Candidate and Party Strategies in Two-Stage Elections Beginning with a Primary

James Adams  
Professor  
Department of Political Science  
University of California at Davis  
One Shields Avenue  
Davis, CA 95616  
jfadams@ucdavis.edu

Samuel Merrill, III  
Professor Emeritus  
Department of Mathematics and Computer Science  
Wilkes University  
Present address:  
3024 43rd Ct. NW  
Olympia, WA 98502  
smerrill@zhonka.net

A previous version of this paper was presented at the Annual Meeting of the Midwest Political Science Association, Chicago, IL, April 20-23, 2006. We thank Ken Greene, Becky Morton, Hans Noel, Brian Sala, Gilles Serra, Walt Stone, John Zaller, Elizabeth Zechmeister, and three anonymous referees for helpful suggestions. All remaining errors are the responsibility of the authors.
In the United States and Latin America, candidates for national and state-level office frequently must win primary elections in order to advance to the general election. We model policy and valence issues for office-seeking candidates facing such two-stage elections. We determine a Nash equilibrium for the candidates’ optimal strategies, and we find that holding a primary is likely to increase a party’s chances of winning the general election, particularly in situations where valence issues that involve the candidates’ campaigning skills and that are not known prior to the campaign are more salient than policy issues. Furthermore, we find that primary elections are especially likely to benefit parties that expect to be underdogs in the general election. Our conclusions are directly relevant to U.S. politics and by extension to the strategic decisions that many Latin American parties currently confront, about whether it is strategically desirable to hold primaries.

Keywords; Nash equilibrium; primaries; spatial modeling; campaign skills
Most Republican and Democratic nominees for national and state-level office in the United States are selected via primary elections, in which the party’s partisans choose between candidates vying for the party’s nomination. Primaries are also increasingly prominent features of elections held outside the United States, since many political parties in Latin America have recently adopted primaries; furthermore, many additional parties are currently considering switching to (or, in some cases, away from) primaries, so that the question of whether this is the preferred candidate selection mechanism is highly salient to these parties’ elites.¹ This issue has also sparked extensive scholarly debate (see, e.g., Carey, 2003; Carey and Polga-Hecimovich, 2006, forthcoming; Colomer, 2003; De Luca et al., 2002). Yet despite the widespread use of primary elections, there is limited theoretical work analyzing the policy strategies of office-seeking candidates who confront a two-stage election process whose first stage is a party primary.²

The major theoretical prediction on primaries is that the need to compete in a partisan primary may pull office-seeking candidates away from a centrist policy strategy that benefits them in the general election (see, e.g., Aranson and Ordeshook, 1972; Burden, 2000; Owen and Grofman, 2006) – a prediction that has been empirically confirmed by Burden (2001) with re-

¹ Carey and Polga-Hecimovich (2006) report that there are 16 Latin American countries in which at least one political party selects its presidential candidate via a primary. In addition, primaries have been used to select candidates in Bulgaria, South Korea, Iceland, Spain, France, Israel, and Mexico (Serra, 2006b).
² Kanthak and Morton (2003) propose a model that incorporates candidate entry and a cost-and-benefits analysis by candidates that motivate them to promote turnout. Using data from U.S. Congressional elections in the 1980s, their model predicts that voting in a primary significantly increases the total vote in the general election and the vote share in the general election for the party holding the primary.
spect to American elections. The spatial logic underlying this prediction is summarized by Carey and Polga-Hecimovich (2006):

…this argument has become conventional journalistic wisdom regarding U.S. primaries: Primaries attract hardcore partisans, who tend to come from the ideological extremes of the two big parties, and these voters in turn choose candidates of limited appeal to the middle-of-the-road voters who dominate the general electorate (2006, page 530).

The empirical and theoretical literature on primaries raises several questions: How do office-seeking candidates balance the centrifugal incentive to appeal to their primary electorates against the centripetal motivation to court the general electorate, if the candidates must present the same policies at both stages of the election process? Furthermore, do circumstances exist in which holding a primary actually enhances a party’s prospects in the general election, and if so what are these circumstances?3

To address the above issues we analyze the strategic policy incentives that primary elections create for office-seeking candidates. We are unaware of any theoretical work on primary elections that analyzes situations where candidates simultaneously consider the electoral implications of policy dimensions and of non-policy-related “valence” factors. Valence factors include the candidates’ personal images along such dimensions as charisma, competence, empathy, integrity, as well as campaigning skill – factors that are widely believed to influence election outcomes (Grose et al., 2004; Grose, 2005; Mondak, 1995; Serra, 2006a; Stokes, 1963). That is

3 Serra (2006a) raises a parallel puzzle: why do party leaders (as in Latin America) adopt primary elections instead of handpicking candidates loyal to them?
what we present here. Specifically, we develop a two-party spatial model of policy and valence issues for office-seeking candidates, who face a two-stage election process that begins with a primary election and who must commit to a single set of policies for both stages. A crucial feature of our model is that the candidates are uncertain, at the time they commit to their policy positions, about how skillfully they (and their opponents) will campaign, and thus about what their non-policy-related valence images will be at the time of the primary and general elections. This uncertainty captures the empirical observation that primary election candidates’ public images often change sharply between the time they announce their candidacy and the date of the primary election, as voters observe how the candidates present themselves and communicate to the electorate. We consider the situation where there are two parties, both of which hold primaries, and also that in which a single party holds a competitive primary. Our study produces three central conclusions.

First, we show that a Nash equilibrium configuration of candidate policy strategies exists, given quite general assumptions about voters.

Second, we show that, relative to a procedure where party elites handpick a nominee – even one located at the median of the general electorate – holding a primary may increase a party’s chance of winning the general election. This is because in many plausible scenarios the strategic advantage arising from the primary electorate’s ability to select a high-quality nominee – i.e. one whose campaigning skills prove attractive to voters (such as Bill Clinton and Ronald Reagan) – outweighs the strategic disadvantage that the primary pulls the party’s nominee away from the center of the general electorate. This result on the “strategic advantage” of primaries is striking, given that our model (intentionally) specifies conditions that appear unpromising for parties holding primaries: namely, that candidates cannot modify their policies between the pri-
primary and the general election; that primary voters are unmoved by the candidates’ projected
general election prospects⁴; that primary voters may hold extremist policy views compared to the
general electorate. Therefore the information-revealing advantage of holding a primary – lead-
ing to a high-quality nominee located between the median positions of the primary and general
electorates – may offer greater utility to party leaders than handpicking an untested nominee.

Our third, related, conclusion is that primary elections are especially beneficial to “un-
derdog” parties, i.e. to parties whose candidates face an uphill battle in the general election, for
example because the rival party’s candidate enjoys incumbency-related advantages. This is be-
cause in this situation the centrifugal pressure that primary elections exert on candidates to di-
verge from the center of the general electorate is weakened, since the underdog party’s candi-
dates project that – in the event they win the primary – they will face a formidable general elec-
tion opponent.

2. Office-Seeking Candidates Confronting a Two-Stage Election Process: A Valence-Uncertainty Model

The voter distribution

We assume that the general election involves two parties, which we label the Democrats and the
Republicans, with voters’ policy preferences distributed along a Left-Right continuum among
citizens who will vote in the general election, labeled the general electorate. This electorate has
a known median voter position \( m_G \), labeled the general election median. We assume that when

⁴ On the web site http://course.wilkes.edu/merrill, we relax this assumption, permitting primary voters to strategically
support the most viable general election candidate. Our analyses suggest that strategic and expressive voting
(for the preferred primary candidate regardless of electability) support identical equilibrium configurations in candidate
strategies, in situations where the candidates’ campaigning skills are sufficiently salient, relative to their policy
strategies.
multiple candidates seek a party’s nomination the party holds a closed primary\(^5\) to select its nominee (we will consider the cases where both parties hold competitive primaries and also where only one party holds a competitive primary), and that the medians of the policy preferences for the respective parties’ primary electorates are known: the Democratic primary electorate’s median voter position is \(m_D\), which we label the Democratic median, and the Republican primary electorate’s median position is \(m_R\). We further assume that the Democratic median is at least as left-wing – and that the Republican median is at least as right-wing – as the general election median, i.e. that \(m_D \leq m_G \leq m_R\).

### The election context

The basic model incorporates the following assumptions:

1) If a party’s primary election is contested, then exactly two candidates compete in the primary. The Democratic candidates in such contested primaries are labeled \(D_1\) and \(D_2\), while the Republican candidates in contested primaries are labeled \(R_1\) and \(R_2\).

