideological position. Voters choose among parties on the basis of their relative distance from the national party platforms. These assumptions lead to a powerful prediction—the Median Voter Theorem. When electoral competition in a two-party system is waged along one-dimension ideological spectrum, the parties will converge to the ideal point of the median voter. Many variants on this model have been developed, allowing, for example, open entry, multiple numbers of parties, varieties of electoral laws, and strategic voting.¹ This important line of inquiry, however, retains the assumption that political parties act as teams. Every politician within a party works for the common good in her or his party, even at the expense of the individual's own electoral fortunes.

The assumption that parties are teams runs counter to a second line of analytical inquiry in Political Science. Politicians, [David Mayhew \(1974\)](#) famously argued, single mindedly seek election. They are self-interested and act in ways to protect and improve their own positions. As a result, in a system like the United States, legislators are very responsive to the preferences of the voters in their own districts and parties resemble somewhat disorganized groups more than teams. (See also, [Fiorina 1980](#), and [Krehbiel 2000](#)). [Robertson \(1976\)](#) criticizes the application of the Downsian model to the United Kingdom because, even in a parliament with strong parties, the parties are internally divided into factions and individual politicians' interests; they are not teams. Surprisingly, little theoretical research has approached the tension between the rational choice models of parties and the rational choice models of candidates.

Ours is not, of course, the first model to extend the Downsian framework by examining internal party conflicts and organization. There are three important strands of theorizing about internal party decisions. The first line of theory assumes competition across many different districts or constituencies. [Snyder \(1994\)](#) analyzes a model similar to ours, but his model is deterministic. That model yields indifference or flat parts of politicians' revealed preferences and multiple equilibria. Under voter uncertainty, though, politicians running for single-member districts are never indifferent about what position their party takes. Even a little uncertainty on the part of candidates induces politicians to maximize their probabilities and produces strong preference orderings over the party platforms.

¹ There is an extensive theoretical literature predicting platform divergence-based factors such as entry deterrence, politicians' policy preferences, voter abstention, primary elections, valence issues, party activists, and special interest groups. Examples include [Aranson and Ordeshook \(1972\)](#), [Wittman \(1977\)](#), [Wittman \(1983\)](#), [Enelow and Hinich \(1981\)](#), [Aldrich \(1983\)](#), [Palfrey \(1984\)](#), [Bernhardt and Ingberman \(1985\)](#), [Calvert \(1985\)](#), [Aldrich and McGinnis \(1989\)](#), [Cox \(1990\)](#), [Ingberman and Villani \(1993\)](#), [Londregan and Romer \(1993\)](#), [Baron \(1994\)](#), [Snyder \(1994\)](#), [Snyder and Ting \(2002\)](#), [Poutvaara \(2003\)](#), [Schofield \(2004\)](#), [Serra \(2005\)](#), and [Callander \(2005\)](#). With the exceptions of [Snyder \(1994\)](#) and [Snyder and Ting \(2002\)](#), however, these models either explicitly or implicitly deal with a single district or parties with collective preferences.

In fact, the relevant literature goes back much farther, at least to [Robertson \(1976\)](#)'s informal discussion of the logic of party strategy. He argues that one of the crucial functions of political parties is to establish connections across the potentially unrelated races for different legislative seats, that one means of doing this is to provide national party positions on salient issues, and that natural intra-party divisions will arise in deciding what these positions should be. [Austen-Smith \(1984, 1986\)](#) extends and refines Robertson's insights by analyzing game-theoretic models in which candidates announce individual platforms to appeal to their districts, but must also commit to an aggregate party platform that will be implemented as policy.² [Calvert and Isaac \(1981\)](#) model the potential tensions between a party's legislators and the president.³ We will assume a simpler structure. Politicians decide using majority rule within their legislative caucus what platform they wish to implement.

A second strand of the literature assumes that politicians or party activists themselves have policy preferences. [Aldrich \(1983\)](#), [Aldrich and McGinnis \(1989\)](#), [Poutvaara \(2003\)](#), [Gomberg et al. \(2004\)](#), and others treat parties as collections of "activists" or primary voters with policy preferences and endogenize both party membership and party positions. Most recently, [Roemer \(1998, 2002\)](#) models internal party decision making using a solution concept he calls "Party Unity Nash Equilibria." He assumes that parties are composed of three types of actors—militants who care about "publicity," reformists who care about policy outcomes, and opportunists who care only about winning office—and choose their platforms by unanimity rule. In these models, there is just one constituency, and internal party conflict is driven by differences in party members' policy preferences (or lack of such preferences). We will maintain the assumption that politicians solely seek to win their own seats and show that they most want a party platform equal to their own districts' median. This may appear as if politicians have their own preferences, but in fact, they merely want to do what gets them elected.

A third strand of the literature focuses on bargaining by parties after elections are held, in multi-party situations where no party wins an outright majority of the seats. In these situations, the politicians and voters understand that any platforms announced by individual parties prior to the election are unlikely to be enacted in their stated form—instead, these platforms constitute initial positions from which the parties will then negotiate when forming a government. So, while these models do not directly

² [Austen-Smith \(1984\)](#) models two parties that compete for control of an n -member legislature by running candidates in each of n single-member districts. Each candidate chooses an individual platform, and these platforms are then aggregated via a "party constitution" into a single position. This is the position that all members of the party will support once elected. Voters know the party constitutions and vote accordingly. Austen-Smith shows that the parties' positions converge in equilibrium, but individual candidates' platforms do not. He also shows that (pure-strategy) equilibria may fail to exist even if the policy space is unidimensional.

³ [Calvert and Isaac \(1981\)](#) consider a model in which candidates announce their own platforms, but where candidates from the president's party must contend with an exogenous party reputation given by the president's record. They analyze the promises made within each district and show that candidates will overcompensate. This model does not endogenize the overall party label.

conflicts inside *parties*, they do model conflicts inside *governing coalitions*. Early work in this vein includes Austen-Smith and Banks Jeffrey (1988) and Baron and Diermeier (2001), and more recent work includes Baron et al. (2011).

3 Basic model

There are two parties, X and Y , that compete for control of government. The parties choose policy platforms from a policy space $Z = \mathbb{R}$.⁴ Lower case z denotes a generic policy position, and x and y denote the platforms of the parties.

Every candidate for office belongs to one and only one party. All politicians in a party run under that party's label and cannot distinguish themselves from that label. Each politician seeks to win his or her own seat in the legislature. The utility of politician j in party X is $Q_X^j = Pr(j \text{ Wins})$, and the utility of politician j in party Y is $Q_Y^j = Pr(j \text{ Wins})$.⁵

Voters choose between the parties on the basis of two factors. First, voters compare parties' policy platforms. We will make the standard assumption that voters have single-peaked policy preferences, but allow a general functional form. Each voter has an ideal policy platform, denoted as w . Write the spatial component of the utility function, using the lower case u , as $-u(z - w)$. Thus, $-u(d)$ is maximized at 0, is increasing for $d < 0$ and decreasing for $d > 0$. Assume u is strictly convex, so $-u$ is strictly concave. Also, assume u is continuous. Second, a valence term captures short-term stochastic shocks, such as economic fluctuations or scandals. These shocks arise exogenously; neither party can influence their occurrence or values. The valence shocks are realized after parties announce their policies. Each party may have separate valence shocks, written v_X and v_Y . Let $v = v_x - v_y$ designate the net valence advantage of party X . The random variable v may take on any value in \mathbb{R} and has cumulative density $G(v)$. For now, we make no additional assumptions about G (e.g. it can be discontinuous); we will add assumptions where necessary below.

The overall utility that a voter receives, denoted with an upper case U , is the sum of their valence and spatial preferences. The utility a voter receives if party X holds office is $U(x, w, v_x) = v_x - u(x - w)$, and the utility a voter receives if party Y holds office is $U(y, w, v_y) = v_y - u(y - w)$. The differential in utility between party X and Y is denoted $\Delta U(w; x, y, v) = v - u(x - w) + u(y - w)$. A voter chooses party X if ΔU is positive and chooses party Y if ΔU is negative.

The policy preference of the indifferent or pivotal voter is of particular interest. Denote the "cut-point", c , as the policy such that all voters with $w < c$ vote for the party on the left along the line Z and all voters with $w > c$ vote for the party on the right. The cut-point depends on the values of x , y , and v , and can therefore be written as a function $c(x, y, v)$. The cut-point can be thought of as the point along Z such that any voter with an ideal point at that point is indifferent between the parties. Given

⁴ We could also make Z a closed interval in \mathbb{R} .

⁵ It is possible generalize to more complex objectives, such as seeking to be in the majority. We leave this to future work.

policy positions, the cut-point depends on the random variable v , and the expected value of the cut-point can be written $E[c]$.

Stronger assumptions about the utility function allow for explicit solutions for the cut-point c . With quadratic utilities, which are commonly assumed in the literature,

$$c = \frac{x + y}{2} + \frac{v}{2(y - x)}.$$

In examples below, we will use this characterization of the cut-point. Inspection of this result reveals that the cut-point need not lie between the two parties. For example, suppose that $x = 0$ and $y > 0$. If $v > y^2$, then $c > y$.

Three propositions that characterize electoral competition generally follow from the more generic assumptions. These propositions characterize properties of the cut-point that are not dependent on the electoral system. The first proposition guarantees that the cut-point is well defined, or only one party holds seats.

Proposition 3.1 *For fixed x , y , and v , if $x < y$ then $\Delta U(w; x, y, v)$ is strictly decreasing in w . If $x > y$ then $\Delta U(w; x, y, v)$ is strictly increasing in w .*

The second proposition establishes the monotonicity of the cut-point in terms of the party positions for what one might consider normal politics. If the valence shock is not too large, then if both parties shift in a certain direction, the cut-point between them will shift in that direction. Formally, this is stated as follows.

