When Will Incumbents Avoid a Primary Challenge? Aggregation of Partial Information About Candidates' Valence

Gilles Serra

1 Introduction

Incumbents and other insiders tend to enjoy a comfortable position within their parties. In particular, they frequently have an advantage to secure their party's nomination for a future election. Outsiders who do not necessarily belong to the dominant faction in the party have a much harder time getting their name on the ballot. They are disadvantaged in at least two ways: they might be less well-known than the party grandees they are competing with; and there might not even be a fair competition such as a primary election for them to prove themselves. A question of interest is why parties allow well-known insiders to have such and advantage over lesser-known outsiders. We would imagine an ambitious party that wishes to win elections to find mechanisms for identifying and selecting the best possible candidate, regardless of that candidate's previous standing in the party. One option would be to democratize the nomination process to let fresh outsiders join an open competition where they can display their true campaigning skills. This option is widely available to political parties around the world, though it is not always used. In this paper I explore the conditions under which candidate-selection is democratized, and I show that rational parties who wish to find the most talented candidate may nevertheless shut down the possibility of unknown hopefuls coming forward to display their talents.

Indeed, a political party can use a variety of methods to nominate those who will later compete for office at a given election. Broadly speaking, a candidate-selection method (CSM) can fall in two categories. On one hand, the method could be *open* (or *democratic*) by allowing the participation of all the members, activists and sympathizers of the party in the nomination of candidates. Of all the selection methods that parties can use, the most open and democratic one is the *primary election*. By

G. Serra (✉)
Department of Political Science, Center for Economics Research and Teaching (CIDE), Carretera México-Toluca 3655, Lomas de Santa Fe, Mexico City 01210, Mexico
e-mail: gilles.serra@cide.edu

N. Schofield et al. (eds.), *Advances in Political Economy*,
DOI 10.1007/978-3-642-35239-3_11,

primary election, I refer to the organized competition among aspiring candidates within the same party that culminates in the democratic vote of all party members. On the other hand, the nomination method could be *closed* (or *undemocratic*), consisting of a closed-door decision at the elite level of the party. For example, the nominee for an upcoming presidential or gubernatorial election could be chosen by a handful of party bosses at a private meeting. As argued throughout this paper, the choice matters for the party in terms of its prospects of winning the election; but it also matters for citizens in terms of the quality of candidates they are offered.

Party leaders are for the most part responsible for the way their parties nominate candidates. In most presidential systems, political parties have leeway in choosing their CSM, and it is usually *not* the case that primaries are exogenously imposed on them by the government. In fact, it is common for political parties to have serious deliberations on what CSM to adopt before even discussing which candidates to select. Their adoption of primary elections is most often *voluntary* rather than mandated by law. Throughout Latin America we repeatedly see party elites debating whether to open the nomination process or not. Actually, it is not uncommon for parties to go back and forth between primaries and other CSMs in recurrent elections, which clearly indicates the strategic nature of that choice. In the United States, party elites also have a strong say in choosing whether their nomination will be open and inclusive, or closed and exclusive. They do so by choosing whether to endorse a favored candidate or not. If party leaders decide to rally behind a well-known insider, they will provide her with public endorsements, strategic advice and large amounts of funding to overwhelm any challenger. On the other hand, if party leaders do not identify an insider candidate that satisfies them, they will withhold or divide their endorsements such that a competitive race among several hopefuls takes place. Thus, while parties are "officially" holding a primary election, in practice that primary can be competitive or uncompetitive. In effect, this is equivalent to choosing between a democratic and an undemocratic CSM. Hence, I claim the explanation for the use of primaries around the world lies in the strategic calculations of party leaders

This paper postulates a benefit to party leaders that helps explain why they occasionally allow the use of primary elections within their parties. To be concrete, I claim that primary elections have a practical advantage over elite-centered nominations: *they reveal information about candidates' appeal to voters*. My premise is that a candidate nominated through a primary election can be expected to have higher campaigning skills than a candidate nominated through an elite appointment. This happens because the primary campaigns reveal valuable information about the contenders. Indeed, there is much uncertainty surrounding the individuals seeking to become a party's candidate, often called *pre-candidates*. Their future vote-getting effectiveness is never known for sure. A primary can serve as a "trial" election *within* a party that shares many of the features of the subsequent general election *between* the parties. Pre-candidates must participate in debates, broadcast television advertisements, manage a campaign, and so forth. Thus primaries can reveal how effective the pre-candidates would be in the general election. In that sense, my model provides an "information rationale" for the existence of primary elections.

On the other hand, as mentioned above, primaries might carry several costs to party leaders. In this paper I focus on one oft-mentioned cost: primary elections might push candidates to adopt policies far from the leaders' preferences. Indeed, the party bosses know that primary voters may not quite share their ideology. They might be too extremist or too moderate to be trusted with the selection of the party's candidate. 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

On that basis, I build a spatial voting model that includes a party's choice between a competitive primary election and an elite-centered nomination. The main question is: When does the informational benefit of primaries outweigh the cost of losing control of the candidates' platforms? As the results will indicate, the answer depends on several fundamental variables: the ideology of parties, the ideology of primary voters, the intensity of the primary election, and the quality of insider and outsider candidates.

This model is a continuation of the research in Serra (2011). The main contribution with respect to that research is analyzing the revelation of *partial* information rather than *full* information, by which I mean that primary elections only reveal part of the information needed to assess a contender, but his or her ability to perform well in the general election would still not be known in full. To be concrete, I assume the contenders' performances within the party are interpreted as "noisy signals" that can be interpreted as forecasts of their performance if they were nominated to compete against another party. In this sense, the model falls in the tradition of modeling voting as a process to *aggregate* information—a tradition initiated by Condorcet (1785), Austen-Smith and Banks (1996), Feddersen and Pesendorfer (1998).

Several new results are found with this modeling choice. Two new variables can be studied more precisely. The ability of primaries to reveal valuable information, which I call the *quality* of primaries; and the reputation of the insider candidate as proficient vote-getter, which I call the *prior belief* about the insider's skill. Regarding the quality of primaries, I find that a party can benefit from stiff competition in its primary election. This result stands in contrast with an oft-mentioned view that parties should ensure their primaries are light and cordial. Regarding the prior belief held about the skill of candidates, I find that an insider might have a good enough reputation to prevent a primary election altogether. This result would help explain why many incumbents are able to be re-nominated for a subsequent election without being opposed inside their parties. Both results are new in the literature on primary elections as far as I can tell.

