Soc Choice Welfare 26:547–569 (2006)
DOI 10.1007/s00355-006-0087-1

ORIGINAL PAPER

Guillermo Owen · Bernard Grofman

Two-stage electoral competition in two-party contests: persistent divergence of party positions

Received: 22 March 2004 / Accepted: 2 March 2005 / Published online: 11 May 2006

Abstract Models of party competition building on Downs (1957) have recognized that there are centrifugal and centripetal forces in party competition; but one such force, the existence of party primaries, has been remarkably neglected in recent literature. We consider party/candidate policy divergence in two-party competition in one dimension where there is a two-stage electoral process, e.g., a primary election (or caucus) among party supporters to select that party's candidate followed by a general election. We develop a model in which (some or all) voters in the primary election are concerned with the likelihood that the primary victor will be able to win the general election and being concerned with that candidate's policy position. This model is similar in all but technical details to that given in an almost totally neglected early paper in *Public Choice* Coleman (1971) 11:35–60, but we offer important new results on electoral dynamics for candidate locations. In addition to accounting for persistent party divergence by incorporating a more realistic model of the institutions that govern elections in the U.S., the model we offer gives rise to predictions that match a number of important aspects of empirical reality such as frequent victories for incumbents and greater than otherwise expected electoral success for the minority party in situations where that party has its supporters more closely clustered ideologically than the supporters of the larger party (in particular, with a concentration of voters between the party mean and the population mean).

A much earlier version of this paper was presented at the Annual Meeting of the Public Choice Society, Long Beach, California, March 24–26, 1995. We are indebted to Dorothy Green and to Clover Behrend-Gethard for bibliographic assistance and to Nicholas Miller and the late Peter Aranson for helpful suggestions.

G. Owen
Department of Mathematics, Naval Post-Graduate School, Monterey, CA, USA

B. Grofman (✉)
School of Social Sciences, University of California, Irvine, CA, USA
E-mail: bgrofman@uci.edu

1 Introduction

The simple Downsian model of two-party competition over a single-issue dimension for a single office predicts party convergence to the policy position on that dimension espoused by the median voter if parties/candidates are motivated solely by office-seeking (Downs 1957) and voters solely by policy proximity to candidate positions. While Downs was certainly correct to call attention to centripetal forces in party competition over a single ideological dimension, there are also important centrifugal forces that need to be taken into account. Even in the U.S., the country whose presidential elections inspired Downs to propose it, the simple Downsian model fails to predict the persistent party divergence we observe.[1]

The need to model centrifugal forces in two-party competition[2] has been well recognized in literature of the past two decades, with a variety of different approaches offered to explain candidate and party divergence. These include the role of ideologically committed party activists/interest groups who are a major source of campaign resources (Aldrich 1983, Baron 1994, Morton 1993); candidates who have policy preferences that they wish to see implemented and not just a desire to win an election (Wittman 1973, 1977, 1983) directional rather than proximity-based voting (Merrill 1993, Merrill and Grofman 1997, Rabinowitz and Macdonald 1989; discounting of candidate positions Grofman 1985, Merrill and Grofman 1999; multiple dimensions of issue competition Schofield 1996; non-policy related motivations for candidate support such as those that give rise to reputational effects and incumbency advantage or to partisan bias (Adams 1999, Adams et al. 2006, Berhardt and Ingberman 1985, Feld and Grofman 1991; and strategic calculations such as concern for future entry Brams 1980, Brams and Merill 1991, Palfrey and Erikson 1994, or policy balancing across multiple contests Alesina and Rosenthal 1995.[3]

The approach we take focuses on the fact that in the United States two-stage election processes such as a primary (or caucus) followed by a general election are

[1] Similarly, it is an indubitable fact that candidates of opposite parties in many recent U.S. elections (e.g., the 1992 senatorial contests in California, the 1964 and 2004 presidential election, to name two blatant examples) cannot be characterized as tweedledum–tweedledee. Moreover, numerous scholars have shown that, when a given U.S. constituency elects members of opposite parties (e.g., when a congressional seat changes hands to a member of the opposite parties, or in states which are simultaneously represented by senators of opposite parties), the difference in voting records between the office-holders of different parties (as judged, say, by Americans for Democratic Action (ADA) scores or similar roll-call measures) can be huge (Bullock and Brady 1983, Fiorina 1974, Grofman et al. 1990, Polle and Rosenthal 1984).

[2] We might also note that when we try to model U.S. politics in two dimensions rather than in one dimension, we still do not observe party convergence. As part of the European party manifestos project, Robertson (Budge et al. 1987, Robertson 1987, p. 69. Fig 3.1) factor-analyzed party platforms to generate a two-dimensional issue space for the U.S., 1948–1980. He found that the Democratic and Republican parties in the United States remained in distinct areas of that issue space. The puzzle of non-divergence is actually even more general. Macdonald, Listhaug and Rabinowitz (Macdonald et al. 1991) show that spatial locations of Western European political parties exhibit a missing center, i.e., a tendency toward the absence of parties in the center of the space of voter ideal points. However, we shall not deal with competition among more than two parties in this essay.

[3] For more general reviews of the literature on spatial competition inspired by Downs see Enelow and Hinich (Adams et al. 2006; Enelow and Hinich 1984; Enelow and Hinich 1990; Grofman 1993, 1996, 2004, 1999).

common. Remarkably, while this point is central to two models of two-party divergence offered in the early l970s, (Aranson and Ordeshook 1972; Coleman 1971, 1972), the implications of such two-stage processes for party/candidate convergence and the size of party support coalitions have been almost entirely neglected in the past two decades—indeed, there have been almost no recent citations to either the Coleman (1971, 1972) or the Aranson and Ordeshook (Aranson and Ordeshook 1972) work. This neglect is especially puzzling given the rise of the "new institutionalism" (Shepsle 1979; Shepsle and Weingast 1984) during this same period.[4]

In the Aranson and Ordeshook model, candidates are assumed to develop expectations about the probability of victory in the primary election (P_1) and the general election (P_2) as a function of the policy position they associate themselves with, and are posited to choose a spatial location so as to maximize $P_1 \times P_2$. The Aranson and Ordeshook (1972) model of two-stage election processes makes candidate choices the focus of their modeling. In contrast, in the Coleman (1971, 1972) model, the focus is on voter motivations. In the Coleman model some (or all) voters in the primary election are concerned with the likelihood that the primary victor will be able to win the general election and with that candidate's policy position and choose among candidate's accordingly. Roughly speaking, Coleman assumed that voters maximize a function that can be thought of as the benefit derived from selecting a party representative whose location is close to their own ideal point discounted by the likelihood that such a candidate will be elected in the general election.

The model we offer is similar in spirit to some of the work on strategic voting/bandwagon effects,[5] and a direct extension of Coleman's (1971, 1972) work. We go beyond that work to offer new results on electoral dynamics for candidate locations in a sequence of two-stage elections and in showing the link between the nature of voter ideological distributions and the likely outcomes of elections. Like Aranson and Ordeshook (Aranson and Ordeshook 1972, Coleman (1971, 1972) supposes that a candidate is constrained to offer roughly the same ideological position in the general as in the primary. We take the same view in the model given below.[6]

---

[4] However, Coleman's work first appeared in an early issue of *Public Choice*, shortly after the journal changed its name from *Papers on Non Market Decision-Making* at a time before many libraries subscribed to this subsequently well-known journal. The 1972 articles of Coleman and Aranson and Ordeshook are book chapters in an excellent edited volume that deserves to be far better known, but that was published by a firm that shortly, thereafter, went out of business.

