ELECTORAL CONTROL AND SELECTION
WITH PARTIES AND PRIMARIES*

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

Michael M. Ting
Department of Political Science and SIPA
Columbia University

September 28, 2009

Abstract

We develop a model of intra-party candidate selection in an environment with partisan electoral competition and voter uncertainty. Candidates for office belong to parties, which are factions of ideologically similar candidates. Each party’s candidates for a general election can be selected either by a “centralized” mechanism that effectively randomizes over possible candidates, or by a voter in a primary election. The voter cares about ideology and valence, and both primary and general elections may reveal candidate valences, thus helping the voter. Our main theoretical result is that while primaries raise the expected quality of a party’s candidates, they may hurt the \textit{ex ante} preferred party in a competitive electorate by increasing the chances of revealing the opposing party’s candidates as superior. Thus primaries are adopted in relatively extreme districts where a clear favorite party exists. An empirical analysis of the adoption of direct primaries and the competitiveness of primary elections across U.S. states supports these predictions.

JEL D72

*We thank Anna Bassi and panel participants at the 2005 meeting of the American Political Science Association and 2009 APSA State Politics conference for helpful comments.
1 Introduction

Scholars have long argued that meaningful political parties are necessary for a successful democracy. Parties serve a variety of purposes in different systems. The list includes: providing long-lived organizations through which relatively short-lived politicians can formulate policies, make credible promises to voters, solve collective action problems, and pursue politics as a career; providing low-cost information to voters about the likely policy goals or ideologies of politicians; providing low-cost information about which politicians are responsible for current policy outcomes; providing voters with distinct policy choices; organizing legislative activity to solve collective choice problems.\footnote{See, for example, Wilson (1885), Schattschneider (1942), Ranney (1951, 1962), Fiorina (1980), Alesina and Spear (1988), Harrington (1992), Cox and McCubbins (1993), Aldrich (1995).}

Fiorina (1980) states one of the main arguments in support of “strong,” or at least “cohesive,” parties: “The only way collective responsibility has ever existed, and can exist given our institutions, is through the agency of the political party; in American politics, responsibility requires cohesive parties.” A World Bank report on political accountability echoes this, with an even broader, world-wide recommendation:

> Effective sanctions on politicians can be enhanced most effectively through a meaningful degree of political competition in the electoral process... Political competition is most effective in promoting accountability when it is channeled through organizations that provide broad constituencies with vehicles — such as mass-based political parties and interest groups — to express their collective demands to political leaders.\footnote{See http://www1.worldbank.org/publicsector/anticorrupt/politicalaccountability.htm.}

While the benefits of stable and cohesive parties seem clear, researchers have paid less attention to the costs. This paper studies certain costs, in particular, those due to a loss in electoral selection that can occur under cohesive parties and imperfectly informed voters, especially in ideologically extreme constituencies. The basic idea is simple: If it is costly for voters to replace an incumbent politician — because the alternative is a politician from
the opposing political party — then voters may find themselves “stuck” with low-quality politicians even when they know the politicians are of low quality. The costs in our model are thus similar in spirit to the costs of ethnic divisions studied by Acemoglu, Robinson, and Verdier (2004) and Padrò i Miquel (2007), although we emphasize selection rather than moral hazard problems, and partisan rather than ethnic divisions.3

The paper also considers one way to reduce these costs, namely primary elections. Our theoretical approach is to consider how voters would act when confronted by candidates of uncertain ideology and quality. Each candidate’s party label provides imperfect information about her ideological position, but provides no information about her quality or valence. As in many models of electoral competition, the valence dimension could represent a characteristic (e.g., skill, charisma, or personal background) that a voter considers valuable in an elected official. The valence might also serve as a reduced form for shirking in office. In a general election, the valence of both parties’ candidates are revealed with positive probability. Thus a moderate voter may vote against her ex ante preferred party if valences are revealed, while an extreme voter has no alternative but to elect the candidate of the preferred party.

Primary elections help to improve voter choices by adding another opportunity for valences to be revealed. In a primary election, an extreme voter can reject a low-valence candidate from her preferred party. In addition, the party also benefits by having higher-valence candidates elected to office. A moderate voter similarly benefits from having more options. But in this case, the increased likelihood of valence revelation may hurt the preferred party. This could happen because primaries reduce the voter’s need to rely on her prior beliefs about the expected utility from electing her preferred party’s candidates.4

---


4Our work diverges from previous models of primaries which have focused on extensions of the Downsian framework (e.g., Coleman 1971, 1972, Aronson and Ordeshook 1972, Meirowitz 2005, Owen and Grofman
In section 3 we derive the main result of the model, which follows directly from the preceding logic. If the voter (or equivalently, the electoral district) is relatively extreme in the sense of having unambiguous favorite party, then primaries will benefit that party. These constituencies can restrict attention to one party, and use the primary election rather than the general election to select politicians. By contrast, the opportunity for primaries to reveal damaging information about a favored party’s candidates induces a preference for no primaries when the district is moderate. Since a favored party is more likely to hold office \textit{ex ante}, the model predicts that primaries are more likely to be introduced where inter-party competition is weakest.

The introduction of primaries allows voters to be more selective in the types of politicians they elect. The median voter in any constituency is strictly better off than in a world without primaries. However, we also find that no type of constituency unambiguously benefits most from having primary elections. If primaries are costless, then all constituencies would prefer to use them. But if primaries are expensive, then the induced preference of a constituency depends on both the marginal informational contribution of the primary election and its willingness to elect a sufficiently “good” politician from the “wrong” party.

The model further predicts that in the world with primaries, all but the most moderate constituencies are dominated by one party. Without primaries, somewhat left-leaning constituencies sometimes elect politicians from the rightist party, and vice versa. Thus, primaries may lead to the appearance of a more polarized electorate.

Finally, in section 4 we present empirical results from elections in the U.S. that are broadly consistent with certain comparative static predictions of the model. We do not view this evidence as a “test” of the model — the model is too stylized, and its predictions to stark, to be used as the basis for empirical work. However, the patterns in the data are so clear that they suggest the need for further theoretical work on primary and general elections

---

2004). Closer to our work are Caillaud and Tirole (1999), Jackson, Mathevet, and Mattes (2007), Castanheira, Crutzen, and Sahuguet (2008), and Serra (2009), who study the incentives of parties to introduce democratic governance. In Jackson \textit{et al.} candidates are distinguished only by ideology, while in Castanheira \textit{et al.} they are distinguished only by valence. Like Serra’s model, ours distinguishes candidates according to both. None of these models arrive at the same predictions as those here.
and the interaction between them, perhaps along the lines developed here.

2 The Model

We develop a simple model of political competition in a single constituency or district. The election is an open-seat race between two political parties, labeled $L$ and $R$. Each party $i$ has two potential candidates, labeled $C_{i1}$ and $C_{i2}$. The potential candidates’ partisan affiliations are public information. Each potential candidate may choose whether to declare entry into the election. Before the general election, each party $i$ with at least one declared candidate generates a single candidate $C_i$. Without primaries, this selection process is random. If there are primaries, then a voter (M) who represents the district’s median voter chooses which candidate will represent the party in the general election. M also chooses the winner of the general election. The winner of the general election implements a policy $x$ from the convex, compact set $X \subseteq \mathbb{R}$.

Parties play two roles in the model. First, the favored party in the district (i.e., the party most likely to hold office ex ante) chooses whether primaries will precede the general election. This choice reflects the favored party’s greater opportunity over the long run to have office-holders in position to decide on whether to implement primaries. Second, they screen the ideological preferences of their members. Candidates can credibly communicate only their party affiliations to the voter. The ideal point of each $C_{ij}$, $y_{C_{ij}}$, is drawn independently according to the common knowledge probability density $g_i(y)$ with support $\Gamma_i \subseteq X$. The mean of the party $i$ ideal point distribution is $\mu_i = \int_X yg_i(y)dy$, where $\mu_L < \mu_R$.

In addition to her party affiliation and ideal point, each candidate also has a valence parameter $v_c \in [-1/2, 1/2]$. The valence of each candidate is drawn independently according to the uniform distribution on $[-1/2, 1/2]$. Unlike ideal points, valence may be revealed in the course of an election campaign, either in the primary or general election. All players are symmetrically informed about valence parameters. The revelation of valence might therefore correspond to a shock in the preferences of voters over heterogeneous (but known) candidate

\footnote{Our results follow with minor modifications for discrete distributions.}
features, such as experience in government, veteran status, or sex.

Parties have lexicographic preferences, maximizing first the probability of winning the election in the district and second the expected valence of its elected candidates. Each potential candidate \( C_{ij} \) has an additive utility function \( u_{C_{ij}}(\cdot) \), consisting of three components. She receives utility \( u(|y_{C_{ij}} - x|) \) from policy \( x \), where \( u : \mathbb{R}_+ \to \mathbb{R}_- \) is continuous, strictly decreasing, and concave with \( u(0) = 0 \), and \( y_{C_{ij}} \in X \) is her ideal point. In addition, candidates care about holding office and receive one unit of utility for winning the general election, and zero otherwise. Finally, declaring entry in the election imposes a cost \( k \in [0, 1] \) on each candidate. The voter cares about policy and valence. Her utility is \( u_M(x, w; y_M) = u(|y_M - x|) + v_w \), where \( y_M \) is her ideal point and \( v_w \) the valence of the election winner \( w \).

The sequence of moves is as follows. All actions are observable unless otherwise noted.

1 **Election Rules.** The party with the ideal point distribution preferred \textit{ex ante} by \( M \) chooses whether there will be primary elections \((e = 1)\) or not \((e = 0)\).

2 **Candidate Draws.** Nature selects \( y_{CL1}, y_{CL2}, y_{CR1}, \text{and } y_{CR2} \), unobserved by \( M \).

3 **Declarations.** Each potential candidate \( C_{ij} \) simultaneously chooses whether to enter the election \((d_{C_{ij}} = 1)\) or not run \((d_{C_{ij}} = 0)\).

4 **Valence Revelation (Primaries).** If \( e = 1 \), Nature reveals all \( v_{C_{ij}} \) with probability \( \pi_p \).

5 **Primary Elections or Candidate Selection.** If \( e = 1 \), then for each party \( i \), \( M \) casts a vote \( r^i_p \) choosing \( C^i \) from the set of party \( i \)'s declared candidates. If \( e = 0 \), then for each party \( i \), \( C^i \) is chosen with equal probability from the set of party \( i \)'s declared candidates. If a party \( i \) has no declared candidates, then \( C^i = \emptyset \).