2) The candidates select a single policy position that is fixed during the primary and the general election. That is, candidates cannot readjust their policies following the primary in order to enhance their appeal to the general electorate. The Democratic candidates’ positions are labeled \(d_1\) and \(d_2\) for candidates \(D_1\) and \(D_2\), respectively; the Republican candidates’ positions are labeled \(r_1\) and \(r_2\) for candidates \(R_1\) and \(R_2\), respectively.

3) Each candidate selects a single position that maximizes his joint probability of winning the primary and general election, i.e. each candidate selects the position that maximizes his probability of winning office.

---

\(^5\) A closed primary election is one where only officially-registered members of the party are eligible to vote.
Our specification that contested primaries involve exactly two candidates simplifies the analysis and the presentation of our results; on the web site (http://course.wilkes.edu/merrill) we show that our results extend – with stronger assumptions – to multicandidate primary elections. The assumption that candidates cannot shift their positions between the primary and the general election is also a simplification, but one that is intuitively plausible: for while real world candidates do at times attempt to “tack towards the center” following their primary nomination, they are hampered by the policy promises they have made during the primary. Furthermore, candidates’ post-primary attempts to dramatically shift their policy images leave them open to the charge of political opportunism or “flip-flopping,” that can damage their general election prospects.  

The voting model

The voting model – which is similar to the one developed in Adams and Merrill (forthcoming 2008) – specifies that voters evaluate candidates based on their policy positions plus a valence component. Specifically, for each voter $i$ with policy preference $x_i$, the policy distance component of $i$’s evaluation of candidate $j$ is given as $g(j, x_i)$, where $g(j, x_i)$ represents $i$’s policy utility for $j$’s position $j$. Probably the two best known special cases for the policy distance

---

6 Burden (2001, Table 7.1) presents empirical analyses that support our assumption, concluding that incumbent congressmen barely alter their legislative behavior between the primary and the general election, shifting roughly .03 policy units closer to the center of the general electorate along a -1 to +1 scale. If candidates are permitted to shift their post-primary positions for the general election, we find that the optimal strategies may shift a corresponding amount, but that the probability of winning office changes only slightly, and that our central substantive conclusions – namely that an equilibrium in candidate strategies will exist, and that “weak” parties benefit disproportionately from holding primaries – extends to this more general model.
component are those in which voters have either linear policy losses, i.e. where policy utilities are given as \( g(j, x_i) = -a|j - x_i| \), or quadratic policy losses for which policy utilities are given as \( g(j, x_i) = -a(j - x_i)^2 \), where in either case \( a \) is a strictly positive parameter that represents the salience of the policy dimension relative to the valence dimension. There is no abstention.

The valence component of voters’ candidate evaluations captures dimensions “on which parties or leaders are differentiated not by what they advocate, but by the degree to which they are linked in the public’s mind with conditions, goals, or symbols which almost everyone approves or disapproves” (Stokes, 1992: 143), such as competence, integrity, leadership ability, and charisma. We assume that the valence component of voter \( i \)’s evaluation of candidate \( j \) is identical for all voters and has two subcomponents: the candidate’s measured valence characteristics, \( V_j \), which the candidates know at the time they select their policy strategies, and which we label the candidate’s pre-campaign valence image; and unmeasured characteristics \( \epsilon_j \) that capture the candidates’ campaigning abilities, which determine how skillfully the candidates present themselves to the public during the election campaign, and which the candidates do not know at the time they choose their policy positions.\(^7\) These unmeasured characteristics are labeled campaign valence images. Thus we have:

\[
\text{Voter } i \text{'s utility for candidate } j = g(j, x_i) + V_j + \epsilon_j .
\]

We assume that the unmeasured component \( \epsilon_j \) for each candidate is identical across voters, and that for each candidate \( j \), \( \epsilon_j \) is identical for the primary and the general election, i.e. the candi-

\(^7\) In this study we assume that the known valence component \( V \) is identical in the primary and the general election even though the constituencies are not the same.
dates reveal their campaigning skills during the primary campaign, and that the \( e_j \)'s are independent over candidates.\(^8\)

Our distinction between the measured and unmeasured components of valence plausibly captures the information environment candidates confront. At the time the candidates commit to their policies they are likely to have formed general impressions about their own – and their opponents’ – valence images, based upon the candidates’ public reputations for competence, integrity, and leadership ability (the candidates form such estimates via their contacts with constituents, public opinion polls and private polling, by studying media coverage, and so on). Furthermore, candidates possess additional information about each competitor’s fund-raising success, the qualities of their campaign organizations and (in some cases) their past performances as campaigners, factors that are crucial for projecting how effectively each candidate will campaign in the current election. Such information forms the basis for the candidate’s pre-campaign valence image \( V_j \). However, one factor that cannot be easily projected in advance is how skillfully each candidate will prove to be at campaigning – a factor that may be crucial to the public’s

\(^8\) We note that Serra (2006a) presents theoretical results on primaries with valence issues where, as in our model, a key assumption is uncertainty about the campaign skills of potential primary contenders; Serra, however, focuses entirely on party elites’ decisions about whether to hold a primary, not on the primary candidates’ policy strategies (which are fixed). In a separate paper, Serra (2006b) determines equilibrium policy strategies for primary candidates, but in a model different from ours – one that assumes polarization of parties and uncertainty about voter preferences but does not employ our key assumption of uncertainty about the quality of potential contenders’ campaigns. Jackson, Mathevet, and Mattes (2007) determine an equilibrium defined by specifying that in each party the party’s nominee cannot be beaten in a head-to-head vote with some other potential nominee, given the other party’s nomination. In contrast to our assumptions, they do not permit candidates to move their policy positions nor do they assume uncertainty about campaign quality.
evaluation of the candidate’s valence characteristics. At the level of presidential primaries, for instance, candidates such as Howard Dean, John Glenn, Edmund Muskie, and Ted Kennedy were widely projected – in advance of the campaign – to be formidable candidates, yet none of them lived up to their press clippings, i.e. each ran an unexpectedly poor campaign and proved less attractive to voters than most analysts had projected. By contrast candidates such as Bill Clinton and Ronald Reagan proved more adept campaigners than many observers anticipated. Note that in these cases the unexpected success or failure of the candidate turned primarily on factors relating to the candidate’s campaigning skills – which in turn influenced perceptions of his charisma, leadership ability, and competence – rather than on the candidate’s policy positions. In addition, we emphasize that the uncertainty that candidates experience with respect to their relative campaigning skills applies even in situations where the candidates are well-known to each other; indeed, most of the presidential candidates mentioned above had well-established reputations as campaigners (in some cases these reputations were positive, in other cases negative), yet in all of these examples the candidate’s previous “track record” proved a poor indicator of his ability to present himself as an attractive candidate in the current campaign. The unmeasured valence component \( \varepsilon_j \) captures this uncertainty.

In addition, we assume that all primary voters vote sincerely, a strong assumption that is contradicted by empirical studies suggesting that some primary voters and party activists consider the candidates’ prospective appeal in the general election (Abramowitz, 1989; Stone and

---

9 We note that Cohen, Karol, Noel, and Zaller (2001) advance an alternative interpretation, namely that candidate success in U.S. presidential primaries (during the 1980-2000 period) was largely determined by elites’ pre-primary political endorsements, an argument suggesting that party elites play a dominant role in choosing candidates even when the selection mechanism is a primary. We do not consider this complication here, because our interest is in analyzing candidate strategies in genuinely competitive primaries.
Abramowitz, 1983). As indicated above, we demonstrate elsewhere (see the reference in footnote 4) that our central conclusions extend to such strategic primary voting scenarios, provided that the candidates’ campaigning skills are sufficiently salient to the voters, relative to the candidates’ policy strategies. Here we simply note that if political parties benefit from holding primaries even in the extreme case where all primary voters ignore the candidates’ appeal to the general electorate, then this conclusion surely applies to situations where primary voters strategically account for this factor.

Given the assumptions of sincere voting in the primary and full voter turnout, the following remarks immediately follow:

**Remark 1.** In a two-candidate primary election, the winning candidate is the one who is preferred by the party’s median primary election voter.

**Remark 2.** In the two-candidate general election, the winning candidate is the one who is preferred by the median general election voter.