Proposition 3.2 *Let $x < y$, $x' < y'$, $x' > x$, and $y' > y$, then there exists $\epsilon > 0$ such that if $|v| < \epsilon$ then $c(x', y', v) > c(x, y, v)$ or all voters vote for party X .*

Large valence shocks result from dramatic changes in the economy, foreign policy catastrophes, and large scandals, such as Watergate, that might affect an entire party. Normal times correspond to situations when the valence shocks are not too large. Of course, normal times are relative to the particular electoral alignment. If the parties are very close to each other, then a large valence shock might eliminate most of the disadvantaged party's legislative seats. This is why if v is large, increasing the platforms of both parties (but decreasing the distance between them, or more generally if there is enough asymmetry in u) can actually decrease the cut-point. For example, consider the case of quadratic utility. Let $x = 0$, $y = 100$, $x' = 99$, $y' = 101$, $v = -400$. Then, by the formula given above, $c(x, y, v) = 50 - 2 = 48$ while $c'(x', y', v) = 100 - 100 = 0$. This example shows how diverging party platforms can help insulate parties from disastrous scandals.

Third, when neither party has an advantage a priori, the expected cut-point is simply the average of the two party's positions. We say the function $G - 1/2$ is *odd* if $G(v) = 1 - G(-v)$ for all v .⁶ Note that if G is differentiable with density function $g = G'$, then $G - 1/2$ is odd if and only if g is symmetric about 0.

Proposition 3.3 *If u is symmetric and $G - 1/2$ is odd, then $E[c] = \frac{x+y}{2}$.*

⁶ Thus, $\phi(v) \equiv G(v) - 1/2 = 1/2 - G(-v) = -\phi(-v)$ for all v , so ϕ satisfies the usual definition of an odd function.

Proposition 3.3 offers an important reference case. Most models of spatial voting assume symmetric spatial preference functions, especially the quadratic, and assume that the mean of the valence shock is zero. When preference functions are symmetric, the parties can anticipate that the ideal point of the voter who is indifferent between the parties lies half way between the two parties' announced policy positions.

The obverse of the proposition is also of interest. A party gains electoral advantages from asymmetries in the distribution of valence shocks or in the spatial utility function. For example, if the mean of the valence shock is non-zero and both parties are converged to the same point, then voters prefer the party with the advantage in the valence shock.

The above definitions and results refer to properties of the model without respect to time. In what follows, we will consider the strategies of the parties across two election cycles. Each cycle is, in turn, divided into two parts. In the first part of an election cycle, the platform is determined. In the second part, the valence shock is realized and therefore the cut-point is determined. The two time periods are indexed with the subscript $t = [0, 1]$. They differ slightly. In the first period, the parties' policy positions x_0 and y_0 are exogenously determined. Then, the valence shock v_0 is realized, which determines the cut-point $c_0 = c(x_0, y_0, v_0)$ and a division of seats between parties X and Y. The division of seats equals the fraction of voters or districts below the cut-point, c_0 . In the second time period, the elected officials within each of the two parties choose their parties' policy platforms x_1^* and y_1^* via majority rule. The asterisks indicate that these platforms represent the endogenous, equilibrium outcome of intra-party decision making. After the parties choose their platforms, the valence shock v_1 is realized, which yields a cut-point $c_1 = c(x_1^*, y_1^*, v_1)$ as an election result. Wherever possible, we will drop the subscript t to avoid clutter.

4 Platform choice in the single-member district system

In districted systems, candidates run and individuals vote within particular districts; votes are not transferable across districts; the candidate who wins the most votes in a district wins the seat; and the total number of seats won is the total number of districts won.

Let the median voter of a district be designated by m , and let $H(z)$ be the CDF of district medians. Assume H is continuous. The cut-point for a given election and the median voter of a given district jointly determine which party wins that district for that election. (Since H is continuous, we can ignore district medians exactly on the cut-point since they constitute measure zero) For example, if $x < y$ then all districts such that $m < c$ choose the candidate of party X and all districts such that $m > c$ choose the candidate of party Y. Given voters' assumed strategies, candidates' expected utility functions are easily described. A candidate, j , from party X has expected utility

$$Q_X^j(x; y, m) = \text{Prob}(\Delta U(m; x, y, v) \geq 0)$$

Similarly, the expected utility of a candidate, j , from party Y is

$$Q_Y^j(y; x, m) = \text{Prob}(\Delta U(m; x, y, v) \leq 0)$$

Let $\tilde{x}(y, m)$ be the most-preferred platform of party X 's candidate in a district with median at m , and let $\tilde{y}(x, m)$ be the most-preferred platform of party Y 's candidate. As the following proposition shows, $\tilde{x}(y, m) = m$ and $\tilde{y}(x, m) = m$.

Proposition 4.1 *In equilibrium, if v has full support, the ordinal preferences on party X 's platform of an incumbent in party X who represents a district with median voter m are described by $u(x - m)$; i.e., the preferences of a voter with median point at m for policy x . Similarly, the ordinal preferences on party Y 's platform by an incumbent in party Y are described by $u(y - m)$. In general, even if v does not have full support, platform m maximizes the incumbent's utility.*

Proposition 4.1 immediately implies that it is natural to describe incumbents as having an "induced ideal point" equal to his or her district median. Uncertainty, which enters through the term v , produces the well-defined preference function described in Proposition 3.1. Legislators will have strong orders over the platform choices because their probability of winning declines smoothly as the cut-point moves away from the district's median voter's ideal point. The amount of uncertainty affects the shape of the preference function of each legislator, but even a little uncertainty will generate a well-defined preference function in which legislators have strong preference orders over alternative policy platforms.⁷ A legislator's ordinal preferences, then, will be the same as the median voter of his or her district, but the legislator and voter would not have the same cardinal utility—that will depend on the probability function, G .⁸

This result is distinct from two other lines of thinking. First, uncertainty guarantees a well-ordered preference function and prevents multiple equilibria. In a deterministic model (Snyder 1994), the induced utility function of incumbents may have "flat spots" or "regions of indifference." The intuition is that, for a given value of y , if valence "jumps" at some points then there are regions of x that induce the same probability of winning for a candidate with given m . Full support of v guarantees that incumbent preferences are well behaved. The equilibrium derived below would still exist, but there may be others.

Second, Proposition 4.1 differs qualitatively from the assumption that politicians have their own personal preferences over policy, as in Wittman (1977, 1983), Calvert (1985), and Roemer (2002). In those models, the politicians in a party have common policy preferences and want policy to move in that direction. Uncertainty may lead to divergence, but the degree of divergence is a function of the variance of the distribution of the shocks.⁹

⁷ It is worth noting that candidate *cardinal* utility is not the same as a representative voter's cardinal utility. Candidate utility is filtered through the distribution of potential valence shocks. Only the ordinal preferences are preserved. While this distinction does not matter for purposes of this paper, it may be important when thinking about extensions (such as platforms maximizing average incumbent utility) or welfare analysis of the results. The loss of cardinal information also implies that thinking of our legislature as a "citizen" legislature made of representative citizens with policy preferences is not quite correct.

⁸ It is also worth noting that the induced ideal points of candidates are *independent* of the other party's platform. This does not mean, of course, that a candidate is *indifferent* about the other party's platform. In fact, all candidates in a party prefer the other party's platform to be as extreme as possible, as that increases the probability of winning.

⁹ In this model, the irrelevance of degree of uncertainty relies on the assumption that candidates do not have direct policy preferences. Suppose one extended the model to give incumbents a direct preference for

Proposition 4.1 ensures that incumbent preferences are well defined, single-peaked, with ideal points equal to the district medians the incumbents represent. The assumption of majority rule within parties yields an explicit characterization of the equilibrium platforms of the parties and leads to the prediction that when parties are not unified teams they will diverge.

We can now make use of the two period structure introduced at the end of section 3. There are two time periods 0, 1 such that x_0 and y_0 are determined exogenously and x_1^* and y_1^* are determined via majority rule by the incumbents from period 0.¹⁰ Assume that $0 < H(c_0) < 1$ so neither party was completely wiped out in the period 0 election.

Assume the inverse of $H(z)$, $H^{-1}(\theta)$ exists for $\theta \in (0, 1)$. By the Median Voter Theorem, each party’s platform in period 1 (x_1^* , or y_1^*) will be the median of the induced ideal points of candidates of that party who won in period 0.

Proposition 4.2 *Suppose $x_0 < y_0$, $0 < H(c_0) < 1$, and v has full support. Then $x_1^*(c_0) = H^{-1}(H(c_0)/2)$ and $y_1^*(c_0) = H^{-1}((H(c_0) + 1)/2)$ are the unique core outcomes of majority-rule bargaining among incumbents.*

A symmetric result holds if $x_0 > y_0$.

Proposition 4.2 has three immediate implications for the platforms chosen by the parties.

- $x_1^* \neq y_1^*$: Party platforms are diverged.
- If $x_0^* < y_0^*$, then $x_1^* < y_1^*$: Party order is preserved over time.
- $x_1^* \leq H^{-1}(1/2) \leq y_1^*$: Parties take positions on opposite sides of the median district.

With the introduction of a stability concept, one can derive a very simple characterization of the party platforms: one party will locate at the 25th percentile of the distribution of voter ideal points and the other at the 75th percentile.

Consider the case when cut-point in period 0 is $H^{-1}(1/2)$. That is, districts left of the median voted for one party while the other districts voted for the other party; each party won 50% of the vote. Then, the party platforms for period 1 are $x_1^* = H^{-1}(1/4)$ and $y_1^* = H^{-1}(3/4)$. The left party platform is the first quartile of districts, and the right party platform is the third quartile. If $H - 1/2$ is an odd function, then there is a sense in which this outcome is stable.

