

Why Primaries?

The Party's Tradeoff between Policy and Valence*

Gilles Serra[†]
Oxford University (Nuffield College)

November 5, 2007

Abstract

We elaborate a theory to explain why and when political parties choose to hold primary elections. Party leaders face a trade-off between primary elections and elite-centered selections. The benefit of a primary is revealing the campaigning skills of candidates. Its cost is the ideological extremism that primary voters might induce on candidates. We find that primary elections are more likely when the party leadership and the potential primary voters are ideologically similar (which is consistent with the recent empirical research by Meinke, Stanton and Wuhs (2006)). Intriguingly, our model predicts that parties with extremist ideologies are more likely to be internally democratic. For intermediate values in the parameters, parties have multiple equilibria in their decision to adopt primaries or not. And finally, parties display a significant degree of contagion, meaning that a party's adoption of a primary will influence the other party to adopt a primary as well.

*This paper benefitted from being presented at Nuffield College in October 2007, the annual meeting of the American Political Science Association in September 2007, and Harvard University in April 2007. I thank Christopher Avery, Robert Bates, Indridi Indridason, John Patty, Brian Richardson, Kenneth Shepsle and James Snyder for their insightful comments. I also thank the Institute of Quantitative Social Sciences for its financial support. All remaining errors are my own.

[†]Postdoctoral Prize Research Fellow at Nuffield College in Oxford University, gilles.serra@nuffield.ox.ac.uk .

1 Introduction

An institution whose origin needs better understanding is the primary election. Indeed, political parties can set different rules for the nomination of the person who will run as their candidate in a given election. Broadly speaking, a candidate-selection method can fall in two categories: *open* (or *democratic*), where the party's rank-and-file members and sympathizers take a vote to nominate their candidate; and *closed* (or *undemocratic*), consisting of a closed-door decision at the elite level of the party.¹

This paper studies the endogenous adoption by political parties of one type of candidate-selection method (CSM) over another. Of all the selection methods that parties can use, the most salient one is the primary election, not least because it has been increasingly frequent around the world.² Indeed, primary elections are gaining popularity among reformers and party leaders, as testified by recurrent journalistic editorials and public speeches. As a result, the frequency of open CSMs is on the rise. They are already prevalent in the U.S., where they have seen a startling increase in recent decades. In 1968, only sixteen states and the District of Columbia were holding primaries; whereas by 1996 the Democratic party was holding primaries in thirty-six states and the Republican party in forty-three (Morton (2006)). Primary elections are seeing a surge in Latin America as well: Carey and Polga-Hecimovich (2006) report that the percentage of candidates nominated by primaries increased from 3% in the 1980s to 4% in the 1990s to 12% in the 2000s, thus increasing fourfold in two decades. The number of democratic candidate selections has also increased in Europe since 1960 with the introduction of membership ballots (Bille (2001)); and parties in other countries like Taiwan have also experimented with primaries in the recent past (see Scarrow (2005) for a survey).

Party leaders themselves are for the most part responsible for this trend whereby parties are voluntarily adopting primary elections. In most party systems around the world, political parties have leeway in choosing their CSM, and it is therefore *not* the case that primaries are exogenously imposed

¹There is of course some variety within these two extremes: for example, CSMs can further be classified as *semi-open* or *semi-closed* (as described in Kanthak and Morton (2001)). But, as our results will show, a division in two broad categories is sufficient for the points this paper wishes to make.

²By primary election we refer to any organized competition among aspiring candidates within the same party that culminates in the democratic vote of all party members.

on parties. As a matter of fact, there is abundant evidence that political parties have serious deliberation on what CSM to adopt before they discuss which candidate to select. Recently, the UMP in France debated for several months whether to adopt its first primary in 2007; so did a coalition of left-wing parties in Italy in 2005. The PAN in Mexico held a heated convention before deciding to adopt an American style primary election in 2006.³

Even in the United States, where the usual claim is that primaries are exogenously imposed on parties by state law, parties have actually had a large degree of freedom in designing their CSM. For example the Democratic Party had a period of introspection and laborious internal discussions during the McGovern-Fraser reforms in the 1970s which culminated in its adoption of primary elections for most offices in most states (Abramowitz and Stone (1984), Bartels (1988), Geer (1989), Aldrich (1980b)). Even today, the CSM is not irrevocably fixed, but rather parties can re-interpret or change the law to make their CSM more or less open, which led Meinke, Staton and Wuhs to characterize the American CSMs as rather fluid: they claim that "state parties can and do change their rules repeatedly over time, often moving from a more open to a more closed rule (or vice versa) and then back again" (Meinke et al. (2006), p. 183). And even without changing any rule or charter, even when primary elections are irrevocably imposed on parties by state law, party leaders still have means to impose their favorite candidate for example by endorsing her and providing her with enough resources to prevent any challenge in the primary –which would explain why most primary elections in America have been uncontested in the last hundred years (Ansolabehere, Hansen, Hirano and Snyder (2005)). So at the very least, the Democratic and the Republican leaders can choose between a competitive and a non-competitive primary.

Thus, any theory of the adoption of competitive primary elections must put party leaders at the center of the decision process. This presents a paradox however: why would party leaders decentralize what could be a closed process where they could have their pick of candidates? Why do party elites delegate the nomination to the party's foot soldiers which might have different preferences?

Our answer begins by claiming that primary elections have an advantage over elite-centered nominations: *they reveal information about the candidates'*

³In all three cases, the winner of the primary election went on to win the general election.

appeal to voters. Our theoretical model describes this information revelation feature in detail, and studies the conditions under which such a feature is attractive enough for party leaders to adopt a primary election.

Party leaders can see a drawback in primary elections, however: primary voters might push candidates to adopt policies far from the leaders' preferences.⁴ Indeed, several empirical studies have found that primary voters have extreme ideological views (Key (1956), Lengle (1981), Polsby (1983)), and that primary elections make candidates adopt extremist platforms (King (1999), Burden (2001, 2004)). It should be noted that several other empirical studies have found the opposite results: that primary voters do *not* have extremist ideological views (Ranney (1968), Geer (1988), Kaufmann, Gimpel and Hoffman (2003)), and that primary elections do *not* make candidates adopt extremist platforms (Gerber and Morton (1998), Ansolabehere, Snyder and Stewart (2001)). Given that the empirical literature is mixed, in this paper we do not take a position on whether primary voters are extremist or moderate, but rather we study both situations by treating the preferences of primary voters as a continuous parameter that can be located anywhere in the political spectrum. We can thus make comparative statics on that parameter.

The main point is that party leaders face a trade-off between the costs and benefits of a primary election. The results in this paper reveal that the party leaders' decision is not trivial and relies heavily on the decision made by the leaders of the other party. The emergence of primary elections must therefore be understood as a game-theoretic rather than a decision-theoretic phenomenon.

The paper adds to the formal literature on institutions. There has been virtually no formal theory studying the *origins* of primary elections. But there is a growing literature studying their *consequences*.⁵ Some of that lit-

⁴Party leaders could see additional drawbacks that this paper does not focus on. Ezra (2001) argues that primaries might harm candidates in the general election by draining their energy and resources. Ansolabehere, Hirano and Snyder (2004) argue that the loyalty of legislators in the U.S. Congress toward their parties decreased following the adoption of the American direct primary. Serra (2007a) argues that party leaders in Latin America may lose the patronage and cronyism that result from directly handpicking a candidate from their inner circle.

⁵For example see Aranson and Ordeshook (1972), Coleman (1972), Aldrich (1980a), Owen and Grofman (2004), Morton and Williams (2001), Buchler (2005), Meirowitz (2005), Adams and Merrill (2006), Owen and Grofman (2006), Snyder and Ting (2006).

erature compares primaries to other CSMs (Chen and Yang (2002), Jackson, Mathevet and Mattes (2006), Kang (2007)). But to the best of our knowledge, our work is the first to model the decision of which CSM to adopt.

As we elaborate later, this paper also draws on the literature on *valence* (Stokes (1963)). This model is closest to the theoretical models that postulate a valence parameter, i.e. a parameter that adds utility on top of policy to all voters.⁶

There is another paper that combines both literatures, the one on valence and the one on primaries, Adams and Merrill (2006), which focuses on the policies adopted by candidates and whether they diverge from the median voter. They believe, like we do, that primaries tend to select the candidates with highest valence. But their paper does not model the decision to adopt a primary or not, and it has several modelling differences with our paper that we will point out in footnotes subsequently.

The rest of our paper is developed as follows: Section 2 describes the information revelation mechanism by which primaries can reveal candidates' appeal to voters. Section 3 sets up the model of electoral competition. Section 4 solves the last stages of the election, after parties have nominated their candidates. Sections 5 and 6 solve the first stages of the election, that is the nomination of candidates, under two different assumptions about the technology available to parties: in Section 5 only one of the parties has the technology to adopt a primary, and in Section 6 both parties have the technology to adopt a primary. And Section 7 offers some conclusions. The Appendix of this paper is divided in two sections. The first section develops the information-revelation mechanism in detail, and the second section contains the proofs of all the results in the paper.

2 The benefits of a primary election

One of the concerns of a political party is nominating a candidate with good campaigning skills. The premise in this paper is that primaries have two features that help a party choose among the individuals seeking to become its candidate (often called *pre-candidates*). First, they reveal some information

⁶Such a parameter is found in Enelow and Hinich (1982), Cukierman (1991), Londregan and Romer (1993), Adams (1999), Ansolabehere and Snyder (2000), Groseclose (2001), Schofield (2004).

about the personal characteristics of candidates that would be useful (or detrimental) in the general election. In effect, a primary election is a full blown election *within* a party which shares many of the features of the subsequent general election *between* the parties: for example the pre-candidates need to run a campaign, debate other candidates, run advertisements on television, and be scrutinized by journalists. Primaries are therefore a good testing ground for the performance that aspiring candidates would have during the general election.

A second feature of primaries is opening the door to unknown or marginalized party members who might have a large appeal to voters but would not be identified through an inside-track nomination. Illustrations of such surprise candidates abound. John F. Kennedy in the U.S., Carlos Menem in Argentina and Ségolène Royal in France are examples of non-mainstream politicians who were able to prove their appeal to voters thanks to a primary. All three surprisingly obtained their parties' nomination, and went on to have a strong showing in the general election.

To be precise, let us assume that a given party, which we will generically call party K , needs to nominate a candidate for an upcoming election. Candidates are characterized by a parameter θ denoting how appealing their non-policy attributes would be to voters in that election. Parameters such as θ have been called "valence parameters" and can be given many interpretations like charisma, honesty, competence and the ability to raise campaign funds (for an overview see Ansolabehere and Snyder (2000) and Groseclose (2001)). In the context of this paper, θ is best interpreted as the candidate's *campaigning skill*. The parameter θ can take a maximum value of Θ , and we say that a candidate with skill Θ is *high-skilled*. At the time that K needs to nominate a candidate, however, it is uncertain whether K 's pre-candidates are high-skilled or not. We call θ_K the campaigning skill of the candidate ultimately nominated by K .

Before selecting a candidate, the leaders of party K need to select a candidate-selection method (CSM). There exist two CSMs: The default CSM would be for the leaders to directly nominate an insider candidate in a closed negotiation at the elite level (such as a caucus or a back-room deal). Alternatively they could implement a primary election where an outsider candidate has a chance to run, and the decision to choose the nominee is delegated to the party's rank and file (RAF). So K must choose its candidate-selection method c_K with $c_K \in \{primary, elite\}$, where *primary* refers to the adop-

tion of a primary election and *elite* refers to an elite-centered selection. We define π_K as probability that the nominee will be high-skilled if K uses an elite selection. In other words, $\pi_K \equiv P(\theta_K = \Theta|elite)$.

Our claim is that a primary election increases K 's probability of nominating a high-skilled candidate above π_K in the amount S_K . We therefore define $S_K \equiv P(\theta_K = \Theta|primary) - P(\theta_K = \Theta|elite)$, and we call S_K the *skill bonus* of a primary. An important implication of this definition is that

$$P(\theta_K = \Theta|primary) = \pi_K + S_K$$

We devote a section in the appendix to develop an information-revelation model from which S_K is derived. In formal-theoretic parlance, we derive S_K from "microfoundations". The force behind the primary skill bonus is the information revealed during the primary campaigns about the campaigning skills of the pre-candidates: Primary voters interpret the performance of a candidate in the primary election as a *forecast* of how well they would perform in the general election.

Those forecasts are imperfect however, meaning that the true skills of candidates are not *fully* revealed during the primary, but they are only *partially* revealed.⁷ To be concrete we assume that a candidate's performance in the primary has a probability q of "being correct", that is, of accurately reflecting the true skill of that candidate, with $q \in (\frac{1}{2}, 1)$. As we describe in the appendix, primary voters take the performances they observe in the primary to update their prior beliefs about the candidates' skills (using Bayes rule).