Finally, we assume that for each candidate $j$ the unknown component $\varepsilon_j$ that represents the candidate’s campaign valence image, is selected independently over candidates from a type 1 extreme value distribution. This assumption, which has been employed extensively both in empirical studies of voting behavior (Schofield and Sened, 2005, 2006; Whitten and Palmer, 1996) and in spatial models of elections (Merrill and Adams, 2001; Schofield and Sened, 2005, 2006), implies that voters’ choice probabilities can be represented via a *logit* function.\(^{10}\)

\(^{10}\) Specifically, in the situation where, for instance, the candidates $D_1$ and $D_2$ contest the Democratic primary, the probability $PD(D_1; d_1, d_2)$ that the median Democratic primary voter votes for candidate $D_1$ when candidates...
3. Candidate Strategies and Office-Seeking Equilibrium: Illustrative Examples

We now analyze candidates’ policy strategies in two-stage elections. We consider the questions: How do office-seeking candidates balance the conflicting strategic imperatives of appealing to their primary and their general election constituencies?, and, How does holding a competitive primary election affect the party nominee’s general election prospects?

Candidate strategies when only one party holds a competitive primary election

We begin with the case where only one party holds a competitive primary election. This scenario occurs in American elections when a candidate from one of the two parties runs unopposed (or faces only token opposition) in his party’s primary – a common situation for incumbent members of Congress. This scenario simplifies the primary candidates’ strategic calculations, since they know their general election opponent’s identity in advance.\(^\text{11}\) In the Appendix we show that when the candidates \(D_1\) and \(D_2\) contest the Democratic primary while \(R_1\) runs un-

\[
D_1 \text{ and } D_2 \text{ locate at } d_1 \text{ and } d_2, \text{ respectively} \quad \text{– which by Remark 1 is } D_1 \text{’s probability of winning the primary} \quad \text{is given by the logit probability function}
\]

\[
PD(D_1; d_1, d_2) = \frac{\exp(g(d_1, m_D) + V_{D1})}{\sum_{j=1}^2 \exp(g(d_j, m_D) + V_{Dj})} = \frac{1}{1 + \exp[(g(d_2, m_D) + V_{D2}) - (g(d_1, m_D) + V_{D1})]}
\]

See Train (2003, Chapter 3) for a proof that the logit model implies choice probabilities of the functional form given by the above equation.

\(^\text{11}\) Calculation, however, of a candidate’s probabilities of winning the primary and the general election are complicated because they are not independent of each other, intuitively, because a candidate’s victory in a competitive primary provides information about his campaigning abilities (i.e. about his campaign valence image \(\epsilon_j\)). The probabilities \(P(A)\) and \(P(B)\) of two events A and B are independent if the probability \(P(A&B)\), that both A and B occur, is equal to \(P(A)\times P(B)\).
opposed in the Republican primary, the probability $P(D_1) = P(D_1; d_1, d_2, r_1)$ that candidate $D_1$ wins both the primary and the general election while $D_1, D_2,$ and $R_1$ adopt the strategies $d_1, d_2$ and $r_1$, respectively – i.e. the probability that $D_1$ wins office – is

$$P(D_1) = \frac{1}{1 + \exp[(g(d_1, m_D) + V_{d_1}) - (g(d_1, m_D) + V_{d_1})] + \exp[(g(r_1, m_G) + V_{r_1}) - (g(d_1, m_G) + V_{d_1})]} . \quad (2)$$

and the probability $P(D_2) = P(D_2; d_1, d_2, r_1)$ that candidate $D_2$ wins office is similarly given as

$$P(D_2) = \frac{1}{1 + \exp[(g(d_1, m_D) + V_{d_1}) - (g(d_2, m_D) + V_{d_2})] + \exp[(g(r_1, m_G) + V_{r_1}) - (g(d_2, m_G) + V_{d_2})]} . \quad (3)$$

We now present illustrative examples that convey intuitions about how competitive primary elections influence office-seeking candidates’ policy strategies, and also about the circumstances in which political parties benefit from competitive primaries. For these examples we assume a voter distribution on a 1-7 left-right scale where the median general election voter is located at $m_G = 4$, while the median Democratic primary voter is located at $m_D = 2$ and the median Republican primary voter at $m_R = 6$. In our initial examples the candidate $R_1$ runs unopposed in the Republican primary, and locates at the position of the median general election voter,

---

These voter locations are intended to roughly capture the NES seven-point Liberal-Conservative scale, for which the midpoint (4) is designated “moderate” and plausibly approximates the median general electorate position $m_G$ in many constituencies. Furthermore the location 2 on the NES scale is designated “liberal” and plausibly approximates the median Democratic primary voter’s position $m_D$, while the location 6 is designated “conservative” and approximates the Republican primary median position $m_R$. 

---

12 These voter locations are intended to roughly capture the NES seven-point Liberal-Conservative scale, for which the midpoint (4) is designated “moderate” and plausibly approximates the median general electorate position $m_G$ in many constituencies. Furthermore the location 2 on the NES scale is designated “liberal” and plausibly approximates the median Democratic primary voter’s position $m_D$, while the location 6 is designated “conservative” and approximates the Republican primary median position $m_R$. 

---

13
i.e. \( r_1 = m_g = 4 \), a position that maximizes her probability of winning office.\(^{13}\) For the illustrations we assume that voters have quadratic policy loss functions and that the policy salience coefficient is \( a = 1 \), i.e. \( g(j, x_j) = -(j - x_j)^2 \), although the equilibrium theory is not limited to this specific case. Finally, we initially assume that the candidates in our examples are “evenly matched” in terms of their pre-campaign valence images \( V \).

The Democratic candidates’ policy strategies when there is a competitive primary. When neither party has a competitive primary, each candidate – by Downsian logic – locates at the position of the overall median voter and each has an equal chance of winning the general election (see Figure 1). We now introduce a second Democratic candidate, \( D_2 \), who competes with \( D_1 \) for the party nomination. Intuitively, we expect the strategic imperative of winning the primary to draw the candidates \( D_1 \) and \( D_2 \) away from the general median position \( m_g = 4 \), and, furthermore, that this centrifugal pressure will depress the Democratic Party’s chances of winning the general election. In fact, only the former intuition is correct. Figure 2 illustrates the resulting equilibrium configuration in candidates’ office-seeking strategies: the Democratic candidates now pair at the position \( d_1^* = d_2^* = 3.26 \), i.e. the candidates’ equilibrium positions are shaded away from the general median \( m_g = 4 \) in the direction of the Democratic median \( m_D = 2 \). Also shown is candidate \( D_1 \)’s probability of winning the primary election and his probability \( P(D_1) \) of winning office (i.e. \( P(D_1) \) represents \( D_1 \)’s probability of winning both the primary and general elections)

\(^{13}\) See the Appendix for a proof that the position that maximizes the Republican candidate's probability of winning the general election under the conditional logit model, regardless of which Democratic candidate wins the primary and what position the Democratic nominee takes, is \( m_g \). Note that \( R_1 \) wins the general election if and only if the median general election voter votes for \( R_1 \).
as a function of his policy strategy, with the candidates $D_2$ and $R_1$ fixed at their equilibrium positions.\(^{14}\) This figure reveals that while $D_1$ maximizes his primary election probability by locating at $m_D=2$, his optimal office-seeking position is $d_1^* = 3.26$, for which $P(D_1)$ is approximately 0.27. Since $D_2$’s equilibrium probability $P(D_2)$ of winning office is also about 0.27, it follows that the probability that one of these two Democratic candidates wins office is approximately 0.54.\(^{15}\) This exceeds the 0.50 probability that the Democrats win office in the event that $D_1$ runs unopposed in the primary and locates at the general election median $m_G=4$.

How, in this example, can the competitive primary pull the Democratic candidates away from the center of the general electorate, yet simultaneously increase the probability that a Democratic candidate wins the general election? The central intuition is that having two Democratic candidates contest the primary increases the probability that the Democrats will nominate a high-quality candidate – i.e. one with high valence attributes. Because the pre-campaign components of valence are identical for both Democratic candidates (as well as for the Republican candidate), the Democrat who demonstrates higher valence in the primary campaign – by campaigning more skillfully – will be nominated, so that in turn this candidate’s overall valence will have a higher expected value than that of the Republican. We label this effect, which we support theoretically below, the competitive primary advantage effect.