Define platforms x', y' as “zero-valence stable” if whenever $v_0 = v_1 = 0$, $x_0 = x'$, and $y_0 = y'$ then $x_1^* = x_0$ and $y_1^* = y_0$.¹¹

Proposition 4.3 (“Quartile Voter Theorem”) *Suppose u is symmetric and $H - 1/2$ is odd. Then $x = H^{-1}(1/4)$ and $y = H^{-1}(3/4)$ are zero-valence stable.*

Footnote 9 continued

a particular platform in addition to being reelected. The less uncertainty, the more a candidate could afford to support a policy that deviates from the median voter of her district. Whether this would lead to more or less divergence depends on whether incumbents tend to be more moderate or extreme than their districts.

¹⁰ We will assume throughout that $x_0 \neq y_0$. If $x_0 = y_0$, the party that receives the favorable shock will win all of the seats. It is unclear, then, what a party is if it has no seats.

¹¹ Whether platforms are in any sense dynamically stable over time for a particular stochastic process, defining v is an interesting question beyond the scope of this paper.

This result provides an interesting analogy to the Median Voter Theorem. When parties are not unified teams but their members must run under a common label, and when the distribution of district median ideal points is symmetric, the cut-point between the parties will divide the electorate equally at 1/2, but the party platforms will not converge. Rather, one will locate at the ideal point of the 25th percentile voter, and the other will locate at the ideal point of the 75th percentile voter, which correspond to the medians within each of the parties.

5 Comparative statics and empirical implications

The model carries predictions about the equilibrium platforms or policies of the parties, the shares of the votes, and the effects of changes in the exogenous features. We will focus on two such features, the valence shock v , and the distribution of district medians H .

5.1 Effects of short-term forces

Party platform choices depend on the random valence term v , which reflects economic times, scandals, and other factors that indicate the ability of the party to produce commonly shared benefits to the voters. Realizations of v_0 affect the cut-point between the parties at time 0, c_0 . One can use changes in c_0 , then, to study the effects of short-run national forces on the positions of the parties and their electoral fortunes.

An interesting and subtle implication of the model is that short-term forces and party positioning interact in their effects on aggregate vote shares or seat shares. To see this clearly, consider the case of quadratic utilities. As noted earlier, the cut-point will be of the form $c = \frac{x+y}{2} + \frac{v}{2(y-x)}$. The first term is just the mid-point between the parties, but the second term is the valence shock weighted by the distance between the parties. Consideration of this last term indicates that the more converged the parties are the more a valence shock will affect changes in the cut-point and, thus, party platform strategies in later elections. However, if parties take highly divergent positions, short-term forces will have muted effects on the cut-point and thus, on fluctuations in the division of the vote.

To our knowledge, this point has not been appreciated in the voluminous empirical literature on economic voting. It suggests that the effect of the economy on aggregate vote shares of parties is magnified by the ideological distance between the parties. Depending on party platform choice, the economy may matter a lot or little.

One can also analyze the formula for c to study how valence directly affects positioning. Four comparative-static type results deserve emphasis.

Proposition 5.1 *Restrict attention to values of c_0 such that $0 < H(c_0) < 1$, $0 < H^{-1}(H(c_0)/2) < 1$, and $0 < H^{-1}((H(c_0) + 1)/2) < 1$. Suppose $x_0 < y_0$, H is strictly increasing for all values z such that $0 < H(z) < 1$, and v has full support, then:*

- x_1^* and y_1^* are strictly increasing in c_0 . (Past good performance causes a party to moderate; past bad performance causes a party to move to the extreme.)

- If $c'_0 > c_0$, then $\exists \epsilon$ such that if $|v_1| < \epsilon$ then $c_1^*(c'_0, v_1) > c_1^*(c_0, v_1)$. (For small valence shocks, past good performance causes good current performance.)
- If u is symmetric and $G - 1/2$ is odd, then $E(c_1^*)$ is increasing in c_0 . In particular, $E(c_1^*) = (1/2)[H^{-1}(H(c_0)/2) + H^{-1}((H(c_0) + 1)/2)]$. (Past good performance causes good future performance, on average.)
- Suppose that H is continuously differentiable, $v_1 = 0$ and u is symmetric, and denotes the first derivative of H as h . The derivative of c_1 with respect to c_0 is $c'_1 = (1/4)h(c_0)[\frac{1}{h(H^{-1}(H(c_0)/2))} + \frac{1}{h(H^{-1}(H(c_0)/2+1/2))}]$.

The first part of Proposition 5.1 indicates that party platform choices depend on past party performance. If times are good for one party, say the party to the right, both parties will move away from that party’s direction, to the left, and the party to the right can expect to gain seats in the next election. Parties that do well tend to moderate, while parties that do poorly tend to move to the extreme. As a party shrinks, the few remaining incumbents tend to come from more extreme districts. In their pursuit of their own individual reelection, these incumbents pull their party further to the extreme.

The second and third parts of Proposition 5.1 describe a force that can create longevity for the majority party. As noted in Proposition 3.3, symmetry in u and G means that one can expect an even split in the shares of seats won. But, as the second and third results suggest, the actual division of seats depends on the values of the shocks. If a party enjoyed a positive shock in period 0 and the party platforms were at the first and third quartile, then the advantaged party will win a majority. If v_1 is not sufficiently negative, then that party will win a majority again in period 1.

While a party gains an advantage from past electoral success, the fourth part of Proposition 5.1 shows that under many circumstances, the advantage is less than the original surge. For example, when the distribution of ideal points is uniform, the derivative of c_1 with respect to c_0 equals $1/2$. Hence, the period 1 cut-point, if $v_1 = 0$, corrects half of the gain from the period 0 shock. While the formula above is not always less than 1 for all distributions, it is for several commonly assumed distributions, such as the uniform, normal, and logistic.

This result is consistent with the observed mid-term seat loss observed in the United States, France, and other nations. Suppose during a typical election $v \approx 0$. However, if a party wins the presidency in a given year, that is evidence that the valence shock was in favor of that party for that year. One should, therefore, expect a coattail effect of more members of that party winning legislative seats that year. However, in the upcoming election, the valence shock will tend back toward its typical value of zero, and some of the legislative gains will be lost. This will be observed as a mid-term seat loss—although, according to the model, the loss will typically be less than the coattail effect.

While this paper does not model long-term strategic consequences of repeated play, the above process suggests that a party could be wiped out by repeated negative shocks. There are several possible reactions to this prediction. First, in many specific cases, these shocks would have to be very large. For example, if H is uniform, a party could repeatedly sustain negative valence shocks equal to one quarter of the seats indefinitely without being wiped out. Second, as a party grows smaller, the assumption that platforms are determined by incumbents and only incumbents become more strained. This assumption was made to clarify and simplify the argument of the paper, but could

of course be relaxed. An obvious extension to the model is to give some weight to challengers in the party and have that weight increase as the number of challengers relative to incumbents increases. This will cause a badly beaten party to eventually begin to moderate and thus, bounce back. Finally, one could keep the model as is and embrace the prediction. Two-party democracies are not necessarily perfectly stable for all times. Just as in the long-run we are all dead, perhaps two-party systems, with sufficiently pernicious H and G distributions, are doomed to eventually experience the slow collapse of one of the parties. In particular, a sufficiently perverse gerrymander could lead to the complete and nearly inevitable destruction of one of the parties.

These results should be contrasted with the Downsian model. In that framework, because the platforms are the same, shocks cause one party to win all seats, and shocks will change all seats or none.

The dependence of current party performance on past party performance in the model has wide-reaching implications for many empirical studies of elections. At the very least, one should expect autocorrelation over time of party electoral performance. Studies of elections over multiple years (such as measures of swing-ratios) that ignore such auto-correlation may understate standard errors. Studies that use lagged election results as a proxy for omitted independent variables should observe a positive correlation between current and past elections. However, such a correlation does not imply that the lagged election results are actually correlated with the omitted variable.

The cross-time dependence of cut-points predicted by the model suggests at least two direct empirical tests. At the very least one should look for and measure cut-points, and check to see whether they follow an auto-regressive process. If one is willing to take the model very seriously, one could also test the specific formula $E(c_1^*) = (1/2)[(H^{-1}(H(c_0)/2) + H^{-1}((H(c_0) + 1)/2))]$.

Of course, to perform the regressions suggested earlier, one would need estimates of c and H . Here is one possible approach. Label the parties as X and Y and assume $x < y$. Let \hat{m}_i be the estimated ideology of district i .¹² A natural estimator for H would be the empirical distribution of districts $\widehat{H}(z)$; the fraction of districts with ideology less than z . To estimate c pick a non-decreasing loss function $L(z)$. Then, $\hat{c} = \operatorname{argmin}_z \sum_i (L(|z - \hat{m}_i|) \delta_i(z))$ where $\delta_i(z) = 1$ if $\hat{m}_i < z$ and district i voted for party Y or $\hat{m}_i > z$ and district i voted for party X . Otherwise $\delta_i(z) = 0$. In other words, district i contributes loss if it votes for the “wrong” party. In this paper’s stylized model, districts never vote for the wrong party, and so such a correction would be unnecessary.