Deriving S_K in such a way leads to some intuitive properties that will buttress the rest of the paper. The main properties of the primary skill bonus, which are proved in the appendix, are stated below. In all cases, $\bar{\pi}$ refers to a constant larger than one-half that depends on the parameters of the model.

Property 1 The primary skill bonus S_K is strictly positive for $\pi_K \in (0, \bar{\pi})$ and zero for $\pi_K \in [\bar{\pi}, 1)$.

⁷Adams and Merrill (2006) also postulate that primaries reveal information about the pre-candidates' true valence, but in their model information is fully revealed in the primary election, and there is no additional information in the general election. In contrast, in our model the information is only partially revealed in the primary, and there is some additional information in the general election.

This result is a formal statement of the claim that primaries bring an electoral advantage to the parties that use them, at least when the elite-selection is not expected to be particularly effective in delivering a high-skilled candidate.

Property 2 The primary skill bonus S_K is strictly decreasing with π_K for $\pi_K \in (0, \bar{\pi})$.

It makes intuitive sense that S_K would decrease with π_K , because the whole advantage of primaries is to improve upon the skill of the candidate that would be nominated through an elite selection, i.e. the insider candidate. As the skill of the insider candidate is expected to be higher, it becomes less likely that a primary will improve upon it. The message of this result is that the electoral advantage brought by primaries is larger the less appealing the insider candidate is to begin with.

Property 3 For every $q \in (\frac{1}{2}, 1)$ there exists a $q' > q$ such that for every $\pi_K \in (0, 1)$, when q increases to q' , S_K will increase.

An increase in q can be interpreted as an improvement in the information-revelation feature of primaries. This improvement could arise because the primary campaigns became longer, or because the media paid more attention to them, or because they included more challenges like debates on television and so on. According to the above there will be a positive effect on S_K if large enough improvements in q can be achieved. So as intuition would suggest, an improvement in the primaries' technology to forecast the candidate's performance in the general election will make those primaries more attractive.

In sum, the benefit to party leaders of adopting a primary election is a larger probability of nominating a candidate with a high campaigning skill –we called that extra probability the primary skill bonus. Primaries might carry a cost however, in terms of the policy that candidates are induced to adopt. That cost is described in detail in the following section. As we will see, party leaders are therefore forced to carry out a cost-benefit analysis in choosing whether to hold primary election or not.

3 Set-up of the electoral competition

The election occurs in a unidimensional space where x is the policy implemented, thus $x \in \mathbb{R}$.

3.1 Parties

There are two parties competing in this election, labeled party A and party B . Both parties consist of a leadership (or "elite") and a rank and file (or "membership"). Each party's leadership needs to choose a candidate-selection method. There are two CSMs to choose from: an elite selection or a primary election. We call c_A the CSM that A adopts and c_B the CSM that B adopts, with $c_A, c_B \in \{primary, elite\}$.

The respective leaders of A and B will be referred to as AL and BL . They are policy-motivated and have ideal policy points X_{AL} and X_{BL} respectively. A 's leaders have a distinct ideology from B 's leaders so that $X_{AL} \neq X_{BL}$. We normalize the ideal point of the median voter in the general election to zero, and without much loss of generality we assume $X_{AL} < 0 < X_{BL}$. The utility functions of A 's leadership and B 's leadership are

$$U_{AL}(x) = -|X_{AL} - x| \tag{1}$$

$$U_{BL}(x) = -|X_{BL} - x| \tag{2}$$

The respective RAFs of A and B are also policy-motivated. To simplify the analysis we will assume that each RAF has a "median member" whose preferences are decisive in the primary election. We call AR the median member of A 's RAF, and BR the median member of B 's RAF. We will call X_{AR} and X_{BR} their ideal points and assume $X_{AR} < 0 < X_{BR}$. The utility functions of A 's RAF and B 's RAF are

$$U_{AR}(x) = -|X_{AR} - x| \tag{3}$$

$$U_{BR}(x) = -|X_{BR} - x| \tag{4}$$

In general we will have $X_{AL} \neq X_{AR}$ and $X_{BL} \neq X_{BR}$ so there is a tension between the policy preferences of a party's leadership and its RAF. It will be useful to have a measure of the *ideological distance* between a party's elite and its primary voters, so we define $\lambda_A \equiv X_{AR} - X_{AL}$ as the "ideological

distance in party A " and $\lambda_B \equiv X_{BR} - X_{BL}$ as the "ideological distance in party B ".

3.2 Candidates

Each candidate is characterized by a parameter θ denoting that candidate's campaigning skill, where θ can take two values: a high value of Θ , with $\Theta > 0$, or a low value of zero.⁸ To focus on the interesting cases, we will assume that Θ is sufficiently large to make a difference in the election; technically we will assume that Θ is strictly larger than $-X_{AL}$, $-X_{AR}$, X_{BL} and X_{BR} .⁹

If the CSM is an elite selection, the party has only one candidate to choose from: an insider.¹⁰ In contrast, if the CSM is a primary election, the party has two candidates (or rather, "pre-candidates") to choose from: an insider and an outsider.¹¹ We call AI , BI , AO , BO the insider and outsider pre-candidates of A and B respectively, and we call θ_{AI} , θ_{BI} , θ_{AO} , θ_{BO} their campaigning skills. We call θ_A and θ_B the skills of the candidates that are finally nominated by A and by B .

The exact values of the candidates' campaigning skills are uncertain before the election. However the parties have some prior information about the skill of their insider candidate. That information could come from previous performance in office, from past elections, or from polls. According to that information, the probabilities that AI and BI are high-skilled are π_{AI} and π_{BI} respectively, with $\pi_{AI}, \pi_{BI} \in (0, 1)$. On the other hand, there is no prior information about the outsider candidates, so they are believed to have a probability of one-half of being high-skilled.¹² All this information is

⁸Assuming that campaigning skills can take only two values is mostly done for presentation purposes. Assuming instead that the campaigning skill could take a continuum of values would greatly complicate the Bayesian updating that we study in the appendix, and would not lead to different conclusions than the ones we obtain. In related research we did study the case where skill can take a continuum of values, and found very similar results (Serra 2007b).

⁹Indeed, for smaller values of Θ the valence dimension loses influence in the election, and the results becomes trivial. For that reason we only report those results in footnotes.

¹⁰We can think of that "only" insider as the most skilled among all the available insiders.

¹¹The model could easily be extended to having more than one outsider. The main effect would be a larger primary skill bonus.

¹²Alternatively, we could assume that there is *some* prior information about the *pool* of outsider candidates. The results would all hold, with only notational adjustments, if we assumed that AO and BO had probabilities $\pi_{AO}, \pi_{BO} \neq \frac{1}{2}$ of being high-skilled.

common knowledge.

The true campaigning skills of A and B 's candidates are revealed when they have to campaign to win the general election, and thus θ_A and θ_B are fully known by the time the general electorate votes between A and B . We call δ the difference in skill between B 's candidate and A 's candidate, that is, $\delta \equiv \theta_B - \theta_A$. Note that δ can take three values: $\delta \in \{-\Theta, 0, \Theta\}$.

Candidates are in charge of formulating policy platforms to compete in the general election. We call those platforms x_A and x_B for A 's candidate and B 's candidate respectively, with $x_A, x_B \in \mathbb{R}$. An important assumption is that a candidate will behave as a perfect agent of those who nominated her, also called her "selectorate". The interpretation is that competition between pre-candidates within a party leads them to accept any policy commitment. So, in striving to win the nomination, candidates are forced to cater to the preferences of their selectorate. To be concrete, the candidates of party B will adopt the preferences $U_{BL}(x)$ if B uses an elite selection, and they will adopt the preferences $U_{BR}(x)$ if B uses a primary election –and similarly for party A .¹³

3.3 The general electorate

The electorate is policy motivated. To simplify the analysis we will assume that there is a "median voter", which we call M , whose preferences are decisive in the election. We normalize her ideal point to zero.

In addition to the policy implemented x , the electorate also cares about the skill θ of the winning candidate. The utility function of M is given by

$$U_M(x, \theta) = -|x| + \theta$$

M will vote for the party whose candidate maximizes her utility. If the candidates of both parties give her the same utility, we will make the following indifference assumption:¹⁴

A1: If M is indifferent between the two parties, she will vote for

¹³Note however that primary voters and party leaders would choose policy strategically, not sincerely, and therefore candidates will also choose policy strategically, not sincerely.

¹⁴This assumption ensures the existence of equilibria between parties. Without it, an equilibrium might not exist, but the behavior of parties would still converge ever so closely to the equilibria described in the text.

the one whose candidate has the highest skill. If both candidates have the same skill, she will randomize equally between the two.

3.4 Timing and solution concept

The timing of this election is the following:

1. **The selection of the candidate-selection method:**

Party leaders in both parties simultaneously choose their CSM.

2. **The nomination campaigns:**

Candidates adopt the preferences of their electorate.

If the CSM is a primary election, the skills of the pre-candidates in that party are *partially* revealed.

3. **The nomination:**

Both parties simultaneously select their candidates.

4. **The early stage of the general-election campaigns:**

Candidates' skills θ_A and θ_B are *fully* revealed.

5. **The late stage of the general-election campaigns:**

Candidates design and announce their platforms x_A and x_B .

6. **The general election:**

The median voter elects A or B .

A consequential feature of this timing is that the candidates' skills θ_A and θ_B are revealed *before* the policy platforms x_A and x_B are chosen. This captures the idea that both candidates will measure each other's skill through their initial performance in the general-election campaign, while remaining vague on the issues (stage 4); and once they have a precise idea of each other's skill they will formulate an unambiguous policy platform (stage 5). The median voter will then take into account the observed skills and policy

platforms to vote for the candidate offering the best combination of the two (stage 6).¹⁵

The game must be solved by backward induction and the solution concept is subgame-perfect equilibrium (SPE) in pure strategies.¹⁶

4 The general election

In this section we solve the general election, that is, Stages 4, 5 and 6 of the game. Therefore we start the analysis as if parties had already resolved their respective nominations.

Stage 4 does not involve any decision; the skills of both candidates are fully revealed to everyone. The first decision is made in Stage 5, when candidates must locate their platforms in the unidimensional political spectrum. The goal of both candidates is to maximize their payoff from the policy implemented after the election (recalling that the candidates have acquired the policy preferences of their electorate). In Stage 6, once parties have chosen their platforms x_A and x_B , the median voter elects A or B to office. The equilibria of this election is given in the following result, where we denote by X_A and X_B the ideal points of the candidates in parties A and B . (In later sections, X_A and X_B will be replaced by the ideal points of each party's electorate.)

Theorem 1 *The equilibrium strategies and equilibrium outcomes of this election for given values of θ_A and θ_B are given in the following table, where $\delta \equiv \theta_B - \theta_A$*

¹⁵Note that this timing differs from previous papers with uncertainty about candidates' valence such as Adams (1999), Londregan and Romer (1993), and Adams and Merrill (2007). The fact that valence is revealed before policy makes our model a *valence-policy model* as defined in Serra (2007b). Most previous models have the reverse sequence thus making them *policy-valence models*.

¹⁶The restriction to pure strategies does not change the results in general. In fact there is only one case where a mixed strategy equilibrium exists, and we describe it in a footnote.

Value of δ	Nash equilibrium/ equilibria	Winning platform	Winning party
Θ	$x_A \in \mathbb{R}$ $x_B = X_B$	X_B	B
0	$x_A = 0$ $x_B = 0$	0	A or B with equal probability
$-\Theta$	$x_A = X_A$ $x_B \in \mathbb{R}$	X_A	A

There are several comments to make about this table. First note the results when $\delta = 0$, that is, when there is no skill difference between the candidates: both parties converge completely to the median voter's ideal point. As soon as $\delta \neq 0$ however, the candidate with highest skill is able to diverge from the median voter toward her ideal point, and still win the election based on her superior skill.¹⁷

The main take-away point is that the policy implemented is favorable to B when $\delta = \Theta$, and favorable to A when $\delta = -\Theta$. In other words, a party obtains a better policy the higher the skill of its nominee with respect to the skill of the other party's nominee.

In the next two sections we take a step back in the election to study the nomination process within parties. In other words we incorporate Stages 1, 2 and 3 of the game to the analysis we just made of Stages 4, 5 and 6. We study two cases, depending on whether only one party or both parties have the technology needed to hold primary election.¹⁸ Section 5 corresponds to only one party having the ability to adopt a primary election, and Section 6 to both parties having the ability to adopt a primary election.

¹⁷The advantaged party diverges all the way to her ideal point because of our assumption that Θ is strictly larger than $-X_{AR}$, $-X_{AL}$, X_{BR} and X_{BL} . For lower values of Θ we can prove that the advantaged party still diverges away from the median voter's ideal point, though not as much.

¹⁸Indeed running a primary requires material resources like booths, ballots etc.; an effective organization including well trained election officers; a developed party structure which includes an accurate membership roster; and a credible and authoritative arbiter to ensure an orderly process. So not every party might have the ability to run a primary.