\(^{14}\) For quadratic-loss utility, $d_1^*$ (and hence $d_2^*$) is obtained by solving $\frac{d_1^* - m_D}{m_G - d_1^*} = \exp\left[\alpha(d_1^* - m_G)^2 + V_{R1} - V_{D1}\right]$ numerically, from which the probability that $D_1$ wins office can be computed by simplifying equation 2 in the text to obtain the formula $P(D_1) = \frac{m_G - d_1^*}{(m_G - d_1^*) + (m_G - m_D)}$. The probability $PD(D_1; d_1, d_2)$ is given in footnote 10.

\(^{15}\) The probability that at least one of the Democrats wins office is twice the probability that either wins because the events that each win are mutually exclusive.
Primary elections and the strategic advantage of the weaker party. Next, we examine the situation where the Democratic Party is “weak,” in the sense that the Republican candidate $R_1$’s pre-campaign valence image is superior to those of both Democratic candidates. For this example we specify that candidate $R_1$’s pre-campaign valence advantage is $V_{R_1} = (V_{D_1} + 2) = (V_{D_2} + 2)$; substantively, the magnitude of this measured advantage is such that if candidate $D_1$ ran unopposed in the Democratic primary and paired with $R_1$ at the general median position $m_G = 4$, the probability $P(D_1)$ that $D_1$ would defeat $R_1$ in the general election would be only about 0.12.

Figure 3 illustrates the resulting equilibrium configuration in candidates’ office-seeking strategies: the Democratic candidates now pair at $d_1^* = d_2^* = 3.77$, i.e. their positions are shaded only slightly away from the median general electorate position $m_G = 4$. In equilibrium the candidates $D_1$ and $D_2$ each have 0.102 probabilities of winning office, so that the probability that one of these Democratic candidates is elected is approximately 0.20. This 0.20 probability of a Democratic victory exceeds the 0.12 probability that the Democrats would win the election in the event that $D_1$ (or $D_2$) ran unopposed in the Democratic primary and located at the center of the general electorate.

Why, in this example, do $D_1$ and $D_2$ locate closer to the center of the general electorate than they did in the earlier scenario, in which the Democratic candidates’ pre-campaign valence images were equal to that of the Republican? The logic that motivates the Democrats’ policy moderation in this “weaker party” example is that they recognize that the primary election winner will confront a formidable general election opponent, i.e. one who possesses a superior pre-campaign valence image. Since winning office requires winning both the primary and general elections, the stronger the general election opponent the Democratic candidates expect to face,
the less willing they are to present leftist policies that depress their policy appeal to the general electorate. This strategic incentive to moderate benefits the weaker party in that the closer the primary candidates locate to the center, the more competitive the party’s nominee will be in the general election. We label this effect, which we support theoretically below, the \textit{weaker party’s competitive primary advantage}.

\[\text{FIGURES 1-3 ABOUT HERE}\]

\textit{Strategies when both parties hold primaries: primaries again benefit the weaker party}

Next, we consider the situation in which both political parties hold competitive primary elections. In this scenario the candidates’ strategic calculations are complicated by uncertainty over the identity of the rival party’s nominee. In a web supplement posted at (http://course.wilkes.edu/merrill) we present closed-form representations of the candidates’ probabilities of winning office in these situations for which each primary involves two candidates. Here, we present a two-party primary election scenario which illustrates the candidates’ strategic calculations, and which moreover suggests that “weaker” parties are more likely to benefit from competitive primary elections than are strong parties. For this exercise we analyze the scenario discussed in the previous example, except that we now introduce a second Republican candidate, $R_2$, whose measured valence characteristics $V_{R_2}$ match those of candidate $R_1$. This captures situations in which one party is dominant in the election constituency and both parties hold competitive primary elections.

Figure 4 illustrates the resulting equilibrium configuration in candidates’ strategies. Due to the competitive primary election the Republican candidates locate at $r_{1*} = r_{2*} = 5.10$, i.e. the primary pulls them sharply away from the center of the general electorate, towards the Republi-
can primary median position $m_R = 6$; meanwhile the Democratic candidates’ equilibrium positions are relatively moderate at $d_1^* = d_2^* = 3.66$. In equilibrium the Democratic candidates’ probabilities of winning office are 0.144 each, and thus the probability that one of these Democratic candidates is elected is approximately 0.29. Note that this 0.29 probability exceeds the 0.20 probability that the Democrats would win the election in the scenario outlined earlier in Figure 3, in which the Republicans did not hold a primary and candidate $R_1$ located at the general election median position. Thus in this two-primary example the weaker Democratic Party again benefits from holding a primary, while the stronger Republican Party does not.

The strategic logic that drives the stronger party’s primary election candidates – in this case the Republicans – far from the center of the general electorate is that they project that the primary is the highest hurdle to winning office, because they recognize that the rival party’s candidates possess inferior pre-campaign valence images. Such noncentrist positioning entails substantial costs in the general election – costs that are likely to outweigh the valence benefit the primary confers by allowing the strong party’s primary voters to select the superior campaigner. Thus, as each of the strong party’s candidates pursues his/her personal objective of winning, these candidates undermine the likelihood that a member of their party wins office. This is illustrated in Figure 4, where the positioning of the strong Republican Party’s candidates is more radical than is that of the weak Democrats.

4. Office-Seeking Equilibrium in Two-Stage Elections: Theoretical Results

We now demonstrate that the strategic logic suggested by the preceding examples generalizes to a large class of two-stage elections. We begin by stating and proving an existence and
uniqueness theorem for a Nash equilibrium in office-seeking candidates’ policy strategies, for
the situation where one party holds a competitive primary election while a single candidate runs
unopposed in the rival party’s primary. The proof of this theorem, along with those of Theorems
2-4 below, are given in the web supplement (http://course.wilkes.edu/merrill); this web supple-
ment also presents extensions of the theoretical results to scenarios involving multicandidate
primaries, i.e. primary elections that feature at least three candidates.16

Theorem 1. If there are two Democratic candidates $D_1$ and $D_2$ and one Republican candidate $R_1$,
and if the policy distance component of the voter’s utility is concave and peaks at the voter’s
ideal point,17 then

(a) there exists a Nash equilibrium in candidates’ office-seeking strategies, and

(b) if $V_{D_1} = V_{D_2}$, then there is exactly one paired Nash equilibrium, i.e., for which $d_1^* = d_2^*$.

Outline of Proof. We show that $P(D_1; d_1, d_2, r_1)$ (for fixed $d_2$ and $r_1$) and $P(D_2; d_1, d_2, r_1)$ (for
fixed $d_1$ and $r_1$) are strictly quasi-concave on the (convex and compact) interval $[m_D, m_C]$ and
that they peak in this interval. It then follows from Theorem 5.3 in McCarty and Meirowitz
(2007: 108) that there exists a set of party strategies that constitute a Nash equilibrium.

---

16 A Nash equilibrium is a set of policy positions (strategies) for each candidate such that no candidate can increase
expected utility by deviating unilaterally. The results on multicandidate primaries apply under the additional as-
sumption that in each party the winning primary candidate is the one preferred by the median primary voter.

17 We say that a function $g$ is concave and peaks at $x_0$ if it is continuous, and if for all $x$ in the domain of $g$ for
which $x \neq x_0$, $\frac{\partial^2 g}{\partial x^2}(x) \leq 0$ and $g(x_0) > g(x)$. Note that if $g$ is concave and peaks at $x_0$, then $g$ is strictly in-
creasing on the left of $x_0$, and strictly decreasing on the right.
The concavity condition of Theorem 1 is satisfied by both the linear and quadratic policy loss metrics. In words, the theorem states that if exactly one of the two parties (here the Democrats) holds a primary, a Nash equilibrium in office-seeking candidates’ strategies exists. The theorem further states that if the two Democratic candidates have equal pre-campaign valence images, then exactly one paired equilibrium configuration exists. This is consistent with the examples presented in Figures 2-3, in which the two Democratic candidates had equal pre-campaign valence images and their equilibrium positions were identical.

Our second theorem relates to the examples on strong and weak parties presented above:

**Theorem 2.** Suppose there is one Republican candidate $R_1$ and two Democratic candidates $D_1$ and $D_2$, whose equilibrium positions $d_1^*$ and $d_2^*$ lie strictly between $m_D$ and $m_G$. Then a unilateral increase (decrease) in $V_{R1}$ gives both Democratic candidates incentives to shift unilaterally toward (away from) $m_G$.