[5] There is a considerable literature about strategic voting in multicandidate contests (e.g., Black 1978, Cain 1978 in which voter support for any candidate depends not just upon that candidate's (relative) policy positioning but also upon the voter evaluations of that candidate's (relative) likelihood of electoral success Similar ideas are also found in the literature in comparative politics on the psychological effects in Duverger's Law 1958, Fedderson 1992, Riker 1982. Brams' work on bandwagon dynamics (see, e.g., discussion in Brams 1978) makes use of the idea that actors make choices to join coalitions that are in part based on the likelihood that the candidate whose coalition they choose to join will be able to win.

[6] In a planned follow up paper we will look at what happens in a two-stage election process if some voters are "expressive" voters in the sense of Glazer 1993 and Glazer et al.1998, i.e., vote for candidates based not simply on their policy platforms but also on the basis of which (types of) voters are expected to be in the candidate's support coalition. (The term "expressive voting" has been used in the literature in various ways. The definition of expressive voting offered in Brennan and Lomasky 1993 is related to but not identical with that offered by Glazer.)

The model we offer gives rise to predictions that match empirical reality in a variety of ways. Not only does it predict persistent nonconvergence of party positions (as do several of the models we mentioned earlier) whenever voters have even some concern for policy outcomes and not just the success of candidates of "their" party, but it also leads us to expect frequent victories for incumbents and provides an explanation for why a minority party may do better in elections than one would otherwise expect when its supporters are more ideologically concentrated than those of the majority party.[7] Moreover, unlike many of the other models that generate centrifugal pressures, it achieves its results simply by making realistic assumptions about the nature of political institutions in the United States, and what we regard as plausible assumptions about voter motivations. In particular, it incorporates a parameter, α, that is a measure of ideological concern, i.e., concern for selecting in the primary a candidate who can win as opposed to selecting a candidate who mirrors the voter's own position.

We would also emphasize that there is a key difference in our model between the results for the case of an open seat, and for the case of a seat where an incumbent is running for reelection. In case of an open seat, there is generally a symmetric result: both parties tend to nominate candidates, more or less, equidistant from the population median. In contrast, in the case of an incumbent, the tendency for the other party will be to nominate someone close to the center if the incumbent is far away and some distance from the center if the incumbent is near the center.[8] This feature of our model, with centrist policies on the part of an incumbent not being matched by centrist policies on the part of a challenger, is quite distinct from most other work in the Downsian tradition (see, however, Adams et al. 2006).

While our operationalization of the model is in terms of a particular utility function, the basic intuitions about convergence and the effects of party distributions we offer hold for a very wide class of functional forms (see Assumption 1 below). Moreover, we offer results for both the case where primary elections in each party are simultaneous and the case where elections are nonsimultaneous (or one party's result can be known in advance, as when an incumbent is expected to win her party's primary), and for the cases with and without an incumbent. Like Cox 1997, our desire is to build a more institutionally rich approach to understanding electoral competition—one that melds the spatial modeling literature inspired by Downs 1957 with issues of concern to students of comparative election systems interested in institutional effects (see, e.g., Lijphart 1984, 1992, and the various essays in Lijphart and Grofman 1984, and Grofman and Lijphart 1986.

---

[7] Other work showing the relative advantages of being a concentrated minority includes Miller 1996 and Merrill et al.1999.

[8] Incumbents are tied to positions they have taken in the past. Also, it is rare for an incumbent—with advantages of name recognition, and, for long-time incumbents, almost certainly a well-developed campaign organization—to lose a party primary.

2 The basic model

2.1 The space of voters and parties

We assume that voters are distributed along a left–right spectrum and located at positions t, where $0 \leq t \leq 1$. The left-wing party (D) occupies positions from 0 to s; the right-wing party (R), from s to 1. Here, s is a fixed number, $0 < s < 1$, possibly, but not necessarily, equal to 1/2. There is a distribution function $F(t)$, monotone nondecreasing, with $F(0) \geq 0$ and $F(1) = 1$. $F(t)$ is the fraction of voters located at some position $u \leq t$.

The median position of the distribution is m, defined by $F(m) = 1/2$. We will also talk about the two party medians, d and r, defined by $F(d) = F(s)/2$ and $F(r) = (F(s) + 1)/2$. In case F is continuous and strictly increasing, these three equations determine m, d, and r uniquely. For discontinuous or for not-strictly increasing functions, slight modifications must be made to the definitions of m, d, and r; we dispense with the details.

2.2 The basic model with and without an incumbent in place

2.2.1 Incumbent

We consider first what happens when there is an incumbent.

Suppose the R party has an incumbent located at position $y \in [s, 1]$. A voter in the D party, located at position t, would be happiest if D could win the election with a challenger at position t. However, it is not certain such a candidate will win. Thus, in a primary election, the voter might prefer a candidate "closer to the center" than t, i.e., located at some $x \in [t, m]$, on the grounds that such a candidate has a greater probability of winning the general election.

Following Coleman 1972, we assume that the voter at t has utility function $u(x, t)$ for victory by a candidate located at position x, and similarly $u(y, t)$ for victory by the incumbent (at y). Let $p(x, y)$ be the probability that a candidate at x will defeat one at y in the general election. Then

$$g(x, y, t) = u(x, t)p(x, y) + u(y, t)(1 - p(x, y)) \tag{1}$$

is voter t's expected utility if the D party nominates a challenger at x. Thus, in the primary, the voter's best hope is to choose a candidate that maximizes the function g.

The following assumption seems reasonable:

Assumption 1(a)–(d) We assume that, (a) for fixed t, the function $u(x, t)$ is a continuous, unimodal function, with a maximum at $x = t$ and hence, strictly increasing in x for $x < t$ and strictly decreasing for $x > t$. (This corresponds to the idea that the voter wants to elect someone as close as possible to his own position.) We also assume that, (b) for fixed y, $p(x, y)$ is strictly increasing in x for $x < y$, and strictly decreasing for $x > y$. We also assume (c) this function is continuous in x except possibly at $x = y$. Finally, we assume (d) that $p(x, y) > 0$ for all x and y.

Lemma 1 Under assumptions 1(a)–(d), the function $g(x, y, t)$ is continuous in x.

Proof The proof to this and subsequent omitted proofs are given in Appendix A.

Lemma 2 Suppose that $y > t$. Then, under assumptions 1(a)–1(d), the function $g(x, y, t)$ will achieve its maximum (in x, for fixed y and t) at some $x^* \in [t, y)$. We call this x^*, voter t's *preferred primary position*.

Suppose, then, that y is fixed. An individual member of D, located at position t, finds that, in the primary election, his preference is for a candidate at position x^*, as described in Lemma 2. Note that x^* varies with the individual. It is then possible to give a distribution for the positions x^*. However, as this x^* depends not only on t and on y, but also on the (individual) utility and (subjective) probability functions, $u(x, t)$ and $p(x, y)$—which might be very different for different members of D—this distribution is difficult to calculate. However, it may be seen that, as $t \le x^* < y$ for each individual, then the median q^* of the distribution must satisfy $d \le q^* < y$, where d is the median for party D.

At this point a standard argument states that, in the D party primary, a candidate A at q^* will defeat any candidate B at a different position. The idea is that B chooses $x < q^*$, then all voters whose preferred primary position is to the right of q^* will prefer A; if $x > q^*$, then all voters with preferred positions to the left of q^* will prefer A. As q^* is the median position (so the argument goes), A will always be preferred by a majority of the D party voters. Thus, the tendency will be for the D party to nominate a candidate at $q^*(y)$.