5.2 Implications for theories of gerrymandering and electoral system bias

One of the functions of gerrymandering is to alter legislative district lines so as to benefit one party over another. Most of the extensive literature on gerrymandering focuses on a particular definition of this concept—electoral system bias. A system is unbiased if, in a hypothetical election in which the parties split the vote evenly, the

¹² Alternatively, one could use a “corrected” estimate for m that accounts for district-specific factors other than ideology, such as incumbency.

parties win equal shares of the seats. Deviations in seat shares from 50–50, when the votes are split 50–50, are taken as the degree of bias. This is a characteristic of the function H . Empirical research implements this definition by regressing seat shares on vote shares and then examining the predicted seat shares of a party when its vote share equals .5. This idea is used to compare electoral systems broadly.

The model suggests an alternative way to think about the overall degree of electoral bias produced by gerrymandering. As noted earlier, most analysts think of partisan gerrymandering only in terms of the direct mapping from votes to seats, or more precisely what is the value of $H(c)$ when the parties divide the electorate evenly.

Spatial models that predict convergence are uninformative about this question. In Downs's model, where parties are teams, gerrymandering has no affect on electoral outcomes.¹³ The parties converge to the national median and elections in all districts end in ties, regardless of H . The same is true in models where there are no parties, and candidates are free to choose whatever position they want within their districts.¹⁴

In this model, gerrymandering can have two affects. First, it alters the *seat shares* received by each party directly by changing the shape of H . Second, altering the shape of H can indirectly affect the share of seats won by a party because it can lead parties to adopt new *platforms*. The full effect of the gerrymander, then, would be the change in seats of the party that resulted from the change in H and the change in platforms.

The first effect is the effect most political scientists look for. Does the shape of H created by a districting plan imply a disadvantage for one of the parties, holding the parties fixed at their positions.

The second effect, however, has been missed in the voluminous formal literature on districting and gerrymandering. A gerrymander can force the opposing party to adopt a platform that makes it a permanent minority. Consider the possible effect of racial districting. Racial gerrymandering may change the makeup of Democratic incumbents by creating a disproportionate number of extremely left-wing districts in that party. That, in turn, might change the platform of the Democratic party as well, most likely shifting it to the left. Thus, the party balance can depend on the shape of how districts are distributed, in addition to the median (or average) district.

To illustrate this more precisely, consider the following examples.

Example 5.1 Let voter utility be symmetric and let $H(z) = z$ where $z \in [0, 1]$.

The uniform distribution implies that the zero-valence stable cut-point is $1/2$, the parties will locate at $1/4$ and $3/4$, and the parties will each win half of the seats.

Example 5.2 Let voter utility be quadratic, and let $H(z) = z^2$ where $z \in [0, 1]$.

As shown in the appendix, the cut-point corresponding to the unique zero-variance stable platforms (given $x < y$) is approximately .653. The left party locates at .462;

¹³ Non-spatial models of voter behavior, such as voters with fixed partisan loyalty, can predict electoral outcome changes from gerrymanders. For an example, see [Shotts \(2001\)](#).

¹⁴ Note in a Downsian model, Gerrymanders can still effect policy outcomes by moving the median legislator. [Shotts \(2002\)](#) provides a formal model describing gerrymandering given purely ideological legislators. [Epstein and O'Halloran \(1999\)](#) provide an empirical case of the median legislator being moved to accommodate racial minority-majority districts. In their example, the median legislator becomes, perversely, less friendly to minority rights.

the right party is at .845; and the right party wins approximately 57% of the seats (i.e., $H(c) \approx 0.427$).¹⁵

Assume, for the sake of simplicity, that the underlying distribution of voters' ideal points is the same in both of these circumstances but that a clever political cartographer managed to reshape district lines so that the initial uniform distribution of district medians became quadratic. Doing so has several effects.

First, it changes the quantiles of the distribution of district medians. In Example 4.1., the median district is at .707, instead of .5, the 25th percentile district is at .5, and the 75th percentile is at .866.

Second, relative to these quantiles, the left party is more extreme and the right party is more moderate. The left party is .207 units below the median district, while the right party is .159 units above the median.

Third, the right party gains an electoral advantage. The cut-point between the two parties lies to the left of the median. The left party becomes something of a permanent minority, as its zero-order stable seat share fell from 50 percent to 43 percent. This, then, is the effect of redistricting on the electoral system.

This analysis also highlights some difficulties with the traditional seats/votes measure of bias. Interpreting the relationship between seats/votes requires mapping the distribution of voters as well as districts. There is the further problem with interpreting the hypothetical case when the vote is evenly divided. Suppose the parties have taken diverging and asymmetric positions around the median. Then, only a large and asymmetric shock would produce the 50–50 division of the vote, but such a shock would necessarily lead the parties to alter their policy platforms (in equilibrium). Hence, the hypothetical 50–50 division of the electorate would be out of sample in two respects: not observed in the data and very unlikely to be observed as it is not in equilibrium.

An example demonstrates a situation where this is clearly the case. Assume that voters' ideal points are uniform, as in Example 4.0, and unchanged by redistricting. Call the distribution of voters' preferences $F(z)$. However, redistricting alters the distribution of district medians from uniform on Z to quadratic, $H(z) = z^2$. In the preredistricting state, the parties locate at the quartiles and each party wins half of the vote and half of the seats. Under the gerrymander, as already shown, the cut-point between the parties is .653, and the left party will receive 43 percent of the seats. This would be an extreme gerrymander: Party X wins 65 percent of the vote and 43 percent of the seats.

Now consider the traditional definition of party bias. Holding fixed the parties' positions under Example 4.1, what is the division of the seats occurs if the division of the votes equals .5? In the terms of the model, there must be a valence shock that makes the vote division equal, holding the parties at platforms $x = .43$ and $y = .84$. This would require a large negative valence shock. Again assuming uniform voter

¹⁵ Example 5.2 is by no means unique. Another good example is a piecewise uniform district density with support between 0 and .75, median at 0.5, and density of 1 in $[0, 0.5)$ and 2 in $(0.5, 0.75]$. In other words, Example 5.1 with the right most districts "compressed." Analogous examples can be constructed where the left most districts have been "expanded." In this example, it is easy to show that the unique zero-valence stable platforms for quadratic utility are 0.2 and 0.6, the corresponding cut-point is 0.4, and $H(0.4) = 0.4$. That is, the left party with its long tail of districts becomes relatively more extreme, the right party relatively more moderate, and the left party receives only 40% of the vote.

ideal points, a cut-point of .5 would produce vote shares of $F(.5) = .5$, and party X 's seat share would fall to just .25. This would mean an enormous "bias" of .25. The zero-order stable definition leads to a much smaller bias of .07.

One could argue that the measure offered above—the expected or zero-order seat loss associated with the change in H —is preferable to the traditional definition of bias. The enormous bias predicted by the traditional measure is not usually in empirical analyses an in-sample value. But, as our discussion highlights, there is a conceptual problem as well: Traditional bias does not account for the fact that the parties have changed their positions to accommodate the new districting map.

As discussed earlier, the model also provides a specific formula for the effects of gerrymandering: $E(c_1^*) = (1/2)[H^{-1}(H(c_0)/2) + H^{-1}((H(c_0) + 1)/2)]$. In this case, H is what is being varied. One could either use the formula as a way to predict the consequences of a proposed future gerrymander or test the model by testing the formula against historical gerrymanders. Note that in any given election with a given cut-point in the prior election, only certain parts of H impact the current cut-point. For example, if a gerrymander only tinkered with districts in the extreme left tail while leaving the rest of the distribution the same, the model would only predict an impact when the previous election went strongly against the left party. A gerrymander that has a small impact one year might have a major impact the next. Or, districting plans that only operate in the tail of the distribution of one party's districts and change the party median little, may have no substantial effect on electoral competition.

The gerrymander suggested by these examples is severe in another sense. It substantially altered the policy outcomes relative to the preference of the median voter. Assuming, for the sake of argument, that the median is at $z = .5$. In Example 5.1, policy will be either $z = .25$ or $.75$, depending on which party wins. In Example 5.2, the right party will win and policy will be at $z = .845$. This is very extreme relative to the median voter's preferences, and all voters but those to the right of the $z = .8 (= (.845 + .75)/2)$ will be worse off.

This contrast raises a deeper welfare issue. In assessing gerrymandering, one should first ascertain whose benefit is at stake. If one cares first and foremost about elites—about elected officials and parties—then the definition of bias as seat loss may make sense. However, if the ultimate goal is representational bias and we care about voter utility, then the relevant comparison is not between seats and votes or even seat loss but how far party platforms deviate from the median.¹⁶

6 Extension: platform choice in the party list system

In the list system, the parties offer a list of candidates running under their label, and the entire national electorate votes for one of the two parties. Parties win shares of seats equal to their shares of the vote. The number of seats won by the party equals the number of seats times the share of seats it deserves. An individual candidate within

¹⁶ For an example of thinking about bias in terms of voter welfare, see Coate and Knight (2005). They show that in some circumstances, a biased seats/votes curve can actually increase voter welfare. However, in their model, party platforms are fixed. The model indicates that to fully assess the welfare implications of bias, one must also account for how the shape of districts impacts party platforms.

a party wins a seat if that candidate's rank on the list is higher than total number of seats the party won in the election.

Party list systems are often viewed as having a very different sort of politics than districted systems because the candidates face very different electorates and are chosen through a very different sort of mechanics. Surprisingly, an analogous logic and results characterize politicians' induced preferences and party positioning in party list and districted systems.

Slightly different notation is required to describe the electoral competition in list systems. Parties win seats in proportion to their votes in the national electorate, rather than the fraction of districts won. As before, the cut-point c defines the ideal point along Z of the voter indifferent between the two parties, X and Y . The cumulative density of the voters' ideal points evaluated at the cut-point c determines the fraction of seats won by the left party X , and the right party, Y , wins the remaining seats. Denote the cumulative density of voters ideal points as $F(z)$, and assume $F(z)$ is continuous and has inverse for $\tau \in (0, 1)F^{-1}(\tau)$.¹⁷ The left party wins a share s of votes (and of seats) equal to $s \equiv F(c)$, and the right party wins $1 - F(c)$ share of votes.