6 **Valence Revelation (Generals).** If \( e = 0 \), or \( e = 1 \) and \( v_{CL} \) and \( v_{CR} \) are unknown to \( M \), then Nature reveals \( v_{CL} \) and \( v_{CR} \) with probability \( \pi_g \).

7 **General Election.** \( M \) casts a vote \( r_g \in \{C^L, C^R\} \) for one of the party candidates.

8 **Policy Choice.** The general election winner chooses \( x \in X \).
In a world without primaries, the voter has no role in selecting either party’s general election candidate. She effectively faces a random draw from each population of potential candidates. With primaries, the voter can elect any candidate from either party.\(^6\) Voters cannot distinguish between challengers if Nature does not reveal candidate valences. In this environment, candidate selection in effect remains random. As in the general election, in the primary election M chooses between challengers whose valence may be revealed if the race becomes “hot.”

We derive Perfect Bayesian equilibria in pure strategies. This consists of the favored party’s optimal choice of election rule \(e^\ast\), each potential candidate’s declaration \(d^\ast_{c}\), and an elected candidate’s optimal policy \(x^\ast\) (i.e., her ideal point). Additionally, M must cast optimal votes \(r^L_p\), \(r^R_p\), and \(r^g_{c}\). These votes depend on M’s beliefs about each candidate’s valence and ideal point parameters. Since there can be no signaling of valence, beliefs about valence are trivial when valence is revealed by Nature, or are equivalent to their prior beliefs otherwise. Posterior beliefs about the ideal point of each candidate following her declaration are represented by the set \(\gamma(d^\ast_{C\cup}) \subseteq \Gamma^i\), which identifies ideological types associated with declaration \(d^\ast_{C\cup}\). The measure of such types follows straightforwardly from \(g^i\). Out of equilibrium, M simply assumes that a party \(i\) candidate’s ideal point is drawn according to \(g^i(\cdot)\).

To rule out a number of uninteresting cases, we assume that potential candidates break ties in favor of not declaring. We also require that M choose at every information set the candidate that would yield the highest expected payoff if she were to achieve office. This refinement simply rules out non-intuitive equilibria where M chooses inferior candidates from the party that is expected to lose the general election, and does not affect the winner of the general election. It is also consistent with an environment in which non-median primary voters vote for candidates with the outcome of the general election in mind.

\(^6\)Note that the voter can vote in both primaries. This might correspond to an open primary, although the main results would hold even if primary voters were partisans, as such voters would also choose the candidate with the highest valence.
3 Main Results

It is clear that any winning candidate \( w \) will choose her ideal policy upon election, and hence \( x^* = y_w \). The game therefore reduces to M's choices at the general election and possibly the primary election. M's assessment of a candidate in any election depends in part on the expected policy utility from each party's distribution. For each party \( i \), this value is:

\[
\pi^i = \int_{\Gamma_i} u(|y_M - y|)g^i(y)dy. \tag{1}
\]

The assumptions on \( u(\cdot) \) and \( g^i(\cdot) \) ensure that there exists some \( d \) such that \( u(|y_M - d|) = \pi^i \). Thus M is indifferent among the policies \( y_M + d, y_M - d \), and a random draw from the party \( i \) distribution of ideal points. They additionally ensure that \( \pi^i \) is decreasing in \( |y_M - \mu^i| \).

Since many of our results will focus on M's preferences, it will be convenient to classify the voter (or, equivalently, the district type) according to her \textit{ex ante} preferences over the parties. Let \( \hat{u} = u_R - u_L \) be M's relative expected preference for party \( R \)'s candidates. Denote by \( m \) the ideal point at which \( \hat{u} = 0 \). In the absence of knowledge about valence, a voter for which \( y_M = m \) will be indifferent between the two parties, with the tie broken in favor of party \( R \). We call a party \textit{favored} if M prefers the expected policy ideal point of its candidates. (Recall that the favored party chooses whether to hold primary elections.) To denote extreme voters, let \( \overline{m} \) and \( m \) be the voter ideal points at which \( \hat{u} = 1 \) and \( \hat{u} = -1 \), respectively. Note that the symmetry of \( u(\cdot) \) implies that \( \overline{m} - m = m - \overline{m} \). Then for any \( y_M \geq \overline{m} \) (respectively, \( y_M \leq m \)), M will prefer party \( R \) (respectively, \( L \)) regardless of valence. Much of the subsequent discussion therefore focus on moderate districts, where \( y_M \in (m, \overline{m}) \).

Our first result establishes a simple but important property of equilibria of the game. In equilibrium, the declaration choices of potential candidates cannot be be informative of ideal points. This occurs because the candidates’ expected payoffs do not depend on their ideal points.\footnote{The result would not hold if either \( k \) varied with ideal points, or if ideal points could be revealed prior to an election.} This allows ideological types within a party to pool their declaration strategies. The voter therefore does not update her beliefs beyond her priors, which were based on party
affiliation alone. This simplifies the subsequent analysis because it fixes the expected policy utility of any party \( i \) candidate at \( \pi_i \).

**Lemma 1** In equilibrium, for all \( C^{ij} \) and \( d_{C^{ij}} \), \( \gamma_{C^{ij}}(d_{C^{ij}}) = \Gamma_i \).

**Proof.** All proofs are in the Appendix.

### 3.1 No Primaries

Consider first the non-primaries world. Since each party’s candidate is chosen randomly from among the set of declared candidates, the parameters of each party’s general election candidate are effectively random draws from the party’s respective distributions if at least one of its candidate enters. A second declared election candidate cannot affect a party’s chances of victory. If no candidate enters the election for a party, then the party loses the election with certainty.

Suppose initially that each party generates at least one candidate. If valence is not revealed in the general election, then M simply chooses the candidate whose party’s ideal points are closest in expectation. Lemma 1 then implies that \( r^*_y = R \) if and only if \( \pi_R \geq \pi_L \).

If Nature reveals the valence parameters, then the condition for voting for \( C^R \) becomes \( \pi_R + v_{C^R} \geq \pi_L + v_{C^L} \). It is therefore possible in a moderate district where \( |\hat{u}| < 1 \) that a high valence differential would cause M to vote for the ideologically more distant party. For any \( y_M \in [m, \overline{m}] \) (equivalently, \( \pi_R > \pi_L \)), party \( R \)'s probability of victory can be calculated by integrating over realizations of \( v_{C^R} \):

\[
\int_{-1/2}^{1/2} \min\{1, v + \hat{u} + 1/2\} dv = \int_{-1/2}^{1/2-\hat{u}} v + \hat{u} + 1/2 \ dv + \int_{1/2-\hat{u}}^{1/2} dv = 1 + 2\hat{u} - \hat{u}^2.
\]

Likewise, for \( y_M \in (m, m) \), party \( R \)'s probability of victory is \((1 + \hat{u})^2/2\). Thus the ex ante
probability of a party $R$ victory is:

$$\rho^R_g = \begin{cases} 
0 & \text{if } \hat{u} < -1 \ (y_M \leq m) \\
\pi_g \frac{1+2\hat{u}+\hat{u}^2}{2} & \text{if } \hat{u} \in [-1, 0) \ (y_M \in (m, m)) \\
1 - \pi_g \left(1 - \frac{1+2\hat{u}-\hat{u}^2}{2}\right) & \text{if } \hat{u} \in [0, 1] \ (y_M \in [m, m]) \\
1 & \text{if } \hat{u} > 1 \ (y_M \geq m).
\end{cases} \quad (2)$$

Given (2), the number of party $R$ candidates willing to enter is then:

$$n^R = \begin{cases} 
0 & \text{if } k \geq \rho_g^R \\
1 & \text{if } k \in \left[\rho_g^R/2, \rho_g^R\right) \\
2 & \text{if } k < \rho_g^R/2.
\end{cases} \quad (3)$$

The expressions for $\rho_g^L$ and $n^L$ are symmetric. These clearly imply an intuitive comparative static: within each party, more candidates will enter as the electoral environment becomes more favorable. At the extremes, where $|\hat{u}| \geq 1$, the favored party will have one candidate if $1 > k \geq 1/2$ and two if $k < 1/2$, while the other will have none. They also imply that a party’s election prospects are decreasing in the revelation probability $\pi_g$ if it is favored, but increasing if it is unfavored.

### 3.2 Primaries

With primaries, M has an additional chance to receive information about candidate valence. If Nature does not reveal the valence of primary candidates, then the general election candidates $C^L$ and $C^R$ are again random draws from their respective party distributions. The game remains identical to the game without primaries. By contrast, the revelation of the primary candidates’ valence parameters may change each party’s election prospects. When valences are revealed, M chooses the candidate with the higher valence in each party. Thus $r_i^* = C^{a_1}$ iff $v_{C^{a_1}} \geq v_{C^{a_2}}$.

The revelation of information in primaries has two effects. First, it increases the expected valence of each party’s candidate. The $n$-th order statistic of a sample of $n$ independent draws from $U[0, 1]$ is distributed according to $\beta(n, 1)$, which has mean $n/(n+1)$. Primaries therefore raise the expected valence of each party’s chosen candidate from zero to $\pi_p/6$. (Increasing the number of primary candidates beyond two would increase the expected valence further.)
Second, the revelation of valence gives each party an additional opportunity to lose the election in an ideologically favorable district. The clearest example of this takes place when $y_M$ is just to the right of $m$. Suppose that an equal number of potential candidates from both parties enter the election. In this environment, M essentially chooses the candidate with the highest expected valence in each election. With probability $\pi_p$, the valence of both parties’ candidates are revealed and both party $R$ candidates will lose with a probability of about $1/2$. This is strictly higher than party $R$’s probability of losing when valences are not revealed during the primaries, which by (2) is approximately $\pi_g/2$. Thus in a moderate district, party $R$ can do strictly worse with primaries than without.

The next result characterizes the entry strategies in all pure strategy equilibria. To simplify the presentation we focus on right-leaning districts ($\hat{u} > 0$), as the results for left-leaning districts are symmetric. Since all candidates within a party are ex ante identical, it is convenient to state the result in terms of the equilibrium number of candidates from each party, denoted by the pair $(n_R, n_L)$. Thus $n_R = 1$ implies that either $C^{R1}$ or $C^{R2}$ enters the election.