Unfortunately, there is a flaw in this argument. The problem is that the function $g(x, y, t)$ need not be unimodal in x. To see this, we might consider a D voter located at a position t very close to the R party incumbent's position at y. (For example, let $t = 0.49$ and $y = 0.51$.) Such a voter may well be so satisfied with the incumbent (even though nominally of the opposite party) that he would prefer his own party (D) to nominate a candidate with little chance of winning the general election. Thus, the voter at position $t = 0.49$, though his own preferred primary position may be $x^* = 0.50$, would prefer a D candidate at $x = 0.01$ to one at $q^* = 0.35$ because the candidate at 0.01 will almost certainly guarantee the election of the incumbent. We will call this event, *strategic primary voting.*

We will, however, assume that this event is so unlikely that we may safely disregard it. More exactly, though such voters may exist, we will assume that the likelihood that they would alter the outcome of the primary election is extremely low. In Appendix B we give reasons in support of this assumption. Moreover, we will subsequently argue that, if this happens with any frequency, the given D voters would probably migrate to the R party. In summary, we feel safe in assuming that our original argument will hold: a D candidate at position q^* will (in the primary) defeat a candidate at any other position.]

The results we state will be based on combining Assumptions 1(a)–(d) with Assumptions 2, 3, and 4 below. A further assumption, that of complete information, appears as Assumption 5 in the next section of the paper.

Assumption 2 Strategic primary voting, though it may exist, will not be prevalent enough to change the outcome of an election.

Because this assumption is intended only to apply to the somewhat peculiar case of some voters in one party preferring the candidate of the other party to their own party's candidate, we will leave Appendix B to a fuller discussion of the meaning of this assumption in terms of the distribution of voters.

To solve specifically for a location, q^*, we need to make some further assumptions about the two functions $u(x, t)$ and $p(x, y)$. The results we obtain below depend on the specific form of these two functions, but the basic intuitions derived here will be valid for a wide number of functional forms.

Assumption 3 We will assume that the voter utility function has the form

$$u(x, t) = e^{-\alpha|x-t|} \tag{2}$$

where α is a parameter representing the importance given to ideology. (In effect, we will see later on that $x^*(y)$ tends to be further away from the party median when α is small; i.e., small values of α means lesser importance is given to ideological purity.)

Assumption 4 For the function $p(x, y)$, we will assume that voters will vote in the general election for the candidate closer to their position. Hence, if there was no chance element, the candidate closer to the median position, m, would certainly win the general election. Thus, for $x < y$, the D candidate would win if $x + y > 2m$, and the R candidate would win if $x + y < 2m$. However, because of unforeseen (chance) events, the electorate may shift ideologically from one side to the other. We assume that this shift can be expressed as a random change in m, normally distributed, with mean 0 and known variance, σ^2. (The parameter σ represents in some way the *volatility* of the electorate.) If this is so, then for $x < y$, the D candidate will win with probability

$$p(x, y) = \Phi\left(\frac{x + y - 2m}{2\sigma}\right) \tag{3}$$

where Φ is the (cumulative) standard normal distribution function, i.e.,

$$\Phi(u) = \int_{-\infty}^{u} \varphi(\nu) d\nu$$

and

$$\varphi(\nu) = \frac{1}{\sqrt{2\pi}} e^{-\nu^2/2}.$$

Of course we have assumed that t, the voter's position, varies among the members of party D. Now, the parameters α and σ are subjective parameters, and thus, could well vary among members of the party. We will, however, make the simplifying assumption that α and σ are common. Thus, D party members, though they differ in their preferences, assign similar importance to ideology, and have similar ideas as to the volatility of the electorate.

Under these assumptions, we can prove

Theorem 1 *Assuming α and σ fixed, then, for a given position y of the incumbent, there exists a position x*(y) such that the utility-maximizing D voter located at t will prefer the D candidate to be in the position max{t, x*(y)}. [Note that, in this theorem, x*(y) does not depend on t.]*

Thus, in a primary election among D voters, with an R incumbent at y, those D voters to the left of $x^*(y)$ will mask their true feelings and vote for a candidate at x^* (if such a candidate exists); those to the right of x^* will reveal their feelings by voting for a candidate at their own position. Then if $x^* < d$, a candidate at d will defeat any other candidate. If $x^* > d$, a candidate at x^* will defeat any other.

Of course, there is no guarantee that such a candidate will exists. However, we feel that this gives rise to a *tendency* to nominate such a candidate. By this, we mean that candidates close to this position have greater likelihood of appearing, and if they appear, of winning the nomination.[9] Hence, with an R incumbent at y, the tendency is for the D party to nominate a candidate located at $q^*(y) = \max\{x^*(y), d\}$.[10]

Unfortunately, computation of $q^*(y)$ can be quite complicated. We can, however, answer a related (and perhaps more interesting) question: which of the two candidates is more likely to win?

From Eq. 2, we see that the candidate closer to m will probably (i.e., with probability greater than 1/2) win the general election. Thus, the question is whether the quantity $y - m$ is larger than $m - x^*(y)$. If, in fact, $y - m$ is larger, then the challenger is closer to the mean than is the incumbent. In such a case, we say that there is a *centralizing tendency*, and the challenger will probably win. If, on the other hand, $m - x^*(y)$ is the larger, then there is a *polarizing tendency* and the incumbent will probably win.

[9] We also believe our results are close to empirical reality in that, for example, we know that U.S. senators of the same party from the same state customarily look virtually identical to one another in voting behavior as judged by aggregate roll-call measures such as ADA Grofman et al. 1995. This evidence suggests that, in a given constituency, we can expect that certain candidate locations are advantaged in the primary of a given party; and ceteris paribus we would expect that the "privileged" locations correspond to the preferences of the median voter in that primary.

[10] Also, while our model posits two-party competition (and, thus, two-candidate competition) in the general election, and is couched in terms of two-candidate competition in the primary, it can be extended to the case of multiple primary candidates where the winner of the primary can be expected to capture a near majority of the primary electorate. In a multicandidate plurality election, it is possible for there to be a candidate location at which a candidate could (under the assumptions we use about a two-stage electoral process and vote motivations therein) defeat any rivals in the primary were there to be a paired competition (i.e., for there to be a Condorcet winner), and yet, the plurality winner of that primary might be located elsewhere. This could occur, for example, if there is more than one candidate located near to (but not precisely at) the vote-maximizing location. Nonetheless, our results are considerably more robust than might first appear, in that voters to the left (right) of the optimum candidate location in the Democratic (Republican) primary, would vote for the candidate at that location and, given the realities of contemporary U.S. politics, there are going to be a lot of such voters (usually near a majority). Thus, if we extend the assumptions of our modeling to multiparty competition, we should get far more stability for a two-stage process with forward-looking voters than is found in the one-stage multicandidate simulation results of Cooper and Munger 2000.

Theorem 2 There will be a centralizing tendency (and the challenger will probably win) if y is greater than the smaller one of the quantities

$$2m - d \quad \text{and} \quad m - \frac{\log\left(1 - \alpha\sigma\sqrt{2\pi}\right)}{2\alpha}. \tag{4}$$

There will be a *polarizing tendency* if y is smaller than both of these two quantities.We should note that, in the second of the expressions (Eq. 4), the logarithm (if it exists) will always be negative. On the other hand, the logarithm does not exist for negative numbers or zero. In such a case, the second quantity is to be treated as $+\infty$ and so the condition for a centralizing tendency is merely that $y > m - d$.