In the party list case, $F(z)$ plays the same role as $H(z)$ in the districted case. Nevertheless, it is important to remember the distinction. $F(z)$ is a description of voter ideal points across an entire nation; $H(z)$ is a description of median voter ideal points, district-by-district.

We will represent the positions on each party's list as the interval $[0, 1]$, where 0 denotes the top of the list and 1 denotes the bottom. A candidate at the λ th position on the list wins a legislative seat if and only if the candidate's party receives a vote (and seat) share larger than the candidate's position on the list, i.e., $\lambda < s$. The λ th candidate on party X 's list has the following utility function

$$Q_X = Prob(s \geq \lambda)$$

The λ th candidate on party Y 's list has the following utility function

$$Q_Y = Prob(1 - s \geq \lambda)$$

The above equations define candidates' preferences over party platforms. As in the single-member district model, candidates' expected utilities can be written in terms of voters' utility functions. For $x < y$, $\lambda \in (0, 1)$:

$$Q_X(x, y, \lambda) = Prob(\Delta U(x, y, v, F^{-1}(\lambda)) \geq 0)$$

and

$$Q_Y(x, y, \lambda) = Prob(\Delta U(x, y, v, F^{-1}(1 - \lambda)) \leq 0).$$

Symmetric equations hold for $x > y$.

¹⁷ Unless the policy space Z is bounded, $F^{-1}(\tau)$ may be undefined at $\tau = 0$ and $\tau = 1$. Since these represent measure zero worth of candidates on the party list, one can safely ignore these cases.

Let $\tilde{x}(y, \lambda)$ be the most-preferred platform of a candidate who is at position λ on party X 's list, and let $\tilde{y}(x, \lambda')$ be the most-preferred platform for a candidate at position λ' on party Y 's list. As the following proposition shows, $\tilde{x}(y, \lambda)$ and $\tilde{y}(x, \lambda')$ are well defined. We require G to be strictly increasing for the results in this section (v has full support).

Proposition 6.1 *For all $\lambda \in (0, 1)$ and y , $Q_X(x, y, \lambda)$ is either double-peaked or single-peaked in x . Furthermore, if u is symmetric around 0, and $F - 1/2$ is an odd function around a given point z with $y > z$, then there is a unique global maximum attained at $\tilde{x}(y, \lambda) = \max(F^{-1}(s), y)$.*

A symmetric result holds for the case $y \leq 0$, and an analogous proposition holds for $Q_Y(x, y, \lambda)$. The preferences of the person exactly at the top and bottom of the list are somewhat pathological, but they constitute a set of measure zero and so are ignored.¹⁸

This result has an interesting implication for understanding politicians' preferences under list systems. There is a seeming divide among political scientists studying districted and list systems. Those studying districted systems, such as the United States, frequently assert that politicians just want to win office, while those studying list systems often maintain policy-oriented politicians. Proposition 5.1 suggests that the difference is more apparent than real. Election-oriented politicians' positions on the list induce policy-oriented preference functions.

Candidates near the top of their party's list have the strongest preference for adopting extreme policies. The intuition is straightforward and reflects the trade-off between a party's expected vote share and the variance of the party's vote share. Consider a candidate at the 25th percentile on the party X 's list ($\lambda = 1/4$). This candidate wins a seat as long as X wins at least 25% of the vote. If X chooses a platform that diverges considerably from Y 's platform, then there will be a subset of voters with a strong policy-based preference for party X . If $x < y$, for example, then voters with $z < (x + y)/2$ all prefer X to Y on policy grounds, and this preference grows more intense the smaller is z . Voters with a strong policy-based preference for party X are very likely to vote for X , since it will require a large, negative valence shock to cause them to change their minds and vote for Y . Thus, even if party X is unlikely to win a majority of the vote, it might be *very* likely to win 25 percent or more of the vote. On the other hand, if X 's platform is close to Y 's, then there is a higher degree of uncertainty about X 's vote share, since the valence shocks dominate voter choices. Although X expected vote share may be higher, the variance of its vote share will also be higher, and the probability that X wins at least 25% of the vote may fall. As shown in Proposition 5.1, if $y > 0$ then the probability that X wins at least 25 percent of the vote is maximized when x is equal to the 25th percentile of the voter ideal-point distribution.

By contrast, candidates low on their party's list all want the same platform, a "me too" platform equals to that of the other party. Such candidates actually *welcome*

¹⁸ Unless one requires the support of F to be bounded, the person on the top of the list wants infinite divergence; the person at the bottom wants full convergence and is indifferent among all other platform choices.

some degree of electoral uncertainty, since their only chance of winning a seat in the legislature lies in a favorable valence shock for their party.

Suppose that each party's incumbent legislators are placed at the top of the party's list. Recall that each party's platform is chosen by simple majority rule among the party's incumbents. Denote the current period as period 1, and let s_0 be party X's seat share from the election in period 0. Assume that $s_0 \in (0, 1)$, so neither party was completely wiped out in period 0's election. Although some incumbents do not have single-peaked preferences, their preferences satisfy a weaker condition that guarantees the existence of a unique majority winner. The equilibrium platforms are simply characterized, as follows.

Proposition 6.2 *Suppose u is symmetric around 0, and $F - 1/2$ is an odd function around a given point z . Suppose further that F^{-1} and u are differentiable, $s_0 \neq 0.5$, and $0 < s_0 < 1$. Then, the equilibrium core platforms for the current election satisfy (i) $x_1^* = F^{-1}(s_0/2)$ and $y_1^* = F^{-1}(s_0/2 + 1/2)$, or (ii) $x_1^* = F^{-1}(1 - s_0/2)$ and $y_1^* = F^{-1}((1 - s_0)/2)$.*

According to Proposition 6.2, each party's platform for the election at time 1 is equal to the median of the ideal points of the voters who supported the party in the election at time 0. In case (i) $x_1^* < 0 < y_1^*$, and in case (ii) $y_1^* < 0 < x_1^*$. Note that the equilibrium platforms are always on opposite sides of the overall median ideal point.¹⁹

Proposition 6.2 is nearly identical to Proposition 4.2 (as long as the cut-point is never *exactly* at the median,) and so most the conclusions that derive from Proposition 4.2 (divergence, past performance effects current platforms and performance, and performance depends on the entire distribution of voters) apply, with one major exception: There is no guarantee that party ordering will be the same. If one was to exogenously impose party ordering then the standard predictions would apply. If one does not impose the restriction exogenously, then there is nothing about the model that prevents party's in part-list systems from "swapping" left-right positions.

Although similar results on party-positioning hold for list and districted systems, the model also helps clarify the differences between them.

First, induced preferences are well defined and concave for all legislators in the districted system, but oddities arise in the list system. Most notably, those legislators on the bottom of the list want perfect convergence exactly, as that plus a favorable shock is their only hope of winning a seat. The analogous legislators in districted systems are those in moderate districts or districts leaning toward the other party. Those in moderate districts want moderate platforms, while those in the "wrong districts" (left party politicians in districts right of the median and vice versa) want their party to be more conservative than the median district.

Second, electoral competition in the list system depends only on the distribution of voters' preferences, not the intermediary mapping of voters' preferences into districts and then into seats. Propositions 3.2 and 5.2 show that the platforms adopted by the parties depend on the relevant electoral distribution functions—either of voters'

¹⁹ If the cut-point is at the median, there exists a convergent equilibrium as well as two divergent equilibria. The intuition is that the pivotal incumbent has a 50 percent chance of winning in both the convergent and divergent equilibria.

ideal points or of district median ideal points. If a particular country were to switch from a list system to a districted system or vice versa, party platforms may change significantly depending on the mapping of voters into districts. If, for example, voter ideology was driven by income, the distribution of ideology in the entire nation may have a thick right tail. A districted system, on the other hand, might be, district median by district median, more balanced. As Example 3.1 illustrates, a right-skewed distribution can result in a small right party. It is possible that this result may give some intuition for why socialist parties are relatively weak in districted systems and strong in party list systems.

7 Concluding remarks

The Median Voter Theorem has provided a focal point for the spatial theory of politics. Two parties, acting as teams to control government, will converge to the ideal point of the median voter in the electorate. But, parties are not teams—at least not in the strict sense that Downs and others have expressed it. Rather parties are collections of politicians motivated by their personal desire to hold office as well as grander ideological goals. Legislators benefit from and enjoy personal benefits from holding their own seat in the legislature. Starting from this assumption, we have arrived at a strikingly different conclusion.

When parties are not teams, parties will represent the median of their elected officials, rather than the median of the electorate as a whole. This result applies equally to single-member district systems and party list systems, which suggests that party divergence is a universal phenomenon driven by the personal interests of politicians. Politicians' desire to hold office will lead parties to diverge. The degree of divergence expected is approximately the inter-quartile range within the electorate.

Electoral systems themselves determine whether parties are teams. The relevant question is this. If all politicians seek to win office, are their induced preferences over policy identical within each party? In neither districted nor list systems was this true. Hence, in neither system should the team assumption be taken lightly. There are, however, some electoral systems that would produce parties as teams. Consider, for example, a system in which all legislators are elected at large from the nation as a whole, voters have as many votes as positions, and voters must vote for all candidates. The theory predicts parties as teams in this circumstance because all candidates will have the same *induced* policy preferences.

We hope our model provides a good starting point for the empirical and theoretical work on parties that does not assume unified times. Empirically, the model generates several predictions that agree with the basic facts of electoral competition. Most notably, elections are not ties in all districts, and the parties are diverged. The simple spatial model in which parties are teams or in which candidates compete district-by-district do not generate these predictions, but instead predict that all elections end in ties, which is clearly false.