5 One party can adopt a primary

In this section we assume that only party B has the necessary technology to implement a primary election should it wish to do so.

5.1 Payoffs

The case where B adopts an elite selection would lead to a situation in which both parties nominate their insider candidates. Thus the probabilities that A 's candidate and B 's candidate are high-skilled are given by π_{AI} and π_{BI} , respectively. A primary election has two differences: first, the probability that B 's nominee is high-skilled increases from π_{BI} to $\pi_{BI} + S_B$, where S_B is the skill bonus of adopting a primary in B . As elaborated in Section 2, that skill bonus comes from the fact that an outsider candidate joins the primary race, and that the primary campaigns reveal information about pre-candidates' skills. And second, it would be the RAF's preferences, and not the leadership's preferences, that would determine B 's policy platform, and thus it would be X_{BR} and not X_{BL} that B 's candidate would adopt if it had skill advantage over A 's candidate.

These remarks are summarized in the following result, where we call $EU_{BL}(c_B)$ the expected utility of B 's leadership from adopting c_B as its CSM.

Lemma 1 *The expected utility of B 's leadership for each value of c_B , $EU_{BL}(c_B)$, is*

$$\begin{aligned} EU_{BL}(elite) &= -(X_{BL} - X_{AL}) \pi_{AI} (1 - \pi_{BI}) \\ &\quad - X_{BL} [\pi_{AI} \pi_{BI} + (1 - \pi_{AI}) (1 - \pi_{BI})] \end{aligned} \quad (5)$$

$$\begin{aligned} EU_{BL}(primary) &= -(X_{BL} - X_{AL}) \pi_{AI} (1 - (\pi_{BI} + S_B)) \\ &\quad - X_{BL} [\pi_{AI} (\pi_{BI} + S_B) + (1 - \pi_{AI}) (1 - (\pi_{BI} + S_B))] \\ &\quad - |X_{BL} - X_{BR}| (1 - \pi_{AI}) (\pi_{BI} + S_B) \end{aligned} \quad (6)$$

Armed with these results, the leadership in party B can measure the consequences of choosing one CSM over the other.

5.2 The optimal selection of a CSM

The leadership in party B will choose the value of c_B by comparing equations 5 and 6 and choosing the CSM that yields the highest expected utility. In case both equations are equal we will make the following indifference assumption:

A2: If a party's leadership is indifferent between a primary election and an elite selection, it will choose an elite selection.

By glancing at those two equations we can see the trade-off that B 's leadership faces in choosing a primary election over an elite selection: the benefit of a primary is that the probability of nominating a low-skilled candidate decreases (due to the primary skill bonus S_B). The cost is that the payoff from having a skill advantage also decreases (due to the ideological distance $X_{BL} - X_{BR}$). Put differently, a primary makes losing less likely but makes winning less attractive.

A primary will therefore be adopted if and only if $EU_{BL}(\text{elite}) < EU_{BL}(\text{primary})$. That condition leads to the following instructive result, recalling that $\lambda_B \equiv X_{BR} - X_{BL}$:

Theorem 2 *The leadership of party B will adopt a primary election if and only if*

$$-\Lambda_B < \lambda_B < \Lambda_B$$

with $\Lambda_B \equiv \frac{S_B[X_{BL}(1-\pi_{AI})-X_{AL}\pi_{AI}]}{(1-\pi_{AI})(\pi_{BI}+S_B)}$.

The intuition behind this result is that B 's leadership will delegate the nomination to its RAF only when the RAF's ideology is close enough to its own. The same intuition can be obtained from Figure 1. If B 's RAF is so ideologically extreme that $\Lambda_B \leq \lambda_B$, then the leadership will not adopt a primary election.

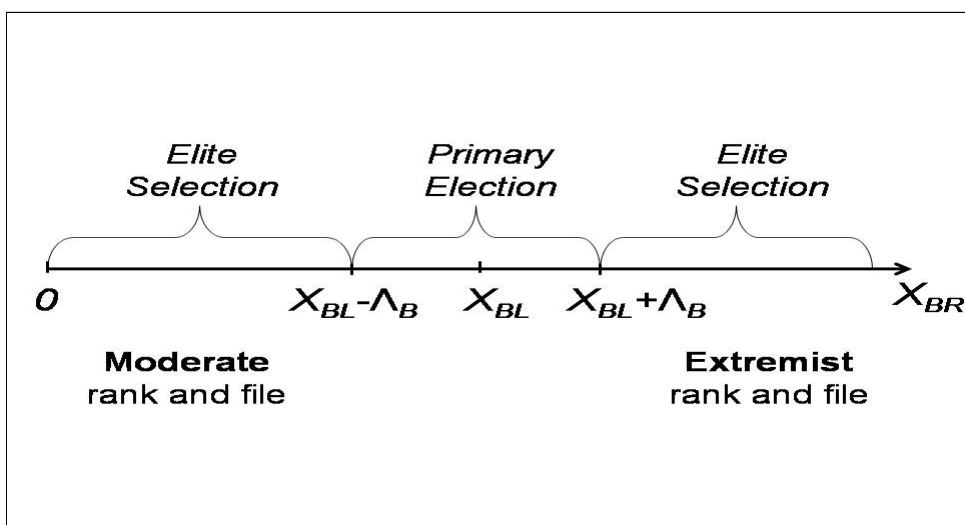


Figure 1

The theorem also describes a striking flip-side to this result: that the party's leadership might find its RAF too *centrist* to adopt a primary. For certain values of the parameters it is possible to have $\lambda_B \leq -\Lambda_B \leq 0$, in which case a primary election will not be adopted because the primary voters are considered too moderate.¹⁹

Given that the region of RAF ideologies for which a primary is adopted is determined by the threshold Λ_B , we can interpret Λ_B as indicating the *likelihood that B will adopt a primary*. Indeed for a larger Λ_B it is more likely that the ideological distance between B 's elite and RAF will be acceptable to adopt a primary. Then a way of phrasing this theorem is that the likelihood of opening the CSM decreases with the ideological distance between the party's leadership and the primary voters.²⁰

¹⁹Such an interval will exist as long as X_{BL} is large enough, specifically $X_{BL} > \frac{-X_{AL}S_B\pi_{AI}}{\pi_{BI}(1-\pi_{AI})}$. Otherwise, there will not exist an interval where the RAF is "too moderate".

²⁰If we consider smaller values of Θ , say $\Theta < X_{BL}, X_{BR}$, the results degenerate: B always chooses a primary election for any value of the parameters. The reason is the following. When B adopts a primary, the probability of nominating a low-skilled candidate decreases (as was mentioned above). But the payoff from winning the election remains the same (instead of decreasing). That is because when B has a valence advantage, that valence advantage is so small that B 's elite and RAF would both choose exactly the same (moderate) policy in order to win the election. So B 's elite does not face any cost in delegating the nomination to its RAF, and therefore it always chooses a primary election.

It should be noted that this result is compatible with recent research by Meinke et al. (2006). In that paper the authors measure the likelihood that parties will open the delegate selection rules in the United States as a function of the ideological distance between the party elite and the potential primary voters. Their strong and statistically significant results indicate that "as the preferences of state party leaders and the voting public diverge, party leaders do become less willing to open their selection processes."

5.3 Comparative statics

To gain insight on what makes the adoption of primary elections more likely, we now study how Λ_B changes with the parameters in the election. In the following result, $\bar{\pi}$ is a constant equal to $\frac{q^2}{1-2q+2q^2}$ as calculated in the appendix.

Theorem 3 *For $\pi_{BI} \in (0, \bar{\pi})$ the threshold Λ_B is:*

1. *Strictly positive*
2. *Strictly increasing with S_B*
3. *Strictly decreasing with π_{BI}*
4. *Strictly decreasing with π_{AI}*
5. *Strictly decreasing with X_{AL}*
6. *Strictly increasing with X_{BL}*

For $\pi_{BI} \in [\bar{\pi}, 1)$ the threshold Λ_B is always equal to zero.

The first four results of this theorem are quite intuitive. As long as the party considers its insider candidate to be relatively weak (meaning that $\pi_{BI} < \bar{\pi}$), it has a strictly positive likelihood of adopting a primary.²¹ When the primary skill bonus S_B increases, indicating that primaries have become better at revealing the skills of candidates, primaries will be more attractive. When the probability that the insider candidate is high-skilled π_{BI} is larger, perhaps because some polls indicate her surge in popularity, then the party

²¹This result has a similar flavor to the main result in Adams and Merrill (2006) that "weak" parties have the most to gain from a primary election.

leadership is more inclined to appoint her directly through an elite selection. And when the probability that A 's candidate is high-skilled increases, then party A becomes more of a threat to party B having a higher probability of winning the election and implementing its leftist ideal policy, which gives B an incentive to adopt a primary election.

The fifth result is more surprising. All things equal, an increase in X_{AL} indicates that the A has become more centrist. According to the theorem, B then becomes less inclined to adopt a primary. The reason is that a moderation of party A makes its victory less painful for B . So when A 's ideal point becomes more moderate, B can better afford to lose the election.

The most important result is the sixth one: that primaries are more likely when X_{BL} increases meaning that party B becomes more extremist.²² The reason is that the more extremist the preferences of a party, the more painful a defeat would be for that party; and therefore the party is more inclined to adopt a primary to improve its probability of winning. A clear empirical prediction of this result is that extremist parties should, all things equal, be internally more democratic. The qualifier "internally" is important in this claim. Communist and fascist parties, which are at the extreme left and extreme right, might very well stand for autocracy. But the prediction of this model is that they will use more democratic ways of selecting their candidates, such as primary elections.

6 Both parties can adopt primaries

In this section we assume that both parties in the election, A and B , have the necessary technology to implement a primary election should they wish to do so. To shorten the analysis we will focus on the case where the rank-and-file members of both parties are more extremist than their leadership (which is, in any case, the situation most frequently mentioned in the literature). Hence we assume that $X_{AR} \leq X_{AL}$ and $X_{BL} \leq X_{BR}$.

6.1 Payoffs

²²If B becomes so extremist that $\Theta < X_{BR}, X_{BL}$, then the use of a primary actually becomes a certainty.

Stage 1 has four possible outcomes, and thus four sets of payoffs to calculate. Those outcomes correspond to the combinations of decisions (c_A, c_B) that parties A and B will make about their respective CSM, where $c_A, c_B \in \{elite, primary\}$. We will call $EU_{AL}(c_A, c_B)$ and $EU_{BL}(c_A, c_B)$ the expected utilities of A 's leadership and B 's leadership from adopting c_A and c_B as their CSMs. They are calculated in the following result.

Lemma 2 *The expected utilities of A 's leadership and B 's leadership for each pair c_A, c_B are given by*

$$\begin{aligned} EU_A(elite, elite) &= -(-X_{AL}) [\pi_{AI}\pi_{BI} + (1 - \pi_{AI})(1 - \pi_{BI})] \\ &\quad - (-X_{AL} + X_{BL})(1 - \pi_{AI})\pi_{BI} \end{aligned}$$

$$\begin{aligned} EU_B(elite, elite) &= -X_{BL} [\pi_{BI}\pi_{AI} + (1 - \pi_{BI})(1 - \pi_{AI})] \\ &\quad - (X_{BL} - X_{AL})(1 - \pi_{BI})\pi_{AI} \end{aligned}$$

$$\begin{aligned} EU_A(elite, primary) &= -(-X_{AL}) [\pi_{AI}(\pi_{BI} + S_B) + (1 - \pi_{AI})(1 - \pi_{BI} - S_B)] \\ &\quad - (-X_{AL} + X_{BR})(1 - \pi_{AI})(\pi_{BI} + S_B) \end{aligned}$$

$$\begin{aligned} EU_B(elite, primary) &= -(X_{BR} - X_{BL})(\pi_{BI} + S_B)(1 - \pi_{AI}) \\ &\quad - X_{BL} [(\pi_{BI} + S_B)\pi_{AI} + (1 - \pi_{BI} - S_B)(1 - \pi_{AI})] \\ &\quad - (X_{BL} - X_{AL})(1 - \pi_{BI} - S_B)(\pi_{AI}) \end{aligned}$$

$$\begin{aligned} EU_A(primary, elite) &= -(X_{AL} - X_{AR})(\pi_{AI} + S_A)(1 - \pi_{BI}) \\ &\quad - (-X_{AL}) [(\pi_{AI} + S_A)\pi_{BI} + (1 - \pi_{AI} - S_A)(1 - \pi_{BI})] \\ &\quad - (-X_{AL} + X_{BL})(1 - \pi_{AI} - S_A)\pi_{BI} \end{aligned}$$

$$\begin{aligned} EU_B(primary, elite) &= -X_{BL} [\pi_{BI}(\pi_{AI} + S_A) + (1 - \pi_{BI})(1 - \pi_{AI} - S_A)] \\ &\quad - (X_{BL} - X_{AR})(1 - \pi_{BI})(\pi_{AI} + S_A) \end{aligned}$$

$$\begin{aligned}
EU_A(\text{primary}, \text{primary}) &= -(X_{AL} - X_{AR})(\pi_{AI} + S_A)(1 - \pi_{BI} - S_B) \\
&\quad - (-X_{AL})[(\pi_{AI} + S_A)(\pi_{BI} + S_B) + (1 - \pi_{AI} - S_A)(1 - \pi_{BI} - S_B)] \\
&\quad - (-X_{AL} + X_{BR})(1 - \pi_{AI} - S_A)(\pi_{BI} + S_B)
\end{aligned}$$

$$\begin{aligned}
EU_B(\text{primary}, \text{primary}) &= -(X_{BR} - X_{BL})(\pi_{BI} + S_B)(1 - \pi_{AI} - S_A) \\
&\quad - X_{BL}[(\pi_{BI} + S_B)(\pi_{AI} + S_A) + (1 - \pi_{BI} - S_B)(1 - \pi_{AI} - S_A)] \\
&\quad - (X_{BL} - X_{AR})(1 - \pi_{BI} - S_B)(\pi_{AI} + S_A)
\end{aligned}$$

Armed with these results, the leadership in party A and party B can measure the consequences of choosing one CSM over the other.