In words, the theorem states that if exactly one of the two parties (here the Democrats) holds a competitive primary, then when the Republican candidate’s pre-campaign valence image improves, the Democratic candidates have office-seeking incentives to moderate their policy positions relative to the median general election voter’s position (the weaker party’s competitive primary advantage effect); and, when the Republican candidate’s pre-campaign valence image deteriorates, the Democratic candidates are motivated to shift to more radical positions. This theoretical result is consistent with the illustrative example presented in Figure 3 above.

---

18 In addition, we note that Theorem 1 holds regardless of whether $R_1$ is free to take any position – in which case if he is a utility optimizer he locates at the position of the median general election voter $m_G$ – or, alternatively, when $R_1$’s position is fixed at some alternative location, as may be the case if $R_1$ is an incumbent whose legislative voting record makes it impossible for him to credibly locate at $m_G$. 

20
Finally, we present two additional results that extend Theorems 1-2 to the case where both parties hold primaries. Both theorems concern the special case where the candidates’ pre-campaign valence images are equal within parties but may differ across parties (i.e. where $V_{D1} = V_{D2}$ and $V_{R1} = V_{R2}$, but not necessarily $V_{D1} = V_{R1}$).

**Theorem 3.** If both parties hold primaries and if $V_{D1} = V_{D2}$ and $V_{R1} = V_{R2}$, there exists a paired equilibrium in candidates’ office-seeking strategies, i.e., one such that $d_1^* = d_2^*$ and $r_1^* = r_2^*$.

**Theorem 4.** Suppose both parties hold primaries, that $V_{D1} = V_{D2}$ and $V_{R1} = V_{R2}$, and that the parties are located at a paired equilibrium configuration, with the equilibrium positions $d_1^* = d_2^*$ located strictly between $m_D$ and $m_G$. Then a unilateral increase (decrease) in $V_{R1}$ and/or $V_{R2}$ gives both Democratic candidates incentives to shift unilaterally toward (away from) $m_G$.

Theorem 3, which proves existence of a paired equilibrium provided that the candidates’ pre-election valence images are equal within parties, extends Theorem 1 from the one-primary to the two-primary scenario. Although we have not proved uniqueness of the equilibrium guaranteed by Theorem 3 for the case in which there are two candidates in each party, our numerical results (see below) suggest that the equilibrium is typically unique. The theorem is illustrated by the example presented earlier in Figure 4, in which both parties’ candidates paired with each other in equilibrium. Theorem 4 similarly extends the results in Theorem 2, on weak and strong parties, to the two-primary scenario.

**Parties’ equilibrium positions for two-stage elections: Numerical results**

Tables 1-2 report equilibrium configurations for two-stage election contests beginning with a primary, for scenarios where one party holds a competitive primary (Table 1) and when
both parties hold competitive primaries (Table 2). The tables also report each candidate’s equilibrium probability of winning office. For these computations we again set the median positions of the Democratic, Republican, and general electorates at $m_D = 2$, $m_R = 6$, and $m_G = 4$, respectively. The top row of results in Table 1, for instance, reports the equilibrium configuration for the example illustrated earlier in Figure 3, in which only the Democrats hold a primary, the policy salience coefficient is set at $a=1$ (see column 1), and the Democratic candidates are “weak” in that $V_{R_1} = V_{D_1} + 2 = V_{D_2} + 2$ (column 2). For this scenario, columns 3-4 report that at equilibrium the Republican candidate $R_1$ locates at $r_1^* = 4.00$ while the Democratic candidates pair at $d_1^* = d_2^* = 3.77$, and that the Democratic candidates’ equilibrium probabilities of winning office are $P^*(D_1) = P^*(D_2) = .102$. Finally, the RHS column in Table 1 reports whether the Democratic party benefits from holding a competitive primary election: this is based on a comparison of the equilibrium probability that the Democrats win office when they hold a competitive primary, compared to the probability they win office if a single Democrat – running unopposed in his party’s primary – located at the median general election position $m_G=4$. Generally, the Democrats are more likely to benefit from holding a competitive primary if the uncertainty about campaign skill is not too low relative to the salience of policy, i.e., when the policy-salience parameter $a$ is not too high.

We located an equilibrium for every scenario that we investigated, and, consistent with Theorems 1 and 3, these equilibrium configurations invariably find the primary election candidates paired with each other. Second, consistent with Theorems 2 and 4, we compute that as the Republican candidates’ pre-campaign valence images improve relative to Democrats’ images, the Democratic candidates’ equilibrium positions become more moderate. Finally, consistent
with our illustrative arguments, note that in all the scenarios for which the Democrats were specified as the “weak” party, holding a primary election benefits the Democrats.

[TABLES 1-2 ABOUT HERE]

5. Connections to Previous Research

In our model of two-stage electoral competition, the strategic trade-off candidates confront is how to balance the imperatives of appealing to their primary and general election constituencies. And, the strategic question party elites confront is whether the benefits primaries confer by identifying skilled, charismatic, campaigners offsets the fact that elites thereby cede control over the candidate selection process to the voters (see Serra, 2006a). We believe these trade-offs are indeed critical considerations for real world candidates and party elites. At the same time our model omits additional strategic factors that other scholars have identified,19 and thus we do not provide a comprehensive accounting of politicians’ primary-related considerations. In particular, we do not claim that our model explains why American parties adopted pri-

19 Additional features that may enhance the electoral benefits of holding primaries include: that primaries may unify the party (Carey, 2003; Poire, 2002); that candidates have the option of announcing ambiguous policy platforms in the primary, thereby preserving their strategic options in the general election (Meirowitz, 2005); that holding a primary can enhance the party’s reputation for democratic “openness” and transparency (see Carey and Polga-Hecimovich, 2006; De Luca et al., 2002); that primary voters select candidates in part based on their projected competitiveness in the general election (Stone and Abramowitz, 1983). Additional factors that can depress the electoral benefits of primaries include negative campaigning in the primary that harms the nominee in the general election (Carey and Polga-Hecimovich, forthcoming); that competitive primaries may force candidates to spend precious campaign resources which would otherwise be available for the general election; and, that primaries may benefit candidates who campaign well among the party faithful but lack campaign skills that appeal to a broader electorate.
maries, and so in the context of American politics we see our central contribution as delineating office-seeking candidates’ strategic incentives in two-stage elections.

By contrast, in Latin America primary elections are by no means established as the preferred candidate-selection mechanism, and there is an ongoing, often heated debate about their effect on candidate viability (see Beer, 2003; Carey and Polga-Hecimovich, 2006, forthcoming; Colomer, 2003; Scarrow, 2005; Serra, 2006a;). We must be cautious in extrapolating our model to these political contexts, because most Latin American countries feature multiparty systems. Nevertheless, some Latin American countries – notably Venezuela, Uruguay, Argentina, and arguably Costa Rica – have featured two dominant parties during periods in the twentieth century ranging from one to several decades, and, furthermore there are additional countries – notably Mexico – where subnational political competition frequently revolves around two dominant parties. Our results are plausibly relevant to these scenarios.

In this regard, recent empirical studies report findings that suggestively support our theoretical results, particularly those on the weaker party’s primary advantage effect. In an analysis of all elections to the Chamber of Deputies in Argentina held during the period 1983-2001 – as noted by Serra (2006a) – De Luca et al. (2002) report that opposition parties (defined as parties that did not run the province to which the district belonged) were significantly more likely to hold primaries than were governing parties, a finding that squares with our argument that parties

---

20 See Ware (2002) for an account of the factors that led American parties to adopt primaries. In addition, we note that it is not feasible, in the American context, to empirically evaluate our hypotheses about the electoral benefits of holding primaries in comparison to alternative candidate selection mechanisms, since the Democrats and Republicans adopted primaries at roughly the same time.