As a theoretical matter, the model has been simplified in several respects in order to clarify the importance of the team assumption. Many analytical extensions are possible and desirable. They produce a wider range of predictions, but are cases of

the general results here. First, we have assumed only two parties, following much of the literature. Multiple parties or a system with entry is much more complicated and requires assumptions about coalition formation and voter sophistication. Second, one can expand the party beyond the incumbent members of the legislature to incorporate Presidents, non-incumbent candidates, multi-body legislatures, and activists. Some of these extensions, such as a President, will create pressures for convergence, while others create centripetal forces. Third, politicians might have objectives beyond winning their own seats, including their own policy preferences and desire for more power within the legislature. Again, some of these extensions, such as winning a majority, push the model in the direction of the team assumption, while others create more dissension within the party. Fourth, a variety of technical complications is likely of interest, including longer time horizons and multiple dimensions.

While we have focussed primarily on positive implications of the model, there are normative implications as well. For example, in the world of our model, most incumbents are significantly better off than they would be if parties were forced to maximize seat shares, or if parties did not have to run under a common banner. In those cases, in equilibrium, each incumbent would only have a 0.5 chance of winning each election. In our framework, all but the most marginal incumbents will have a greater chance of winning their own personal seat.

There is a welfare implication for voters as well. When parties are not teams, the parties diverge, and one of the two parties' policy platforms will win and become law. The winning platform will deviate considerably from the median, and thus, a majority of the voters would prefer a policy that is closer to the median. There is little voters can do to change this situation. The majority of politicians within each party support the divergent positions that the parties take: those positions maximize the politicians' chances of winning office. Voters cannot achieve the same degree of coordination. They are left with a policy that is less than optimal for a large majority.

Appendix

Proof of Proposition 3.1. We will only show the case $x < y$ here. Consider two values z and z' , where, without loss of generality, $z' > z$. Then either $z' - x > z - x \geq z' - y > z - y$ or $z' - x > z' - y > z' - y > z - y$. We need to show that

$$\Delta U(z; x, y, v) > \Delta U(z'; x, y, v),$$

that is,

$$\begin{aligned} v - u(z - x) + u(z - y) &> v - u(z' - x) + u(z' - y) \\ u(z' - y) + u(z - x) &< u(z - y) + u(z' - x). \end{aligned}$$

Define α and α' implicitly by

$$\begin{aligned} z' - y &= \alpha(z - y) + (1 - \alpha)(z' - x) \\ z - x &= \alpha'(z - y) + (1 - \alpha')(z' - x). \end{aligned}$$

Adding the two relations above and manipulating yields

$$(\alpha + \alpha' - 1)(z - y) = (\alpha + \alpha' - 1)(z' - x).$$

Since $(z' - x) > (z - y)$ this implies $\alpha + \alpha' = 1$. By strict convexity

$$\begin{aligned} u(z' - y) &< \alpha u(z - y) + (1 - \alpha)u(z' - x) \\ u(z - x) &< \alpha' u(z - y) + (1 - \alpha)u(z' - x). \end{aligned}$$

Adding the two relations above and using the result $\alpha + \alpha' = 1$ yields

$$u(z' - y) + u(z - x) < u(z - y) + u(z' - x)$$

as desired. □

To prove Proposition 3.2 we will first prove two lemmas.

Lemma 3.1a *Let $x < x' < y$. $\exists \epsilon > 0$ such that if either $c(x, y, v) > x'$ or $|v| < \epsilon$ then either $c(x', y, v) > c(x, y, v)$ or all voters vote for party X.*

Proof of Lemma 3.1a. There are two cases. Case one is $c(x, y, v) > x'$. Then, the voter with ideal point $c(x, y, v)$ strictly prefers x' to x and therefore $\Delta U(c(x, y, v); x', y, v) > 0$. By Proposition 2.1, ΔU is strictly decreasing in ideal points, and therefore, the solution z^* to $\Delta U(z^*; x', y, v) = 0$ must be greater than $c(x, y, v)$ or not exist. But if z^* exists, it is by definition the cut-point $c(x', y, v)$. If it does not exist, then $\Delta U(z; x', y, v) > 0$ for all z . That is all voters prefer party X.

Case two is $c(x, y, z) \leq x'$. If $v = 0$, a voter with ideal point x' strictly prefers party X to Y. That is $\Delta U(x'; x', y, 0) > 0$. By continuity of ΔU with respect to v , there exists $\epsilon > 0$ such that $\forall |v| < \epsilon, U(x'; x', y, v) > 0$. Then for such v by a similar argument as case one, $c(x', y, v) > x'$ or all voters support party X. But in this case, $c(x', y, z) > x' \geq c(x, y, z)$. □

Proof of Lemma 3.1a. Let $x < y < y'$. If $c(x, y, v) < y$ then either $c(x, y', v) > c(x, y, v)$ or all voters vote for party X. Furthermore, $\exists \epsilon > 0$ such that if $|v| < \epsilon$ then $c(x, y, v) < y$ and therefore either $c(x', y, v) > c(x, y, v)$ or all voters vote for party X.

Proof of Lemma 3.1b. First, consider the case where $c(x, y, v) < y$. Then, the voter with ideal point $c(x, y, v)$, given v , strictly prefers y to y' , is indifferent between x and y (by definition of cut-point) and therefore, strictly prefers x to y' . Therefore, $\Delta U(c(x, y, v); x, y', v) > 0$. By Proposition 2.1, ΔU is strictly decreasing in ideal points, and therefore, the solution z^* to $\Delta U(z^*; x, y', v) = 0$ must be greater than $c(x, y, v)$ or not exist. But if z^* exists, it is by definition the cut-point $c(x, y', v)$. If it does not exist, then $\Delta U(z; x, y', v) > 0$ for all z . That is all voters prefer party X.

For the second part, if $v = 0$ and party X's platform is at x and party Y's platform is at y , then a voter with ideal point at y strictly prefers party Y to X and a voter with ideal point x strictly prefers party X to Y. Therefore, by Proposition 2.1, $c(x, y, 0)$

exists and $x < c(x, y, 0) < y$. By continuity of $\Delta U(y; x, y, v)$ and $\Delta U(x; x, y, v)$ with respect to v , $\exists \epsilon$ such that $\forall |v| < \epsilon$ a voter with ideal point at x still strictly prefers x to $c(x, y, v)$ and a voter with ideal point at y still strictly prefers y to $c(x, y, v)$. Therefore $x < c(x, y, v) < y$. \square

Note, Lemmas 3.1a and 23.1b imply that as long as the original cut-point lies between the two platforms, and platform shifts do not cross the cut-point, then the cut-point is monotonic in platforms. Roughly speaking, we need either small v combined with potentially large platform changes, or small- to medium-sized v (such that the cut-point stays between the parties) combined with small changes in platforms to ensure monotonicity.

Proof of Proposition 3.2. Case one is $y' - y < x' - x$. Let $\delta = y' - y$. Since $\Delta U(z; x, y, v) = \Delta U(z + \delta; x + \delta, y + \delta, v)$, $c(x + \delta, y + \delta, v) = c(x, y, v) + \delta$. Therefore, $c(x + \delta, y', v) > c(x, y, v)$. For this case, $x' > x + \delta$. Then, apply Lemma 2.1a to show $\exists \epsilon$ such that $\forall v$ such that if $v > -\epsilon$ then $c(x', y', v) > c(x + \delta, y', v)$ or all voters vote for party X .

Case two is $x' - x < y' - y$. Let $\delta = x' - x$. Since $\Delta U(x, y, v, z) = \Delta U(x + \delta, y + \delta, v, z + \delta)$, $c(x + \delta, y + \delta, v) = c(x, y, v) + \delta$. Therefore, $c(x', y + \delta, v) > c(x, y, v)$. For this case, $y' > y + \delta$. Then, apply Lemma 2.1b to show $\exists \epsilon$ such that $\forall v$ such that if $v < \epsilon$ then $c(x', y', v) > c(x', y + \delta, v)$

Case three is $x' - x = y' - y$. Let $\delta = x' - x$. Since $\Delta U(x, y, v, z) = \Delta U(x + \delta, y + \delta, v, z + \delta)$, $c(x + \delta, y + \delta, v) = c(x, y, v) + \delta$. Therefore, for this case, $c(x', y', v) > c(x, y, v)$. \square

Proof of Proposition 3.3. Recall that a cut-point is defined as the solution to $-u(z - x) + u(z - y) + v = 0$. Since u is symmetric, when $v = 0$, the solution is $z = (x + y)/2$. For $v \neq 0$, write the solution as $c(x, y, v) = (x + y)/2 + \delta_+$. Then, $-u((x - y)/2 + \delta_+) + u((y - x)/2 + \delta_+) + v = 0$. But by symmetry of u , the following must hold

$$\begin{aligned} -u((y - x)/2 - \delta_+) + u((x - y)/2 - \delta_+) + v &= 0 \\ u((y - x)/2 - \delta_+) - u((x - y)/2 - \delta_+) - v &= 0 \end{aligned}$$