In general, we see that, for this type of utility function, there is a polarizing tendency if y is close to the median m, but there is centralizing tendency for larger values of y. (In any case, D will probably not move left of its median position d.) Heuristically, we interpret this as follows: if the R incumbent is very conservative, then the left wing of the D party will place great importance on winning the general election and might be willing to sacrifice its leftist principles in the hope that a moderate D challenger may win. If, on the other hand, the R incumbent is very moderate, the left wing of the D party will see little to be gained if a moderate D candidate is elected, and so will be more insistent on getting a leftist D challenger—even though such a challenger might have a relatively small probability of winning.

In case there is an incumbent from D, located at position x, this analysis can be applied, mutatis mutandis, to the selection of the R nominee. In general, we will find that, for a given x, there is a corresponding $y^*(x)$ such that, in the R primary, all voters to the left of y^* will vote their own preferences, while all those to the right of y^* will mask their preferences and vote for a candidate at y^*. The tendency will then be for the R party to nominate a challenger at $\min\{r, y^*(x)\}$, where r is the R party median.

The condition for a polarizing vs a centralizing tendency is essentially the same here, and Eq. 4 takes the form

$$x > \max\left\{2m - r, m + \frac{\log\left(1 - \alpha\sigma\sqrt{2\pi}\right)}{2\alpha}\right\} \tag{5}$$

with the understanding once again that if the logarithm in Eq. 5 does not exist, then Eq. 5 reduces simply to $x > 2m - r$.

2.2.2 Open seat

We consider next what happens when there is no incumbent, with each party having a primary, held simultaneously. It is of course difficult to determine just what one party's voters know about the likely winner of the opposition primary. We will however assume that, in effect, there is a very good system of publicly released polls, so that voters in each primary vote with good information as to the opposition's likely nominee. Thus, each party nominates its candidate with an eye to the opposition nominee.

More precisely, we make the following assumption:

Assumption 5 Each party's voters know the induced preferences of the other side. In particular, they know the position of the other side's median voter, and they know that the other side has a tendency to nominate a candidate at this voter's induced preference.

We look here for equilibrium positions, ($x\#$,$y\#$), satisfying

$$x\# = \max\{d, x^*(y\#)\} \tag{6}$$

$$y\# = \min\{r, y^*(x\#)\} \tag{7}$$

where x^* and y^* are as above.

Now, for these equations, there are, essentially, four possibilities, as there are two alternatives in each of the two equations. Two are of special interest and depend to a large extent on the size of α and σ, and on the distance from the population median, m, to the two party medians, d and r. Recall that α is small if individuals place little importance on ideological purity, and that σ is small if the electorate is very rigid. Then we have the following result: if the electorate is very rigid, then the parties will tend to nominate candidates close to the population median, m, at equal distance from m. If the electorate is volatile, then the parties will tend to nominate candidates at their party medians, d and r.

Specifically,

Theorem 3 If

$$\alpha\sigma\sqrt{2\pi} \le \min\left\{1 - e^{\alpha(m-r)}, 1 - e^{\alpha(d-m)}\right\}, \tag{8}$$

then

$$x\# = m - \mu, \quad y\# = m + \mu \tag{9}$$

where

$$\mu = -\frac{\log\left(1 - \alpha\sigma\sqrt{2\pi}\right)}{2\alpha}. \tag{10}$$

If, on the other hand,

$$\alpha\sigma\sqrt{2\pi} \ge \max\left\{1 - e^{\alpha(m-r)}, 1 - e^{\alpha(d-m)}\right\}, \tag{11}$$

then $x\# = d$ and $y\# = r$.

It may be of interest to note that, as the product $\alpha\sigma$ goes to 0, μ is asymptotically equivalent to $\sigma\sqrt{\pi/2}$. Thus, if the electorate is very rigid, or if the party members have relatively little interest in ideological purity, Eq. 11 will take the form

$$x\# = m - \sigma\sqrt{\pi/2}, \quad y\# = m + \sigma\sqrt{\pi/2}. \tag{12}$$

Consider next the situation that arises if the two parties' primaries are held on different dates. Without loss of generality, let us assume that the *R* party primary is held first. In that case, *D*'s problem—once the *R* nominee is known—will be as in case (a) above, where an incumbent exists, and the tendency will be for *D* to nominate a candidate located at the position $q^* = \max\{d, x^*(y)\}$, where $x^*(y)$ is as discussed above.

Consider then the decision to be made by an *R* voter. Whatever the position *y* of the *R* nominee, he must expect that the *D* party will nominate someone at $q^*(y)$. Then, for a voter at position *t*, his expected utility is given by

$$h(y, t) = g(q^*(y), y, t) \tag{13}$$

and thus, he will look for the *y* that maximizes $h(y, t)$.

Then we have the following:

Theorem 4 Assuming the *R* party primary is held first, the tendency will be for *R* to nominate a candidate at the same position *y#* as in the case of simultaneous primaries. If this happens, then the *D* party tendency will be to nominate a candidate at the same *x#* as in the case of simultaneous primaries.

In other words, *simultaneous primaries and staggered primaries should give the same results*, though in practice the second party to choose will be better able to take advantage of the first party's errors.

Example 1 Consider the following very symmetric situation. Let $F(t) = t$; i.e., there is a uniform distribution of voters along the spectrum so that $m = 1/2$. Assume also that the two parties divide the spectrum equally, with *D* from 0 to 1/2 and *R* from 1/2 to 1. In this case $d = 1/4$ and $r = 3/4$. Let now $\sigma = 0.05$, and $A = 4$. We then obtain $\mu = 0.174$, so $x\# = 0.326$ and $y\# = 0.674$. Over the long run, there will be a tendency for the parties' nominees to approach these points. Note also that an *R* incumbent between 0.5 and 0.674 will likely win, whereas one between 0.674 and 1 will probably lose.

Example 2 Let $\alpha = 10$, and let all other parameters be as in Example 1. In this case, the tendency will be toward the two party medians: $x\# = 0.25$ and $y\# = 0.75$.

Example 3 Suppose next that $\alpha = 2$, and let everything else be as in Examples 1 and 2. Then we will have $\mu = 0.44$, so that $x\# = 0.356$ and $y\# = 0.644$.

We see from these three numerical examples that, as α decreases (so that ideological purity becomes less important) the likely candidates will move toward the median of the distribution.

Example 4 Finally, suppose $\sigma = 0.025$, while everything else is as in Example 1. In this case, we will have $\mu = 0.072$, so $x\# = 0.428$ and $y\# = 0.572$. We see that a decrease in σ, corresponding to a greater rigidity in the electorate, will also cause the candidates to move toward the median.

2.3 Dynamics of expected party shares in two-stage elections

We consider next the question of the expected split in the electorate, s–i.e., is s arbitrary, or will there be a tendency for this s to move to some "natural" equilibrium position over the long run? There are actually two possibilities, depending on whether $\alpha\sigma$ is small or large (as discussed above).

2.3.1 An example to illustrate sequential dynamics of two-stage elections

Consider the following example.

Example 5 Suppose once again that there is a uniform distribution of the population ($F(t) = t$), but that, for whatever reason, $s = 0.6$; i.e., the D party has 60% of the population. Then the two party medians will be at $d = 0.3$ and $r = 0.8$.

$\alpha\sigma$is large. Suppose, first, that $\alpha\sigma$ is large. In that case, as seen in the earlier discussion of the basic model, the tendency is for the two parties to nominate candidates at the party medians. Then in the general election, the population will probably split at the midpoint between these, that is, 0.55. The D party will win, but its share of the vote is only 55%, even though 60% of the voters are affiliated with it. The reason is that, with such a large party, the moderate wing of the party will begin to feel more in tune with the R party. How long they will remain within the D party, is of course, another question.