21 Over the past fifteen years, subnational elections in northern Mexico have been dominated by the candidates from the PRI and the PAN, while elections in the South have been dominated by candidates of the PRI and the PRD.
have added incentives to hold primaries when they expect to face a general election opponent who enjoys incumbency-related advantages. Furthermore, our arguments also plausibly illuminate empirical findings on election outcomes reported by Carey and Polga-Hecimovich (forthcoming). In a study of every democratic presidential election in Latin America from the late 1970s to 2004, Carey and Polga-Hecimovich conclude that, among candidates from opposition parties, candidates selected via primaries substantially outperformed those selected via alternative procedures: specifically, the authors estimate that, other factors being equal, opposition party candidates selected in primaries achieved a general election “bonus” of nine percentage points in the popular vote relative to the support for other opposition party candidates, a statistically significant effect. These results, again, are consistent with our arguments that parties have added incentives to hold primaries when they expect to face a general election opponent who enjoys incumbency-related advantages. By contrast, the authors estimate that candidates from the incumbent party who were selected in primaries gained no significant electoral bonus, compared to incumbent-party candidates selected via alternative methods.\(^ {22}\)

Finally, although Carey and Polga-Hecimovich do not attempt to parse out the reasons why opposition party candidates selected via presidential primaries outperformed their general election opponents, the authors’ discussion of the 1988 Argentinean presidential election highlights a case where a primary election identified a high-quality candidate from an opposition party, who would probably not have been nominated via an alternative selection procedure:

Primaries may simply be more effective than elite-driven search processes in identifying candidates with broad popular appeal…Carlos Menem’s emergence in the

\(^{22}\) Specifically, the authors estimate that incumbent-party candidates selected in primaries achieved a general election “bonus” of only one percentage point, a statistically insignificant effect.
1988 Partido Justicialista (PJ) primary in Argentina is an example. As governor of tiny La Rioja Province in Western Argentina, relative unknown Menem faced Buenos Aires governor Antonio Francisco Cafiero in the race for the PJ nomination. Menem’s flamboyant and hard-nosed style of campaigning proved effective. His upset victory over Cafiero attracted wide-spread media attention to his personal charisma and political canniness, and he went on to win the general election by 15% of the vote (Carey and Polga-Hecimovich, forthcoming, page 12).

This real-world example exactly mirrors our theoretical arguments about the expected “valence benefit” that primary elections confer, by helping parties identify strong campaigners who will prove attractive to the general electorate.

The weak party’s primary advantage result also bears an interesting relationship to remarkable work by Schofield and Sened (2005, 2006), which analyzes the strategic implications of valence issues in multiparty elections – i.e. elections involving three or more parties – without primaries. The authors report theoretical and empirical results that in such elections, valence-advantaged parties have electoral incentives to locate near the center of the voter distribution, while valence-disadvantaged parties are motivated to present policies that diverge from the center of the general electorate – a pattern the authors demonstrate captures the patterns of multi-party competition in Israel and various European polities. However our theoretical results demonstrate that in the presence of party-sponsored primaries, this strategic policy incentive is reversed. And the reason for this surprising distinction is that the intra-party competition we

---

23 Roughly speaking, the strategic logic underlying the authors’ results is that when valence-advantaged parties present centrist policies, valence-disadvantaged parties must flee the center in order to avoid competing with these strong parties strictly on valence issues – a competition the valence-disadvantaged parties will lose.
have modeled gives candidates from strong parties particularly compelling strategic incentives to appeal to the center of their primary electorates, at the expense of the general electorate. Thus our results suggest that in party systems that feature significant amounts of intra-party competition, valence dimensions can have fundamentally different implications for policy strategies than they do when political competition is primarily between parties, not within them.

Finally, although we have focused here on office-seeking candidates and parties, our theoretical results are also relevant to policy-seeking parties, i.e. parties that seek office in order to implement their preferred policies as opposed to proposing policies in single-minded pursuit of office (see Greene, 2007; Groseclose, 2001; Londregan and Romer, 1993; Wittman, 1983). To the extent that our model captures strategic features of real-world politics, it suggests that by holding primary elections, political parties can have the best of both worlds: namely, primaries create strategic incentives for office-seeking candidates to present policies that are shaded in the direction of the primary electorate’s sincere policy preferences, thus conferring a “policy benefit” to policy-oriented party members; and, holding a primary may simultaneously enhance the likelihood that the party wins office, an outcome that is valued both by office-seeking elites (who value holding office per se) and by policy-seeking elites (who value holding office because it allows the office-holder to implement his pre-election policy promises).

6. Conclusion

Despite the widespread use of primary elections in the United States and (increasingly) outside the U.S., there is little theoretical work on the policy strategies of office-seeking candidates who confront a two-stage election. We have developed a spatial model in which candidates compete in a primary followed by a general election, and where moreover the candidates’ electoral appeal is based on a combination of their policy positions, their valence-related reputa-
tions with respect to factors such as competence, charisma, and integrity, and the candidates’ campaigning skills; these latter abilities may prove critical to electoral success, but the candidates cannot project them with certainty at the time they announce their policy positions.

We have shown that our two-stage election model supports an equilibrium in office-seeking candidates’ strategies, under quite general conditions. Furthermore, we have presented theoretical results and numerical calculations that suggest a paradoxical conclusion: namely, that although primary elections draw candidates away from the center of the general electorate, political parties can nevertheless benefit from holding primaries. This is because primary elections allow parties to identify high-quality candidates who will prove to be effective campaigners in the general election. Furthermore, we have shown that “weak” parties have additional incentives to hold primaries, because weak parties’ candidates have strategic motivations to locate closer to the center of the general electorate than do strong parties’ candidates. Empirical work on Latin American politics by Carey and Polga-Hecimovich (2006, forthcoming), De Luca et al. (2002), Poire (2002), and others supports our theoretical arguments.

Our results on the potential electoral benefits of primaries are especially striking given that our theoretical set-up appears, at first glance, to entail bleak prospects for parties holding primaries. In our model the candidates cannot moderate their policies between the primary and the general election; primary voters vote myopically for their preferred candidate, without regard to whether he or she will prove competitive in the general election; and, our numerical examples involve scenarios where the primary electorates’ policy preferences differ dramatically from those of the general electorate. These assumptions capture considerations that have prompted many prominent scholars to argue that primary elections impose a handicap on a party’s pursuit of office: namely, that a radical primary electorate composed of hardcore partisans may short-
sightedly select a candidate of limited general electoral appeal (see Colomer, 2003; V.O. Key, 1947; Polsby, 1983;). Our demonstration, that parties may benefit from holding primaries even in a strategic environment that constitutes a “worst-case scenario” for primary election candidates, strongly suggests that our conclusions extend to alternative settings that relax the strong assumptions in our model. Other aspects that relate to primaries have been left for later work, such as uncertainty over the distribution of voters’ policy preferences; multiple policy dimensions; scenarios where candidates have some leeway to modify their positions between the primary and general elections; scenarios where the candidates intrinsically value winning the primary, because this enhances their reputation and thereby benefits them in future elections; the threat of entry of additional primary candidates (a prospect that constrains the positioning of candidates already in the race); and analysis of the sequential primaries that are used in the American presidential nomination process. In addition, we hope to extend our theoretical results to a more general model that relaxes our assumption of extreme value random utility.

The above caveats notwithstanding, we believe we have provided a theoretical rationale for the increasingly widespread use of primary elections throughout the world. We have shown how primary elections can enhance political parties’ prospects in the general election, and, furthermore, that weak parties in particular are likely to benefit from holding primaries. Finally, our results suggest that primary elections may enhance parties’ abilities to achieve both their office-seeking and their policy-seeking objectives: namely, primaries create strategic incentives for office-seeking candidates to present policies that are shaded in the direction of the party members’ sincere policy preferences, and, simultaneously, holding a primary can enhance the likelihood that the party wins office.
Appendix

A.1. Model 1: Two Democrats and one Republican

Derivation of the probabilities of winning office:

Writing

\[ M_D(D_j) = g(d_j, m_D) + V_{D_j}, \quad j = 1, 2 \]
\[ M_G(D_j) = g(d_j, m_G) + V_{D_j}, \quad j = 1, 2, \text{ and} \]
\[ M_G(R_1) = g(r_1, m_G) + V_{R_1} \]

equation 2 in the text can be written as

\[ P(D_1) = \frac{e^{M_D(D_1)}}{e^{M_D(D_1)} + e^{M_D(D_2)} + e^{[M_G(R_1) - M_G(D_1) + M_D(D_1)]}}, \]

which applies to the model with two Democrats who compete in a primary and one Republican who runs unopposed in the Republican primary. For simplicity, we have written \( P(D_1) \) for \( P(D_1; d_1, d_2, r_1) \), the probability that \( D_1 \) wins the primary and the general election while \( D_1, D_2, \) and \( R_1 \) assume the strategies \( d_1, d_2 \) and \( r_1 \), respectively. Note that for any fixed value of the campaign valence image \( \varepsilon_{D_1} \), the probability that \( D_1 \) wins both the Democratic primary and the general election is given by

\[
P[M_D(D_1) + \varepsilon_{D_1} > M_D(D_2) + \varepsilon_{D_2} \text{ and } M_G(D_1) + \varepsilon_{D_1} > M_G(R_1) + \varepsilon_{R_1}] \\
= P[\varepsilon_{D_2} < \varepsilon_{D_1} + M_D(D_1) - M_D(D_2) \text{ and } \varepsilon_{R_1} < \varepsilon_{D_1} + M_G(D_1) - M_G(R_1)] \\
= F_{D_2}[\varepsilon_{D_1} + M_D(D_1) - M_D(D_2)]F_{R_1}[\varepsilon_{D_1} + M_G(D_1) - M_G(R_1)] \\
= F_{D_2}[\varepsilon_{D_1} + W_D]F_{R_1}[\varepsilon_{D_1} + W_G],
\]
where $F_{D_2}$ and $F_{R_1}$ are the cumulative distribution functions for the campaign valence images $\varepsilon_{D_2}$ and $\varepsilon_{R_1}$ of $D_2$ and $R_1$, respectively, and for convenience, we have written

$$W_D = M_D(D_1) - M_D(D_2)$$

and

$$W_G = M_G(D_1) - M_G(R_1).$$

Equality holds in equation A1 because $\varepsilon_{D_2}$ and $\varepsilon_{R_1}$ are independent.\(^{24}\)

Since $\varepsilon_{D_2}$ and $\varepsilon_{R_1}$ follow extreme value distributions of type 1 and writing $s = \varepsilon_{D_1}$ for convenience, we observe (extending the argument in Train, 2003: 78-79) that the overall probability that $D_1$ wins both the Democratic primary and the general election is given by

$$\int_{-\infty}^{\infty} \exp\left[-e^{-s}e^{-W_D}\right] \exp\left[-e^{-s}e^{-W_G}\right] \exp(-e^{-s})e^{-s} ds$$

$$= \int_{-\infty}^{\infty} \exp\left[-e^{-s}\left(e^{-W_D} + e^{-W_G} + 1\right)\right] e^{-s} ds$$

$$= \int_{0}^{\infty} \exp\left[-t\left(e^{-W_D} + e^{-W_G} + 1\right)\right] dt,$$

where we have substituted $t = e^{-s}$ and $dt = -e^{-s} ds$. It follows that the last expression is

\(^{24}\) Note that, although winning the primary and winning the general election are not independent, given a specific value of the random variable reflecting $D_1$’s campaign valence image, then the campaign valence image of the other Democrat and of the Republican are independent. The probabilities that $D_1$ wins both elections – each conditional on the specific value of $D_1$’s campaign valence image – are then integrated with respect to the probability density for those campaign valence images. Equivalently, equation 2 can be derived by expressing the probability that $D_1$ wins both elections as the product of the probability that $D_1$ wins the primary and the probability that $D_1$ wins the general election conditional on having won the primary.
\[
\exp\left[-t(e^{-W_D} + e^{-W_G} + 1)\right]_0^\infty = \frac{1}{e^{-W_D} + e^{-W_G} + 1} = \frac{1}{e^{-[M_D(D_1) - M_D(D_2)]} + e^{-[M_G(D_1) - M_G(D_1)]} + 1},
\]

which is equivalent to the expression in equation 2. The derivation of equation 3 is similar.

To see that \( R_1 \) maximizes her probability of winning the general election by locating at the position of the general election median voter, note that the probability that \( R_1 \) wins office is the sum of the probability that \( R_1 \) wins the general election with \( D_1 \) as opponent and the corresponding probability with \( D_2 \) as opponent. The first of these probabilities is the probability that voter \( m_D \) prefers \( D_1 \) to \( D_2 \) and voter \( m_G \) prefers \( R_1 \) to \( D_1 \). Thus the probability

\[
P(R_1 D_1) = P(R_1 D_1; d_1, d_2, r_1) \quad \text{that} \quad R_1 \quad \text{wins the general election with} \quad D_1 \quad \text{as opponent is given by}
\]

\[
P(R_1 D_1) = P[M_D(D_1) + \varepsilon_{D_1} > M_D(D_2) + \varepsilon_{D_2} \quad \text{and} \quad M_G(R_1) + \varepsilon_{R_1} > M_G(D_1) + \varepsilon_{D_1}]
\]

\[
= \frac{e^{M_D(D_1)}}{e^{M_D(D_1)} + e^{M_D(D_2)}} - \frac{e^{M_D(D_1)}}{e^{M_D(D_1)} + e^{M_D(D_2)} + e^{M_G(R_1) - M_G(D_1) - M_D(D_1)}}, \quad (A2)
\]

and a similar formula for \( P(R_1 D_2) \). The term \( M_G(R_1) \) appears only once in each of these formulas. Thus, given any possible set of positions for \( D_1 \) and \( D_2 \), the sum \( P(R_1 D_1) + P(R_1 D_2) \) is maximized by maximizing \( M_G(R_1) \). The latter term, given by \( M_G(R_1) = g(r_1, m_G) + V_{R_1} \), is in turn maximized by choosing \( r_1 = m_G \).
References


Train, Kenneth. 2003. *Discrete Choice Methods with Simulation*. Cambridge: Cam-

Table 1. Equilibrium Configurations When Only One Party Holds a Primary

<table>
<thead>
<tr>
<th>Policy salience coefficient (1)</th>
<th>$R_1$’s measured valence advantage ($V_{R}-V_{D}$) (2)</th>
<th>$R_1$’s office-seeking equilibrium position, with no primary (3)</th>
<th>$D_1$ and $D_2$’s equilibrium positions, when the party holds a contested primary (4)</th>
<th>Does holding a primary benefit the Democrats? (5)</th>
</tr>
</thead>
<tbody>
<tr>
<td>$a=1$</td>
<td>$2$</td>
<td>$r_1^* = 4.00$ $P^*(R_1) = .796$</td>
<td>$d_1^* = d_2^* = 3.77$ $P^<em>(D_1) = P^</em>(D_2) = .102$</td>
<td>Yes</td>
</tr>
<tr>
<td>$a=1$</td>
<td>$1$</td>
<td>$r_1^* = 4.00$ $P^*(R_1) = .627$</td>
<td>$d_1^* = d_2^* = 3.54$ $P^<em>(D_1) = P^</em>(D_2) = .187$</td>
<td>Yes</td>
</tr>
<tr>
<td>$a=1$</td>
<td>$0$</td>
<td>$r_1^* = 4.00$ $P^*(R_1) = .462$</td>
<td>$d_1^* = d_2^* = 3.26$ $P^<em>(D_1) = P^</em>(D_2) = .269$</td>
<td>Yes</td>
</tr>
<tr>
<td>$a=1$</td>
<td>$-1$</td>
<td>$r_1^* = 4.00$ $P^*(R_1) = .333$</td>
<td>$d_1^* = d_2^* = 3.00$ $P^<em>(D_1) = P^</em>(D_2) = .333$</td>
<td>No</td>
</tr>
<tr>
<td>$a=1$</td>
<td>$-2$</td>
<td>$r_1^* = 4.00$ $P^*(R_1) = .237$</td>
<td>$d_1^* = d_2^* = 2.77$ $P^<em>(D_1) = P^</em>(D_2) = .382$</td>
<td>No</td>
</tr>
<tr>
<td>$a=2$</td>
<td>$2$</td>
<td>$r_1^* = 4.00$ $P^*(R_1) = .803$</td>
<td>$d_1^* = d_2^* = 3.78$ $P^<em>(D_1) = P^</em>(D_2) = .099$</td>
<td>Yes</td>
</tr>
<tr>
<td>$a=2$</td>
<td>$1$</td>
<td>$r_1^* = 4.00$ $P^*(R_1) = .657$</td>
<td>$d_1^* = d_2^* = 3.59$ $P^<em>(D_1) = P^</em>(D_2) = .172$</td>
<td>Yes</td>
</tr>
<tr>
<td>$a=2$</td>
<td>$0$</td>
<td>$r_1^* = 4.00$ $P^*(R_1) = .523$</td>
<td>$d_1^* = d_2^* = 3.37$ $P^<em>(D_1) = P^</em>(D_2) = .239$</td>
<td>No</td>
</tr>
<tr>
<td>$a=2$</td>
<td>$-1$</td>
<td>$r_1^* = 4.00$ $P^*(R_1) = .417$</td>
<td>$d_1^* = d_2^* = 3.18$ $P^<em>(D_1) = P^</em>(D_2) = .292$</td>
<td>No</td>
</tr>
<tr>
<td>$a=2$</td>
<td>$-2$</td>
<td>$r_1^* = 4.00$ $P^*(R_1) = .333$</td>
<td>$d_1^* = d_2^* = 3.00$ $P^<em>(D_1) = P^</em>(D_2) = .333$</td>
<td>No</td>
</tr>
</tbody>
</table>