In addition to these new results, many of the previous results in Serra (2011) are corroborated. In particular, this paper also finds that primaries are more likely when there is congruence between the elite and the mass membership of the party; and primaries are more appealing to the party that is most disadvantaged given its valence and policies.

The rest of the paper is developed as follows: Sect. 2 briefly summarizes the theoretical literature that relates to my model. Section 3 introduces a spatial vot-

ing model between two parties that will serve to study the general election. It is a variant of the Downsian voting model, with an additional dimension corresponding to the candidates' valence. In Sect. 4, I take a step back in the electoral process, and I study the nomination that takes place inside a party before the general election. Section 5 develops a signaling mechanism for primary voters to update their beliefs about pre-candidates based on their performance in the primary campaigns. Section 6 introduces a cost of adopting primaries based on the lack of congruence between the elite and the mass in the party. In Sect. 7, I derive a number of conditions for a party to hold a competitive primary election, which is the purpose of this paper. Finally, Sect. 8 discusses the main results and suggests some interpretations of relevance to democratic theory. The Appendix contains all the proofs of the results in this paper.

2 Previous Theories of the Adoption of Primary Elections

The paper adds to the formal literature on primary elections. Most authors have studied the consequences of primaries, rather than their causes. Several papers in that literature share common aspects with this one, especially those comparing different candidate-selection methods (CSM). Owen and Grofman (2006) compare primaries with different degrees of divergence between the party mean and the population mean. Jackson et al. (2007) study three different nomination processes: an arbitrary appointment by a party leader, a primary election, and a spending competition between candidates. In Castanheira et al. (2010), parties select their internal organization possibly including intra-party competition. Cho and Kang (2008) compare open and closed primary elections.

Another set of papers that relate to my model, are those that have paid attention to informational aspects of primaries. In Caillaud and Tirole (2002) and Castanheira et al. (2010), the use of primaries provides information about the credibility and trustworthiness of the party. In Meirowitz (2005), primaries allow candidates to acquire information about voters' preferences. Then there is a set of papers where primaries reveal information about the valence of primary contenders.

For instance, Adams and Merrill (2008) postulate that primary elections may allow a party to identify a high-quality nominee. The authors find, as I do, that weak parties benefit from primaries more than strong parties do. In spite of those similarities, our models have important differences because the focus of their paper is the candidates' choice of platforms, while the focus of my paper is the parties' choice of candidates.

Another closely related paper is Snyder and Ting (2011) who also studies a party's decision to hold a primary election or not. As in my model, parties compete both in terms of ideology and valence. Snyder and Ting also assume that primaries increase the expected valence of the nominee. A main difference is the alternative CSM. If a party does not hold a primary, Snyder and Ting assume that the nominee will be chosen at random among all the willing pre-can-

didates. In contrast, I assume the party elite will choose an insider candidate in a smoke-filled room. Another difference is that both parties are bound to use the same CSM by state law, whereas in my model parties can have different CSMs.

Kselman (2012) develops a model where aspirants must compete in a primary election to obtain their party's nomination. In his model, candidates enjoy a type of valence that serves as a bonus for parties that are office-seeking. Interestingly, this type of valence is particularistic in the sense that only a subset of voters benefit from it.

Finally, this paper is related to the literature on *endogenous valence*. Some other papers have also allowed the agents in their models to affect the valence parameter are Ashworth and de Mesquita (2009), Schofield and Sened (2005), Schofield (2007), Carrillo and Castanheira (2008), Callander (2008), Meirowitz (2008), Schofield et al. (2008).

The model in this paper is one of the few that combines both literatures, the one on valence and the one on primaries. As in Adams and Merrill (2008), Snyder and Ting (2011), and Serra (2011), the premise here is that primaries help parties by revealing the valence of their candidates. Unlike those papers, however, this paper develops a signaling mechanism to reveal partial rather than full information.

3 General Election Between the Two Parties

In this section I focus on the competition between two parties without any reference to primary elections. In essence, this corresponds to the "general election" that occurs after all parties have already completed their nomination cycle. This will be a *valence-policy model*, meaning that it will have two dimensions. First, the election occurs in a left-right policy spectrum. I denote by x the policy implemented, with $x \in \mathbb{R}$. Second, there is a dimension corresponding to valence, which is described in detail below. The valence dimension is denoted by v, with $v \in \mathbb{R}_+$. The model I present here is an application of the more general model developed in Serra (2010).

3.1 Parties

There are two parties competing in this election, labeled party L and party R. Following the Wittman-Calvert-Roemer tradition, I assume that parties are *policy-motivated*, meaning that they care about the policy implemented after the election (Wittman 1973; Calvert 1985; Roemer 2001). Parties L and R have ideal policy points X_L and X_R, respectively. The two parties have distinct ideologies so that $X_L \neq X_R$. I normalize the ideal point of the median voter in the general election

to zero, and without much loss of generality I assume $X_L < 0 < X_R$. The utility functions of L and R are

$$U_R(x) = -|X_R - x|$$
$$U_L(x) = -|X_L - x|$$

In later sections I will specify two separate groups within party R with different ideal points X_{RE} and X_{RM}. For this section, however, it is sufficient to think of X_R as the generic ideal point of R. At this stage it is useful to define a few concepts. By a party's *extremism* I will mean how far its ideal point is from the median voter's ideal point. Concretely, party R's extremism will be measured by $|X_R|$, and party L's extremism will be measured by $|X_L|$.[1]

Finally, parties formulate policy platforms to compete in the election, and they do so strategically in order to maximize their expected utility. I call those platforms x_L and x_R, with $x_L, x_R \in \mathbb{R}$.

3.2 Candidates

All candidates are characterized by a parameter v denoting how appealing their non-policy attributes are to voters in that election. Parameters such as v have been called "valence parameters" and can be given many interpretations (for an overview see Schofield (2007) and Adams et al. (2009)). In the context of this paper, v is best interpreted as the candidate's *campaigning skill*. It can take two values: a low value normalized to zero corresponding to a low-skilled candidate, and a high value of V corresponding to a high-skilled candidate. Hence $v \in \{0, V\}$. I label v_L and v_R the skills of candidates in parties L and R, respectively. To focus on the interesting cases, I will assume that valence is sufficiently salient to make a difference in the election; technically I will assume that the valence of a high-skilled candidate is strictly larger than the extremism of both parties, meaning that $|X_L|$, $|X_R| < V$.[2] Indeed, for smaller values of V, the valence dimension loses influence in the election and the results become trivial. I report these results in footnotes, and I refer the reader to Serra (2011) for a fuller analysis of a lower salience of valence.