**Proposition 1** Number of candidates. *In districts where $\hat{u} \geq 0$, the set of possible entrants*
in pure strategy equilibria is:

\[
(n^*_R, n^*_L) = \begin{cases} 
(0, 1) & \text{if } k \geq \frac{1}{2} \text{ and } \\
& k \geq \pi_p \left( \frac{1+2\hat{u}-\hat{u}^2}{2} \right) + (1 - \pi_p)\rho^R_g
\end{cases}
\]

\[
(1, 0) & \text{if } k \geq \frac{1}{2} \\
& k \geq \pi_p \left( \frac{1+3\hat{u}-\hat{u}^3}{3} \right) + (1 - \pi_p)\rho^R_g
\]

\[
(0, 2) & \text{if } k \in \left[ 1 - \pi_p \left( \frac{1+2\hat{u}-\hat{u}^2}{2} \right) - (1 - \pi_p)\rho^R_g, \frac{1}{2} \right] \text{ and } \\
& k \geq \pi_p \left( \frac{1+3\hat{u}-\hat{u}^3}{3} \right) + (1 - \pi_p)\rho^R_g
\]

\[
(1, 1) & \text{if } k \in \left[ \frac{1}{2} - \pi_p \left( \frac{1+2\hat{u}-\hat{u}^2}{2} \right) - (1 - \pi_p)\rho^R_g, \frac{1}{2} \right] \text{ and } \\
& k \geq \pi_p \left( \frac{1+3\hat{u}-\hat{u}^3}{6} \right) + (1 - \pi_p)\rho^R_g
\]

\[
(1, 2) & \text{if } k \in \left[ \frac{1}{2} - \pi_p \left( \frac{1+3\hat{u}-\hat{u}^3}{6} \right) - (1 - \pi_p)\rho^R_g, \frac{1}{2} \right] \text{ and } \\
& k \geq \pi_p \left( \frac{1+3\hat{u}-\hat{u}^3}{6} \right) + (1 - \pi_p)\rho^R_g
\]

\[
(2, 0) & \text{if } k \in \left[ 1 - \pi_p \left( \frac{2+3\hat{u}-3\hat{u}^2+\hat{u}^3}{3} \right) - (1 - \pi_p)\rho^R_g, \frac{1}{2} \right] \text{ and } \\
& k \geq \pi_p \left( \frac{2+3\hat{u}-3\hat{u}^2+\hat{u}^3}{3} \right) + (1 - \pi_p)\rho^R_g
\]

\[
(2, 1) & \text{if } k \in \left[ \frac{1}{2} - \pi_p \left( \frac{2+3\hat{u}-3\hat{u}^2+\hat{u}^3}{3} \right) - (1 - \pi_p)\rho^R_g, \frac{1}{2} \right] \text{ and } \\
& k < \pi_p \left( \frac{2+3\hat{u}-3\hat{u}^2+\hat{u}^3}{6} \right) + (1 - \pi_p)\rho^R_g
\]

\[
(2, 2) & \text{if } k < \frac{1}{2} - \pi_p \left( \frac{2+3\hat{u}-3\hat{u}^2+\hat{u}^3}{6} \right) - (1 - \pi_p)\rho^R_g.
\]

In the statement of Proposition 1, the order of presentation of each \((n^*_R, n^*_L)\) pair is roughly monotonic in the range of values of \(k\) supporting the pair. Thus, \((0, 1)\) requires the highest levels of \(k\), and \((2, 2)\) the lowest. While the conditions supporting each pair are occasionally cumbersome, they simply reflect the trade-off between a potential candidate’s cost of entry and her chances of achieving office. For example, \((n^*_R, n^*_L) = (2, 0)\) requires that \(k < 1/2\), so that each party \(R\) entrant can receive expected utility \(1/2 - k > 0\). Likewise, the lower bound on \(k\) is the probability that the party \(L\) candidate wins when \((n_R, n_L) = (2, 1)\). This value deters each party \(L\) potential candidate from entering.

It is clear from (4) that the equilibrium number of entrants is often non-unique. It may also be locally non-monotonic in \(k\) (and \(\hat{u}\) as well). While only one potential candidate enters for \(k\) sufficiently high, and all potential candidates enter for \(k\) sufficiently low, the number of entrants within each party can vary widely over moderate values \(k\).
To derive stronger predictions about party primary strategies, we examine a simple equilibrium selection rule. For any \( k \), we select the equilibrium with the maximum number of entrants from the favored party. This refinement would be consistent with an environment in which party \( R \)'s potential candidates were allowed to declare interest first, and party \( L \)'s potential candidates reacted to these announcements.

Using the derivations from the proof of Proposition 1, it is easily shown that this refinement eliminates the possibility that \((n^*_R, n^*_L) = (1, 2), (0, 2)\) or \((0, 1)\). It also allows \((n^*_R, n^*_L) = (1, 2)\) only when \( \hat{u} \) or \( \pi_p \) is sufficiently low. All of the other expressions in Proposition 1 remain unchanged. Thus, the number of party \( R \) candidates is always at least the number of party \( L \) candidates. Additionally, the number of party \( L \) entrants is weakly decreasing in \( \hat{u} \), but the number of party \( R \) candidates is weakly increasing or constant. Figure 1 illustrates two examples of the predicted number of candidates under this refinement.

This refinement also allows us to identify the role of voter uncertainty over the candidate’s ideological positions. Suppose that due to incumbency or extensive media coverage, the distribution of the party \( R \) candidates’ ideal points was degenerate. The effect of this knowledge would be identical to that of shifting \( \overline{\mu}^R \), and consequently \( \hat{u} \). By the concavity of \( u(\cdot) \), a party \( R \) candidate located at \( y_{CR} = \mu^R \) would make the voter more favorably inclined toward party \( R \) than when she was uncertain over \( y_{CR} \) (i.e., effectively increasing \( \hat{u} \)). The revelation of a relatively moderate candidate would therefore increase the extent to which party \( R \) is favored, decrease the number of party \( L \) entrants, while the revelation of an extremist would increase the number of party \( L \) entrants.

Our main result establishes that the favored party will choose to implement primaries in ideologically extreme electorates. Primary elections disproportionately benefit heavily favored parties because its voters are unable to take advantage of high valence candidates from the unfavored party. The result is again restricted to electorates where \( \hat{u} \geq 0 \), although identical results for left-leaning districts obviously hold by symmetry. Part (i) has a stronger comparative static that assumes the equilibrium with the maximum number of favored party
candidates. The prediction is that ideological extremity is both necessary and sufficient.\(^8\) Parts (ii)-(iii) arrive at sufficient conditions that apply more broadly.

**Proposition 2** Primaries in extreme districts. *In districts where \( \hat{u} \in [0, 1) \):*

(i) In an equilibrium where the maximum number of favored party (R) candidates enter, there exists \( m^R \in [m, \overline{m}) \) such that \( e^* = 1 \) if and only if \( y_M \geq m^R \).

(ii) In all equilibria where \( n^*_R > 0 \), there exists \( m^{R'} \in [m, \overline{m}) \) such that \( e^* = 1 \) if \( y_M \geq m^{R'} \).

(iii) In all equilibria where \( n^*_R, n^*_L > 0 \), if \( \pi_g < 2/3 \), then there exists \( m^{R''} \in (m, \overline{m}) \) such that \( e^* = 0 \) if \( y_M \leq m^{R''} \). \( \blacksquare \)

A simple corollary of this result is that in the most centrist districts, an *unfavored* party would stand to gain from introducing primaries. This may occur in an environment where primary elections are imposed by an outside institutional player, such as a court. But since politicians from unfavored parties would be *ex ante* less likely to be elected into positions that would allow them to introduce primaries, it follows that primaries are most likely to be introduced first in relatively extreme constituencies.

Thus, Proposition 2 predicts that the effects of party representation across multiple districts will depend on whether primaries are implemented in a decentralized or centralized way. Suppose that there are many ideologically diverse districts, and that each could choose independently whether to introduce primaries. In this environment, the ability to introduce primaries would strictly reduce the unfavored party’s prospects in districts that are somewhat more moderate than \( \overline{m} \) and \( m \). Both parties’ fortunes in other districts would be remain the same, either because moderate districts will not adopt primaries or because extreme districts would never vote for the unfavored party. The overall reduction in the number of competitive districts would increase the perceived polarization of the electorate.\(^9\) If, however, primaries

---

\(^8\)The result can also be shown to hold for values of \( \hat{u} \) supporting any given combination of entrants.

\(^9\)These results are analogous to those of introducing party competition in the first place. Suppose that each challenger were drawn according to either \( g^L \) or \( g^R \) with equal probability, and that M did not know the distribution from which each challenger chosen. This “non-partisan” world puts extremist candidates at a disadvantage in ideologically compatible districts. However, moderate candidates might not see their electoral prospects change by much as their probability of election would remain near 1/2. Thus right-wing candidates benefit from both partisan competition and primaries in extreme districts, but not in moderate ones.
were imposed “nationally,” then the results would be more ambiguous as the unfavored party would benefit in the most moderate districts.

The following example illustrates the equilibrium and the effect of voter knowledge of a candidate’s ideal point.

*Example.* Let all players have quadratic utility, with \( u(x) = -x^2 \). Each party has uniformly distributed candidate ideal points, with party \( L \) candidates in \([-2, -1]\) and party \( R \) candidates in \([1, 2]\). Thus M’s expected utility from a zero-valence candidate from party \( L \) is: 
\[
\int_{-2}^{-1} - (y_M - x)^2 \, dx = -y_M^2 - 3y_M - 7/3.
\]
Likewise, her expected utility from a party \( R \) candidate is 
\[
-y_M^2 + 3y_M - 7/3.
\]
The expected difference in utility between two candidates drawn from their respective party distributions is then \( \hat{u} = 6y_M \).

Suppose that \( k = 0 \), so that two candidates from each party enter in equilibrium, and consider a district where \( y_M \geq 0 \), so that \( R \) is the favored party. Without primaries, \( C^R \) wins if valences are unrevealed. But if valences are revealed, \( C^L \) can win if \( v_{CL} - v_{CR} > 6y_M \). This cannot occur if \( y_M > 1/6 \), as any such voter automatically elects \( C^R \). Otherwise, \( C^L \) wins with probability \( (1 - 6y_M)^2/2 \). Party \( R \)'s *ex ante* probability of winning is therefore 
\[
\rho_g^R = 1 - \pi_g (1 - 6y_M)^2/2.
\]

With primaries, each party’s chances of winning remain the same if valences are not revealed at the primary stage. But if valences are revealed, party \( R \) wins if the valence of its best candidate exceeds that of party \( L \)'s by at least \( \hat{u} \). The density of the second order statistic of party \( L \)'s valences is \( 2v + 1 \ (v \in [-1/2, 1/2]) \), and so the probability of party \( R \) winning is:
\[
\int_{-1/2}^{1/2} (1 - \max\{0, v - \hat{u} + 1/2\})^2 (2v + 1) \, dv = \int_{-1/2}^{-1/2+\hat{u}} (2v + 1) \, dv + \int_{-1/2+\hat{u}}^{1/2} (1 - (v - \hat{u} + 1/2)^2) (2v + 1) \, dv
\]
\[
= 1/2 + 8y_M - 36y_M^2 + 216y_M^4.
\]
Note that since party \( R \)'s total probability of victory is simply a linear combination of the preceding expression and \( \rho_g^R \), the probability of valence revelation in the primaries \( \pi_p \) does...
not affect the induced preferences over primaries. To check when party \( R \) would prefer primaries, it is sufficient to compare the two probabilities of victory. It is then easily verified that party \( R \) prefers primaries if and only if \( y_M \) is greater than some threshold \( m^R \). This threshold depends on \( \pi_g \), with primaries becoming better as \( \pi_g \) increases. At \( \pi_g = 1 \), \( m^R = 0 \), while at \( \pi_g = 0.5 \), \( m^R = 0.097 \).