6.2 Equilibrium

A and B are playing a simultaneous-move game at Stage 1 of the election. Our equilibrium concept, SPE, imposes that A and B play a Nash equilibrium (NE) of the reduced game at this stage, where the payoffs are the expected utilities calculated in the above lemma.

As it turns out, the Nash equilibria that the parties will coordinate on depend crucially on the ideological distances between each party's leadership and RAF, λ_A and λ_B . Given our assumption in this section that the RAFs are more extremist than their leadership, we have that $\lambda_A \leq 0$ and $\lambda_B \geq 0$. As the following theorem indicates, the equilibrium of this election hinges on whether the magnitudes of λ_A and λ_B are small, intermediate or large in a sense made precise below.

Theorem 4 *In this election, the equilibrium pairs (c_A, c_B) for each value of λ_A and λ_B are:*

1. $(\text{primary}, \text{primary})$ if $\Lambda_A < \lambda_A \leq 0$ and $0 \leq \lambda_B < \Lambda_B$
2. $(\text{primary}, \text{primary})$ if $\Lambda_A < \lambda_A$ and $\Lambda_B \leq \lambda_B < \widetilde{\Lambda}_B$
3. $(\text{primary}, \text{primary})$ if $\widetilde{\Lambda}_A < \lambda_A \leq \Lambda_A$ and $\lambda_B < \Lambda_B$
4. $(\text{primary}, \text{elite})$ if $\Lambda_A < \lambda_A$ and $\widetilde{\Lambda}_B \leq \lambda_B$
5. $(\text{elite}, \text{primary})$ if $\lambda_A \leq \widetilde{\Lambda}_A$ and $\lambda_B < \Lambda_B$

6. *(primary, primary)* or *(elite, elite)* if $\widetilde{\Lambda}_A < \lambda_A \leq \Lambda_A$ and $\Lambda_B \leq \lambda_B < \widetilde{\Lambda}_B$
7. *(elite, elite)* if $\widetilde{\Lambda}_A < \lambda_A \leq \Lambda_A$ and $\widetilde{\Lambda}_B \leq \lambda_B$
8. *(elite, elite)* if $\lambda_A \leq \widetilde{\Lambda}_A$ and $\Lambda_B \leq \lambda_B < \widetilde{\Lambda}_B$
9. *(elite, elite)* if $\lambda_A \leq \widetilde{\Lambda}_A$ and $\widetilde{\Lambda}_B \leq \lambda_B$

where

$$\begin{aligned}\Lambda_A &\equiv \frac{S_A(X_{AL}(1 - \pi_{BI}) - X_{BL}\pi_{BI})}{(\pi_{AI} + S_A)(1 - \pi_{BI})} \\ \Lambda_B &\equiv \frac{S_B(X_{BL}(1 - \pi_{AI}) - X_{AL}\pi_{AI})}{(\pi_{BI} + S_B)(1 - \pi_{AI})} \\ \widetilde{\Lambda}_A &\equiv \frac{S_A(X_{AL}(1 - \pi_{BI} - S_B) - (X_{BL} + \lambda_B)(S_B + \pi_{BI}))}{(\pi_{AI} + S_A)(1 - \pi_{BI} - S_B)} \\ \widetilde{\Lambda}_B &\equiv \frac{S_B(X_{BL}(1 - \pi_{AI} - S_A) - (X_{AL} + \lambda_A)(\pi_{AI} + S_A))}{(\pi_{BI} + S_B)(1 - S_A - \pi_{AI})}\end{aligned}$$

These equilibria can be depicted in the space (λ_A, λ_B) , as done in Figure 2. We can draw several conclusions from this figure and from the theorem.

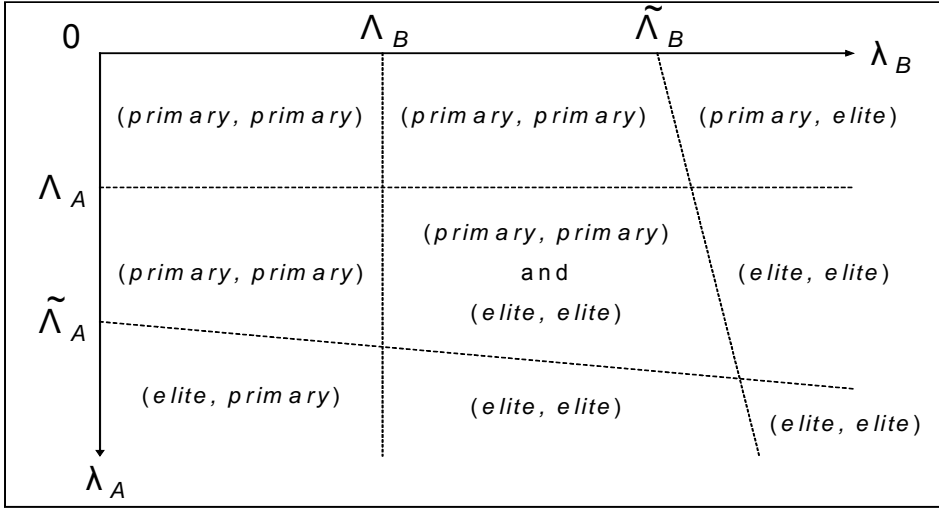


Figure 2

First, the use of primary elections is associated with low ideological distance (that is, values of λ_A and λ_B close to zero). For example when the

ideological distance between B 's leadership and RAF is small, meaning that $\lambda_B < \Lambda_B$, party B will always adopt a primary election for any decision that A makes about adopting a primary or not.

Second, for intermediate values of λ_A and λ_B the outcome is uncertain, in the sense that there are multiple equilibria. When $\widetilde{\Lambda}_A < \lambda_A \leq \Lambda_A$ and $\Lambda_B \leq \lambda_B < \widetilde{\Lambda}_B$ there are two Nash equilibria in pure strategies that parties could coordinate on (and, as indicated in the appendix, there is a third Nash equilibrium in mixed strategies). In contrast, all other cases where λ_A or λ_B are either small or large have a unique Nash equilibrium. This theorem thus makes the interesting prediction that institutional stability does not necessarily have a monotonic relationship to the structural parameters: the level of internal democracy within parties is predictable and stable for low and for large ideological distance between leaders and members, but is uncertain and unstable for intermediate values of such distance.

Third and most importantly, there is a large degree of *contagion* among parties. By "contagion" we refer to different parties adopting the same CSM (as used by De Luca, Jones and Tula (2002)). This result is relevant for empirical studies. Indeed, it is sometimes claimed that a party democratizing its candidate selection will influence other parties to democratize as well (see De Luca et al. (2002) for a discussion). This paper provides a theoretical basis for that claim. In seven of the nine possible cases listed in the theorem, parties end up adopting (*primary, primary*) or (*elite, elite*). Even in the intermediate case where parties have multiple equilibria, those equilibria display $c_A = c_B$. Only in cases of extreme asymmetry between parties do we observe them adopting different CSMs. A clear empirical prediction of this model is thus a large correlation in the parties' level of internal democracy.

7 Conclusions

Primary elections are a frequent method used by political parties around the world to select their candidates –and increasingly so. Some benefit, private or public, needs to be pointed out if we are to explain the increasing popularity of primary elections. In particular, some benefit needs to be identified for the party leaders who are responsible of choosing whether to adopt a primary or not. The premise in this paper is that primary elections can serve as a mechanism to reveal information about the candidates' personal appeal to voters. In particular, by forcing candidates to run a primary campaign before

the general election campaign, the candidates reveal their campaigning skills and the primary voters can select them accordingly. A second feature of primary elections is opening the door to previously unknown or marginalized candidates, thus expanding the pool of pre-candidates that the party can choose from.

An implication of those two features is that a primary election will increase the expected valence of the party's nominee. However, this benefit will be offset by the possibility that primaries will push candidates to adopt policies that the party leaders disapprove of. Therefore party leaders face a trade-off between choosing a primary election versus an elite-centered selection method.

According to the results in this paper, that trade-off depends on parameters in systematic, but sometimes surprising ways. Quite intuitively, party leaders are more likely to adopt primary elections when the candidates they have otherwise available are expected to have low campaigning skills; when primaries are highly effective at revealing accurate information about the candidates' appeal to voters; when the potential primary voters are not too extremist; and when the opposing party has a strong candidate. More surprisingly, party leaders are also more likely to adopt primaries when the potential primary voters are *not too moderate*, when the opposing party is *centrist*, and most intriguingly, when those leaders have *extremist* policy preferences.

When studying the game-theoretic interaction between the two parties deciding which CSM to adopt, we reach two additional conclusions. First, intermediate values in the parameters lead to institutional instability, meaning that there are multiple equilibria which parties could coordinate on. And second, parties are expected to display a significant degree of contagion, meaning that a party's adoption of a certain CSM will influence the other party to adopt the same CSM. An empirical prediction of this result is thus a positive correlation between parties' adoption of primary elections.

In essence, this paper offers an explanation for why and when political parties decide to adopt primary elections, and points out several of the consequences of that decision. More broadly, this paper adds to the literature on endogenous institutions which stipulates that the rules governing political action are not always imposed, but are rather chosen by the political actors themselves. Such is the case the case of candidate-selection rules, whose origin lie in the strategic calculations of the selectors themselves.

8 Appendix: Primaries as an information-revelation mechanism

8.1 The model

Assume that a given party, which we will generically call party K , needs to nominate a candidate for an upcoming election. Candidates are characterized by a parameter θ denoting the candidate's *campaigning skill*. A candidate's campaigning skill can take two values: a high value denoted by Θ corresponding to a high-skilled candidate, and low value normalized to zero corresponding to a low-skilled candidate.

Before choosing its candidate, K 's leadership must choose its candidate-selection method c_K with $c_K \in \{primary, elite\}$, where *primary* refers to the adoption of a primary election and *elite* refers to an elite selection. If $c_K = elite$, party K has only one pre-candidate to choose from called the *insider* candidate and denoted by KI . If $c_K = primary$, party K has two pre-candidates to choose from which are the *insider*, KI , and an *outsider* denoted by KO . We call θ_{KI} and θ_{KO} the campaigning skills of KI and KO respectively, and we call θ_K the skill of the candidate that is finally nominated by K . The exact values of KI 's and KO 's campaigning skills are uncertain ex-ante, but the party has some prior information about the skill of its insider candidate KI . According to that information, candidate KI has a probability π_{KI} of being high-skilled, with $\pi_{KI} \in (0, 1)$. On the other hand, there is no prior information about candidate KO , so she is believed to have a probability of one-half of being high-skilled. All this information is common knowledge.

Party leaders and party members want to maximize the expected skill of the party's nominee, $E(\theta_K)$. Note that the expected value of K 's nominee is proportional to her probability of being high-skilled, given that

$$E(\theta_K) = P(\theta_K = \Theta) \cdot \Theta$$

And therefore it is equivalent for party K to maximize $E(\theta_K)$ or to maximize $P(\theta_K = \Theta)$, which is what we will focus on in the rest of this section.

If party leaders choose an elite-centered selection they will directly nominate KI and therefore $P(\theta_K = \Theta|elite) = \pi_{KI}$. If however they choose to adopt a primary election the candidate KO will join the race, and the nom-

ination is delegated to the party's RAF who will vote between KI and KO .

If there is a primary election, party members interpret the performance of a candidate in the primary-election campaign as a *forecast* of how well they would perform in the general-election campaign.