In this model, candidates do not have policy preferences of their own. Rather, they will adopt the policy preferences of their party. To be exact, the candidate will behave as if having the exact utility function of the party that nominated her. She will announce the platform designed by her party during the campaigns, and she will implement such platform in case she wins the election.

[1] Of course, note that $|X_R| = X_R$ and $|X_L| = -X_L$.

[2] This is equivalent to assuming that $-V < X_L$ and $X_R < V$.

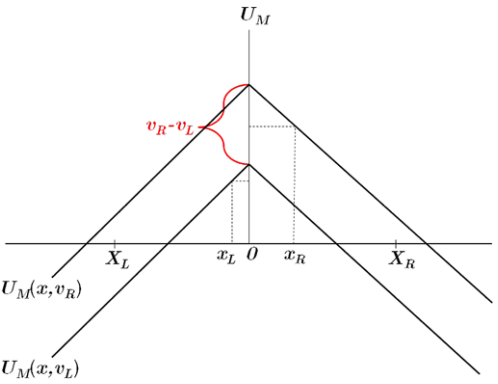

Fig. 1 The effect of a valence advantage for R over L

3.3 The General Electorate

The electorate cares about the policy implemented after the election. To simplify the analysis, I will assume that there is a median voter, which I call M, whose preferences are decisive in the election. I normalize her ideal point to zero.

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

$$U_M(x, v) = -|x| + v$$

M will vote for the party whose candidate maximizes her utility. I make the following indifference assumptions. If M is indifferent between the two parties, she will vote for the one whose candidate has the highest skill. If both candidates have the same skill, she will randomize equally between the two.

It is worth looking more closely at how the median voter makes her decision in this kind of model. As elaborated in Serra (2010), M's appreciation for a candidate decreases with the distance between her ideal point and that candidate's platform, and increases with the candidate's valence. In essence, the valence parameter v "shifts up" the utility function of M. An example of how M evaluates R and L is illustrated in Fig. 1, where it is assumed that $v_L < v_R$ and $|x_L| < |x_R|$. In the case depicted in this figure, candidate R is strictly preferred to candidate L in spite of having a more extremist platform. Candidate R is able to win the election because her higher score in the valence dimension more than compensates her extremism in the policy dimension.

3.4 Timing and Solution Concept

The timing of this election is the following:

1. **Assessment of the candidates' skills**: Parties announce their candidates who start campaigning. The candidates' campaigning skills v_L and v_R are observed.

2. **Assessment of the policy platforms**: Candidates announce their platforms x_L and x_R.
3. **The general-election vote**: The median voter elects L or R.

Stage 1 does not involve any decision: the candidates are revealed to voters, along with their valence attributes. The first decision is made in Stage 2 where each candidate must announce and promote her platform taking the other candidate's platform into account. In Stage 3, once candidates' skills, v_L, v_R, and platforms, x_L, x_R, have been observed and assessed, the median voter elects L or R to office. All this information is common knowledge. The game must be solved by backward induction and the solution concept is subgame-perfect equilibrium (SPE) in pure strategies. It will be important to recall that a SPE requires that all strategies form a Nash equilibrium (NE) in every subgame.

3.5 Results of the General Election

Before stating the main results of this section, some important variables should be defined. I call Δv the difference in skill between R's candidate and L's candidate. To be concrete, $\Delta v \equiv v_R - v_L$. Note that Δv can take three values: $\Delta v \in \{-V, 0, V\}$. I call x_L^* and x_R^* the equilibrium strategies of parties L and R, and x^* the winning platform. These parameters will determine the results of the general election, as indicated in the main theorem on this section. It must be remember that valence was assumed to be salient enough that $|X_L|$ and $|X_R|$ are smaller than V, which implies that $-V < X_L$ and $X_R < V$.

Theorem 1 *The equilibrium strategies and equilibrium outcomes of this election for given values of* v_L, v_R, V, X_L *and* X_R *are given in Table 1, where* $\Delta v \equiv v_R - v_L$.

There are several comments to make about Table 1.[3] First note the results when $\Delta v = 0$, that is, when there is no skill difference between the candidates. Both par-

Table 1 Equilibrium outcomes of the general election

Value of Δv	Equilibrium platforms x_R^* and x_L^*	Winning platform x^*	Winning party
V	$x_R^* = X_R$ $x_L^* \in \mathbb{R}$	X_R	R
0	$x_R^* = 0$ $x_L^* = 0$	0	R or L with equal probability
$-V$	$x_R^* \in \mathbb{R}$ $x_L^* = X_L$	X_L	L

[3]The proofs of all the results come in the Appendix.

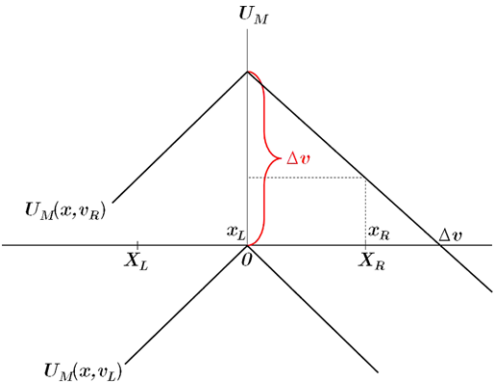

Fig. 2 Equilibrium platforms x_L^* and x_R^* when there is a valence advantage for R over L

ties converge completely to the median voter's ideal point. However, when $\Delta v \neq 0$ the candidate with highest skill is able to diverge from the median voter toward the ideal point of her party, and still win the election based on her superior skill. So the policy implemented is biased toward R when $\Delta v > 0$, biased toward L when $\Delta v < 0$, and unbiased when $\Delta v = 0$. In fact, given the assumption that valence is salient enough, the party with the highest-skilled candidate is able to pull policy all the way to its ideal point.[4] Such equilibrium is illustrated in Fig. 2, which depicts the case where $0 < X_R < \Delta v$.

4 The Nomination Process

In this section, I take a step back in the election process to study the nomination of candidates within a party. At this stage, the identity of each party's candidate is still unknown. Consequently, the exact values of the candidates' campaigning skills are uncertain. However, there exist some prior beliefs about these skills based on some information about parties and their potential candidates. According to that information, the probabilities that L's candidate and R's candidate will be high-skilled are π_L and π_R respectively, with $\pi_L, \pi_R \in (0, 1)$. In other words, $\pi_L \equiv P(v_L = V)$ and $\pi_R \equiv P(v_R = V)$. Those prior beliefs before the election campaigns are common knowledge among voters and parties.