Finally, consider an environment in which the party \( R \) candidate’s ideological position was known \textit{ex ante}. Then M’s utility from voting for party \( R \) is \(- (y_M - y_{CR})^2 \), which implies that party \( R \)’s valence advantage is \(- y_{CR}^2 + (3 + 2y_{CR})y_M + 7/3 \). This implies a party \( R \) valence advantage higher than \( \hat{u} \) if \( y_{CR} = 1.5 \), and if \( y_{CR} = 1 \) (resp., \( = 2 \)) and \( y_M < 4/3 \) (resp., \( > 5/3 \)). Under these conditions, party \( L \) still has two entrants (since \( k = 0 \)), but party \( R \)’s probability of victory increases relative to the case where M is uncertain over \( y_{CR} \).

The example illustrates that the probability of valence revelation in the general election affects \( m^R \), and hence the types of districts that would adopt primaries. The intuition for this is that a favored party is at risk in a “moderate” district when \( \pi_g \) is high. This creates an incentive to increase the number of draws through primaries. By contrast, when \( \pi_g \) is low, the favored party is more inclined to allow the voter to act on her prior beliefs about the distributions of ideal points.

We observe finally that from the perspective of voter welfare, it can be shown that all voters benefit from primaries. Primary elections strictly increase the probability of valence revelation, and this can never hurt the voter. However, whether partisan or moderate voters benefit more depends on \( \pi_g \), which determines the added value of revealing more information in the primaries. To see this, consider two districts with voter preferences at \( \hat{u} = 0 \) and \( \hat{u} = 1 \). Recall that the \( k \)-th order statistic from \( k \) draws from a uniform distribution is distributed according to \( \beta(k, 1) \), with expected value \( k/(k + 1) \). In the former district, the voter is indifferent between parties and receives expected valence \((1 - \pi_g)(0) + \pi_g(1/6)\) without primaries. With primaries, she receives \( \pi_p(3/10) + (1 - \pi_p)[(1 - \pi_g)(0) + \pi_g(1/6)] \), for an expected gain of \( \pi_p(3/10 - \pi_g/6) \). In the latter district, the voter always chooses the party \( R \) candidate and receives expected valence \( 0 \) without primaries. With primaries, the
expected valence is $\pi_p/6$, which exceeds the expected gain enjoyed by the moderate district if $\pi_g > 0.8$. In this case, if primaries are costly to implement, then “partisan” districts might be more likely to hold them.\textsuperscript{10} By contrast, the moderate district will benefit more from primaries when $\pi_g$ is low, which increases the expected marginal informational contribution of the primary election.

4 Evidence From U.S. Primaries

In this section we use data from the U.S. states to examine several hypothesis generated by the model above. The data used cover primary and general elections for almost all states over the period 1888-2006.\textsuperscript{11} Details about the data, including sources, can be found in Ansolabehere and Snyder (2002) and Ansolabehere et al. (2006).

Proposition 2 shows that extremist constituencies have more to gain from primaries than moderate constituencies. By implication, extremist constituencies should be more likely to use primaries than moderate constituencies, and there should be more competition in their primaries. The U.S. states provide a crude laboratory for testing this hypothesis. Most states passed some kind of primary law during the period 1898-1915. Some of these laws made primaries mandatory and some did not; some covered all statewide elected offices and some were more limited; and some provided for closed primaries and some were more open.

We ask the question: what is the correlation between one-party domination and primary activity? More specifically, were states that passed strong, competition-enhancing, primary laws during this period likely to be dominated by one party (a sign of “extremism”), and were than the states that passed “weak” laws, or no primary laws, states where the two major parties were more evenly matched (a sign of “moderation”)?

We divide states into two classes, Primary States and Other States, defined as follows.

\textsuperscript{10}It is not clear why primary elections should be any more expensive than general elections. One possibility is that the low-information environment leads to an especially large amount of (mainly wasteful) political campaigning.

\textsuperscript{11}We drop LA after 1974 since it uses a unique, non-partisan, runoff system in place of a partisan general election.
Call a primary election “competitive” if the winner wins with no more than 60% of the vote. Consider the nominations for all partisan elected statewide executive offices and U.S. Senate seats held during the period 1900-1930. Let $D_i$ be the fraction of Democratic nominees for these elections who were chosen in a “competitive” primary in state $i$, and let $R_i$ be the analogous fraction for Republicans. Then $i$ is a Primary State if and only if $\max\{D_i, R_i\} > .10$. Most states fall into this category, but the following do not: CT, DE, IN, KY, NM, NY RI, UT. This is the set of Other States – i.e., those that made little or no use of primaries. Some of these states did not hold any primaries during the period 1900-1930, or even have a primary law (CT, NM, RI, UT), others excluded many offices from their primary law (IN, NY), and others had barriers to entry and/or or party machines that severely limited primary competition (IN, KY, NY).

To measure the degree of inter-party competition we use the variable Vote Margin, defined as follows. Let $V_{ij}$ be the Democratic share of the two-party vote in race $j$ in state $i$. Let $\mathcal{J}_i$ be the set of all races for partisan elected statewide executive offices and U.S. Senate seats held in state $i$ during the period 1890-1916 and prior to the regular use of direct primaries. Let $J_i$ be the number of races in $\mathcal{J}_i$, and let $\bar{V}_i = (1/J_i) \sum_{j \in \mathcal{J}_i} V_{ij}$ be the average Democratic vote share over the races in $\mathcal{J}_i$. Then $\text{Vote Margin}_i = |\bar{V}_i - .5|$. High values of $\text{Vote Margin}$ mean that one party is heavily favored in the state, while low values indicate a competitive situation.

Table 1 presents the results of a simple analysis correlating Primary State and Vote Margin. The table is divided into two panels. Both panels show clearly that the states that did not pass strong primary laws during the Progressive era tended to those in which voters were relatively evenly divided among the two major parties – i.e., states that were

---

12If a primary is uncontested – or if the nomination is made by caucus or convention – then the winner’s vote percentage is 100%, so the primary is not competitive.

13Two other states – MI and SD – are also excluded many statewide offices from their primary law, but had such competitive primaries in the offices covered by the law that they easily fall into the Primary State category. Also, ID held no primaries from 1920-1930, but had such competitive primaries during the period 1910-1918 that it, too, easily falls into the Primary State category.

14We drop races where a minor-party candidate or independent candidate received more than 15% of the vote. Two states – AZ and OK – introduced the direct primary at statehood. For these states we use the elections between statehood and 1916. The results are similar if we drop them from the analysis.
“moderate” from a partisan point of view. The first panel shows a simple difference-in-means.

The second panel shows another cut of the data. The median value of Vote Margin is .068. Call the 24 states with Vote Margin below the median Two-Party States, and call the 24 states with Vote Margin above the median One-Party States. The second panel shows the two-by-two contingency table, and corresponding chi-square test. All but one of the One Party States passed a strong primary law during the Progressive era. A majority of the Two Party States did as well, but a much larger percentage almost 30% – did not.

Proposition 2 shows that the introduction of direct primaries benefits the favored party in a state, causing it to win even more elections than before. (In fact, taken literally the proposition implies that the introduction of the direct primary essentially wipes out the minority party.) Again, the U.S. states provide a crude laboratory for testing this hypothesis. The idea is to look just before and just after the introduction of direct primaries in a state, and see what happens to the average vote-share of the party that was advantaged at the time the primary was adopted.

As above, let $V_{ij}$ be the Democratic share of the two-party vote in race $j$ in state $i$. Let $J_i^0$ be the set of all races for partisan elected statewide executive offices and U.S. Senate seats held in state $i$ during the 10 years just prior to the regular use of direct primaries. And, let $J_i^1$ be the set of all races for partisan elected statewide executive offices and U.S. Senate seats held in state $i$ during the 10 years just after the regular use of direct primaries. Let $J_i^0$ be the number of races in $J_i^0$, and let $\bar{V}_i^0 = (1/J_i^0) \sum_{j \in J_i^0} V_{ij}$ be the average Democratic vote share over the races in $J_i^0$. Similarly, let $J_i^1$ be the number of races in $J_i^1$, and let $\bar{V}_i^1 = (1/J_i^1) \sum_{j \in J_i^1} V_{ij}$ be the average Democratic vote share over the races in $J_i^1$.

Since we are studying changes in the vote, we must control for the effects of party tides. To do this, let $\bar{V}_{Ni}^0$ be the average Democratic vote share nationwide across all races during the 12 years just before the regular use of direct primaries in state $i$, and let $\bar{V}_{Ni}^1$ be the average Democratic vote share nationwide across all races during the 12 years just after the

\footnote{Again, we drop races where a minor-party candidate or independent candidate received more than 15% of the vote.}
regular use of direct primaries in state \( i \). The variables we analyze are \( \text{Rel Dem Vote Before Primaries}_i = \bar{V}_i^0 - \bar{V}_{Ni}^0 \) and \( \text{Change in Rel Dem Vote}_i = (\bar{V}_i^1 - \bar{V}_{Ni}^1) - (\bar{V}_i^0 - \bar{V}_{Ni}^0) \).

Table 2 presents the results. The coefficient on \( \text{Rel Dem Vote Before Primaries} \) shows clearly that the party that was favored by voters prior to the use of primaries tended to do even better after primaries were introduced. On average, the Democratic vote share increased in states that were already leaning Democratic, and the Republican vote share increased in states that were already leaning Democratic. A simple counting exercise shows the same pattern: The Democratic vote-share increased in 8 of the 10 most Democratic states, and the Republican vote-share increased in 8 of the 10 most Republican states.