A candidate's performance in the primary can itself reflect high skill or low skill. We assume that a high-skilled candidate has a larger probability of having a high-skilled performance in the primary. To be concrete, we denote by s_j the performance of candidate j , with $j = KI, KO$; and we say that $s_j = high$ if j 's performance in the primary was high skilled, and $s_j = low$ if j 's performance in the primary was low skilled. We assume that a candidate's performance in the primary has a probability q of accurately forecasting the performance she would have in the general election, with $q \in (\frac{1}{2}, 1)$. In other words, s_{KI} and s_{KO} have probability q of "being correct". Essentially, s_{KI} and s_{KO} are (independently distributed) random variables whose distribution depend on θ_{KI} and θ_{KO} in the following way:

$$\begin{aligned} P(s_j = high|\theta_j = \Theta) &= P(s_j = low|\theta_j = 0) = q \\ P(s_j = high|\theta_j = 0) &= P(s_j = low|\theta_j = \Theta) = 1 - q \\ j &= KI, KO \end{aligned}$$

Once the party members observe the candidates' performances, they can update their beliefs about KI 's and KO 's skills using Bayes rule. The candidates' performances are public, and therefore all the RAF members observe the same s_{KI} and s_{KO} . Thus all the RAF members of K update their prior beliefs with the same information.²³

Given its interest in winning the general election, the RAF will vote for the candidate that is believed to have the highest skill. If both candidates are believed to have the same skill thus making the RAF indifferent, we will make the following indifference assumption for convenience.

A3: When a party member is indifferent between KI and KO she will vote for the one whose *prior* probability of being high-skilled was largest. If both had the same prior, she will randomize equally.

²³In related research we have studied the case where the estimates are not public but private, and thus each primary voter has a different s_{KI} and s_{KO} . We found that the effects of holding a primary are the same as here, but slightly weaker.

8.2 Results

We define $S_K \equiv P(\theta_K = \Theta|primary) - P(\theta_K = \Theta|elite)$, which we call the *skill bonus* of a primary. An important implication of this definition is that

$$P(\theta_K = \Theta|primary) = \pi_{KI} + S_K$$

which means that S_K can also be interpreted as the extra probability of having a high-skilled candidate that a primary brings above an elite selection.

The exact values of $P(\theta_K = \Theta|primary)$, $P(\theta_K = \Theta|elite)$ and S_K are calculated in Section 2 of this appendix. Their most salient feature is the following result. In all results below, the two constants $\underline{\pi}$ and $\bar{\pi}$ are given by $\underline{\pi} = \frac{(1-q)^2}{1-2q+2q^2}$ and $\bar{\pi} = \frac{q^2}{1-2q+2q^2}$.

Proposition 1 *The primary skill bonus S_K is strictly positive for $\pi_{KI} \in (0, \bar{\pi})$ and zero for $\pi_{KI} \in [\bar{\pi}, 1)$.*

8.3 Comparative statics

We now study how S_K changes with respect to its two main determinants: the prior about the insider candidate's skill, π_{KI} , and the accuracy of the candidates' performances q . We first describe the comparative statics with respect to π_{KI} .

Proposition 2 *The primary skill bonus S_K is strictly decreasing with π_{KI} for $\pi_{KI} \in (0, \bar{\pi})$, and constant (equal to zero) for $\pi_{KI} \in [\bar{\pi}, 1)$.*

Now we describe the comparative statics with respect to q , which are still intuitive but less straightforward. As the following result shows, for intermediate values of π_{KI} the primary skill bonus is increasing with q . But surprisingly, it can be insensitive to *small* changes in q when π_{KI} is very large or very small.

Proposition 3 *The effect on the primary skill bonus S_K of a marginal increase in q is strictly positive for $\pi_{KI} \in [\underline{\pi}, \bar{\pi}]$, but is null for $\pi_{KI} \in (0, \underline{\pi})$ and $\pi_{KI} \in (\bar{\pi}, 1)$.*

There is a sense, however, in which q has a positive effect on the primary skill bonus *for any* value of π_{KI} if *large enough* improvements in q can be achieved. This is underlined by the result below.

Proposition 4 For every $q \in (\frac{1}{2}, 1)$ there exists a $q' > q$ such that for every $\pi_{KI} \in (0, 1)$, when q increases to q' , S_K will increase.

9 Appendix: Proofs

9.1 Theorem 1

Proof. This table comes directly from the table in Theorem 1 of Serra (2007b) when δ can take only three values, $\delta \in \{-\Theta, 0, \Theta\}$. In the context of this paper we must recall our assumption that $-X_{AL}, -X_{AR}, X_{BL}, X_{BR} < \Theta$, which implies that $X_B \leq \Theta$ and $-\Theta \leq X_A$. All the results follow. ■

9.2 Lemma 1

Proof. First we calculate $EU_{BL}(elite)$. Given the prior beliefs about AI and BI 's skills, the distribution of possible values of δ is

$$\delta = \begin{cases} \Theta & \text{with probability } (1 - \pi_{AI}) \pi_{BI} \\ 0 & \text{with probability } \pi_{AI} \pi_{BI} + (1 - \pi_{AI}) (1 - \pi_{BI}) \\ -\Theta & \text{with probability } \pi_{AI} (1 - \pi_{BI}) \end{cases}$$

which, according to Theorem 1, leads to the following distribution of equilibrium platforms:

$$x^* = \begin{cases} X_{BL} & \text{with probability } (1 - \pi_{AI}) \pi_{BI} \\ 0 & \text{with probability } \pi_{AI} \pi_{BI} + (1 - \pi_{AI}) (1 - \pi_{BI}) \\ X_{AL} & \text{with probability } \pi_{AI} (1 - \pi_{BI}) \end{cases}$$

where x^* refers to the platform that wins in equilibrium.

Accordingly, the expected utilities of B 's leaders are given by

$$\begin{aligned} E(U_{BL}(x^*) | elite) &= -|X_{BL} - X_{BL}| (1 - \pi_{AI}) \pi_{BI} \\ &\quad - |X_{BL} - 0| [\pi_{AI} \pi_{BI} + (1 - \pi_{AI}) (1 - \pi_{BI})] \\ &\quad - |X_{BL} - X_{AL}| \pi_{AI} (1 - \pi_{BI}) \end{aligned}$$

which is equivalent to the result stated in the theorem.

We now calculate $EU_{BL}(primary)$. We first calculate the new distribution of δ , noting that $P(\theta_B = \Theta | primary) = \pi_{BI} + S_B$ as stated in Section

2. This gives

$$\delta = \begin{cases} \Theta & \text{with probability } (1 - \pi_{AI})(\pi_{BI} + S_B) \\ 0 & \text{with probability } \pi_{AI}(\pi_{BI} + S_B) + (1 - \pi_{AI})(1 - (\pi_{BI} + S_B)) \\ -\Theta & \text{with probability } \pi_{AI}(1 - (\pi_{BI} + S_B)) \end{cases}$$

which, according to Theorem 1, leads to the following distribution of equilibrium platforms (remembering that it is X_{BR} and not X_{BL} that is implemented when B has the highest skilled candidate.)

$$x^* = \begin{cases} X_{BR} & \text{with probability } (1 - \pi_{AI})(\pi_{BI} + S_B) \\ 0 & \text{with probability } \pi_{AI}(\pi_{BI} + S_B) + (1 - \pi_{AI})(1 - (\pi_{BI} + S_B)) \\ X_{AL} & \text{with probability } \pi_{AI}(1 - (\pi_{BI} + S_B)) \end{cases}$$

Accordingly, the expected utility of B 's leadership is given by

$$\begin{aligned} E(U_{BL}(x^*)) &= -|X_{BL} - X_{BR}|(1 - \pi_{AI})(\pi_{BI} + S_B) \\ &\quad - |X_{BL} - 0|[\pi_{AI}(\pi_{BI} + S_B) + (1 - \pi_{AI})(1 - (\pi_{BI} + S_B))] \\ &\quad - |X_{BL} - X_{AL}|\pi_{AI}(1 - (\pi_{BI} + S_B)) \end{aligned}$$

which is equivalent to the expression of $EU_{BL}(\textit{primary})$ given in the lemma. ■

9.3 Theorem 2

Proof. Given our indifference assumption A3, B will adopt a primary election if and only if $EU_{BL}(\textit{elite}) < EU_{BL}(\textit{primary})$. Recalling that $\lambda_B \equiv X_{BR} - X_{BL}$, it is just a matter of algebra to prove that

$$\begin{aligned} EU_{BL}(\textit{elite}) &< EU_{BL}(\textit{primary}) \Leftrightarrow \\ 0 &< -|\lambda_B|(1 - \pi_{AI})(\pi_{BI} + S_B) \\ &\quad - X_{BL}[\pi_{AI}S_B + (1 - \pi_{AI})(-S_B)] \\ &\quad - (X_{BL} - X_{AL})\pi_{AI}(-S_B) \Leftrightarrow \\ 0 &< -|\lambda_B|(1 - \pi_{AI})(\pi_{BI} + S_B) + S_B[X_{BL}(1 - \pi_{AI}) - X_{AL}\pi_{AI}] \end{aligned}$$

Solving for $|\lambda_B|$ in the expression above tells us that the leadership of party B will adopt a primary election if and only if $|\lambda_B| < \frac{S_B[X_{BL}(1 - \pi_{AI}) - X_{AL}\pi_{AI}]}{(1 - \pi_{AI})(\pi_{BI} + S_B)}$.

By defining $\Lambda_B \equiv \frac{S_B[X_{BL}(1-\pi_{AI})-X_{AL}\pi_{AI}]}{(1-\pi_{AI})(\pi_{BI}+S_B)}$, that inequality is equivalent to $-\Lambda_B < \lambda_B < \Lambda_B$. ■

9.4 Theorem 3

Proof. First note that $\Lambda_B = 0$ whenever $S_B = 0$. And we know from Proposition 1 that $S_B = 0$ for $\pi_{BI} \in [\bar{\pi}, 1)$. Therefore Λ_B equals zero in that interval, and will not change with any of the parameters.

We also know from Proposition 1 that $S_B > 0$ for $\pi_{BI} \in (0, \bar{\pi})$, and thus $\Lambda_B > 0$ in that interval. To derive the comparative statics in the theorem we differentiate Λ_B with respect to the parameters.

To study the effect of S_B we note that $\frac{\partial \Lambda_B}{\partial S_B} = \frac{[X_{BL}(1-\pi_{AI})-X_{AL}\pi_{AI}]\pi_{BI}}{(1-\pi_{AI})(\pi_{BI}+S_B)^2}$ which is strictly positive, and therefore Λ_B is strictly increasing with S_B .

To study π_{BI} we must note that it has two effects on Λ_B : a direct effect, and an indirect effect through its effect on S_B . In total, we have that $\frac{d\Lambda_B}{d\pi_{BI}} = \frac{\partial \Lambda_B}{\partial \pi_{BI}} + \frac{\partial \Lambda_B}{\partial S_B} \frac{\partial S_B}{\partial \pi_{BI}}$. We can calculate that $\frac{\partial \Lambda_B}{\partial \pi_{BI}} = \frac{-S_B[X_{BL}(1-\pi_{AI})-X_{AL}\pi_{AI}]}{(1-\pi_{AI})(\pi_{BI}+S_B)^2}$ which is strictly negative. On the other hand we just calculated that $\frac{\partial \Lambda_B}{\partial S_B}$ is strictly positive, and we know from Theorem 2 that $\frac{\partial S_B}{\partial \pi_{BI}}$ is strictly negative. We therefore have that $\frac{d\Lambda_B}{d\pi_{BI}} < 0$ and Λ_B is strictly decreasing with π_{BI} .

To study the effect of π_{AI} we note that $\frac{\partial \Lambda_B}{\partial \pi_{AI}} = \frac{S_B(\pi_{BI}+S_B)(-X_{AL})}{((1-\pi_{AI})(\pi_{BI}+S_B))^2}$ which is strictly positive, and therefore Λ_B is strictly increasing with π_{AI} .

To study the effect of X_{AL} we note that $\frac{\partial \Lambda_B}{\partial X_{AL}} = \frac{S_B[-\pi_{AI}]}{(1-\pi_{AI})(\pi_{BI}+S_B)}$ which is strictly negative, and therefore Λ_B is strictly decreasing with X_{AL} .

To study the effect of X_{BL} we note that $\frac{\partial \Lambda_B}{\partial X_{BL}} = \frac{S_B(1-\pi_{AI})}{(1-\pi_{AI})(\pi_{BI}+S_B)}$ which is strictly positive, and therefore Λ_B is strictly increasing with X_{BL} . ■

9.5 Lemma 2

Proof. We will just calculate the expected utilities for B , given that the results for A are symmetric. First note that if $c_A = elite$, B faces the same situation as in the previous section, and therefore $EU_B(elite, elite)$ and $EU_B(elite, primary)$ are given by equations 5 and 6 respectively.

From the point of view of party B , there are only two difference if A adopts a primary instead of an elite selection: first that the probability that A 's candidate is high-skilled increases by S_A , and second that the ideal point of A 's candidate is that of A 's RAF rather than that of A 's leadership.

Therefore $EU_B(\textit{primary}, \textit{elite})$ is obtained by replacing π_{AI} by $\pi_{AI} + S_A$, and X_{AL} by X_{AR} in the expression for $EU_B(\textit{elite}, \textit{elite})$.