The rest of this paper seeks to study the ability of party R to increase π_R by choosing a CSM over another. Indeed, choosing to hold a primary election could affect π_R positively under circumstances specified below. There could be a cost, however, in terms of the policy implemented by the candidate after a primary. Solving party R's cost-benefit analysis is the final goal of this research. I eschew in

[4]This ideal point depends on which group controls policy within the party. In this section we have called X_L and X_R the generic ideal points of parties L and R. In later sections, however, party R's ideal point will be given by $X_R = X_{RE}$ if the leaders control policy, or $X_R = X_{RM}$ if the members control policy. In other words, what we mean by "party" will vary according to the CSM.

this paper the parallel decision of party L who might also be pondering whether to choose a primary election. Such analysis is being done in a separate paper, and here I simply assume that party L has already chosen a candidate by any method. In other words, π_L is taken as an exogenous parameter. In any case, remember that the actual campaigning skills of L and R's candidates are revealed when they start campaigning to win the election. Thus v_L and v_R are fully known when voters decide who to vote for.

4.1 Party Members Versus Party Leaders

Party R consists of an "elite" (or "leadership") and a "membership" (or "rank and file"). The elite of R will be referred to as RE. This leadership is policy-motivated and has an ideal policy point X_{RE}, with $X_{RE} > 0$. The utility function of R's elite is

$$U_{RE}(x) = -|X_{RE} - x|$$

The rank and file (RAF) of R is also policy-motivated. To simplify the analysis, I will assume that the RAF has a median member whose preferences are decisive in the primary election. I call RM the median member of R and I call X_{RM} her ideal point, with $X_{RM} > 0$. The utility function of RM is

$$U_{RM}(x) = -|X_{RM} - x|$$

In general, we will have $X_{RE} \neq X_{RM}$, so there will be a tension between the policy preferences of a party's leadership and its RAF. It will be useful to measure the divergence, if any, between a party's establishment and its primary voters. With that purpose, I define d_R as the *internal divergence* in party R, where $d_R \equiv |X_{RM} - X_{RE}|$. An interesting interpretation of d_R is as the *congruence* (or lack thereof) between R's elite and mass membership. Higher levels of the internal divergence d_R indicate a lower elite-mass congruence inside the party. Note that d_R can take any non-negative value: $d_R \geq 0$.

Parties are also responsible for formulating policy platforms to compete in the election. More precisely, parties are in charge of indicating the policy platforms they wish their candidates to follow in each circumstance. If party R uses a leadership selection, then its leaders formulate the policy strategies to be followed by its candidate. If, instead, party R uses a primary election, then its candidate will follow the policy strategies desired by the RAF. Note that both the leadership and the RAF think strategically. This implies that they would not passively impose their ideal points on the candidate, but rather, they will design a strategy that maximizes their expected utility taking into account the behavior of the rival party in the general election.

Table 2 The objective of party R's candidate

After an elite selection:	$\max_{x_R} U_{RE}(x) = -\lvert X_{RE} - x\rvert$
After a primary election:	$\max_{x_R} U_{RM}(x) = -\lvert X_{RM} - x\rvert$

4.2 Primary Election Versus Elite Endorsement

Before selecting a candidate, the leadership of party R needs to choose a candidate-selection method (CSM). There exist two methods: an elite endorsement or a primary election. The default CSM would be for the leadership to directly nominate or endorse an insider candidate. Alternatively, it could hold a competitive 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. I call m_R the method that R's leaders choose, with $m_R \in \{elite, primary\}$. Following standard language in the party-politics literature, I will call *selectorate* the group in charge of selecting a party's candidate. If $m_R = elite$, the selectorate is the party's leadership. If $m_R = primary$, the selectorate is the party's RAF. In the former case, $X_R = X_{RE}$. In the latter case, $X_R = X_{RM}$.

Candidates adopt the policy preferences of their selectorate. In other words, they behave as perfect agents of whichever group inside their party nominated them. Therefore, depending on whether the CSM is a primary election or an elite endorsement, the nominee will inherit the preferences of either *RM* or *RE*, respectively. This is summarized in Table 2.

The interpretation is that in striving to win the nomination, the pre-candidates are forced to cater to the wishes of those selecting them. In exchange for having their names on the ticket, they have to yield on policy by making concrete commitments to those in charge if the nomination. Those commitments are credible because parties have effective ways of enforcing their candidates' promises.

4.3 Insiders Versus Outsiders

An important difference across nomination rules is the number of aspirants who have a realistic chance of getting their party's nomination. When a party elite chooses to endorse someone without further consultation, it is usually because there is a trusted insider who has previously emerged as the natural nominee. In contrast, when a party decides to allow a truly competitive primary election, it is opening the door to outside aspirants who might have previously been unknown or ignored. This empirical observation motivates the following assumptions.

Any individual who is officially contesting the party's nomination will be referred to as a *pre-candidate*. If $m_R = elite$ then party R has only one pre-candidate to choose from, which I call the *insider* and I denote by *RI*. If $m_R = primary$ then party R has two pre-candidates to choose from, which consist of the insider, *RI*, and an outsider denoted by *RO*. Hence, by adopting a primary, the party is expanding the pool of candidates that it can choose from.

I call v_{RI} and v_{RO} the campaigning skills of *RI* and *RO* respectively, and I call v_R the campaigning skill of the candidate who is finally nominated by R. As I mentioned before, a candidate's skill can take two values, 0 or V. However, the exact values of the pre-candidates' campaigning skills are uncertain ex-ante. The party has some prior information about the probability that its insider candidate, *RI*, is high-skilled or low-skilled. That information could come from previous performance in office, from past elections, or from polls. According to that information, *RI* has a probability π_{RI} of being high-skilled, with $\pi_{RI} \in (0, 1)$. On the other hand, the party has *no* prior information about the outsider candidate. The party believes that the outsider candidate *RO* has a probability of one-half of being high-skilled, hence $\pi_{RO} = \frac{1}{2}$.

4.4 Timing

The timing of the nomination is the following:

1. **The selection of the candidate-selection method**: The leaders of party R choose a nomination process.
2. **The nomination contest**: If the CSM is a primary election, the pre-candidates commit to pursuing the policy interests of *RM* and some information about their skills is revealed. If the CSM is an elite endorsement, the pre-candidates commit to pursuing the policy interests of *RE* and no information is revealed.
3. **The nomination decision**: Party R selects its candidate.

After this nomination, the game is played exactly as described in the previous section, i.e. the three stages of the nomination are followed by the three stages of the general election. All this information is common knowledge.