Continuing with Example 5, it seems reasonable to expect that, if the two candidates are, in fact, chosen to lie at $x = d = 0.3$ and $y = r = 0.8$, then, the D party members located in the interval $0.55 \leq t \leq 0.6$ will *eventually* become disenchanted with D and hence, migrate to the R party. Thus, after a certain while we can expect that the party split will be at $s = 0.55$.

This is not the end of the story, however. The point is that, with this change in the split point s, the two party medians will also change, so that d will now be at 0.275 while r moves to 0.775. As $\alpha\sigma$ is still assumed to be large, the candidates will be chosen at the two party medians, and the vote split will now be at 0.525. But this means that D voters in the interval $0.525 \leq t \leq 0.55$ will be unhappy with their party, and once again we can expect them to migrate to R. This process will continue until the parties split at $s = 0.5$.

In the more general case, we look for a split point, s, which will cause party members to vote for their own party's nominee. Still assuming that $\alpha\sigma$ is large, the nominees will tend to be located at the two party medians, d and r, respectively. We then have the relations

$$F(d) = \frac{F(s)}{2} \tag{14}$$

$$F(r) = F(d) + \frac{1}{2} \tag{15}$$

$$s = \frac{d+r}{2}. \tag{16}$$

These three equations can be solved for the three unknowns. Thus, the split point s, which had been assumed as a given (exogenous) in the previous section of this paper, has a natural (endogenous) value.

$\alpha\sigma$ is small. Suppose, on the other hand, that $\alpha\sigma$ is small. Then, by Theorem 5, the tendency will be for the parties to nominate candidates at positions $x^* = m - \mu$ and $y^* = m + \mu$. If in fact the nominees are at these two positions, then everyone to the left of m will vote D (because such voters are closer to x^* than to y^*), and everyone to the right of m will vote R. In such case, there will be (as mentioned above) a very close election. More interesting is that, as people generally try to be in the party with the more compatible candidates, we can expect that the parties will split at the population median; i.e., the tendency will be for a natural party split at $s = m$. The parties will tend to be of equal size.

2.3.2 Additional examples of election dynamics for two-stage elections

Example 6 Suppose the population distribution is given by $F(t) = t^2$. In this case, there is a concentration of voters on the right side of the spectrum, and indeed, we find the median to be located at $m = 0.707$. As mentioned above, we obtain a natural value $s = m = 0.707$ if $\alpha\sigma$ is small.

Suppose however that $\alpha\sigma$ is large in this example. Eqs. 14, 15, 16 now take the form

$$\begin{aligned} s^2 &= 2d^2 \\ r^2 &= d^2 + \tfrac{1}{2} \\ s &= \tfrac{d+r}{2} \end{aligned}$$

These have the solution $d = 0.462$, $s = 0.653$, $r = 0.845$. Note that the split point is considerably to the left of the median; i.e., the R party will tend to be considerably larger than D. Thus, the concentration of voters on one side of the spectrum will tend to help the party on that side.

Example 7 Suppose

$$F(t) = \begin{cases} 0.2 + t, & \text{for } 0 \leq t \leq 0.8 \\ 1, & \text{for } t \geq 0.8 \end{cases}$$

This corresponds to a large block of voters (20%) at the left extreme and otherwise a uniform distribution for the remaining voters. In this case, the median voter is at $m = 0.3$. Application of Eqs. 14, 15, 16 will give us $s = 0.3$, $d = 0.05$, and $r = 0.55$. Thus, the parties are of equal size, and concentration of voters at the left end of the spectrum does not help D.

Example 8 Suppose

$$F(t) = \begin{cases} 0.3 + t, & \text{for } 0 \leq t \leq 0.7 \\ 1, & \text{for } t \geq 0.7. \end{cases}$$

In this case, the median voter is at $m = 0.2$, and Eqs. 14, 15, 16 will give us $s = 0.233$, $d = 0$, and $r = 0.467$. In this case, D will be the larger party and we conclude that the advantage lies, not so much in having a concentration of voters, but in having a concentration of voters *between the party mean and the population mean*.

Conclusion

The aim of this paper has been to extend the standard Downsian model of party competition to make it more realistic both in its assumptions and in its results. To make it more realistic in its assumptions, we have allowed for an important institutional feature of many U.S. elections, a two-stage process with both a primary and a general. Moreover, by allowing primary voters to care *both* about what policies candidates espouse and the likelihood that a candidate will win the general election, we have avoided the peculiar dichotomy of the standard Downsian approach wherein voters are posited to care only about their policy proximity to the candidates at the same time that candidates are posited to care only about winning.[11]

The simple modification of recognizing the two-stage nature of much of U.S. electoral politics leads us to expectations about party competition that are much more in line with what is actually observed in the electoral arena.

First, in our model we almost always get nonconvergence—with the most likely result, the location of the winning candidates of each party near their own party medians, perhaps slightly shifted toward the preferences of the overall median voter. This theoretically derived expectation squares very well with the empirical studies of candidate/party divergence such as Shapiro et al. 1990. Of course, as noted earlier, there are numerous other models that also give rise to expectations of nonconvergence (see, e.g., Alesina and Rosenthal 1995; but these models do not as

[11] Wittman 1973, 1977 and numerous subsequent authors have modified the standard model to permit candidates to care both about winning and about policy. Models of strategic voting Black 1978, Cain 1978 permit voters to care about candidate's chances of winning.

readily allow for modeling the dynamics of change in the size of party support coalitions, nor are they motivated by a realistic modeling of a key institutional feature of party competition in the U.S., that is, the existence of party primaries. Moreover, our model allows us to account for seemingly perverse outcomes such as that in which a political party paints itself into an ideological corner that dooms it to continued minority status because its voters maintain an emphasis on ideological purity that prevents them from nominating a candidate with policies with more appeal to the median voter than the policies offered by the candidate of the opposing party, and we get different results about the likelihood of convergence in cases with an incumbent than we do for cases involving open seats.

Second, in general, we get incumbency advantage (as long as there is some concern among voters in the "out" party for the policy outcomes that their candidate will espouse). While there has been a great deal written about reasons for incumbency advantage such as name recognition and access to various perquisites of office, as far as we are aware of, none of the standard spatial models of party competition give any particular advantage to incumbents.[12] Indeed, the standard Downsian model applied to multidimensional competition implies that incumbents are always vulnerable to defeat.[13]

Third, our model permits an explanation for the success of a candidate of a party whose affiliates are a minority of the electorate, based on the potential for a minority party whose voters are more ideologically concentrated to field the candidate who is closer to the overall median voter. In particular, we show an advantage to a party with a concentration of voters between the party mean and the population mean (Examples 7 and 8).

Certainly, the model we have offered is far from the last word, but it is a quite a flexible one, as we have illustrated via a number of examples.

Appendix A

Lemma 1 Under Assumption 1, the function $g(x, y, t)$ is continuous in x.

Proof Given the continuity assumptions on both u and p, it is clear that we need worry only about continuity at $x = y$. Now, Eq. 1 tells us that $g(x, y, t)$ is a weighted average of $u(x, t)$ and $u(y, t)$. Thus, $g(x, y, t)$ must approach $u(y, t)$ whenever $u(x, t)$ does so, i.e., whenever x approaches y. But also from Eq. 1,

$$g(y, y, t) = u(y, t). \tag{17}$$

Thus, $g(x, y, t) \to u(y, t)$ as $x \to y$, and g is continuous in x.