$EU_B(\textit{primary}, \textit{primary})$ is obtained by replacing π_{BI} by $\pi_{BI} + S_B$, and X_{BL} by X_{BR} in the expression for $EU_B(\textit{primary}, \textit{elite})$. ■

9.6 Theorem 4

Proof. In order to find the Nash equilibria we need to compute what the optimal strategy is for each party, for each given strategy of the other party.

- If $c_B = \textit{elite}$, what is the optimal c_A ?

Given our indifference assumption A3, we know that A will prefer a primary over an elite selection if and only if $EU_A(\textit{elite}, \textit{elite}) < EU_A(\textit{primary}, \textit{elite})$. By plugging the values calculated in the previous lemma, and solving for λ_A , this inequality is equivalent to $\lambda_A < \Lambda_A$ with $\Lambda_A \equiv \frac{1}{(\pi_{AI} + S_A)(1 - \pi_{BI})} S_A(X_{AL}(1 - \pi_{BI}) - X_{BL}\pi_{BI})$.

- If $c_A = \textit{elite}$, what is the optimal c_B ?

Given our indifference assumption A3, we know that B will prefer a primary over an elite selection if and only if $EU_B(\textit{elite}, \textit{elite}) < EU_B(\textit{elite}, \textit{primary})$. By plugging the values calculated in the previous lemma, and solving for λ_B , this inequality is equivalent to $\lambda_B < \Lambda_B$ with $\Lambda_B \equiv \frac{1}{(\pi_{BI} + S_B)(1 - \pi_{AI})} S_B(X_{BL}(1 - \pi_{AI}) - X_{AL}\pi_{AI})$.

- If $c_B = \textit{primary}$, what is the optimal c_A ?

Given our indifference assumption A3, we know that A will prefer a primary over an elite selection if and only if $EU_A(\textit{elite}, \textit{primary}) < EU_A(\textit{primary}, \textit{primary})$. By plugging the values calculated in the previous lemma, and solving for λ_A , this inequality is equivalent to $\lambda_A < \Lambda_A$ with $\Lambda_A \equiv \frac{1}{(\pi_{AI} + S_A)(1 - \pi_{BI} - S_B)} S_A(X_{AL}(1 - \pi_{BI} - S_B) - (X_{BL} + \lambda_B)(S_B + \pi_{BI}))$.

- If $c_A = \textit{primary}$, what is the optimal c_B ?

Given our indifference assumption A3, we know that B will prefer a primary over an elite selection if and only if $EU_B(\text{primary}, \text{elite}) < EU_B(\text{primary}, \text{primary})$. By plugging the values calculated in the previous lemma, and solving for λ_B , this inequality is equivalent to $\lambda_B < \widetilde{\Lambda}_B$ with $\widetilde{\Lambda}_B \equiv \frac{1}{(\pi_{BI} + S_B)(1 - S_A - \pi_{AI})} S_B (X_{BL} (1 - \pi_{AI} - S_A) - (X_{AL} + \lambda_A) (\pi_{AI} + S_A))$.

With straightforward algebra we can prove that $\widetilde{\Lambda}_A \leq \Lambda_A \leq 0$ (with strict inequalities if $\pi_{AI} < \bar{\pi}$), and that $0 \leq \Lambda_B \leq \widetilde{\Lambda}_B$ (with strict inequalities if $\pi_{BI} < \bar{\pi}$). We can thus partition the space of possible pairs (λ_A, λ_B) in 9 areas, which are listed below. For each of those areas, we look for the possible Nash equilibria.

1. if $\Lambda_A < \lambda_A \leq 0$ and $0 \leq \lambda_B < \Lambda_B$

Both parties have a strictly dominant strategy which $c_A = \text{primary}$ and $c_B = \text{primary}$. Therefore the unique NE is $(\text{primary}, \text{primary})$.

2. if $\Lambda_A < \lambda_A$ and $\Lambda_B \leq \lambda_B < \widetilde{\Lambda}_B$

A has a strictly dominant strategy which is $c_A = \text{primary}$. B 's best response to $c_A = \text{primary}$ is $c_B = \text{primary}$ so the unique NE is $(\text{primary}, \text{primary})$.

3. if $\widetilde{\Lambda}_A < \lambda_A \leq \Lambda_A$ and $\lambda_B < \Lambda_B$

B has a strictly dominant strategy which is $c_B = \text{primary}$. A 's best response to $c_B = \text{primary}$ is $c_A = \text{primary}$ so the unique NE is $(\text{primary}, \text{primary})$.

4. if $\Lambda_A < \lambda_A$ and $\widetilde{\Lambda}_B \leq \lambda_B$

A has a strictly dominant strategy which is $c_A = \text{primary}$. B has a strictly dominant strategy which is $c_B = \text{elite}$. So the unique NE is $(\text{primary}, \text{elite})$.

5. if $\lambda_A \leq \widetilde{\Lambda}_A$ and $\lambda_B < \Lambda_B$

A has a strictly dominant strategy which is $c_A = \text{elite}$. B has a strictly dominant strategy which is $c_B = \text{primary}$. So the unique NE is $(\text{elite}, \text{primary})$.

6. if $\widetilde{\Lambda}_A < \lambda_A \leq \Lambda_A$ and $\Lambda_B \leq \lambda_B < \widetilde{\Lambda}_B$

A 's best response to c_B is $c_A = c_B$, and B 's best response to c_A is $c_B = c_A$. Therefore this is a coordination game in which both parties which to adopt the same strategy. Therefore there are two NE in pure strategies: $(primary, primary)$ and $(elite, elite)$. There is also a NE in mixed strategies in which A adopts $c_A = primary$ with probability

$$\frac{EU_B(elite, primary) - EU_B(elite, elite)}{EU_B(primary, elite) - EU_B(elite, elite) + EU_B(elite, primary) - EU_B(primary, primary)}$$
 and B adopts $c_B = primary$ with probability

$$\frac{EU_A(primary, elite) - EU_A(elite, elite)}{EU_A(primary, elite) - EU_A(elite, elite) + EU_A(elite, primary) - EU_A(primary, primary)}.$$

7. if $\widetilde{\Lambda}_A < \lambda_A \leq \Lambda_A$ and $\widetilde{\Lambda}_B \leq \lambda_B$

B has a strictly dominant strategy which is $c_B = elite$. A 's best response to $c_B = elite$ is $c_A = elite$ so the unique NE is $(elite, elite)$.

8. if $\lambda_A \leq \widetilde{\Lambda}_A$ and $\Lambda_B \leq \lambda_B < \widetilde{\Lambda}_B$

A has a strictly dominant strategy which is $c_A = elite$. B 's best response to $c_A = elite$ is $c_B = elite$ so the unique NE is $(elite, elite)$.

9. if $\lambda_A \leq \widetilde{\Lambda}_A$ and $\widetilde{\Lambda}_B \leq \lambda_B$

Both parties have a strictly dominant strategy which $c_A = elite$ and $c_B = elite$. Therefore the unique NE is $(elite, elite)$.

Although not necessary for this theorem, it is interesting to note that all nine areas can be proved to exist for $\pi_{AI}, \pi_{BI} < \bar{\pi}$. It is quite easy to prove that areas 1 through 8 exist using the fact that Λ_A does not depend on λ_B and Λ_B does not depend on λ_A . Proving that area 9 exists is more laborious but can be done using the fact that the lines defined by $\widetilde{\Lambda}_A$ and $\widetilde{\Lambda}_B$ will always cross in the quadrant where $\lambda_A < 0$ and $0 < \lambda_B$. We omit a detailed proof for space reasons. ■

9.7 Proposition 1

We start by stating and proving a claim regarding the RAF's behavior. We need one more definition: When $s_{KI} \neq s_{KO}$, we say that the RAF "votes according to the performances" if it votes for the candidate whose performance was best.

Claim 1 *Party K's RAF will*

if $\pi_{KI} \in (0, \underline{\pi}]$, always vote for KO

if $\pi_{KI} \in (\underline{\pi}, \frac{1}{2})$, vote according to the performances if $s_{KI} \neq s_{KO}$, and vote for KO if $s_{KI} = s_{KO}$

if $\pi_{KI} = \frac{1}{2}$, vote according to the performances if $s_{KI} \neq s_{KO}$, and randomize between KI and KO if $s_{KI} = s_{KO}$

if $\pi_{KI} \in (\frac{1}{2}, \bar{\pi})$, vote according to the performances if $s_{KI} \neq s_{KO}$, and vote for KI if $s_{KI} = s_{KO}$

if $\pi_{KI} \in [\bar{\pi}, 1)$, always vote for KI

with

$$\underline{\pi} = \frac{(1-q)^2}{1-2q+2q^2} \text{ and } \bar{\pi} = \frac{q^2}{1-2q+2q^2}$$

Proof of Claim 1. Party K's RAF will vote for the candidate that she believes to have highest probability of being high-skilled. The beliefs she holds about each candidate's skill depend on two pieces of information: her prior beliefs, and the information acquired throughout the primary campaign. Given that the RAF is rational, it will update her prior beliefs based on the performances s_{KI} and s_{KO} to form a couple of posterior beliefs about the probabilities that KI and KO are high-skilled. If the RAF uses Bayes Rule to update her prior beliefs after receiving a given estimate, her posterior beliefs will be given by

$$\begin{aligned} P(\theta_{KI} = \Theta | s_{KI} = low) &= \frac{(1-q)\pi_{KI}}{(1-q)\pi_{KI} + q(1-\pi_{KI})} \\ P(\theta_{KI} = \Theta | s_{KI} = high) &= \frac{q\pi_{KI}}{q\pi_{KI} + (1-q)(1-\pi_{KI})} \\ P(\theta_{KO} = \Theta | s_{KO} = low) &= 1-q \\ P(\theta_{KO} = \Theta | s_{KO} = high) &= q \end{aligned}$$

There are four couple of performances (s_{KI}, s_{KO}) that the RAF could receive, which are $(0, 0)$, (Θ, Θ) , $(0, \Theta)$ and $(\Theta, 0)$, We study each of them in turn, along with the decision that the RAF makes upon receiving those couples of estimates.

- If the RAF observes estimates $s_{KI} = low$ and $s_{KO} = low$:

The RAF will vote for KI if $P(\theta_{KO} = \Theta | s_{KO} = low) < P(\theta_{KI} = \Theta | s_{KI} = low)$ which is equivalent (after some algebra) to $\frac{1}{2} < \pi_{KI}$. Then, given our indif-

ference assumption A1, the RAF will vote for KO if $\pi_{KI} < \frac{1}{2}$, will vote for KI if $\frac{1}{2} < \pi_{KI}$, and will randomize equally if $\pi_{KI} = \frac{1}{2}$.

- If the RAF observes estimates $s_{KI} = high$ and $s_{KO} = high$:

The RAF will vote for KI if $P(\theta_{KO} = \Theta | s_{KO} = high) < P(\theta_{KI} = \Theta | s_{KI} = high)$ which is equivalent (after some algebra) to $\frac{1}{2} < \pi_{KI}$. Then, given our indifference assumption A1, the RAF will vote for KO if $\pi_{KI} < \frac{1}{2}$, will vote for KI if $\frac{1}{2} < \pi_{KI}$, and will randomize equally if $\pi_{KI} = \frac{1}{2}$.

- If the RAF observes estimates $s_{KI} = low$ and $s_{KO} = high$:

The RAF will vote for KI (in other words, disregard the candidates' performance) if $P(\theta_{KO} = \Theta | s_{KO} = high) < P(\theta_{KI} = \Theta | s_{KI} = low)$ which is equivalent (after some algebra, and noting that $1 - 2q + 2q^2 > 0$) to $\frac{q^2}{1-2q+2q^2} < \pi_{KI}$. Then, given our indifference assumption A1 (and noting that $\frac{1}{2} < \frac{q^2}{1-2q+2q^2}$), the RAF will vote for KI if and only $\bar{\pi} \leq \pi_{KI}$, with $\bar{\pi} \equiv \frac{q^2}{1-2q+2q^2}$.

- If the RAF observes estimates $s_{KI} = high$ and $s_{KO} = low$:

The RAF will vote for KO (in other words, disregard the candidates' performance) if $P(\theta_{KO} = \Theta | s_{KO} = low) < P(\theta_{KI} = \Theta | s_{KI} = high)$ which is equivalent (after some algebra, and noting that $1 - 2q + 2q^2 > 0$) to $\pi_{KI} < \frac{(1-q)^2}{1-2q+2q^2}$. Then, given our indifference assumption A1 (and noting that $\frac{(1-q)^2}{1-2q+2q^2} < \frac{1}{2}$), the RAF will vote for KO if and only $\pi_{KI} \leq \underline{\pi}$, with $\underline{\pi} \equiv \frac{(1-q)^2}{1-2q+2q^2}$.