5 The Benefit of Primary Elections

In this section, I develop a model of primary elections as a means to acquire some information about the campaigning skills of aspirants. Primaries reveal partial information through a system of noisy signals sent by candidates and processed by primary voters using Bayes rule. This informational mechanism is the main innovation with respect to Adams and Merrill (2008), Serra (2011), Snyder and Ting (2011) and other models postulating that primaries reveal information about candidates. In those models information is *fully* revealed in the primary election, and there is no additional information in the general election. In contrast, in this model the information is only *partially* revealed in the primary, and there is additional information in the general election. As I will show, this realistic assumption leads to new insights about the adoption of primary elections, in particular the possibility that a high-skilled insider might prevent such primaries.

A later section describes a cost of primaries. This will allow studying, in the final section of the paper, the cost-benefit analysis carried out by party leaders when deciding whether to hold a primary election or stick to an elite selection.

5.1 Primaries as a Mechanism to Reveal Information

Here I formalize the informational incentive to adopt primary elections. For party leaders, the benefit is to increase the expected campaigning skill of their nominee. I will call that increase the "primary skill bonus". Primaries achieve this in two ways. (1) The pool of potential nominees is expanded. Concretely, primaries open the door to untested or non-mainstream contenders who can register as pre-candidates hoping to display their skills during the primary campaign. Those outsiders might have a large appeal to voters but would not come to the party's attention through an inside-track elite nomination. And (2) useful information about those pre-candidates is revealed. Specifically, primaries can reveal valuable information about the pre-candidates' assets and resources. Indeed, during the primary campaigns the pre-candidates are tested on how they raise funds, manage a team of supporters, debate other candidates, design political advertisements and give interviews to journalists. So primaries serve as a testing ground for the subsequent general election. In that sense this paper provides an information rationale for democratizing a political party.

Given these differences, each method will have different probabilities of nominating a high-skilled candidate. The value that party leaders are seeking to maximize is $\pi_R \equiv P(v_R = V)$. To do so, they calculate which candidate-selection method m_R maximizes $P(v_R = V|m_R)$, with $m_R \in \{primary, elite\}$.

To calculate $P(v_R = V|elite)$ note that if party leaders choose to select the candidate themselves they would directly nominate *RI*. The probability of nominating a high-skilled candidate would simply be π_{RI}. Hence $P(v_R = V|elite) = \pi_{RI}$.

If, however, they choose to hold a competitive primary election, the candidate *RO* would join the race and the nomination will be delegated to the party's RAF who will decide between *RI* and *RO*. Hence the probability of nominating a high-skilled candidate, $P(v_R = V|primary)$, would depend on the actual skills of these candidates, which are ex-ante uncertain except for the prior beliefs.

The premise in this paper is that primaries will reveal some information about the actual skills of their pre-candidates. This information subsequently helps the party choose the most skilled one. To be more precise, if there is a primary election, a candidate's performance in the primary can itself reflect high skill or low skill. Party members interpret the performance of a candidate in the primary-election campaign as a *forecast* of how well she would perform in the general-election campaign against the other party. Those forecasts are imperfect, however, because the information is "noisy." Hence I assume that the true skills of candidates v_{RI} and v_{RO} are revealed only *partially* if there is a primary election.

To be concrete, I denote by s_j the performance of candidate j in the primary, with $j = RI, RO$. I say that $s_j = high$ if j's performance showed high skill, and $s_j = low$

if j's performance showed low skill. I 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_{RI} and s_{RO} have probability q of "being correct". We can interpret s_j as a noisy signal of candidate j's skill, and we can interpret q as the quality of this signal. More broadly, q is a measure of the effectiveness of primary elections as an information-revelation method.

In sum, the pre-candidates' performances, s_{RI} and s_{RO}, are independently-distributed random variables whose distribution depend on v_{RI} and v_{RO} in the following way:

$$P\left(s_j = high | v_j = 1\right) = P\left(s_j = low | v_j = 0\right) = q$$
$$P\left(s_j = high | v_j = 0\right) = P\left(s_j = low | v_j = 1\right) = 1 - q$$
$$j = RI, RO$$

Once the party members observe the candidates' performances, they can update their prior beliefs about *RI*'s and *RO*'s skills using Bayes rule. This approach to voting based on updated beliefs following a noisy signal has its roots in Condorcet (1785), Austen-Smith and Banks (1996), and Feddersen and Pesendorfer (1998).

The candidates' performances are public, and therefore the values of s_{RI} and s_{RO} are common knowledge. In particular, all the RAF members observe the same s_{RI} and s_{RO}, and hence they update their beliefs based on the same information. Given its interest in winning the general election, the RAF will vote for the candidate who is believed to have the highest skill. When a party member is indifferent between *RI* and *RO*, I assume she will vote for the one whose prior probability of being high-skilled was largest. If both have the same prior, she will randomize equally.

5.2 Primary Voters Update Their Beliefs

These elements allow studying the behavior of primary voters. When $s_{RI} \neq s_{RO}$, I say that a member of party R's rank and file will "vote according to the signals" if her strategy is to vote for the pre-candidate whose signal was highest, meaning, whose performance was best in the primary campaign. On the other hand, if her strategy does not depend on the signals sent during the primary, meaning that performance in the primary is irrelevant, I say that a member of party R will "ignore the signals".

These concepts can be used to describe the RAF's behavior during a primary. As it turns out, their behavior will depend crucially on their prior belief about the insider candidate's valence, π_{RI}. In all the results below, the symbols $\underline{\pi}$ and $\overline{\pi}$ refer to two constants whose values are $\underline{\pi} \equiv \frac{(1-q)^2}{1-2q+2q^2}$ and $\overline{\pi} \equiv \frac{q^2}{1-2q+2q^2}$.

Lemma 1 *In a primary election, for each value of π_{RI}, the rank-and-file members of party R will*

- *if* $\pi_{RI} \in (0, \underline{\pi}]$, *ignore the signals and always vote for RO*
- *if* $\pi_{RI} \in (\underline{\pi}, \frac{1}{2})$, *vote according to the signals if* $s_{RI} \neq s_{RO}$, *and vote for RO if* $s_{RI} = s_{RO}$
- *if* $\pi_{RI} = \frac{1}{2}$, *vote according to the signals if* $s_{RI} \neq s_{RO}$, *and randomize between RI and RO if* $s_{RI} = s_{RO}$
- *if* $\pi_{RI} \in (\frac{1}{2}, \overline{\pi})$, *vote according to the signals if* $s_{RI} \neq s_{RO}$, *and vote for RI if* $s_{RI} = s_{RO}$
- *if* $\pi_{RI} \in [\overline{\pi}, 1)$, *ignore the signals and always vote for RI*.