Lemma 2 Suppose that $y > t$. Then, under Assumption 1, the function $g(x, y, t)$ will achieve its maximum (in x, for fixed y and t) at some $x^* \in [t,y)$.

[12] Feld and Grofman 1991 is a partial exception; they model the consequences of incumbency advantage but do not treat incumbency advantage as endogenous.

[13] In multidimensional voting games, no core is expected McKelvey 1976; in the absence of a core position at which to locate, the incumbent, whose position is presumably more or less frozen, can be defeated by a challenger who offers a policy platform that is located in the incumbent's win-set—at least, if we neglect the complication of partisan biases or incumbency benefit of the doubt affecting voter choices Adams et al. 2006, Feld and Grofman 1991.

Proof Let us rewrite Eq. 1 as

$$g(x, y, t) = (u(x, t) - u(y, t))p(x, y) + u(y, t) \tag{18}$$

which gives us, also,

$$g(t, y, t) = (u(t, t) - u(y, t))p(t, y) + u(y, t)$$

If, now, $x < t$, we have (by Assumptions 1)

$$(u(t, t) - u(y, t))p(t, y) > (u(t, t) - u(y, t))p(x, y) > (u(x, t) - u(y, t))p(x, y)$$

and so

$$g(t, y, t) > g(x, y, t).$$

Thus, g cannot be maximized by any $x < t$.

Note next, as in the proof of Lemma 1, that $g(x, y, t)$ is a weighted average of $u(x, t)$ and $u(y, t)$. Thus, in particular, $g(t, y, t)$ is a weighted average of $u(t, t)$ and $u(y, t)$. Moreover, $g(y, y, t) = u(y, t)$. As $u(t, t) > u(y, t)$, we will have

$$g(t, y, t) > u(y, t),$$

and we see that the maximum cannot be at $x = y$.

Furthermore, note that, for $x \geq y > t$, we will have $u(x, t) \leq u(y, t)$. Thus, $g(x, y, t)$, as a weighted average of $u(x, t)$ and $u(y, t)$, cannot be greater than $u(y, t)$ and is therefore, smaller than $g(t, y, t)$. Thus, the maximum cannot be at any $x \geq y$. Together with the fact that g is continuous in x, this guarantees a maximum at some x^* in the interval (t, y).

Theorem 1 *Assuming $\alpha$and σfixed, then, for a given position y of the incumbent, there exists a position x*(y) such that the utility-maximizing D voter located at t will prefer the D candidate to be in the position max{t, x*(y)}.*

Proof We have (as given above)

$$g(x, y, t) = e^{-\alpha|x-t|}\, \Phi\left(\frac{x + y - 2m}{2\sigma}\right) + e^{-\alpha|y-t|}\, \Phi\left(\frac{2m - x - y}{2\sigma}\right).$$

To maximize this, we recall that, by Lemma 2, (assuming $t < y$), the maximizing x will be some $x^* \in (t, y)$. With this assumption, $|x - t| = x - t$ and $|y - t| = y - t$. Then, differentiation of g gives us

$$\begin{aligned} \frac{\partial g}{\partial x} &= -\alpha e^{-\alpha(x-t)}\, \Phi\left(\frac{x + y - 2m}{2\sigma}\right) \\ &\quad + \frac{1}{2\sigma}\left(e^{-\alpha(x-t)} - e^{-\alpha(y-t)}\right)\varphi\left(\frac{x + y - 2m}{2\sigma}\right) \end{aligned} \tag{19}$$

Setting this derivative equal to zero, we obtain

$$2\alpha\sigma\Phi\left(\frac{x+y-2m}{2\sigma}\right)=\left(1-e^{\alpha(x-y)}\right)\varphi\left(\frac{x+y-2m}{2\sigma}\right) \tag{20}$$

It should be noted that t does not appear in the expression Eq. 20. Thus, it is possible (at least numerically) to solve Eq. 20 for x as a function of y (α and σ being fixed parameters). It may be proved that this x is unique. Let, then, $x^*(y)$ be the solution of Eq. 20. If this $x^* \geq t$, then we find that g increases as x increases from t to x^*, and then decreases from x^* to y. Thus, this $x^*(y)$ will maximize g.

Of course, it may turn out that this expression (Eq. 20) yields $x^* < t$. If so, then the expression for $\partial g/\partial x$ is incorrect (some signs are wrong) and we find that g decreases as x increases from t to y. Then the maximizing value for x will actually be at t. (This happens if t is sufficiently close to y.)

We conclude that *the utility-maximizing D voter located at t will prefer the D candidate to be in the position* $\max\{t, x^*(y)\}$.

Theorem 2 There will be a centralizing tendency (and the challenger will probably win) if y is greater than the smaller one of the quantities

$$2m-d \quad \text{and} \quad m-\frac{\log\left(1-\alpha\sigma\sqrt{2\pi}\right)}{2\alpha}. \tag{21}$$

There will be a polarizing tendency if y is smaller than both of these two quantities.

Proof Assume, first of all, that $y > 2m - d$. We know that (by definition) $g^* \geq d$. Thus,

$$m-q^*\leq m-d=(2m-d)-m<y-m$$

and there is a centralizing tendency.

Suppose, next, that $y < 2m - d$. To better understand the situation here, let us consider the equation

$$\Phi(\nu)=w\varphi(\nu) \tag{22}$$

where Φ and ϕ are the normal distribution and density function (defined above), respectively.

If we restrict this to positive values of w, some analysis will show that Eq. 22 gives ν as a monotone increasing function of w, with $\nu \to -\infty$ as $w \to 0$; $\nu = 0$ when $w=\sqrt{\pi/2}$. As $w \to +\infty$, ν will behave asymptotically as $\sqrt{2\log w}$.

Now, we can see that Eq. 22 is equivalent to Eq. 20 if we set

$$\nu = \frac{x^* + y - 2m}{2\sigma} \tag{23}$$

and

$$w = \frac{1 - e^{-\alpha\left(x^* - y\right)}}{2\alpha\sigma} \tag{24}$$

and of course Eq. 23 can be rewritten as

$$m - x^*(y) = y - m - 2\nu\sigma \tag{25}$$

We see that, if $\nu < 0$, then $m - x^*(y) > y - m$; i.e., there is a polarizing tendency, and the D party will nominate someone who is farther from the median m than the R incumbent. If $\nu > 0$, then there is a centralizing tendency. Now the condition for $\nu < 0$ is that $w < \sqrt{\pi/2}$. Unfortunately, it is not trivial to determine w as (by Eq. 24) it also depends on x^*. However, some slight analysis will show that (due to the monotonicity of ν in w), the sign of ν will be unchanged if we replace x^* in Eq. 25 by $2m - y$ (which would correspond to $m - x^* = y - m$). If so, then we can replace w by

$$w = \frac{1 - e^{2\alpha(m-y)}}{2\alpha\sigma} \tag{26}$$

And we conclude that there is a polarizing tendency if $w < \sqrt{\pi/2}$. We note that this condition can be rewritten as

$$y < m - \frac{\log\left(1 - \alpha\sigma\sqrt{2\pi}\right)}{2\alpha} \tag{27}$$

and (still assuming that $y < 2m - d$) we conclude that *the condition for a polarizing tendency is that Eq. 27 hold. There will be a centralizing tendency otherwise.* (Note that the logarithm in this expression, if it exists at all, is negative. If the logarithm does not exist, then Eq. 27 is assumed to hold.) Note moreover that, *when there is a polarizing tendency, then the incumbent will probably win; with a centralizing tendency, the challenger will probably win.*

Theorem 3 If

$$\alpha\sigma\sqrt{2\pi} \le \min\left\{1 - e^{\alpha(m-r)}, 1 - e^{\alpha(d-m)}\right\}, \tag{28}$$

then

$$x\# = m - \mu, \quad y\# = m + \mu \tag{29}$$

where

$$\mu = -\frac{\log\left(1 - \alpha\sigma\sqrt{2\pi}\right)}{2\alpha}. \tag{30}$$

If, on the other hand,

$$\alpha\sigma\sqrt{2\pi}\max\left\{1 - e^{\alpha(m-r)}, 1 - e^{\alpha(d-m)}\right\}, \tag{31}$$

then $x\# = d$ and $y\# = r$.