The following table summarizes these results:

	$s_{KI} = low$ $s_{KO} = low$	$s_{KI} = high$ $s_{KO} = high$	$s_{KI} = low$ $s_{KO} = high$	$s_{KI} = high$ $s_{KO} = low$
if $\pi_{KI} \in (0, \underline{\pi}]$	Vote for KO	Vote for KO	Vote for KO	Vote for KO
if $\pi_{KI} \in (\underline{\pi}, \frac{1}{2})$	Vote for KO	Vote for KO	Vote for KO	Vote for KI
if $\pi_{KI} = \frac{1}{2}$	Randomize	Randomize	Vote for KO	Vote for KI
if $\pi_{KI} \in (\frac{1}{2}, \bar{\pi})$	Vote for KI	Vote for KI	Vote for KO	Vote for KI
if $\pi_{KI} \in [\bar{\pi}, 1)$	Vote for KI	Vote for KI	Vote for KI	Vote for KI

Which is what the claim states. ■

We can now use the RAF's behavior described in the claim to prove the proposition.

Proof of Proposition 1. We first endeavor to calculate $P(\theta_K = \Theta|primary)$.

We do so by noting that

$$P(\theta_K = \Theta) = \sum_{\theta_{KI}, \theta_{KO}} \sum_{s_{KI}, s_{KO}} P(\theta_K = \Theta|s_{KI}, s_{KO}; \theta_{KI}, \theta_{KO}) \cdot P(s_{KI}, s_{KO}|\theta_{KI}, \theta_{KO}) \cdot P(\theta_{KI}, \theta_{KO})$$

which uses the definition of conditional probability twice.

Given the prior probabilities about the skills of KI and KO , we have that

$$P(\theta_{KI}, \theta_{KO}) = \begin{cases} (1 - \pi_{KI}) \frac{1}{2} & \text{for } \theta_{KI} = 0 \text{ and } \theta_{KO} = 0 \\ (1 - \pi_{KI}) \frac{1}{2} & \text{for } \theta_{KI} = 0 \text{ and } \theta_{KO} = \Theta \\ \pi_{KI} \frac{1}{2} & \text{for } \theta_{KI} = \Theta \text{ and } \theta_{KO} = 0 \\ \pi_{KI} \frac{1}{2} & \text{for } \theta_{KI} = \Theta \text{ and } \theta_{KO} = \Theta \end{cases}$$

And given the probabilities that the estimates about the skills of KI and KO are correct, the value of $P(s_{KI}, s_{KO}|\theta_{KI}, \theta_{KO})$ is given in the following table

Value of $P(s_{KI}, s_{KO} \theta_{KI}, \theta_{KO})$	If $\theta_{KI} = 0$ and $\theta_{KO} = 0$	If $\theta_{KI} = \Theta$ and $\theta_{KO} = 0$	If $\theta_{KI} = 0$ and $\theta_{KO} = \Theta$	If $\theta_{KI} = \Theta$ and $\theta_{KO} = \Theta$
For $s_{KI} = low$ and $s_{KO} = low$	q^2	$(1 - q)q$	$q(1 - q)$	$(1 - q)^2$
For $s_{KI} = low$ and $s_{KO} = high$	$q(1 - q)$	$(1 - q)^2$	q^2	$(1 - q)q$
For $s_{KI} = high$ and $s_{KO} = low$	$(1 - q)q$	q^2	$(1 - q)^2$	$q(1 - q)$
For $s_{KI} = high$ and $s_{KO} = high$	$(1 - q)^2$	$q(1 - q)$	$(1 - q)q$	q^2

In order to calculate $P(\theta_K = \Theta|s_{KI}, s_{KO}; \theta_{KI}, \theta_{KO})$ we use Lemma 1, which tells us who the nominee is for every couple s_{KI}, s_{KO} . In accordance with that lemma, we need to study several intervals for π_{KI} separately.

- If $\pi_{KI} \in (0, \underline{\pi}]$:

The nominee is always KO , and therefore $P(\theta_K = \Theta|s_{KI}, s_{KO}; \theta_{KI}, \theta_{KO}) = P(\theta_{KO} = \Theta) = \frac{1}{2}$. Plugging those values in the addition above leads to

$$P(\theta_K = \Theta) = \frac{1}{2}$$

- If $\pi_{KI} \in (\underline{\pi}, \frac{1}{2})$:

The nominee and her probability of being high-skilled are given by the following table

Value of $P(\theta_K = \Theta | s_{KI}, s_{KO}; \theta_{KI}, \theta_{KO})$:

Values of (s_{KI}, s_{KO})	Nominee	If $\theta_{KI} = 0$ and $\theta_{KO} = 0$	If $\theta_{KI} = \Theta$ and $\theta_{KO} = 0$	If $\theta_{KI} = 0$ and $\theta_{KO} = \Theta$	If $\theta_{KI} = \Theta$ and $\theta_{KO} = \Theta$
(0, 0)	<i>KO</i>	0	0	1	1
(0, Θ)	<i>KI</i>	0	1	0	1
(Θ , 0)	<i>KO</i>	0	0	1	1
(Θ , Θ)	<i>KO</i>	0	0	1	1

Plugging those values in the addition above leads to

$$P(\theta_K = \Theta) = \pi_{KI}q^2 + q - \frac{1}{2}q^2 - \pi_{KI}q + \frac{1}{2}\pi_{KI}$$

- If $\pi_{KI} = \frac{1}{2}$:

The nominee and her probability of being high-skilled are given by the following table

Value of $P(\theta_K = \Theta | s_{KI}, s_{KO}; \theta_{KI}, \theta_{KO})$:

Values of (s_{KI}, s_{KO})	Nominee	If $\theta_{KI} = 0$ and $\theta_{KO} = 0$	If $\theta_{KI} = \Theta$ and $\theta_{KO} = 0$	If $\theta_{KI} = 0$ and $\theta_{KO} = \Theta$	If $\theta_{KI} = \Theta$ and $\theta_{KO} = \Theta$
(0, 0)	<i>KO</i> or <i>KI</i>	0	$\frac{1}{2}$	$\frac{1}{2}$	1
(0, Θ)	<i>KI</i>	0	1	0	1
(Θ , 0)	<i>KO</i>	0	0	1	1
(Θ , Θ)	<i>KO</i> or <i>KI</i>	0	$\frac{1}{2}$	$\frac{1}{2}$	1

Plugging those values in the addition above leads to

$$P(\theta_K = \Theta) = \frac{1}{2}q + \frac{1}{4}$$

- If $\pi_{KI} \in (\frac{1}{2}, \bar{\pi})$:

The nominee and her probability of being high-skilled are given by the following table

Value of $P(\theta_K = \Theta | s_{KI}, s_{KO}; \theta_{KI}, \theta_{KO})$:

Values of (s_{KI}, s_{KO})	Nominee	If $\theta_{KI} = 0$ and $\theta_{KO} = 0$	If $\theta_{KI} = \Theta$ and $\theta_{KO} = 0$	If $\theta_{KI} = 0$ and $\theta_{KO} = \Theta$	If $\theta_{KI} = \Theta$ and $\theta_{KO} = \Theta$
(0, 0)	<i>KI</i>	0	1	0	1
(0, Θ)	<i>KI</i>	0	1	0	1
(Θ , 0)	<i>KO</i>	0	0	1	1
(Θ , Θ)	<i>KI</i>	0	1	0	1

Plugging those values in the addition above leads to

$$P(\theta_K = \Theta) = \pi_{KI}q - \pi_{KI}q^2 + \frac{1}{2}q^2 + \frac{1}{2}\pi_{KI}$$

- If $\pi_{KI} \in [\bar{\pi}, 1)$:

The nominee is always *KI*, and therefore $P(\theta_K = \Theta | s_{KI}, s_{KO}; \theta_{KI}, \theta_{KO}) = P(\theta_{KI} = \Theta) = \pi_{KI}$. Plugging those values in the addition above leads to

$$P(\theta_K = \Theta) = \pi_{KI}$$

We can now calculate our value of interest, S for each value of π_{KI} . We use that values above to calculate $S \equiv P(\theta_K = \Theta | \text{primary}) - P(\theta_K = \Theta | \text{elite})$, remembering that $P(\theta_K = \Theta | \text{elite}) = \pi_{KI}$.

$$S = \begin{cases} \frac{1}{2} - \pi_{KI} & \text{if } \pi_{KI} \in (0, \underline{\pi}] \\ \pi_{KI}q^2 + q - \frac{1}{2}q^2 - \pi_{KI}q + \frac{1}{2}\pi_{KI} - \pi_{KI} & \text{if } \pi_{KI} \in (\underline{\pi}, \frac{1}{2}) \\ \frac{1}{2}q + \frac{1}{4} - \pi_{KI} & \text{if } \pi_{KI} = \frac{1}{2} \\ \pi_{KI}q - \pi_{KI}q^2 + \frac{1}{2}q^2 + \frac{1}{2}\pi_{KI} - \pi_{KI} & \text{if } \pi_{KI} \in (\frac{1}{2}, \bar{\pi}) \\ \pi_{KI} - \pi_{KI} & \text{if } \pi_{KI} \in [\bar{\pi}, 1) \end{cases}$$

which gives, with some algebra and noting the continuity of S at $\pi_{KI} = \underline{\pi}$, $\pi_{KI} = \frac{1}{2}$ and $\pi_{KI} = \bar{\pi}$,

$$S = \begin{cases} \frac{1}{2} - \pi_{KI} & \text{if } \pi_{KI} \in (0, \underline{\pi}] \\ \pi_{KI}q^2 - \pi_{KI}q - \frac{1}{2}q^2 - \frac{1}{2}\pi_{KI} + q & \text{if } \pi_{KI} \in [\underline{\pi}, \frac{1}{2}] \\ -\pi_{KI}q^2 + \pi_{KI}q + \frac{1}{2}q^2 - \frac{1}{2}\pi_{KI} & \text{if } \pi_{KI} \in [\frac{1}{2}, \bar{\pi}] \\ 0 & \text{if } \pi_{KI} \in [\bar{\pi}, 1) \end{cases}$$

which is what the theorem claims.

Now we need to analyze the sign of S . If $\pi_{KI} \in (0, \underline{\pi}]$ we have that $S = \frac{1}{2} - \pi_{KI} > 0 \Leftrightarrow \pi_{KI} < \frac{1}{2}$, but that is satisfied because $\pi_{KI} \leq \underline{\pi}$ and we have already noted that $\underline{\pi} < \frac{1}{2}$. If $\pi_{KI} \in [\underline{\pi}, \frac{1}{2}]$ we have that $S = \pi_{KI}q^2 - \pi_{KI}q - \frac{1}{2}q^2 - \frac{1}{2}\pi_{KI} + q > 0 \Leftrightarrow \pi_{KI} < \frac{2q - q^2}{1 + 2q - 2q^2}$ (noting that $1 + 2q - 2q^2 > 0$) which is satisfied because $\frac{1}{2} < \frac{2q - q^2}{1 + 2q - 2q^2}$. If $\pi_{KI} \in [\frac{1}{2}, \bar{\pi})$ we have that $S = -\pi_{KI}q^2 + \pi_{KI}q + \frac{1}{2}q^2 - \frac{1}{2}\pi_{KI} > 0 \Leftrightarrow \pi_{KI} < \frac{q^2}{1 - 2q + 2q^2}$ which is satisfied because $\bar{\pi} = \frac{q^2}{1 - 2q + 2q^2}$. And finally if $\pi_{KI} \in [\bar{\pi}, 1)$ we have $S = 0$. So we have indeed $S > 0$ for $\pi_{KI} \in (0, \underline{\pi}) \cup [\underline{\pi}, \frac{1}{2}) \cup [\frac{1}{2}, \bar{\pi})$ and $S = 0$ for $\pi_{KI} \in [\bar{\pi}, 1)$, as the theorem claims. ■

9.8 Proposition 2

Proof. We calculate the differential of S with respect to π_{KI} and check its sign. If $\pi_{KI} \in (0, \underline{\pi})$, $\frac{\partial S}{\partial \pi_{KI}} = -1$ which is strictly negative. If $\pi_{KI} \in (\underline{\pi}, \frac{1}{2})$, $\frac{\partial S}{\partial \pi_{KI}} = q^2 - q - \frac{1}{2}$ which is strictly negative for $q \in (\frac{1}{2}, 1)$. If $\pi_{KI} \in (\frac{1}{2}, \bar{\pi})$, $\frac{\partial S}{\partial \pi_{KI}} = -q^2 + 2q - 1$ which is strictly negative for $q \in (\frac{1}{2}, 1)$. So S is decreasing with π_{KI} in all those intervals. S is non-differentiable at $\pi_{KI} = \underline{\pi}$ and $\pi_{KI} = \frac{1}{2}$, but is continuous at both points, and is therefore decreasing just like their neighboring points. Hence S decreases with π_{KI} when $\pi_{KI} \in (0, \underline{\pi}) \cup \{\underline{\pi}\} \cup (\underline{\pi}, \frac{1}{2}) \cup \{\frac{1}{2}\} \cup (\frac{1}{2}, \bar{\pi})$.