But this implies that the cut-point for $-v$ is $c(x, y, -v) = (x + y)/2 - \delta_+$. Therefore, $c(x, y, v) + c(x, y, -v) = x + y = 2c(x, y, 0)$. If $G - 1/2$ is odd, then the expected value of the cut-point is as follows:

$$\begin{aligned} E(c(x, y)) &= \int_{-\infty}^{\infty} c(x, y, v) dG \\ &= \int_0^{\infty} c(x, y, -v) dG(v) + \int_0^{\infty} c(x, y, v) dG(v) \\ &= 2c(x, y, 0) \int_0^{\infty} c(x, y, -v) dG(v) \end{aligned}$$

$$\begin{aligned}
 &= 2c(x, y, 0)(1/2) \\
 &= c(x, y, 0)
 \end{aligned}$$

□

Proof of Proposition 4.1. The expected utility of a candidate from party X with district median m is

$$\begin{aligned}
 Q_X(x, y, m) &= Prob(\Delta U(m; x, y, v) \geq 0) \\
 &= 1 - G(u(x-m) - u(y-m))
 \end{aligned}$$

Since in equilibrium y is treated as a constant, $u(x-m) - u(y-m)$ represents the same preferences as $u(x-m)$. G is a CDF and therefore is monotonic. Any maximum of $-u(x-m)$ is a minimum of $u(x-m)$, a minimum of $G(u(x-m))$ (by monotonicity), and a maximum for $1 - G(u(x-m))$. Since m is a maximum for $-u(x-m)$, it is also a maximum for $Q_X(x, y, m)$. If v has full support, then G is strictly increasing. Now, $u(x-m)$ represents the opposite preferences to $-u(x-m)$, $G(u(x-m))$ represents the same preferences as $u(x-m)$ by strict monotonicity, and $1 - G(u(x-m))$ represents the opposite preferences to $G(u(x-m))$. Therefore, $Q_X(x, y, m)$ represents the same preferences as $-u(x-m)$. Similar arguments apply for $Q_Y(y, x, m)$. □

Proof of Proposition 4.2. Since $0 < H(c_0) < 1$, each party has positive vote share. Since $x_0 < y_0$, every district with median less than c_0 votes for party X in period 0. Since v has full support, by Proposition 3.1, party X ideal points correspond to district medians in equilibrium. Therefore, incumbent members of party X have ideal points given by the CDF $H(z)/H(c_0)$ for $z < H(c_0)$. The median incumbent ideal point is defined by the equation $H(z_m^*)/H(c_0) = 1/2$. Solving yields $z_m^* = H^{-1}(H(c_0)/2)$. Consider any alternative platform x' . If $x' > z_m^*$, then there exists ϵ that defines a majority coalition strictly preferring z_m^* to x' . This coalition consists of all incumbents with districts from $H^{-1}(0)$ to $H^{-1}(H(c_0)/2 + \epsilon) < x$. A similar coalition exists if $x' < z_m^*$. However, there does not exist a majority coalition that prefers any $x \neq z_m^*$. Therefore, z_m^* is the unique core point. A similar argument applies to show $y_1^*(c_0) = H^{-1}((H(c_0) + 1)/2)$. Strict monotonicity of H and H^{-1} implies $x_1^* < y_1^*$. □

Proof of Proposition 4.3. The result is immediate from Proposition 4.2 and the definition of zero-valence stable. □

Proof of Proposition 5.1. • From Proposition 4.2, $x_1^*(c_0) = H^{-1}(H(c_0)/2)$ and $y_1^*(c_0) = H^{-1}((H(c_0) + 1)/2)$. H is strictly increasing and continuous. The inverse of a strictly increasing and continuous function is a strictly increasing function, therefore H^{-1} is a strictly increasing function. The composite function of two strictly increasing functions is also strictly increasing. Therefore, x_1^* and y_1^* are strictly increasing in c_0 .

• From the above result, $x_1^*(c'_0) > x_1^*(c_0)$ and $y_1^*(c'_0) > y_1^*(c_0)$. From Proposition 4.2, $x_1^*(c'_0) < y_1^*(c'_0)$ and $x_1^*(c_0) < y_1^*(c_0)$. $c_1^*(c'_0, v_1) = c^*(x_1^*(c'_0), y_1^*(c'_0), v_1)$ and $c_1^*(c_0, v_1) = c^*(x_1^*(c_0), y_1^*(c_0), v_1)$ by definition. Apply Proposition 3.2 to get the result.

• Apply Proposition 3.3 to the formula from Proposition 4.2.

Since $v_1 = 0$ and u are symmetric, c_1 is the mid-point of the two-party platforms $c_1 = (1/2)(x_1^*(c_0) + y_1^*(c_0))$. From Proposition 4.2, $x_1^*(c_0) = H^{-1}(H(c_0)/2)$ and $y_1^*(c_0) = H^{-1}((H(c_0) + 1)/2)$, therefore $c_1 = (1/2)(H^{-1}(H(c_0)/2) + H^{-1}((H(c_0) + 1)/2))$. Differentiation yields the result. \square

Proof of claims in Example 5.2. Suppose $z \in [0, 1]$ and $H(z) = z^2$. Then the PDF is $h(z) = 2z$, and the inverse is $H^{-1}(\theta) = \sqrt{\theta}$. The average is $\bar{z} = \int_0^1 zh(z)dz = \frac{2}{3}$. The median is determined by the equation $\frac{1}{2} = z_m^2$, so $z_m = \frac{\sqrt{2}}{2}$.

By Proposition 4.2, the definition of zero-variance stable platforms, and the formula for cut-point with quadratic utility,

$$c = \frac{1}{2} \left(\sqrt{\frac{c^2}{2}} \right) + \frac{1}{2} \left(\sqrt{\frac{c^2}{2} + \frac{1}{2}} \right)$$

After some algebra this simplifies to

$$c = \frac{\sqrt{2}}{2} \left(\frac{1}{\sqrt{4 - 2\sqrt{2}}} \right) < \frac{\sqrt{2}}{2} = z_m.$$

The formulas in the text follow immediately. \square

Proof of Proposition 6.1. First, suppose $x = y$. Then $s_X = 1$ if $v > 0$, $s_X = 0$ if $v < 0$, and $s_X = 1/2$ if $v = 0$. Thus, $Q_X(x, y, s) = Prob(v \geq 0) = 1 - G(0)$ for all s .

Next, consider $x < y$. Then

$$\begin{aligned} Q_X(x, y, s) &= Prob(\Delta U(F^{-1}(s); x, y, v) \geq 0) \\ &= Prob(v > u(x - F^{-1}(s)) - u(y - F^{-1}(s))) \\ &= 1 - G(u(x - F^{-1}(s)) - u(y - F^{-1}(s))) \end{aligned} \tag{A1}$$

Next, consider $x > y$. Then

$$\begin{aligned} Q_X(x, y, s) &= Prob(\Delta U(F^{-1}(1-s); x, y, v) \geq 0) \\ &= Prob(v > u(x - F^{-1}(1-s)) - u(y - F^{-1}(1-s))) \\ &= 1 - G(u(x - F^{-1}(1-s)) - u(y - F^{-1}(1-s))) \end{aligned} \tag{A2}$$

Note, $Q_X(x, y, s)$ is always of the form $1 - G(\Delta U)$. Since G is strictly monotonic, we can focus on finding the extreme values for ΔU :

If $x = y$ then $\Delta U = 0$. If $x < y$ then $\Delta U = u(x - F^{-1}(s)) - u(y - F^{-1}(s))$. And if $x > y$ then $\Delta U = u(x - F^{-1}(1-s)) - u(y - F^{-1}(1-s))$.

Each part of ΔU is continuous. Therefore, $\lim_{x \rightarrow y} u(x - F^{-1}(s)) - u(y - F^{-1}(s)) = 0$ for the $x < y$ part of ΔU and $\lim_{x \rightarrow y} u(x - F^{-1}(1-s)) - u(y - F^{-1}(1-s)) = 0$ for the $x > y$ part of ΔU . Since the limits are equal to the value at $x = y$, all of ΔU are continuous.

Furthermore, ΔU has one unique maximum in x for $x \leq y$. The maximum is either at $\tilde{x}(s, y) = F^{-1}(s)$ or at the ‘‘corner’’ $\tilde{x}(s, y) = y$. Similarly, ΔU has one unique

maximum in x for $x \geq y$. The maximum is either at $\tilde{x}(s, y) = F^{-1}(1-s)$ or at the “corner” $\tilde{x}(s, y) = y$.

Without loss of generality, assume the point that around which $F - 1/2$ is odd is $z = 0$. Therefore, $y \geq 0$ and $1 - F(y) \leq 1/2 \leq F(y)$. Divide the interval $[0, 1]$ into $[0, 1 - F(y))$, $[1 - F(y), F(y)]$, and $(F(y), 1]$. These intervals correspond to cases (i), (ii), and (iii), respectively.

The characterizations for cases (ii) and (iii) are easily derived from the results above. For case (ii), where $s \in [1 - F(y), F(y)]$, $Q_X(x, y, s)$ reaches an interior maximum at $x = F^{-1}(s)$ on the interval $(-\infty, y)$, and it reaches a corner maximum at $x = y$ on the interval $[y, \infty)$. Thus, $Q_X(x, y, s)$ is single-peaked, attaining its global maximum at with $\tilde{x}(y, s) = F^{-1}(s)$. For case (iii), where $s > F(y)$, $Q_X(x, y, s)$ reaches a corner maximum at $x = y$ on both $(-\infty, y]$ and $[y, \infty)$. Thus, $Q_X(x, y, s)$ is single-peaked and attains its global maximum at with $\tilde{x}(y, s) = y$.