There are several noteworthy features of this result, the first one being how influential the prior beliefs are: given that each member of R is assumed to be rational and to use all information available to make her decision, she will combine the prior beliefs about the candidates with the new information coming from their performance. However, the prior beliefs might be so compelling that even a Bayesian party member will choose to disregard the candidates' performances. In particular, for high enough values of π_{RI} the RAF will *always* vote for *RI* even if it receives strong indications of the insider's low skill compared with the outsider's high skill. Primary voters will simply not trust that such performances will carry through to the general election. Hence the insider candidate *RI* is immune against an open contest with the outsider *RO*; he will be nominated regardless of their performances. This result is significant as it opens the possibility that any information revealed during the primary election will be useless: primary voters might vote according to preexisting information while completely ignoring the new information.

On the other hand, the results for intermediate values of π_{RI} go in the expected direction: primary voters will take the signals into account, and will vote for the candidate whose performance in the primary campaigns was best. Hence the insider candidate I will indeed be vulnerable to being beaten by the outsider O in an open contest.

Our next task is to quantify the benefit of holding a primary instead of a leadership selection. As I derive below, the bonus of using a primary election is to increase the expected skill of the party's nominee. Hence the value I am looking to find is the difference between $E(v_R|primary)$ and $E(v_R|elite)$.[5] It is easy to see that such difference is given by

$$E(v_R|primary) - E(v_R|elite) = V \cdot S$$
$$\text{with } S \equiv P(v_R = V|primary) - P(v_R = V|elite)$$

The important value is S, which represents the extra probability of having a high-skilled candidate that a primary brings above an elite selection. I call it the *skill bonus* of a primary. Studying S, how large it is and how it changes, is the main task now. Rather than giving the exact value of S, which comes in the Appendix,

[5] We should keep in mind that, even though the actual value of v_R is discreet, the expected value $E(v_R)$ is continuous.

I will focus on the key properties that will buttress the rest of the paper. I start by rephrasing the previous considerations in terms of π_R, which is the variable that party R is seeking to maximize.

Theorem 2 *The probability that R's nominee will be high-skilled, π_R, given R's nomination process, m_R, is given by*

$$\pi_R \equiv P(v_R = V|m_R) = \begin{cases} \pi_{RI} & \text{if } m_R = elite \\ \pi_{RI} + S & \text{if } m_R = primary \end{cases}$$

where S is called the primary skill bonus and is given by $S \equiv P(v_R = V|primary) - P(v_R = V|elite)$.

This demonstrates how the information revealed in primary campaigns is translated into a better nominee in terms of valence. Holding an internal contest will increase the probability of nominating a high-skilled candidate in the amount S. Is that a small or a large benefit? I answer that question in the next subsection.

5.3 What Makes Primaries More Appealing?

I begin by establishing whether primaries have a benefit to party leaders.[6]

Lemma 2 *The primary skill bonus S is strictly positive for $\pi_{RI} \in (0, \overline{\pi})$ and zero for $\pi_{RI} \in [\overline{\pi}, 1)$.*

Primaries therefore do bring a benefit for small enough priors about the insider's skill. When the insider candidate is weak, meaning that π_{RI} is below a certain threshold, forcing her to compete with an outsider candidate increases the excepted skill of the nominee by a strictly positive amount. The reason is that for $\pi_{RI} \in (0, \overline{\pi})$ party members will take a serious look at the outsider candidate's performance in the primary to decide whether she is more convincing than the party insider. This result was expected as it conforms with previous findings in Serra (2011).

The surprising result comes from high priors about the insider's skill: in such case a primary election might not bring any benefit whatsoever. When the insider candidate is strong, meaning that π_{RI} is above a certain threshold, forcing her to compete with an outsider candidate does not increase the expected skill of the nominee at all. The reason is that for $\pi_{RI} \in [\overline{\pi}, 1)$ party members find the insider candidate so compelling that they will vote for her regardless of the outsider candidate's performance in the primary. This result is new with respect to the papers about primaries that I am aware of.

[6] As mentioned before, the symbols $\underline{\pi}$ and $\overline{\pi}$ refer to two constants whose values are $\underline{\pi} \equiv \frac{(1-q)^2}{1-2q+2q^2}$ and $\overline{\pi} \equiv \frac{q^2}{1-2q+2q^2}$.

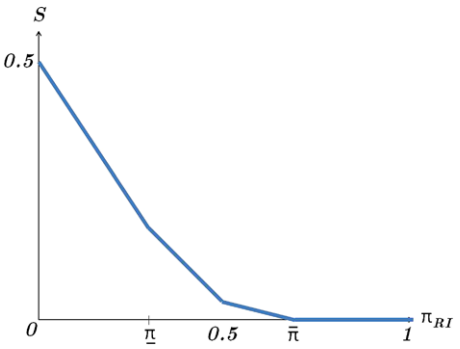

Fig. 3 The primary skill bonus S as a function of the insider's probability of being high-skilled π_{RI}

It is now turn to study how S changes with a change in its two main determinants: the prior about the insider candidate's skill, π_{RI}, and the accuracy of the candidates' performances q. Do they make primaries more or less attractive? I first describe the comparative statics with respect to π_{RI}.

Lemma 3 *The primary skill bonus S is strictly decreasing with π_{RI} for $\pi_{RI} \in (0, \overline{\pi})$, and constant (equal to zero) to any increase in π_{RI} for $\pi_{RI} \in [\overline{\pi}, 1)$.*

Several insights about S can come from the lemma above, most notably that it decreases with π_{RI}. This makes intuitive sense, because the benefit of primaries is to improve upon the skill of the candidate that would be nominated through an elite selection, namely 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. In fact, as mentioned before, this electoral advantage reaches zero once the insider candidate's appeal to voters exceeds a certain threshold labeled $\overline{\pi}$.

The message is that the electoral advantage brought by primaries is larger the less appealing the insider candidate is to begin with. This is clearly seen in Fig. 3, which depicts the value of S as a function π_{RI}.

I can turn now to studying how S changes with q. Remember that we can interpret q as the *quality* of primary elections as an information-revelation method. To be exact, an increase in q improves the accuracy of the performances s_{RI} and s_{RO} as forecasts of future performances in the general election. This improvement could occur 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. In essence, a larger q implies that the primary performance is a better *forecast* of the candidate's campaigning ability in the general election. Intuition would suggest that any improvement in the primaries' technology would make those primaries more attractive. Surprisingly, as the following result shows, this intuition is only correct under certain circumstances.