Notes. For these computations the median position of the Democrats’ primary voter was set to $m_D=2$, and the median position of the general election voter was set to $m_G=4$. The Democratic Party is assumed to hold a competitive primary election, while the Republican candidate $R_1$ runs unopposed in the Republican primary. Voters are assumed to have quadratic policy losses. The RHS column reports whether the Democratic party benefits from holding a competitive primary election: this is based on a comparison of the equilibrium probability that the Democrats win office when they hold a competitive primary, compared to the probability they win office if a single Democrat runs unopposed in his party’s primary and locates at the median general election position $m_G=4$. 
Table 2. Equilibrium Configurations When Both Parties Hold Primaries

<table>
<thead>
<tr>
<th>Policy salience coefficient (1)</th>
<th>Republicans’ measured valence advantage ($V_R - V_D$) (2)</th>
<th>$D_1$’s and $D_2$’s office-seeking equilibrium positions (3)</th>
<th>$R_1$ and $R_2$’s office-seeking equilibrium positions (4)</th>
<th>Does holding a primary benefit the Democrats? (5)</th>
</tr>
</thead>
<tbody>
<tr>
<td>$a=1$</td>
<td>$+2$</td>
<td>$d_1^* = d_2^* = 3.66$</td>
<td>$r_1^* = r_2^* = 5.10$</td>
<td>Yes</td>
</tr>
<tr>
<td></td>
<td></td>
<td>$P^<em>(D_1) = P^</em>(D_2) = 0.144$</td>
<td>$P^<em>(R_1) = P^</em>(R_2) = 0.356$</td>
<td></td>
</tr>
<tr>
<td>$a=1$</td>
<td>$+1$</td>
<td>$d_1^* = d_2^* = 3.51$</td>
<td>$r_1^* = r_2^* = 4.89$</td>
<td>Yes</td>
</tr>
<tr>
<td></td>
<td></td>
<td>$P^<em>(D_1) = P^</em>(D_2) = 0.195$</td>
<td>$P^<em>(R_1) = P^</em>(R_2) = 0.305$</td>
<td></td>
</tr>
<tr>
<td>$a=1$</td>
<td>$0$</td>
<td>$d_1^* = d_2^* = 3.33$</td>
<td>$r_1^* = r_2^* = 4.67$</td>
<td>Yes</td>
</tr>
<tr>
<td></td>
<td></td>
<td>$P^<em>(D_1) = P^</em>(D_2) = 0.250$</td>
<td>$P^<em>(R_1) = P^</em>(R_2) = 0.250$</td>
<td></td>
</tr>
<tr>
<td>$a=1$</td>
<td>$-1$</td>
<td>$d_1^* = d_2^* = 3.11$</td>
<td>$r_1^* = r_2^* = 4.49$</td>
<td>No</td>
</tr>
<tr>
<td></td>
<td></td>
<td>$P^<em>(D_1) = P^</em>(D_2) = 0.305$</td>
<td>$P^<em>(R_1) = P^</em>(R_2) = 0.195$</td>
<td></td>
</tr>
<tr>
<td>$a=1$</td>
<td>$-2$</td>
<td>$d_1^* = d_2^* = 2.90$</td>
<td>$r_1^* = r_2^* = 4.34$</td>
<td>No</td>
</tr>
<tr>
<td></td>
<td></td>
<td>$P^<em>(D_1) = P^</em>(D_2) = 0.356$</td>
<td>$P^<em>(R_1) = P^</em>(R_2) = 0.144$</td>
<td></td>
</tr>
<tr>
<td>$a=2$</td>
<td>$+2$</td>
<td>$d_1^* = d_2^* = 3.58$</td>
<td>$r_1^* = r_2^* = 4.94$</td>
<td>Yes</td>
</tr>
<tr>
<td></td>
<td></td>
<td>$P^<em>(D_1) = P^</em>(D_2) = 0.179$</td>
<td>$P^<em>(R_1) = P^</em>(R_2) = 0.321$</td>
<td></td>
</tr>
<tr>
<td>$a=2$</td>
<td>$+1$</td>
<td>$d_1^* = d_2^* = 3.46$</td>
<td>$r_1^* = r_2^* = 4.80$</td>
<td>Yes</td>
</tr>
<tr>
<td></td>
<td></td>
<td>$P^<em>(D_1) = P^</em>(D_2) = 0.212$</td>
<td>$P^<em>(R_1) = P^</em>(R_2) = 0.288$</td>
<td></td>
</tr>
<tr>
<td>$a=2$</td>
<td>$0$</td>
<td>$d_1^* = d_2^* = 3.33$</td>
<td>$r_1^* = r_2^* = 4.67$</td>
<td>No</td>
</tr>
<tr>
<td></td>
<td></td>
<td>$P^<em>(D_1) = P^</em>(D_2) = 0.250$</td>
<td>$P^<em>(R_1) = P^</em>(R_2) = 0.250$</td>
<td></td>
</tr>
<tr>
<td>$a=2$</td>
<td>$-1$</td>
<td>$d_1^* = d_2^* = 3.20$</td>
<td>$r_1^* = r_2^* = 4.54$</td>
<td>No</td>
</tr>
<tr>
<td></td>
<td></td>
<td>$P^<em>(D_1) = P^</em>(D_2) = 0.288$</td>
<td>$P^<em>(R_1) = P^</em>(R_2) = 0.212$</td>
<td></td>
</tr>
<tr>
<td>$a=2$</td>
<td>$-2$</td>
<td>$d_1^* = d_2^* = 3.06$</td>
<td>$r_1^* = r_2^* = 4.42$</td>
<td>No</td>
</tr>
<tr>
<td></td>
<td></td>
<td>$P^<em>(D_1) = P^</em>(D_2) = 0.321$</td>
<td>$P^<em>(R_1) = P^</em>(R_2) = 0.179$</td>
<td></td>
</tr>
</tbody>
</table>

Notes. For these computations the median position of the Democrats’ primary voter was set to $m_D=2$; the median general election voter was located at $m_G=4$; the median Republican primary voter was located at $m_R=6$. Voters are assumed to have quadratic policy losses. The RHS column reports whether the Democratic party benefits from holding a competitive primary election: this is based on a comparison of the equilibrium probability that the Democrats win office when they hold a competitive primary, compared to the probability they win office if a single Democrat runs unopposed in his party’s primary and locates at the median general election position $m_G=4$. 

39
Figures 1-4: Illustrative Examples of Candidate Strategies in Primary Elections

Figure 1: $D_1$ is Unopposed in the Primary, and $V_{D_1} = V_{R_1}$

Figure 2: $D_1$ and $D_2$ Contest the Primary, and $V_{D_1} = V_{D_2} = V_{R_1}$

Notes: For these computations the median voter positions were set to $m_D = 2$, $m_G = 4$, and $m_G = 6$. Voters were assigned quadratic policy loss functions with the policy salience parameter set to $a=1$. 
Illustrative Examples, Continued

Figure 3: $D_1$ and $D_2$ Contest the Primary, and $V_{R1}=(V_{D1}+2)=(V_{D2}+2)$

![Graph showing the probability of $D_1$ winning the Democratic primary and being elected, with candidate $D_1$'s location on the x-axis and probability on the y-axis.]

Notes: For these computations the median voter positions were set to $m_d = 2$, $m_G = 4$, and $m_g = 6$. Voters were assigned quadratic policy loss functions with the policy salience parameter set to $a = 1$.

Figure 4: Both Parties Hold Contested Primaries, and $V_{R1}=V_{R2}=(V_{D1}+2)=(V_{D2}+2)$

![Graph showing the probability of $D_1$ winning the Democratic primary and being elected, with candidate $D_1$'s location on the x-axis and probability on the y-axis.]

Notes: For these computations the median voter positions were set to $m_d = 2$, $m_G = 4$, and $m_g = 6$. Voters were assigned quadratic policy loss functions with the policy salience parameter set to $a = 1$. 