We can also test the hypothesis that extremist constituencies use primaries more than moderate constituencies by studying the patterns of primary competition over the entire period primaries have been used in the U.S. We can use this data to study within-state variation over time, as well as variation across states.

We consider the period 1926-2006, after primaries had been established for a decade in all of the states that passed laws during the main stage of activity, 1898-1915. We consider races for U.S. Senator and the following statewide executives: governor, lieutenant governor, attorney general, secretary of state, treasurer, and auditor/controller/comptroller. We restrict attention to these because they are the most common elective statewide offices.

The dependent variable is \( \text{Competitive Primary} \), which we define as a primary in which the winner wins with no more than 60% of the vote.

The main independent variables are three dummy variables, \( \text{Democrats Favored} \), \( \text{Republicans Favored} \), and \( \text{Competitive State} \), defined as follows. Let \( J_{it} \) be the set of all of the races listed above held in state \( i \) during the 10-year period \( t-10 \) to \( t-1 \). Let \( J_{it} \) be

\[ \text{Democrats Favored}_i = \begin{cases} 1 & \text{if Democrats were favored in } J_{it} \text{ races} \\ 0 & \text{otherwise} \end{cases} \]

\[ \text{Republicans Favored}_i = \begin{cases} 1 & \text{if Republicans were favored in } J_{it} \text{ races} \\ 0 & \text{otherwise} \end{cases} \]

\[ \text{Competitive State}_i = \begin{cases} 1 & \text{if there were competitive primaries in } J_{it} \text{ races} \\ 0 & \text{otherwise} \end{cases} \]

---

16 The exact period chosen does not really matter. We also considered, for example, entire period 1915-2006, the period 1940-2006 (NM and UT had primaries by then), and the post-war period, 1946-2006.

17 All but 7 states have had an elected attorney general at some point during the period studied; all but 10 states have had an elected secretary of state; all but 9 have had an elected treasurer, and all but 10 have had an elected auditor/controller/comptroller. Also, all but 11 states have had an elected lieutenant governor that was elected separately from the governor.

18 As above, if a primary is uncontested – or if the nomination is made by caucus or convention – then the winner’s vote percentage is 100%, so the primary is not competitive.

19 Again, we drop races where a minor-party candidate or independent candidate received more than 15%
the number of races in $J_{it}$, and let $V_{it} = (1/J_{it}) \sum_{j \in J_{it}} V_{ij}$ be the average Democratic vote share over the races in $J_{it}$. Then Democrats Favored$_{it} = 1$ if and only if $V_{it} > .58$, Republicans Favored$_{it} = 1$ if and only if $V_{it} < .42$, and Competitive State$_{it} = 1$ if and only if $.42 \leq V_{it} \leq .58$.\textsuperscript{20}

Finally, we also include a control variable to distinguish between races where incumbents are running and races where they are not. While our model does not account for this, previous empirical work finds that incumbents often get a “free ride” in the primaries when seeking renomination. We therefore include the variable Incumbent Running, which is 1 if an incumbent is running for re-election in the primary and 0 otherwise.

Table 3 presents the results. The table shows four specifications, two for Democrats (columns 1 and 2) and two for Republicans (columns 3 and 4). For each party, one specification is a pooled cross-section/time-series regression with the main variables of interest plus year fixed-effects (columns 1 and 3). The other is a panel regression that includes state fixed effects in addition to the year fixed-effects (columns 2 and 4). In all specifications the standard errors are clustered by state to deal correct for heteroskedasticity and general within-state autocorrelation. Note that in each specification we include only two of the three variables Democrats Favored, Republicans Favored, and Competitive State, since collectively they are perfectly multi-collinear.

The main hypothesis is that competition should be highest in the primaries of the favored party, in states that have a favored party. The estimates in Table 3 imply that this is the case. For Democrats, the coefficients on Democrats Favored are positive and statistically significant, and also larger than the coefficients on Competitive State. For the specification in column 1 a test of the hypothesis that the coefficient on Democrats Favored is equal to the coefficient on Competitive State yields an F-statistic of 36.49 which is significant at the .001 level. For the specification in column 2 this test yields an F-statistic of 8.74, which is significant at the .004 level. Thus, if the partisan balance in a state increases over time,
say from a *Democrats Favored* to *Competitive State*, then competition in the Democratic primary falls.

Columns 3 and 4 show qualitatively similar results for Republicans. The coefficients on *Republicans Favored* are positive and statistically significant, and also larger than the coefficients on *Competitive State*. The F-tests comparing the coefficients on *Republicans Favored* and *Competitive State* reject the null hypothesis of equality at the .05 level or better. For the specification in column 3 the F-statistic of 18.78, which is significant at the .001 level; and for column 4 the F-statistic of 3.99, which is significant at the .05 level. Note, though, that the estimated effects are uniformly smaller for Republicans than for Democrats.

There are two features of primary and general election competition that our model does not predict. First, the overall *level* of primary competition is low, at least compared to the level of general election competition. Our model predicts that primary elections should be quite competitive.

Second, the level of primary competition has fallen steadily over the past 100 years in races involving an incumbent, but it has remained roughly constant in open-seat races (*e.g.*, Ansolabehere *et al.*, 2006). Our model does predict an overall drop in primary competition, since the number of one-party states has also fallen. For example, for the period 1926-1960, only 33% of the state-year observations fall into the *Competitive State* category, but for the period 1961-2006 almost 55% do. However, a large part of the decline is unexplained. And, since our model does not directly consider races with incumbents, it cannot address the difference between incumbent-contested and open-seat races.

The data suggest that primaries entail substantial costs, and that the costs might have grown over time. What are these costs? First, there are the costs of voting, and of expensive and possibly wasteful campaigning. These are non-trivial – for example, citing costs, several states do not bother to hold primaries for unopposed candidates – but they do not seem huge. Another cost, at least to one party and it supporters, is the possibility that “divisive” primaries can hurt the party’s nominees in the general election. A number of empirical studies
suggest that closely contested primaries damage nominees. Mud-slinging campaigns may leave the nominee with a negative image even among many voters normally loyal to the party, may anger party activists who supported the losing candidate so they refuse to work for the nominee, and may drain the nominee’s funds.

5 Discussion

The model developed here combines a strategic voter with partisan politicians in a simple framework to study the impact of an open party nomination procedure. The voter chooses between candidates whose ideological preferences are roughly identified by parties but whose valences are unknown. Both general and primary elections may reveal their valence parameters. In this environment, primaries can help the voter to select better politician types, but parties may not benefit from allowing voters to do so. If the voter is certain to vote for a favored party, then that party benefits because primaries allow the voter to select the better of its candidates. But in an environment where a voter might vote for either party, the information revealed through primaries might hurt the favored party’s candidates. As a result, a party will implement primaries wherever voters share their policy preferences. This pattern is broadly supported by data from U.S. primaries.

Several natural extensions remain for future work. The first is to add ex post incentives to discipline incumbents for their performance in office to the ex ante incentives to select good politician types examined here. It is likely that in this setting, the introduction of primaries would strengthen the voter’s incentive to select good types, because a voter no longer needs to discipline a “bad” incumbent from a favored party by electing a candidate from an unfavored party. We would therefore expect that primaries would continue to allow the voter to be more selective, especially when a voter matches the ideology of one of the

21 The existing empirical research is mixed, however. Johnson and Gibson (1974), Bernstein (1977), and Kenney and Rice (1987) find that close primaries hurt candidates, but Hacker (1965), Piereson and Smith (1975), Miller, Jewell, and Sigelman (1988), and Atkeson (1998) find they do not. Born (1981) finds that close primaries hurt incumbents but not challengers. Also, none of these studies fully corrects for the endogeneity of divisive primaries.
parties closely. But additionally, it would reduce the ability of incumbent office-holders to shirk.

Another extension concerns the existence of parties. Parties are exogenous in our model, but they clearly help the electoral prospects of certain types of candidates. A more satisfying treatment of parties would consider the affiliation incentives of their membership (i.e., the set of potential candidates). Such a model would provide a more complete understanding of the role that parties play in the presence of strategic voters.
Appendix

Proof of Lemma 1. Suppose to the contrary that for some $C^{ij}$ and $d_{C^{ij}}$, $\gamma_{C^{ij}}(d_{C^{ij}}) \neq \Gamma^{i}$. This implies that $d_{C^{ij}}$ is not constant in realized ideal points $y$. For any $y$ such that $d_{C^{ij}} = 1$, the probability of $C^{ij}$ winning the general election must be at least $k$. Further, since potential candidates break ties in favor of not entering the election, the probability of $C^{ij}$ winning the general election must be strictly greater than $k$. Thus for any type $y'$ such that $d_{C^{ij}} = 0$, $C^{ij}$ could choose $d_{C^{ij}} = 1$ instead. Since the expected payoff for any such $C^{ij}$ does not depend on $y$, a type $y'$ potential candidate would also expect a strictly positive payoff by entering: contradiction.

Proof of Proposition 1. Since $M$ clearly votes for $C^{R}$ when $y_{M} \geq \bar{m}$, it is clear that $n_{L}^{\ast} = 0$ and $n_{R}^{\ast} = 1 (= 2)$ if $k \geq 1/2 (< 1/2)$ for all such $y_{M}$.

For the remainder of the proof, we restrict attention to $y_{M} \in [m, \bar{m})$. Notationally, let $\rho_{p}^{R}(n_{R}, n_{L})$ be the ex ante probability that a party $R$ candidate wins the general election, given $n_{R}$ party $R$ entrants and $n_{L}$ party $L$ entrants.

We first calculate $\rho_{p}^{R}(\cdot, \cdot)$ for each possible configuration of entrants. Since each candidate receives one unit of utility for winning the general election, the expected utility of a party $i$ entrant is simply $\rho_{p}^{i}(n_{R}^{\ast}, n_{L}^{\ast})/n_{i} - k$, and thus a potential candidate will enter if this value is anticipated to be positive. Note that if a party has zero candidates, then that party loses with certainty. We therefore focus on the four nontrivial cases: $(n_{R}, n_{L}) = (2, 2), (2, 1), (1, 2), and (1, 1)$.