For case (i), where $s < 1 - F(y)$, $Q_X(x, y, s)$ is double-peaked, reaching an interior local maximum at $x = F^{-1}(s) < y$ and at $x = F^{-1}(1-s) > y$. All those remains are to show that $Q_X(F^{-1}(s), y, s) > Q_X(F^{-1}(1-s), y, s)$; then $\tilde{x}(y, s) = F^{-1}(s)$ is the global optimum for all s in case (i), and $x = F^{-1}(s)$ is only a local optimum. Substituting, we have

$$Q_X(F^{-1}(s), y, s) = 1 - G(-u(y - F^{-1}(s)))$$

and

$$Q_X(F^{-1}(1-s), y, s) = 1 - G(-u(y - F^{-1}(1-s))).$$

These two equations imply that $Q_X(F^{-1}(s), y, s) > Q_X(F^{-1}(1-s), y, s)$ if and only if $-u(y - F^{-1}(s)) < -u(y - F^{-1}(1-s))$. Since $F - 1/2$ is odd, $1 - F(y) = F(-y)$. Also, $s < 1 - F(y)$ for the case under consideration, so we have $F^{-1}(s) < -y < 0 < y < F^{-1}(1-s)$. Again using the fact that $F - 1/2$ is odd, $-y - F^{-1}(s) = F^{-1}(1-s) - y$. Thus, $y - F^{-1}(s) = 2y + F^{-1}(1-s) - y > F^{-1}(1-s) - y$. Thus, since $-u$ is symmetric and single-peaked, $-u(y - F^{-1}(s)) < -u(y - F^{-1}(1-s))$. Thus, $Q_X(F^{-1}(s), y, s) > Q_X(F^{-1}(1-s), y, s)$ as desired.

Parts (i)-(iii)', for the case $y \leq 0$, are proved in a similar fashion. □

Proof of Proposition 6.2. From equations A1 and A2 of the proof of Proposition 6.1, and the fact that ordinal utility functions are preserved from monotonic transformations, the preferences of candidates in party X can be written as

$$-u(x - F^{-1}(s)) + u(y - F^{-1}(s))$$

if $x < y$ and

$$-u(x - F^{-1}(1-s)) + u(y - F^{-1}(1-s))$$

if $x > y$. Assume $y > F^{-1}(1/2)$. Consider $x_1^* \equiv F^{-1}(s_0/2)$. We will show there exists a majority coalition preferring x_1^* to any other platform x' . There are four cases:

- (i) if $x' \leq x_1^*$, then all candidates in party X with $s' \geq s_0/2$ prefer x_1^* to x'
- (ii) if $x_0 < x' \leq y$, then all candidates in party X with $s' \leq s_0/2$ prefer x_1^* to x'
- (iii) if $y < x' \leq F^{-1}(1 - s_0/2)$, then all candidates in party X with $s' \leq s_0/2$ prefer x_1^* to x'
- (iv) if $x' > y$ and $x' > F^{-1}(1 - s_0/2)$, then all candidates in party X with $s' \geq s_0/2$ prefer x_1^* to x' .

Cases (i) and (ii) are immediate from single peakedness (since $x' \leq y$ and $x_1^* < y$). For cases (iii) and (iv), a candidate in party X at position s strictly prefers x_1^* to x' if

$$-u(x_1^* - F^{-1}(s)) + u(x' - F^{-1}(1-s)) - u(y - F^{-1}(1-s)) + u(y - F^{-1}(s)) > 0.$$

In the proof of Proposition 6.1, we showed $-u(y - F^{-1}(s)) < -u(y - F^{-1}(1-s))$. Therefore, a sufficient condition for the above relation to hold is

$$-u(x_1^* - F^{-1}(s)) + u(x' - F^{-1}(1-s)) \geq 0.$$

At $s = s_0/2$ the above expression is satisfied. Taking derivative with respect to s yields

$$\frac{dF^{-1}}{ds}(s)u'(x_1^* - F^{-1}(s)) + \frac{dF^{-1}}{ds}(1-s)u'(x' - F^{-1}(1-s)).$$

Note that $dF^{-1}/ds > 0$ and $u'(z) < 0$ if $z < 0$ and $u'(z) > 0$ if $z > 0$.

For case (iii) for the proposed coalition $s < s_0/2$, the derivative is negative. Since the condition above is satisfied at $s_0/2$, and the condition is decreasing in s , it must be satisfied for all $s < s_0/2$. Therefore, everyone in the proposed coalition strictly supports x_1^* over x' . For case (iv), the derivative is positive for $s > s_0/2$, and therefore, everyone in the coalition supports x over y .

By a similar argument, one can show if $y < F^{-1}(1/2)$ then $x_1^* = F^{-1}(1 - s_0/2)$ beats all other x' . The proof for y_1^* is also similar. \square

References

- Aldrich, J.H.: A Downsian spatial model with party activism. *Am Polit Sci Rev* **77**, 974–990 (1983)
- Aldrich, J.H., McGinnis, M. D.: A model of party constraints on optimal candidate positions. *Math Comp Model* **12**, 437–450 (1989)
- Aranson, P., Ordeshook, P.: Spatial strategies for sequential elections In: Niemi, R., Weisberg, H. (eds.) *Probability Models of Collective Decision Making*, Columbus: Charles E. Merrill (1972)
- Austen-Smith, D.: Two-party competition with many constituencies. *Math Soc Sci* **7**, 177–198 (1984)
- Austen-Smith, D.: Legislative coalitions and electoral equilibrium. *Pub Choice* **50**, 185–210 (1986)
- Austen-Smith, D., Banks, J.S.: Elections, coalitions, and legislative outcomes. *Am Polit Sci Rev* **82**, 405–422 (1988)
- Baron, D.P.: Electoral competition with informed and uninformed voters. *Am Polit Sci Rev* **88**, 33–47 (1994)
- Baron, D.P., Diermeier, D.: Elections, Governments, and parliaments in proportional representation systems. *Q J Econ* **116**, 933–967 (2001)
- Baron, D.P., Diermeier, D., Fong, P.: A dynamic theory of parliamentary democracy. *Econ Theory* (2011, this issue)

- Bernhardt, D.M., Ingberman, D.E.: Candidate reputations and the 'incumbency effect'. *J Pub Econ* **27**, 47–67 (1985)
- Callander, S.: Electoral competition in heterogenous districts. *J Polit Econ* **113**, 1116–1145 (2005)
- Calvert, R.L.: Robustness of the multidimensional voting model: candidate motivations, uncertainty, and convergence. *Am J Polit Sci* **29**, 69–95 (1985)
- Calvert, R.L., Isaac, R.M.: The inherent disadvantage of the presidential party in midterm congressional elections. *Pub Choice* **36**, 141–146 (1981)
- Cox, G.W.: Centripetal and centrifugal incentives in electoral systems. *Am J Polit Sci* **34**, 903–935 (1990)
- Coate, S., Knight, B.: Socially optimal districting. NBER Working Paper Series Working Paper 11462 (2005)
- Downs, A.: *An Economic Theory of Democracy*. New York: Harper and Row (1957)
- Epstein, D., O'Halloran, S.: A social science approach to race, redistricting, and representation. *Am Polit Sci Rev* **93**, 187–191 (1999)
- Enelow, J.M., Hinich, M.J.: A new approach to voter uncertainty in the Downsian spatial model. *Am J Polit Sci* **25**, 483–493 (1981)
- Fiorina, M.: The decline of collective responsibility in American politics. *Daedalus* **109**, 25–45 (1980)
- Gomberg, A.M., Marhuenda, F., Ortin, I.O.: Equilibrium in a model of endogenous political party formation. *Econ Theory* **24**, 373–394 (2004)
- Hotelling, H.: Stability in competition. *Econ J* **29**, 41–57 (1929)
- Ingberman, D.E., Villani, J.: An institutional theory of divided government and party polarization. *Am J Polit Sci* **37**, 429–471 (1993)
- Krehbiel, K.: Party discipline and measures of partisanship. *Am J Polit Sci* **44**, 212–227 (2000)
- Londregan, J., Romer, T.: Polarization, incumbency, and the personal vote. In: Barnett, W.A., Hinich, M.J., Schofield, N.J. (eds.) *Political Economy: Institutions, Information, Competition, and Representation*, Cambridge: Cambridge University Press (1993)
- Mayhew, D.: *Congress: The Electoral Connection*. New Haven: Yale University Press (1974)
- Palfrey, T.R.: Spatial equilibrium with entry. *Rev Econ Stud* **51**, 139–156 (1984)
- Poutvaara, P.: Party platforms with endogenous party membership. *Pub Choice* **117**, 79–98 (2003)
- Robertson, D.: *A theory of party competition*. London: Wiley (1976)
- Roemer, J.E.: Why the poor do not expropriate the rich: an old argument in new garb. *J Pub Econ* **70**, 399–424 (1998)
- Roemer, J.E.: *Political Competition: Theory and Applications*. Cambridge: Harvard University Press (2002)
- Schofield, N.: Local political equilibria. In: David, A.-S., Duggan, J. (eds.) *Social Choice and Strategic Decisions: Essays in Honor of Jeffrey S. Banks.*, Heidelberg: Springer (2004)
- Serra, G. (2005) Primary elections in two-party systems: the effects of party nominations on public policy. Unpublished manuscript, Harvard University
- Shotts, K.W.: The effect of majority-minority mandates on partisan Gerrymandering. *Am J Polit Sci* **45**, 120–135 (2001)
- Shotts, K.W.: Gerrymandering, legislative composition, and national policy outcomes. *Am J Polit Sci* **456**, 398–414 (2002)
- Snyder, J.M. Jr.: Safe seats, marginal seats, and party platforms: the logic of party platform differentiation. *Econ Polit* **6**, 201–213 (1994)
- Snyder, J.M. Jr., Ting, M.M.: An informational rationale for political parties. *Am J Polit Sci* **46**, 90–110 (2002)
- Wittman, D.: Candidates with policy preferences: a dynamic model. *J Econ Theory* **15**, 86–103 (1977)
- Wittman, D.: Candidate motivation: a synthesis of alternatives. *Am Polit Sci Rev* **77**, 142–157 (1983)