Proof We look here for equilibrium positions, ($x\#$, $y\#$), satisfying

$$x\# = \max\{d, x^*(y\#)\} \tag{32}$$

$$y\# = \min\{r, y^*(x\#)\} \tag{33}$$

To solve this, let us assume, first, that in both Eqs. 32 and 33, the second alternative holds: $x\# = x^*(y\#)$ and $y\# = y^*(x\#)$. Then Eq. 25 must hold, and by symmetry,

$$m - y^*(x) = x - m - 2\nu\sigma \tag{34}$$

must also hold. This means that $\nu = 0$. But then $w = \sqrt{\pi/2}$, and moreover, $y\# - m = m - x\#$. Letting the common value of these two quantities be μ, we have (from Eq. 30)

$$1 - e^{-2\alpha\mu} = \alpha\sigma\sqrt{2\pi} \tag{35}$$

and then

$$x\# = m - \mu, \quad y\# = m + \mu$$

where μ is as in Eq. 30.

Two things could conceivably go wrong with Eqs. 29, 30. For one thing, we might be taking the logarithm of a negative number or 0; for another, the μ obtained might be too large, so that $m - \mu < d$, or $m + \mu > r$. Some analysis shows that, to avoid this, we must have

$$\alpha\sigma\sqrt{2\pi} \leq \min\left\{1 - e^{\alpha(m-r)}, 1 - e^{\alpha(d-m)}\right\}, \tag{36}$$

Assuming, on the other hand, that

$$\alpha\sigma\sqrt{2\pi} \geq \max\left\{1 - e^{\alpha(m-r)}, 1 - e^{\alpha(d-m)}\right\}, \tag{37}$$

then we will have both $m - \mu < d$, and $m + \mu > r$. In this case, we see (by monotonicity of ν) that

$$y^*(d) > y^*(m - \mu) = m + \mu > r$$

and so

$$\min \{y^*(d), r\} = r. \tag{38}$$

By a similar argument,

$$\max \{x^*(r), d\} = d, \tag{39}$$

and so $x\# = d$ and $y\# = r$.

There is, finally, the possibility that $\alpha\sigma\sqrt{2\pi}$ lies between the min and the max of the two numbers in Eqs. 28 and 31. In that case, one of $x\#$ and $y\#$ is at the median point (d or r) while the other is given by Eq. 29. This can only hold if the situation is quite asymmetric. We will omit the details of this last case.

Theorem 4 Assuming the R party primary is held first, the tendency will be for R to nominate a candidate at the same position $y\#$ as in case of simultaneous primaries. If this happens, then the D party tendency will be to nominate a candidate at the same $x\#$ as in the case of simultaneous primaries.

Proof As mentioned above, we must here consider the decision to be made by an R voter. Whatever the position y of the R nominee, he must expect that the D party will nominate someone at $q^*(y)$. Then, for a voter at position t, his expected utility is given by

$$h(y, t) = g(q^*(y), y, t) \tag{40}$$

and thus he will look for the y that maximizes $h(y, t)$.

Differentiating, we obtain

$$\frac{\partial h}{\partial y} = \frac{\partial g}{\partial y} + \frac{\partial g}{\partial x}\left(\frac{dq^*}{dy}\right) \tag{41}$$

Suppose first that $\alpha\sigma$ is small, so that Eq. 28 holds. Then $x\#$, $y\#$ are given by Eqs. 29, 30, and $x\# = x^*(y\#) > d$. Substituting $y = y\#$ and $x = x\#$ in Eq. 41, we see that both $\partial g/\partial y$ and $\partial g/\partial x$ equals 0 and so $\partial h/\partial y = 0$.

Suppose, on the other hand, that $\alpha\sigma$ is large, so that $y\# = r$, $x\# = d$. In this case, $dq^*/dy = 0$ throughout some neighborhood of $y\#$, and so $\partial h/\partial y = \partial g/\partial y$. Thus the same y that maximizes g also maximizes h. But this is precisely $y\#$.

Appendix B. On the problem of strategic primary voting

We are interested here in the problem of voter behavior in primaries. The assumption is that there is an R incumbent at position y. Assume further that, in the D party primary, there is a candidate at $q^*(y)$, and another at a different position l. Under what circumstances will there be a majority for the candidate at l?

As mentioned above, $q^* = \max\{d, x^*(y)\}$. We will therefore consider the two cases (a) $q^* = d$ and (b) $q^* = x^*$.

(a) Suppose, first, that $q^* = x^* \geq d$. As discussed above, all D voters with $t \leq x^*$ will have $x^*(y)$ as their preferred primary position. But as the party median $d \leq x^*$, this means that at least half the D voters will be located at some $t \leq x^*$. It follows that in such a case a candidate at x^* will defeat any other D candidate. Thus, in this case, strategic primary voting will have no effect on the outcome.

(b) Consider, then, the case $q^* = d > x^*(y)$. Suppose there is a D candidate at d, and another at l. There are two sub-cases, depending on whether l is (1) greater than or (2) smaller than d.

(1) We first look at the sub-case $l > d$. As in the discussion of the proof of Theorem 3, above, we note that, for all $t < x^*$, g increases as x goes from t to x^*, and then decreases from x^* to y. Thus all the voters at these positions $t < x^*$ prefer d to l. Similarly, for $x^* \leq t \leq d$, g decreases as x goes from t to y. Thus, all of these voters will also prefer d to l. Thus, at least half the voters prefer d to l, and thus, d will defeat l.

(2) We are left with the sub-case $l < d$. It is in this case that strategic primary voting can be effective. In fact, suppose that $x^* \leq l < d$. As in sub-case (1), we find that for voters located at some $t \leq l$, g decreases as x goes from t to y, Thus, all such voters will prefer l to d. Moreover, some of the voters slightly to the right of s will also prefer *s to d*. Thus, l will be preferred to d by all voters at $t < h$, where h is somewhere between l and d, closer to l than to d. Of course, d is the median of the D party, so the voters to the left of h cannot be a majority. On the other hand, D voters who are close to y might also prefer l to d. A necessary (but not sufficient) condition is that $2t > y + l$.

Thus, in this sub-case, there may be a majority (in the D primary) for the candidate at l. This will be a heterogeneous majority, consisting on the one hand of voters on the left, who would like l to win the general election, and on the other of voters on the right, who merely want l to win the primary so as to guarantee y's victory in the general election. It is not however clear that this type of coalition can last for more than one or two elections: after a while the two sides of the coalition will realize that each is using the other for its own purposes. If the candidate at l goes on to win the general election, the right part of the coalition will be horrified at what it's done, and may well decide to leave the D party for R. If, as is more likely, the candidate at l loses, it is the left wing that will be upset (will in fact complain about the right wingers' cynicism) and refuse to deal with the right wing again.

In general then, we suggest that strategic primary voting, though it may occur once in a while, cannot be viable as a recurring phenomenon.