Lemma 4 *The effect on the primary skill bonus S of a marginal increase in q is strictly positive for $\pi_{RI} \in [\underline{\pi}, \overline{\pi}]$, but is null for $(0, \underline{\pi})$ and $(\overline{\pi}, 1)$.*

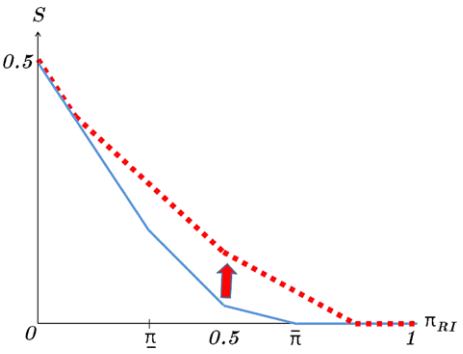

Fig. 4 The effect of increasing the quality of signals q on the primary skill bonus S

The result goes in the expected direction for moderate priors about the insider candidate's skill. For intermediate values of the prior π_{RI}, marginal increases in q will indeed increase S. The reason is that primary voters are unsure about the relative merits of the insider candidate compared to the unknown outsider that will join the race. They will pay close attention to the primary campaigns to nominate the candidate with a better performance. A higher quality of the information revealed will increase the probability of making the right nomination choice. Such an increasing effect is depicted in Fig. 4.

However, for other priors, the quality of a primary elections will bear no impact on its benefit. When the insider candidate is expected to be overwhelmingly competent in the general election, she will be nominated even if her performance in the primary is appalling. Primary voters will trust that her performance in the primary was due to bad luck. On the other hand, when the insider candidate is expected to be overwhelmingly unqualified, she will lose to the outsider candidate even if her performance was better. Primary voters will believe her performance was just a fluke that does not justify giving her a chance in the general election. In sum, for extremely high or extremely low values of π_{RI}, primary voters quickly make up their minds, either to nominate *RI* for sure or to nominate *RO* for sure, regardless of any campaign events that may occur. Improving the quality of primaries by marginally increasing q will have no effect on this decision.

In sum, primaries have two potential benefits: (1) allowing primary voters to replace the insider candidate with an outsider candidate whose prospect are believed to be superior; and (2) using new information revealed during the primary campaigns to discriminate between both candidates. As it turns out, whether those benefits actually occur depends crucially on the prior beliefs about the campaigning skill of the insider candidate. This finding is qualitatively summarized in Table 3.

To summarize this section, the benefit, when there is one, of primary elections is a larger probability of nominating a candidate with a high campaigning skill. I 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 a consequence, the party leadership needs to carry out a cost-benefit analysis when choosing whether to hold a primary election or not.

Table 3 The two potential benefits of a primary election as a function of π_{RI}

Expectation that *RI* is high-skilled, π_{RI}	Benefit of primaries		
	Replacing *RI* with *RO*	Using the information revealed during the primary	Skill bonus of a primary *S*
Low	Yes, for sure	No, information ignored	High
Intermediate	Yes, probably	Yes, taken into account	Low
High	No, never	No, information ignored	Zero

6 The Cost of Primary Elections

As we just saw, the benefit to party leaders of adopting a competitive primary election is to increase the expected skill of their nominee. However, primaries might carry a cost in terms of the policy that candidates are induced to adopt. To be precise, a primary election has two differences with respect to an elite endorsement: first, the probability that *R*'s nominee is high-skilled increases from π_{RI} to $\pi_{RI} + S$. And second, it would be *RM* and not *RE* that *R*'s candidate would have made policy commitments to; and thus it would be the RAF's preferences rather than the leadership's preferences which would determine *R*'s policy platform.

By glancing at Table 4, we can readily see the trade-off that *R*'s leadership faces in choosing a primary election over an elite endorsement. As a benefit, using a primary increases the probability of nominating a high-skilled candidate (due to the primary skill bonus *S*). As a cost, the payoff from having the highest skilled candidate decreases (due to the internal divergence $X_{RM} - X_{RE}$). Put differently, *a primary makes losing less likely but makes winning less attractive*.

The goal now is to find expressions for the expected utility of *R*'s leadership by choosing either a primary election or an elite selection. I call $EU_{RE}(m_R)$ the expected utility of *R*'s leadership from adopting m_R as its CSM. It can be derived from Theorem 1, which gives the outcomes of the election depending on the value $\Delta v \equiv v_R - v_L$. If *L*'s candidate has a skill advantage, she will announce the platform X_L and she will win the election. If *R*'s candidate has a skill advantage, she will announce the platform X_{RE} if she was nominated by an elite appointment or she will announce X_{RM} if she was nominated by a primary election; and either way she will win the election. If *L*'s candidate and *R*'s candidate have the same skill, they will both announce the platform 0 and they will tie in the election. These considerations lead to the following expressions for $EU_{RE}(m_R)$.

Table 4 The trade-off faced by party *R*'s elite

	Probability that *R* wins the election	Utility of *RE* if *R* wins the election
Elite selection	π_{RI}	0
Primary election	$\pi_{RI} + S$	$-\lvert X_{RE} - X_{RM} \rvert$

Lemma 5 *The expected utility of R's leadership for each value of* m_R *is*

$$\begin{aligned} EU_{RE}(m_R = elite) = & -(X_{RE} - X_L)\pi_L(1 - \pi_{RI}) \\ & -(X_{RE} - 0)\big[\pi_L\pi_{RI} + (1 - \pi_{RI})\big] \\ & -(X_{RE} - X_{RE})(1 - \pi_L)\pi_{RI} \\ EU_{RE}(m_R = primary) = & -(X_{RE} - X_L)\pi_L\big(1 - (\pi_{RI} + S)\big) \\ & -(X_{RE} - 0)\big[\pi_L(\pi_{RI} + S) + (1 - \pi_L)\big(1 - (\pi_{RI} + S)\big)\big] \\ & -|X_{RE} - X_{RM}|(1 - \pi_L)(\pi_{RI} + S) \end{aligned}$$

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

7 The Optimal Selection of a CSM

The leadership in party R will choose the optimal rule m_R by comparing $EU_{RE}(m_R = elite)$ and $EU_{RE}(m_R = primary)$. It will choose the CSM that yields the highest expected utility, and if it is indifferent, I assume that it will choose an elite selection. A primary will be adopted if and only if $EU_{RE}(m_R = elite) < EU_{RE}(m_R = primary)$. That condition leads to the following result, recalling that $d_R \equiv |X_{RM} - X_{RE}|$.