Case 1: $(2, 2)$. We calculate $\rho_{p}^{R}(2, 2)$ by integrating over realizations of $\max\{v_{C^{R1}}, v_{C^{R2}}\}$, the highest valence among party $R$’s candidates. This value is represented by the random variable $V_{(2)}$, the second order statistic from two i.i.d. draws from the $U[-1/2, 1/2]$ distribution. Thus $V_{(2)} = V - 1/2$, where $V \sim \beta(2, 1)$. The density of $V_{(2)}$ is $2v + 1$ for $v \in [-1/2, 1/2]$, and 0 elsewhere. Hence:

$$
\rho_{p}^{R}(2, 2) = (1 - \pi_{p})\rho_{g}^{R} + \pi_{p}\int_{-1/2}^{1/2} (\min\{1, \max\{0, v + \hat{u} + 1/2\}\})^{2}(2v + 1)dv. \tag{5}
$$
Observe that \( \hat{u} \geq 0 \) implies that \( v + \hat{u} + 1/2 \geq 0 \). Now letting \( w = v + 1/2 \) and changing variables yields:

\[
\rho_p^R(2, 2) = \pi_p \int_0^1 (\min\{1, w + \hat{u}\})^2 2wdw + (1 - \pi_p)\rho_g^R
\]

\[
= \pi_p \left[ \int_0^{1-\hat{u}} (w + \hat{u})^2 2wdw + \int_{1-\hat{u}}^1 2wdw \right] + (1 - \pi_p)\rho_g^R
\]

\[
= \pi_p \left[ \frac{w^4}{2} + \frac{4\hat{u}w^3}{3} + \hat{u}^2 w^2 \bigg|_{0}^{1-\hat{u}} + w^2 \bigg|_{1-\hat{u}}^1 \right] + (1 - \pi_p)\rho_g^R
\]

\[
= \pi_p \left[ \frac{\hat{u}^4}{6} - \hat{u}^2 \frac{4\hat{u}}{3} + \frac{1}{2} \right] + (1 - \pi_p)\rho_g^R. \tag{6}
\]

Analogously, party \( L \)'s probability of victory is:

\[
\rho_p^L(2, 2) = 1 - \pi_p \left[ \frac{\hat{u}^4}{6} - \hat{u}^2 \frac{4\hat{u}}{3} + \frac{1}{2} \right] - (1 - \pi_p)\rho_g^R.
\]

Case 2: \((2, 1)\). Analogously to the first case, \( \rho_p^R(2, 1) \) can be written as follows:

\[
\rho_p^R(2, 1) = (1 - \pi_p)\rho_g^R + \pi_p \int_{-1/2}^{1/2} \min\{1, \max\{0, v + \hat{u} + 1/2\}\} (2v + 1) dv. \tag{7}
\]

Observe that \( \hat{u} \geq 0 \) implies that \( v + \hat{u} + 1/2 \geq 0 \). Letting \( w = v + 1/2 \) and changing variables yields:

\[
\rho_p^R(2, 1) = \pi_p \int_0^1 \min\{1, w + \hat{u}\} 2wdw + (1 - \pi_p)\rho_g^R
\]

\[
= \pi_p \left[ \int_0^{1-\hat{u}} (w + \hat{u})^2 2wdw + \int_{1-\hat{u}}^1 2wdw \right] + (1 - \pi_p)\rho_g^R
\]

\[
= \pi_p \left[ \frac{2w^3}{3} + \hat{u}^2 w^2 \bigg|_{0}^{1-\hat{u}} + w^2 \bigg|_{1-\hat{u}}^1 \right] + (1 - \pi_p)\rho_g^R
\]

\[
= \pi_p \left[ \frac{\hat{u}^3}{3} - \hat{u}^2 + \frac{2}{3} \right] + (1 - \pi_p)\rho_g^R. \tag{8}
\]

Analogously, party \( L \)'s probability of victory is:

\[
\rho_p^L(2, 1) = 1 - \pi_p \left[ \frac{\hat{u}^3}{3} - \hat{u}^2 + \frac{2}{3} \right] - (1 - \pi_p)\rho_g^R.
\]

Case 3: \((1, 2)\). We calculate \( \rho_p^R(1, 2) \) by integrating over realizations of \( v_{C,R} \). Now \( \rho_p^R \) can be written as follows:

\[
\rho_p^R(1, 2) = (1 - \pi_p)\rho_g^R + \pi_p \int_{-1/2}^{1/2} \min\{1, \max\{0, v + \hat{u} + 1/2\}\} (2v + 1) dv. \tag{9}
\]
Observe that \( \hat{u} \geq 0 \) implies that \( v + \hat{u} + 1/2 \geq 0 \). Letting \( w = v + 1/2 \) and changing variables yields:

\[
\rho_p^R(1, 2) = \pi_p \int_0^1 (\min\{1, w + \hat{u}\})^2 dw + (1 - \pi_p)\rho_g^R
\]
\[
= \pi_p \left[ \int_0^{1-\hat{u}} (w + \hat{u})^2 dw + \int_{1-\hat{u}}^1 dw \right] + (1 - \pi_p)\rho_g^R
\]
\[
= \pi_p \left[ \frac{(w + \hat{u})^3}{3} \bigg|_0^{1-\hat{u}} + w\bigg|_{1-\hat{u}}^1 \right] + (1 - \pi_p)\rho_g^R
\]
\[
= \pi_p \left[ -\frac{\hat{u}^3}{3} + \hat{u} + \frac{1}{3} \right] + (1 - \pi_p)\rho_g^R. \tag{10}
\]

Analogously, party L’s probability of victory is:

\[
\rho_p^L(1, 2) = 1 - \pi_p \left[ -\frac{\hat{u}^3}{3} + \hat{u} + \frac{1}{3} \right] - (1 - \pi_p)\rho_g^R.
\]

Case 4: \((1, 1)\). Following the revelation of valence in the primaries, each candidate’s probability of victory is unaffected by the general election. Party R’s probability of victory conditional upon revelation in the primaries is therefore identical to its probability of victory conditional upon revelation in the generals. Using (2), this probability is \((1 + 2\hat{u} - \hat{u}^2)/2\).

The ex ante probabilities of victory are then:

\[
\rho_p^R(1, 1) = \pi_p \left[ \frac{1 + 2\hat{u} - \hat{u}^2}{2} \right] + (1 - \pi_p)\rho_g^R \tag{11}
\]
\[
\rho_p^L(1, 1) = 1 - \pi_p \left[ \frac{1 + 2\hat{u} - \hat{u}^2}{2} \right] - (1 - \pi_p)\rho_g^R
\]

We now derive best responses for each party’s \(i\)’s potential candidates, given the anticipated entry of \(n_{-i}\) candidates from the opposing party. Note that \(\rho^R(2, 1) > \rho^R(2, 2) > \rho^R(1, 1) > \rho^R(1, 2)\) and \(\rho^L(2, 1) < \rho^L(2, 2) < \rho^L(1, 1) < \rho^L(1, 2)\). Within party R, no potential candidate will enter if \(k \geq \rho_p^R(1, n_L)\), one potential candidate will be willing to enter if \(k < \rho_p^R(1, n_L)\), and both will be willing to enter if \(k < \rho_p^R(2, n_L)\). Using the expressions for \(\rho_p^R(\cdot, n_L)\), the best responses for party R’s potential candidates are then:

Given \(n_L = 0\): \(n_R = 1\) if \(k \geq 1/2\); \(n_R = 2\) if \(k < 1/2\).
Given $n_L = 1$: $n_R = 0$ if $k \geq \rho_p^R(1, 1)$; $n_R = 1$ if $k \in [\rho_p^R(2, 1)/2, \rho_p^R(1, 1))$; $n_R = 2$ if $k < \rho_p^R(2, 1)/2$.

Given $n_L = 2$: $n_R = 0$ if $k \geq \rho_p^R(1, 2)$; $n_R = 1$ if $k \in [\rho_p^R(2, 2)/2, \rho_p^R(1, 2))$; $n_R = 2$ if $k < \rho_p^R(2, 2)/2$.

For party $L$, no potential candidate will enter if $k \geq \rho_p^L(n_R, 1)$, one potential candidate will be willing to enter if $k < \rho_p^L(n_R, 1)$, and both will be willing to enter if $k < \rho_p^L(n_R, 2)/2$. Using the expressions for $\rho_p^L(n_R, \cdot)$, the best responses for party $L$’s potential candidates are then:

Given $n_R = 0$: $n_L = 1$ if $k \geq 1/2$; $n_L = 2$ if $k < 1/2$.

Given $n_R = 1$: $n_L = 0$ if $k \geq \rho_p^L(1, 1)$; $n_L = 1$ if $k \in [\rho_p^L(1, 2)/2, \rho_p^L(1, 1))$; $n_L = 2$ if $k < \rho_p^L(1, 2)/2$.

Given $n_R = 2$: $n_L = 0$ if $k \geq \rho_p^L(2, 1)$; $n_L = 1$ if $k \in [\rho_p^L(2, 2)/2, \rho_p^L(2, 1))$; $n_L = 2$ if $k < \rho_p^L(2, 2)/2$.

We can now combine the best responses to derive pure strategy equilibria. There are two cases. For each, we divide the set of entry costs into six regions. For each region, we identify the set of possible best response pairs $(n_R, n_L)$ as well as any parametric restrictions that may exist.

*Case (i):* $\rho_p^L(2, 2)/2 < \rho_p^L(2, 1) < \rho_p^L(1, 2)/2 < \rho_p^L(1, 1) < 1/2$.

- $k \geq 1/2$: $(1, 0)$ for all such $k$, and $(0, 1)$ for $k \geq \rho_p^R(1, 1)$.
- $k \in [\rho_p^L(1, 1), 1/2)$: $(2, 0)$ for all such $k$, and $(0, 2)$ for $k \geq \rho_p^R(1, 2)$.
- $k \in [\rho_p^L(1, 2)/2, \rho_p^L(1, 1))$: $(2, 0)$ for all such $k$, and $(1, 1)$ if $k \geq \rho_p^R(2, 1)/2$.
- $k \in [\rho_p^L(2, 1), \rho_p^L(1, 2)/2)$: $(2, 0)$ for all such $k$, and $(1, 2)$ if $k < \rho_p^R(1, 2)$ (which always holds) and $k \geq \rho_p^R(2, 2)/2$.
- $k \in [\rho_p^L(2, 2)/2, \rho_p^L(2, 1))$: $(1, 2)$ if $k < \rho_p^R(1, 2)$ (which always holds) and $k \geq \rho_p^R(2, 2)/2$, and $(2, 1)$ if $k < \rho_p^R(2, 1)/2$ (which always holds).

- $k < \rho_p^L(2, 2)/2$: $(2, 2)$ if $k < \rho_p^R(2, 2)/2$ (which always holds), and $(1, 2)$ if $k \in [\rho_p^R(2, 2)/2, \rho_p^R(1, 2))$.

The latter condition can never hold, and thus $(2, 2)$ is the unique prediction.