References

Adams J (1999) Multicandidate spatial competition with probabilistic voting. Public Choice 99:259–274

Adams J, Merrill S, Grofman B (2006) A unified theory of party competition: party strategies and policy representation in France, Britain, Norway and the United States. Cambridge University Press, New York

Aldrich JH (1983) A Downsian spatial model with party activism. Am Polit Sci Rev 77:974–990

Alesina A, Rosenthal H (1995) Partisan politics, divided government and the economy. Cambridge University Press, New York
Aranson P, Ordeshook PC (1972) Spatial strategy for sequential elections. In: Niemi RG, Weisberg H (eds) Probability models of collective decision making. Charles E. Merrill, Columbus, OH, pp 298–331
Baron DD (1994) Electoral competition with informed and uninformed voters. Am Polit Sci Rev 88(1):33–47
Bernhardt MD, Ingberman DE (1985) Candidate reputations and the incumbency effect. J Public Econ 27:47–67
Black J (1978) The multicandidate calculus of voting: application to Canadian Federal Elections. Am J Polit Sci 22:609–638, August
Brams SJ (1978) Presidential election game. Yale University Press, New Haven, CT
Brams SJ (1980) Spatial models of election competition. University Modules in Applied Mathematics
Brams EJ, Merrill S (1991) Final-offer arbitration with a bonus. Eur J Polit Econ 7:79–82
Brennan G, Lomasky W (1993) Democracy and decision: the pure theory of electoral preference. Cambridge University Press, New York
Budge I, David R, Derek H (eds) (1987) Ideology, strategy and party change: spatial analyses of post-war election programmes in 19 democracies. Cambridge University Press, Cambridge, MA
Bullock C, Brady DW (l983) Party, constituency, and roll-call voting in the U.S. Senate. Legis Stud Q 8:29–43
Cain B (1978) Strategic voting in Great Britain. Am J Polit Sci 22(3):639–655
Coleman JS (1971) Internal processes governing party positions in elections. Public Choice 11:35–60, Fall
Coleman JS (1972) The positions of political parties in elections. In: Niemi RG, Weisberg F (eds) Probability models of collective decision making. Charles E. Merrill, Columbus, OH, pp 332–357
Cooper A, Munger MC (2000) The (un)predictability of presidential primaries with many candidates. Public Choice 103(3-4):337–355
Cox G (1997) Making votes count. Cambridge University Press, New York
Downs A (1957) An economic theory of democracy. Harper, New York
Duverger M (1958) Political parties, their organization and activity in the modern state (B. North and R. North, Trans.) Wiley, New York, Methuen, London
Enelow JM, Hinich MJ (1984) The spatial theory of political competition: an introduction. Cambridge University Press, New York
Enelow JM, Hinich MJ (1990) Advances in the spatial theory of voting. Cambridge University Press, New York
Fedderson TJ (1992) A voting model implying Duverger's law and positive turnout. Am J Polit Sci 36:938–962
Feld SL, Grofman B (1991) Incumbency advantage, voter loyalty and the benefit of the doubt. J Theor Polit 3(2):115–137
Fiorina M (1974) Representatives, roll calls and constituencies. Lexington Books, Lexington, MA
Glazer A (1993) Political equilibrium under group identification. In: Grofman B (ed) Information, participation and choice: an economic theory of democracy in perspective. University of Michigan, Ann Arbor, MI, pp 81–92
Glazer A, Grofman B, Owen G (1998) A neo-Downsian model of group-oriented voting and racial backlash. Public Choice 97:23–34
Grofman B (1985) Neglected role of the status quo in models of issue voting. J Polit 47:231–237
Grofman B (1993) Toward an institution-rich theory of political competition, with a supply-side component. In: Grofman B (ed) Information, participation and choice: 'an economic theory of democracy' in perspective. University of Michigan, Ann Arbor, MI, pp 179–193
Grofman B (1996) Downsian political economy. In: Goodin R, Klingemann, HD (eds) New handbook of political science. Oxford University Press, London
Grofman B (2004) Downs and two-party convergence. Annu Rev Pol Sci 7:25–46
Grofman B, Griffin R, Berry R (1995) House members who become senators: learning from a 'natural experiment' in representation. Legis Stud Q 20(4):513–529

Grofman B, Griffin R, Glazer A (1990) Identical geography, different constituencies, see what a difference party makes. In: Johnston RJ, Shelley F, Taylor P (eds) Developments in electoral geography. Croom Helm, London, pp 207–217
Grofman B, Lijphart A (eds) (1986) Electoral laws and their political consequences. Agathon, New York
Lijphart A (1984) Democracies: patterns of majoritarian and consensus government in twenty-one countries. Yale University Press, New Haven CT
Lijphart A (l992) Parliamentary vs. presidential government. Oxford University Press, New York
Lijphart A, Grofman B (eds) (1984) Choosing an electoral system. Praeger, New York
Macdonald SE, Listhaug O, Rabinowitz G (1991) Issues and party support in multiparty systems. Am Polit Sci Rev 85:1107–1131
McKelvey RD (1976) Intransitivities in multidimensional voting models and some implications for agenda control. J Econ Theory 12:472–482
Merrill S (1993) Voting behavior under the directional spatial model of electoral competition. Public Choice 77:739–756
Merrill S, Grofman B (1997) Directional and proximity models of spatial two-party competition: a new synthesis. J Theor Polit 9(1):25–48
Merrill S, Grofman B (1999) A unified theory of voter choice. Cambridge University Press, New York
Merrill S, Grofman B, Brunell T, Koetzle W (1999) The power of ideologically concentrated minorities. J Theor Polit 11(1):57–74
Miller N (1996) Majority rule and minority interests. In: Shapiro I, Hardin R (eds) Political order (Nomos XXXVIII). New York University Press, New York, pp 207–250
Morton R (1993) Incomplete information and ideological explanations of platform divergence. Am Polit Sci Rev 87(2):382–392
Palfrey T, Erikson R (1994) Preemptive spending and entry deterrence by incumbents in congressional elections. Social science working Paper. California Institute of Technology, March
Poole KT, Rosenthal H (1984) The polarization of American politics. J Polit 46:1061–1079
Rabinowitz G, Macdonald SE (l989) A directional theory of issue voting. Am Polit Sci Rev 83:93–121
Riker WH (1982) The two-party system and Duverger's Law: an essay on the history of political science. Am Polit Sci Rev 6(4):753–768
Robertson D (1987) Britain, Australia, New Zealand and the United States, 1946-1981: an initial comparative analysis. In: Budge I, Robertson D, Hearl D (eds) Ideology, strategy and party change: spatial analysis of post-war elections programmes in nineteen democracies. Cambridge University Press, New York, pp 39–72
Schofield N (l996) The heart of a polity. In: Schofield N (ed) Collective decision-making: social choice and political economy. Kluwer-Nijhoff, Boston, pp 183–220
Shapiro CR, Brady DW, Brody JA, Brody RA, Ferejohn JA (1990) Linking constituency opinion and senate voting scores: a hybrid explanation. Legis Stud Q 15(4):599–623
Shepsle K (1979) Institutional arrangements and equilibrium in multidimensional voting models. Am J Polit Sci 23:27–59
Shepsle KA, Weingast BR (1984) Uncovered sets and sophisticated voting outcomes with implications for agenda institutions. Am J Polit Sci 28(1):49–74
Wittman D (1973) Parties as utility maximizers. Am Polit Sci Rev 18:490–498
Wittman D (1977) Candidates with policy preferences: a dynamic model. J Econ Theory 14: 180–189
Wittman D (1983) Candidate motivation: a synthesis of alternative theories. Am Polit Sci Rev 72:78–90