If $\pi_{KI} \in [\bar{\pi}, 1)$, S is constant for all values of π_{KI} (and equal to zero), so an increase in π_{KI} will not affect it. ■

9.9 Proposition 3

Proof. We calculate the differential of S with respect to q and check its sign, remembering that the values of $\underline{\pi}$ and $\bar{\pi}$ are $\underline{\pi} = \frac{(1-q)^2}{1-2q+2q^2}$ and $\bar{\pi} = \frac{q^2}{1-2q+2q^2}$. According to the values of S in Theorem 1, if $\pi_{KI} \in (0, \underline{\pi})$, $\frac{\partial S}{\partial q} = 0$; similarly if $\pi_{KI} \in (\bar{\pi}, 1)$, $\frac{\partial S}{\partial q} = 0$. So in those intervals, S is unresponsive to marginal changes in q .

However, if $\pi_{KI} \in (\underline{\pi}, \frac{1}{2})$, $\frac{\partial S}{\partial q} = 2\pi_{KI}q - \pi_{KI} + 1 - q$ which is strictly positive; if $\pi_{KI} = \frac{1}{2}$, $\frac{\partial S}{\partial q} = \frac{1}{2}$ which is strictly positive; if $\pi_{KI} \in (\frac{1}{2}, \bar{\pi})$, $\frac{\partial S}{\partial q} = -2\pi_{KI}q + \pi_{KI} + q$ which is strictly positive. So in those intervals, S is strictly increasing with marginal increases in q .

To analyze the cases where $\pi_{KI} = \underline{\pi}$ and $\pi_{KI} = \bar{\pi}$, note that $\frac{\partial}{\partial q} \left(\frac{(1-q)^2}{1-2q+2q^2} \right) < 0$, so with a marginal increase in q , $\underline{\pi}$ remains in the interval $\left[\frac{(1-q)^2}{1-2q+2q^2}, \frac{1}{2} \right]$ where we just proved that S is increasing with q . Similarly note that $\frac{\partial}{\partial q} \left(\frac{q^2}{1-2q+2q^2} \right) > 0$, so with a marginal increase in q , $\bar{\pi}$ remains in the interval $\left[\frac{1}{2}, \frac{q^2}{1-2q+2q^2} \right]$ where we just proved that S is increasing with q .

To summarize, S is unresponsive to marginal changes in q for $\pi_{KI} \in (0, \underline{\pi}) \cup (\bar{\pi}, 1)$, and is strictly increasing with q for $\pi_{KI} \in \{\underline{\pi}\} \cup (\underline{\pi}, \frac{1}{2}) \cup \{\frac{1}{2}\} \cup (\frac{1}{2}, \bar{\pi}) \cup \{\bar{\pi}\}$. ■

9.10 Proposition 4

Proof. First assume that $\pi_{KI} \in [\underline{\pi}, \bar{\pi}]$ with $\underline{\pi} = \frac{(1-q)^2}{1-2q+2q^2}$ and $\bar{\pi} = \frac{q^2}{1-2q+2q^2}$. In that case we know from Lemma 2 that $\frac{\partial S}{\partial q} > 0$. So S is strictly increasing when q increases to any $q' > q$.

Now assume that $\pi_{KI} \in (0, \underline{\pi})$. Note that $\lim_{q \rightarrow 1} \frac{(1-q)^2}{1-2q+2q^2} = 0$, so there exists $q'' > q$ such that $\frac{(1-q'')^2}{1-2q''+2q''^2} < \pi_{KI}$, in which case $\pi_{KI} \in [\underline{\pi}'', \bar{\pi}'']$ where $\underline{\pi}'' = \frac{(1-q'')^2}{1-2q''+2q''^2}$ and $\bar{\pi}'' = \frac{q''^2}{1-2q''+2q''^2}$. And given the result above, S is strictly increasing when q'' increases to any $q' > q''$. And given that S cannot have decreased when q increased to q'' , we have that S will increase strictly when q increases to q'' .

A similar logic is used when $\pi_{KI} \in (\bar{\pi}, 1)$ by noting that $\lim_{q \rightarrow 1} \frac{q^2}{1-2q+2q^2} = 1$. ■

References

- [1] Abramowitz, Alan I. and Walter J. Stone (1984), *Nomination Politics: Party Activists and Presidential Choice*, Praeger Publishers, New York, NY.
- [2] Adams, James (1999), "Policy Divergence in Multicandidate Probabilistic Spatial Voting", *Public Choice*, 100: 103-122.
- [3] Adams, James and Samuel Merrill, III (2006), "The Spatial Strategies of Candidates Facing Both a Primary and a General Election: When

Do Primary Elections Help Parties Win Office?", paper presented at the MPSA 64th Annual Meeting.

- [4] Aldrich, John H. (1980a), "A Dynamic Model of Presidential Nomination Campaigns", *American Political Science Review*, Vol. 74, Issue 3, Sep., 651-669.
- [5] Aldrich, John H. (1980b), *Before the Convention: Strategies and Choices in Presidential Nomination Campaigns*, The University of Chicago Press, Chicago.
- [6] Ansolabehere, Stephen, Mark Hansen, Shigeo Hirano and James M. Snyder, Jr., (2005) "The Decline of Competition in U.S. Primary Elections, 1908-2004", typescript (June 2005), forthcoming in *The Marketplace of Democracy*, edited by Michael McDonald and John Samples.
- [7] Ansolabehere, Stephen, Shigeo Hirano and James Snyder (2004), "What did the Direct Primary do to Party Loyalty in Congress?", March 2004, typescript.
- [8] Ansolabehere, Stephen and James M. Snyder, Jr., (2000), "Valence Politics and Equilibrium in Spatial Election Models", *Public Choice*, Vol. 103, no. 3-4, June 2000, pp. 327-36.
- [9] Ansolabehere, Stephen, James M. Snyder, Jr. and Charles Stewart, III (2001), "Candidate Positioning in U.S. House Elections", *American Journal of Political Science*, Vol. 45, No. 1, January 2001, pp. 136-159.
- [10] Aranson Peter H. and Peter C. Ordeshook, (1972), "Spatial Strategies for Sequential Elections," in *Probability Models of Collective Decision Making*, ed. R. G. Niemi and H. F. Weisberg, Charles E. Merrill Publishing Company, Columbus Ohio.
- [11] Bartels, Larry M. (1988), *Presidential Primaries and the Dynamics of Public Choice*, Princeton University Press, Princeton, New Jersey.
- [12] Bille, Lars (2001), "Democratizing a Democratic Procedure: Myth or Reality?", *Party Politics*, Vol. 7, No. 3, pp. 363-380.
- [13] Buchler, Justin (2005), "Competition, Representation and Redistricting", *Journal of Theoretical Politics* 17(4): 431-463.

- [14] Burden, Barry C. (2001), "The Polarizing Effects of Congressional Primaries", in *Congressional Primaries and the Politics of Representation*, Peter F. Galderisi, Marni Ezra and Michael Lyons eds., Rowman & Littlefield Publishers, New York.
- [15] Burden, Barry C. (2004), "Candidate Positioning in US Congressional Elections", *British Journal of Political Science*, 34(2): 211-227.
- [16] Carey, John M. and John Polga-Hecimovich (2006), "Primary Elections and Candidate Strength in Latin America", *Journal of Politics*, Vol. 68, No. 3, August 2006, pp. 530-543.
- [17] Chen, Kong-Pin and Sheng-Zhang Yang (2002), "Strategic Voting in Open Primaries", *Public Choice*, 112: 1-30.
- [18] Coleman, James S. (1972), "The Position of Political Parties in Elections," in *Probability Model of Collective Decision Making*, ed. R. G. Niemi and H. F. Weisberg, Charles E. Merrill Publishing Company, Columbus Ohio.
- [19] Cukierman, Alex (1991), "Asymmetric Information and the Electoral Momentum of Public Opinion Polls", *Public Choice*, 70: 181-213.
- [20] De Luca, Miguel, Mark P. Jones and María Inés Tula (2002), "Back Rooms or Ballot Boxes? Candidate Nomination in Argentina", *Comparative Political Studies*, Vol.. 35, No 4, 413-436.
- [21] Enelow, James M. and Melvin J. Hinich (1982), "Nonspatial Candidate Characteristics and Electoral Competition", *The Journal of Politics*, Vol. 44, No. 1, Feb., 115-130.
- [22] Ezra, Marni (2001), "The Benefits and Burdens of Congressional Primary Elections", in *Congressional Primaries and the Politics of Representation*, Peter F. Galderisi, Marni Ezra and Michael Lyons eds., Rowman & Littlefield Publishers, New York.
- [23] Geer, John G. (1988), "Assessing the Representativeness of Electorates in Presidential Primaries", *American Journal of Political Science*, Vol. 32, No. 4 (Nov., 1988), pp. 929-945.

- [24] Geer, John G. (1989), *Nominating Presidents: An Evaluation of Voters and Primaries*, Greenwood Press, Westport, Connecticut.
- [25] Gerber, Elisabeth R. and Rebecca B. Morton (1998), "Primary Election Systems and Representation", *Journal of Law, Economics, and Organization*, 14: 304-324.
- [26] Groseclose, Tim (2001), "A Model of Candidate Location When One Candidate Has a Valence Advantage", *American Journal of Political Science*, Vol. 45, No. 4, 862-886.
- [27] Jackson, Matthew O., Laurent Mathevet and Kyle Mattes (2006), "Nomination Processes and Policy Outcomes", *Quarterly Journal of Political Science*, Vol. 2, Issue 1, pp. 67-94, 2007.
- [28] Kanthak, Kristin and Rebecca B. Morton (2001), "The Effects of Electoral Rules on Congressional Primaries", in *Congressional Primaries and the Politics of Representation*, Peter F. Galderisi, Marni Ezra and Michael Lyons eds., Rowman & Littlefield Publishers, New York.
- [29] Key, V. O. (1956), *American state politics: An introduction*, New York: Knopf.
- [30] Kaufmann, Karen M., James G. Gimpel and Adam H. Hoffman (2003) "A Promise Fulfilled? Open Primaries and Representation", *Journal of Politics*, Vol. 65, No. 2, May, pp. 457-476.
- [31] Kang, Insun (2007), "Open Primaries and Crossover Voting", paper presented at the 65th Annual Conference of the Midwest Political Science Association.
- [32] King, David C. (1999), "Party Competition and Polarization in American Politics", manuscript, Harvard University.
- [33] Lingle, James (1981), *Representation and presidential primaries*, Westport, CT: Greenwood.
- [34] Londregan, John and Thomas Romer (1993), "Polarization, Incumbency, and the Personal Vote", in *Political Economy: Institutions, Competition and Representation*, W. A. Barnett, M. J. Hinich, and N.J. Schofield (eds.), Cambridge University Press, Cambridge, UK.

- [35] Meinke, Scott R., Jeffrey K. Staton and Steven T. Wuhs (2006), "State Delegate Selection Rules for Presidential Nominations, 1972-2000", *The Journal of Politics*, Vol. 68, No. 1, February, pp. 180-193.
- [36] Meirowitz, Adam (2005), "Informational Party Primaries and Strategic Ambiguity", *Journal of Theoretical Politics*, 17(1): 107-136.
- [37] Morton, Rebecca B. (2006), *Analyzing Elections: New Institutionalism in American Politics*, W. W. Norton & Company.
- [38] Morton, Rebecca B. and Kenneth C. Williams (2001), *Learning by Voting*, University of Michigan Press, Ann Arbor.
- [39] Owen, Guillermo and Bernard Grofman (2004), "Two-stage electoral competition in two-party contests: persistent divergence of party positions", *Social Choice and Welfare*, Heidelberg: June 2006. Vol. 26, Issue 3; p. 547
- [40] Polsby, Nelson W. (1983), *Consequences of Party Reform*, Cambridge: Oxford University Press.
- [41] Ranney, Austin (1968), "Representativeness of primary electorates", *Midwest Journal of Political Science*, 12:224-238.
- [42] Scarrow, Susan (2005), "Implementing Intra-party Democracy", in *Political Parties and Democracy in Theoretical and Practical Perspective*, National Democratic Institute for International Affairs, research paper series.
- [43] Schofield, Norman (2004), "Equilibrium in the Spatial 'Valence' Model of Politics", *Journal of Theoretical Politics*, 16 (4): 447-481.
- [44] Serra, Gilles (2007a), "Primary Elections versus Smoke-Filled Rooms: A Theory of Party Democratization in Latin America", Chapter 2 of Ph.D. dissertation, Harvard University.
- [45] Serra, Gilles (2007b), "A Model of Candidate Selection with Endogenous Valence", Chapter 3 of Ph.D. dissertation, Harvard University.
- [46] Snyder, James M. and Michael M. Ting (2005), "Electoral Control and Selection with Parties and Primaries", typescript.

- [47] Stokes, Donald (1963), "Spatial Models of Party Competition", *American Political Science Review*, 57:368-377.