*Case (ii):* $\rho_p^L(2, 2)/2 < \rho_p^L(1, 2)/2 < \rho_p^L(2, 1) < \rho_p^L(1, 1) < 1/2$. Note that the only
difference between this case and Case (i) is that $\rho^*_p(1, 2)/2 < \rho^*_p(2, 1)$, which occurs when either $\hat{u}$ or $\pi_p$ is sufficiently low and $\pi_g > 0$. The analysis is therefore identical to that of Case (i), except for the following regions:

$k \in [\rho^*_p(2, 1), \rho^*_p(1, 1))$: (2, 0) for all such $k$, and (1, 1) if $k \geq \rho^*_p(2, 1)/2$.

$k \in [\rho^*_p(1, 2)/2, \rho^*_p(2, 1))$: (2, 1) if $k < \rho^*_p(2, 1)/2$ (which now does not always hold), and (1, 1) if $k \geq \rho^*_p(2, 1)/2$.

$k \in [\rho^*_p(2, 2)/2, \rho^*_p(1, 2)/2)$: (1, 2) if $k < \rho^*_p(1, 2)$ (which always holds) and $k \geq \rho^*_p(2, 2)/2$, and (2, 1) if $k < \rho^*_p(2, 1)/2$ (which always holds).

Combining these results and substituting the appropriate values for $\rho^*_p(.)$ yields the result.

\[\bullet\]

**Proof of Proposition 2.** Since we assume (without loss of generality) $y_M \geq m$, party $R$ is the favored party and chooses $e$. Notationally, let $\rho^*_p(n_R, n_L)$ be the *ex ante* probability that a party $R$ candidate wins the general election.

(i) We first identify the possible numbers of candidates as $\hat{u}$ increase, starting with $\hat{u} = 0$. To do this, it is first necessary to list the number of candidates that enter under the refinement that chooses the equilibrium with the maximum number of favored-party entrants. Adapting from the proof of Proposition 1:

\[
(n^*_R, n^*_L) = \begin{cases} 
(1, 0) & \text{if } k \geq \frac{1}{2} \\
(2, 0) & \text{if } k \in \left[ 1 - \pi_p \left( \frac{2+3\hat{u}^2+\hat{u}^4}{3} \right) - (1 - \pi_p) \rho^*_p \frac{1}{2} \right] \\
(2, 1) & \text{if } k \in \left[ 1 - \pi_p \left( \frac{2+3\hat{u}^2+\hat{u}^4}{3} \right) - (1 - \pi_p) \rho^*_p \frac{1}{2} \right] \\
(2, 2) & \text{if } k < \frac{1}{2} - \pi_p \left( \frac{2+3\hat{u}^2+\hat{u}^4}{3} \right) - (1 - \pi_p) \rho^*_p \frac{1}{2} \end{cases}
\]

\[(12)\]

Three observations simplify the characterization of possible paths. First, each of the right-hand side expressions in (12) with the exception of $\pi_p \left( \frac{2+3\hat{u}^2+\hat{u}^4}{6} \right) + (1 - \pi_p) \rho^*_p \frac{1}{2}$ are
non-intersecting and non-increasing. Second, at \( \hat{u} = 1 \), \((n^*_R, n^*_L) = (1,0)\) or \((2,0)\). Third, since \( \rho_p^R(2,1) > \rho_p^L(1,2) \), the requirement that \( k \geq \rho_p^L(1,2)/2 \) for the \((1,1)\) equilibrium in non-binding.

Now given these equilibrium entry strategies, there are six possible paths for \((n^*_R, n^*_L)\) as \( \hat{u} \) increases:

\[
(n^*_R, n^*_L) = \begin{cases} 
(1,0) & \text{if } k \geq \frac{1}{2} \\
(2,0) & \text{if } k \in \left[ \frac{\pi_p}{3} + \frac{(1-\pi_p)\mu_g}{2}, \frac{1}{3} \right] \\
(2,1), (2,0) & \text{if } k \in \left[ \frac{\pi_p}{4} + \frac{(1-\pi_p)\pi_g}{4}, \min \left\{ \frac{1}{2} - \frac{\pi_p}{6} - \frac{(1-\pi_p)\pi_g}{4}, \right\} \\
(1,1), (2,0) & \text{and} \\
(1,1), (2,1), (2,0) & \text{if } k \in \left[ \frac{1}{2} - \frac{\pi_p}{6} - \frac{(1-\pi_p)\pi_g}{4}, \frac{\pi_p}{3} + \frac{(1-\pi_p)\pi_g}{2} \right] \\
(2,2), (2,1), (2,0) & \text{if } k < \frac{\pi_p}{4} + \frac{(1-\pi_p)\mu_g}{4}. 
\end{cases}
\tag{13}
\]

Note that the paths starting with \((1,1)\) do not exist for all values of \(\pi_p, \pi_g\). We now proceed through the six cases in \((13)\).

**Case 1:** \((1,0)\). Clearly, party \( R \) wins the election, and is therefore indifferent between \( e = 0 \) and \( e = 1 \). The result follows trivially.

**Case 2:** \((2,0)\). Again, party \( R \) wins the election with certainty. Since a primary election results in a higher expected valence, the party strictly prefers \( e = 1 \).

**Case 3:** \((2,1), (2,0)\). Observe that since \( \rho_p^R(2,0) > \rho_p^R(2,1) \), given the result of case 2 it is sufficient to show that \( \rho_p^R(2,1) \geq \rho_g^R \) iff \( m \geq m^R \), where \( m^R \in [m, \overline{m}] \).

From the proof of Proposition 1, we have: \( \rho_p^R(2,1) = \pi_p \left[ \frac{\hat{u}^3}{3} - \hat{u}^2 + u + \frac{2}{3} \right] + (1-\pi_p)\rho_g^R \). Party \( R \) chooses primaries \((e^* = 1)\) iff \( \rho_p^R(n_R, n_L) \geq \rho_g^R \). Using \((2)\), \( \rho_p^R(n_R, n_L) \geq \rho_g^R \) iff:

\[
\frac{\hat{u}^3}{3} - \hat{u}^2 + u + \frac{2}{3} \geq 1 - \pi_g \left( 1 - \frac{1 + 2\hat{u} - \hat{u}^2}{2} \right)
\Leftrightarrow \quad \frac{\hat{u}^3}{3} - \left( 1 - \frac{\pi_g}{2} \right) \hat{u}^2 + \left( 1 - \pi_g \right) \hat{u} \geq \frac{1}{3} - \frac{\pi_g}{2}.
\tag{14}
\]

It now suffices to show that \( l(\hat{u}) \geq 1/3 - \pi_g/2 \) iff \( \hat{u} > \tilde{u} \), for some \( \tilde{u} \in (0,1) \). Denote by \( l(\hat{u}) \) the left-hand side of \((14)\). Differentiating \( l(\hat{u}) \) yields \( l' = \hat{u}^2 - (2 - \pi_g) \hat{u} + 1 - \pi_g \) and \( l'' = 2\hat{u} - 2 + \pi_g \). Straightforward calculation reveals that \( l \) has two local extrema on \((0,1]\):
a minimum at \( \hat{u} = 1 \) and a maximum at \( \hat{u}^* = 1 - \pi_g \). Thus \( l \) must be strictly increasing on \( \hat{u} \in [0, \hat{u}^*) \) and strictly decreasing on \( \hat{u} \in (\hat{u}^*, 1) \), with \( l(\hat{u}) > 1/l(1) \) for all \( \hat{u} \in (\hat{u}^*, 1) \).

Note that at \( \hat{u} = 1 \) (i.e., \( y_M = m \)), \( l(\hat{u}) = 1/3 - \pi_g/2 \). There are two cases. First, if \( l(0) \geq 1/3 - \pi_g/2 \), then the above facts imply that \( l(\hat{u}) > 1/3 - \pi_g/2 \) for all \( \hat{u} \in (0, 1) \). Letting \( m^R = 0 \) completes the proof. Second, if \( l(0) < 1/3 - \pi_g/2 \), then there exists some \( \tilde{u} \in (0, \hat{u}^*) \) such that \( l(\tilde{u}) < l(1) \) iff \( \hat{u} < \tilde{u} \). Let \( m^R \) be the value of \( y_M \) satisfying \( \bar{u}^R - \bar{u}^L = \tilde{u} \). The existence of such a \( m^R \in (m, \bar{m}) \) follows from the continuity and monotonicity of \( u(\cdot) \).

**Cases 4 and 5:** (1, 1), (2, 0) and (1, 1), (2, 1), (2, 0). Observe that since \( \rho^R_p(2, 0) > \rho^R_p(2, 1) > \rho^R_p(1, 1) \), given the results of cases 2 and 3 it is sufficient to show that \( \rho^R_p(1, 1) < \rho^R_g \) for all \( m < \bar{m} \).

From the proof of Proposition 1, we have: \( \rho^R_p(1, 1) = \pi_p \left[ \frac{1+2\hat{u} - \hat{u}^2}{2} \right] + (1 - \pi_p)\rho^R_g \). Party \( R \) chooses primaries (\( e^* = 1 \)) iff \( \rho^R_p(n_R, n_L) \geq \rho^R_g \). Using (2), this is equivalent to:

\[
\frac{1+2\hat{u} - \hat{u}^2}{2} \leq 1 - \pi_g \left(1 - \frac{1+2\hat{u} - \hat{u}^2}{2} \right).
\]

Since this expression holds only at \( \hat{u} = 1 \), a primary can never be preferred by the favored party when \( (n_R, n_L) = (1, 1) \) and \( m < \bar{m} \).

**Case 6:** (2, 2), (2, 1), (2, 0). Observe that since \( \rho^R_p(2, 0) > \rho^R_p(2, 1) > \rho^R_p(2, 2) \), given the result of case 3 it is sufficient to show that \( \rho^R_p(2, 2) \geq \rho^R_g \) iff \( m \geq m^R \), where \( m^R \in [m, \bar{m}] \).

From the proof of Proposition 1, we have: \( \rho^R_p(2, 2) = \pi_p \left[ \frac{\hat{u}^4}{6} - \hat{u}^2 + \frac{4\hat{u}}{3} + \frac{1}{2} \right] + (1 - \pi_p)\rho^R_g \). Party \( R \) then chooses primaries (\( e^* = 1 \)) iff \( \rho^R_p(n_R, n_L) \geq \rho^R_g \). Using (2), \( \rho^R_p(n_R, n_L) \geq \rho^R_g \) iff:

\[
\frac{\hat{u}^4}{6} - \hat{u}^2 + \frac{4\hat{u}}{3} + \frac{1}{2} \geq 1 - \pi_g \left(1 - \frac{1+2\hat{u} - \hat{u}^2}{2} \right)
\]

\[
\Leftrightarrow \frac{\hat{u}^4}{6} - \left(1 - \frac{\pi_g}{2}\right)\hat{u}^2 + \left(\frac{4}{3} - \pi_g\right)\hat{u} \geq \frac{1-\pi_g}{2}.
\]

It is sufficient to show that \( l(\hat{u}) \geq (1 - \pi_g)/2 \) iff \( \hat{u} > \tilde{u} \), for some \( \tilde{u} \in (0, 1) \). Denote by \( l(\hat{u}) \) the left-hand side of (16). Note that at \( \hat{u} = 0 \) (i.e., \( y_M = m \)), \( l(0) < (1 - \pi_g)/2 \), and hence \( \rho^R_p(n_R, n_L) < \rho^R_g \). Likewise, at \( \hat{u} = 1 \) (i.e., \( y_M = \bar{m} \)), \( l(1) = (1 - \pi_g)/2 \) and \( \rho^R_p(n_R, n_L) = \rho^R_g \).

Differentiating \( l(\hat{u}) \) yields \( l' = 4/3 - \pi_g - (2 - \pi_g)\hat{u} + 2\hat{u}^2/3 \). Further differentiation yields \( l'' = \pi_g - 2 + 2\hat{u}^2 \) and \( l''' = 4\hat{u} \). Straightforward calculation reveals that \( l'(0) > 0, l'(1) = 0, l' \)
is strictly convex for \( \hat{u} > 0 \), and \( l'' > 0 \) in the neighborhood of \( \hat{u} = 1 \). These facts imply that \( l \) has two local extrema on \( (0, 1) \): a minimum at \( \hat{u} = 1 \) and a maximum at \( \hat{u}^* \in (0, 1) \). Further, since there are only two local extrema, \( l(\hat{u}^*) > l(1) \). Thus \( l \) must be strictly increasing on \( \hat{u} \in [0, \hat{u}^*) \) and strictly decreasing on \( \hat{u} \in (\hat{u}^*, 1) \), with \( l(\hat{u}) > l(1) = (1 - \pi_g)/2 \) for all \( \hat{u} \in (\hat{u}^*, 1) \). Thus since \( l(0) < (1 - \pi_g)/2 \), there exists some \( \tilde{\mu} \in (0, \hat{u}^*) \) such that \( l(\tilde{\mu}) < l(1) \) iff \( \tilde{\mu} < \hat{u} \).

Finally, let \( m^R \) be the value of \( y_M \) satisfying \( \overline{\pi}_R - \overline{\pi}_L = \tilde{\mu} \). The existence of some \( m^R \in (m, \overline{m}) \) follows from the continuity and monotonicity of \( u(\cdot) \).

(ii) We first eliminate two candidates for equilibrium values of \((n^R_0, n^L_0)\). From Proposition 1, at \( \hat{u} = 1 \), \((n^R_0, n^L_0) = (1, 1) \) requires that \( k \in [0, 0) \) and \( k \geq 1/2 \). This is obviously impossible, and since the bounds on \( k \) are continuous in \( \hat{u} \), there exists some neighborhood \((1 - \epsilon', 1 + \epsilon')\) of \( \hat{u} = 1 \) where \((n^R_0, n^L_0) = (1, 1) \) cannot be an equilibrium. By an identical argument, there is no equilibrium in which \((n^R_0, n^L_0) = (1, 2) \) in some neighborhood \((1 - \epsilon'', 1 + \epsilon'')\) of \( \hat{u} = 0 \). Let \( m' \) and \( m'' \) denote the values of \( m \) at which \( \hat{u} = 1 - \epsilon', 1 - \epsilon'' \), respectively.

Now consider equilibria where \((n^R_0, n^L_0) = (1, 0) \) and \((2, 0) \) for some values of \( \hat{u} \). In these cases, party \( R \) chooses \( e^* = 1 \) for all \( \hat{u} \).

For equilibria where \((n^R_0, n^L_0) = (2, 1) \) and \((n^R_0, n^L_0) = (2, 2) \), the proof of part (i) established that \( \rho_p^R(n_R, n_L) \geq \rho_g^R \) for any voter with \( \hat{u} \) sufficiently high. Let \( m''' \) denote the location of a voter with the minimum value of \( \hat{u} \) for which both \( \rho_p^R(2, 1) \geq \rho_g^R \) and \( \rho_p^R(2, 2) \geq \rho_g^R \) are satisfied, and observe that the proof of part (i) implies that \( m''' \in [m, \overline{m}) \).

Thus in any equilibrium with \( n^R_R > 0 \) for all \( \hat{u} \), party \( R \) chooses \( e^* = 1 \) for any district where \( y_M \geq m'''' \equiv \max\{m', m'', m''''\} \).

(iii) From the expressions for \( \rho_p^R(n_R, n_L) \) in (6), (8), (10) and (11) derived in the proof of Proposition 1, we have the following values for \( \rho_p^R(n_R, n_L) \) at \( \hat{u} = 0 \):

\[
\rho_p^R(2, 2) = \frac{\pi_p}{2} + (1 - \pi_p) \left(1 - \frac{\pi_g}{2}\right)
\]

\[
\rho_p^R(2, 1) = \frac{2\pi_p}{3} + (1 - \pi_p) \left(1 - \frac{\pi_g}{2}\right)
\]

\[
\rho_p^R(1, 2) = \frac{\pi_p}{3} + (1 - \pi_p) \left(1 - \frac{\pi_g}{2}\right)
\]
\[ \rho^R_p(1,1) = \frac{\pi_p}{2} + (1 - \pi_p) \left( 1 - \frac{\pi_g}{2} \right) \]

Clearly, \( \rho^R_p(2,1) \) is the maximal value among these expressions. Thus for any equilibrium where \( n^*_R > 0 \) and \( n^*_L > 0 \), party \( R \) strictly prefers no primaries \( e^* = 0 \) at \( \hat{u} = 0 \) if \( \rho^R_p(2,1) < \rho^R_g \). Using (2) and (8), this obtains if: \( 1 - \pi_g/2 > 2/3 \), or \( \pi_g < 2/3 \).

Now note that (6), (8), (10) and (11) imply that all expressions for \( \rho^R_p(n_R, n_L) \) are continuous in \( \hat{u} \). Thus there exists a neighborhood of \( \hat{u} = 0 \) where \( \rho^R_p(2,1) < \rho^R_g \) for all \( \hat{u} \). Choosing some \( m^{R''} \) in this neighborhood completes the proof. \( \blacksquare \)
References


Le Borgne, Eric, and Ben Lockwood. 2001b. “Candidate Entry, Screening, and the Political Budget Cycle.” Unpublished manuscript.


Table 1
Primaries and the Degree of Inter-Party Competition

<table>
<thead>
<tr>
<th></th>
<th># States</th>
<th>Vote Margin</th>
</tr>
</thead>
<tbody>
<tr>
<td>Primary States</td>
<td>40</td>
<td>.112</td>
</tr>
<tr>
<td>Other States</td>
<td>8</td>
<td>.025</td>
</tr>
</tbody>
</table>

$t$-stat. for difference in means: 2.358 ($p < .011$)

<table>
<thead>
<tr>
<th></th>
<th>One-Party States</th>
<th>Two-Party States</th>
</tr>
</thead>
<tbody>
<tr>
<td>Primary States</td>
<td>.958 (23)</td>
<td>.708 (17)</td>
</tr>
<tr>
<td>Other States</td>
<td>.042 (1)</td>
<td>.292 (7)</td>
</tr>
</tbody>
</table>

Chi-square test of independence: 5.400 ($p < .020$)
Table 2  
**Primaries and Changes in the Degree of Inter-Party Competition**  
Dep. Var. = Change in Rel Dem Vote

<p>| | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Rel Dem Vote Before Primaries</td>
<td>.198*</td>
</tr>
<tr>
<td></td>
<td>(.048)</td>
</tr>
<tr>
<td>Constant</td>
<td>−.012</td>
</tr>
<tr>
<td></td>
<td>(.008)</td>
</tr>
<tr>
<td># Observations</td>
<td>38</td>
</tr>
<tr>
<td>R-Square</td>
<td>.322</td>
</tr>
</tbody>
</table>

Standard errors in parentheses.  
* = significant at the .01 level.  
Only Primary States included. Also, AZ and OK are excluded because they held primaries since statehood.
### Table 3

**Primary and Inter-Party Competition**

**1926-2006**

Dep. Var. = Competitive Primary

<table>
<thead>
<tr>
<th></th>
<th>Democratic Primaries</th>
<th>Republican Primaries</th>
</tr>
</thead>
<tbody>
<tr>
<td>Democrats Favored</td>
<td>0.376* (.042)</td>
<td>0.194* (.035)</td>
</tr>
<tr>
<td>Republicans Favored</td>
<td></td>
<td>0.270* (.044)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>0.156* (.043)</td>
</tr>
<tr>
<td>Competitive State</td>
<td>0.143* (.041)</td>
<td>0.089* (.028)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>0.124* (.038)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>0.082* (.031)</td>
</tr>
<tr>
<td>Incumbent Running</td>
<td>-0.314* (.027)</td>
<td>-0.319* (.025)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>-0.230* (.026)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>-0.244* (.028)</td>
</tr>
<tr>
<td># Observations</td>
<td>7,808</td>
<td>7,808</td>
</tr>
<tr>
<td></td>
<td>7,687</td>
<td>7,687</td>
</tr>
<tr>
<td>R-Square</td>
<td>0.152</td>
<td>0.246</td>
</tr>
<tr>
<td></td>
<td>0.084</td>
<td>0.176</td>
</tr>
<tr>
<td>State Fixed Effects?</td>
<td>No</td>
<td>Yes</td>
</tr>
<tr>
<td></td>
<td>No</td>
<td>Yes</td>
</tr>
</tbody>
</table>

Robust standard errors in parentheses, clustered by state.

* = significant at the .01 level.

Year fixed-effects included in all specifications.
Figure 1: Number of entering candidates \((n^*_R, n^*_L)\) as a function of \(\hat{u}, k\). On the top, \(\pi_p = 0.8\) and \(\pi_g = 0.2\). On the bottom, \(\pi_p = 0.1\) and \(\pi_g = 0.9\); note that one entrant from each party is now possible. In both cases, the total number of entrants is decreasing in \(k\). Reducing the relative value of \(\pi_p\) increases entry by party \(L\) because it reduces party \(R\)’s chances of choosing its higher-valence candidate.