Theorem 3 *The leadership of party R will adopt a primary election if and only if*

$$d_R < T$$

with $T \equiv \frac{S[X_{RE}(1-\pi_L) - X_L\pi_L]}{(1-\pi_L)(\pi_{RI}+S)}$.

The intuition behind this result is that R's leadership will delegate the nomination if and only if the RAF's ideology is close enough to its own. In other words, internal party democratization will only ensue from enough elite-mass congruence. How close do primary voters need to be to the party elite? It depends on a certain threshold, T, introduced in the theorem. If the preferences of the elite and the mass of party R are so incongruent that $T \leq d_R$ then the leadership will not adopt a primary election. This could happen for two reasons. On one hand, the RAF could be so far on the right of the leadership that $X_{RE} + T \leq X_{RM}$. In that case the leadership will not adopt a primary election because the primary voters are too *extremist*. On the other hand, the RAF could be so far on the left of the leadership that $X_{RM} \leq X_{RE} - T$. In that case the leadership will not adopt a primary election because the primary voters are too *centrist*.

As it turns out, the first reason (that primary voters might be too extreme) is frequently found in some way or another in scholarly comments about primary elections. Yet the second reason (that primary voters might be too moderate) is equally

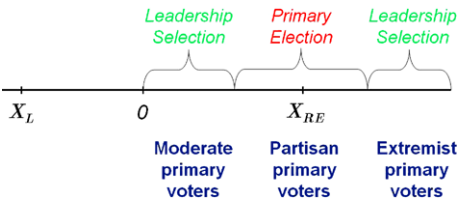

Fig. 5 The candidate-selection method as a function of the ideal point of the median primary voter, X_{RM}

intuitive but is seldom mentioned in the existing literature. The same intuition can be obtained from Fig. 5. For low values of X_{RM} (which I label "moderate primary voters") the party will endorse an insider candidate. For intermediate values of X_{RM} (which I label "partisan primary voters") the party will hold a competitive primary election. For high values of X_{RM} (which I label "extremist primary voters") the party will endorse an insider candidate. Consequently, the CSM has a non-monotonic relationship with the ideal point of the median primary voter.

From the results above it is clear that the threshold T determines how likely primary elections are. The interval $(X_{RE} - T, X_{RE} + T)$ corresponds to the values that X_{RM} should take for the nomination to be delegated to party members. Such interval can therefore be interpreted as the *likelihood that R will adopt a primary*. For a larger T it is more "likely" that the internal divergence between R's establishment and RAF will be lead to a primary. Then a way of phrasing the previous theorem is that *the likelihood of opening the CSM decreases with the internal divergence between the party's leadership and the primary voters*.

7.1 Comparative Statics

We would like to gain insight on what makes the adoption of primary elections more likely. According to the previous theorem, the likelihood of adopting a primary is given by T. Hence, I study how T changes with the parameters in the model. As it turns out, the results will crucially depend on the value of π_{RI}. To be specific, I need to divide two cases. The first case is $\pi_{RI} \in (0, \overline{\pi})$ corresponding to low and intermediate priors, and the second case is $\pi_{RI} \in [\overline{\pi}, 1)$ corresponding to high priors. Recall that $\underline{\pi}$ and $\overline{\pi}$ refer to two constants whose values are $\underline{\pi} \equiv \frac{(1-q)^2}{1-2q+2q^2}$ and $\overline{\pi} \equiv \frac{q^2}{1-2q+2q^2}$.

I start with low and intermediate prior beliefs about the skill of the insider candidate, which corresponds to the situation where primaries are most attractive.

Theorem 4 *Suppose the initial expectation that RI is high-skilled, π_{RI}, is such that $\pi_{RI} \in (0, \overline{\pi})$. Then the threshold T, which determines the likelihood of primaries, is*:

1. *Strictly positive*
2. *Strictly increasing with S*

3. *Strictly decreasing with* π_{RI}
4. *Strictly increasing with* q *if* $\pi_{RI} \in [\underline{\pi}, \overline{\pi})$, *and insensitive to* q *otherwise*
5. *Strictly increasing with* π_L
6. *Strictly decreasing with* X_L
7. *Strictly increasing with* X_{RE}.

The first two results of this theorem corroborate the benefit of primaries. First, I find that $T > 0$. Hence there will always exist a certain distance with the RAF that party leaders can tolerate for delegating it the nomination decision. Second, this threshold increases with the primary skill bonus. The larger the primary skill bonus S, the more likely it is that the elite will forgo appointing the insider in a smoke-filled room.

The third and fourth results decompose the effect of S in its two components, π_{RI} and q. The effect of the expected competence of the insider candidate is intuitive: the more competent the insider candidate is, the less likely that a primary will identify a better candidate, and hence the less attractive primaries are. This effect can be observed in Fig. 6 which depicts how the likelihood of adopting a primary decreases with the prior belief about the insider. The comes from Lemma 3 which established the negative effect of π_{RI} on S, and hence on T.

The effect of q is also intuitive though more complex. As I mentioned, an increase in q can be interpreted as an improvement in the information-revelation feature of primaries. For intermediate values of π_{RI}, an increase in q will increase S as we know from Lemma 3, which in turn will increase T. In other words, a primary election is more attractive for party leaders when its ability to reveal information is larger. This effect can be observed in Fig. 7 which depicts how the likelihood of adopting a primary increases when the quality of primaries increase.

This result contradicts a certain view of primaries in the literature. It is sometimes advised that primary elections should be short and smooth to avoid candidates draining their energy and resources (see for example Ezra (2001)). The theorem above provides a different perspective. A party can actually benefit from having long and challenging primaries, as this would increase the amount of information revealed about pre-candidates (namely q). This result is new in the literature about

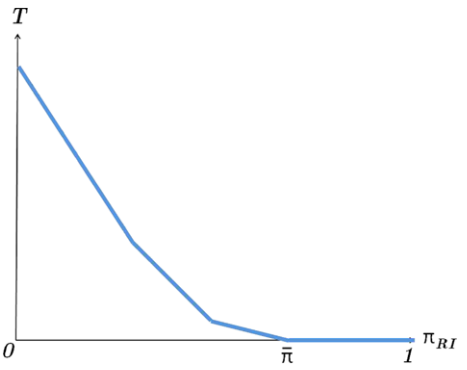

Fig. 6 The likelihood of adopting a primary as a function of the insider's probability of being high-skilled π_{RI} (all things equal)

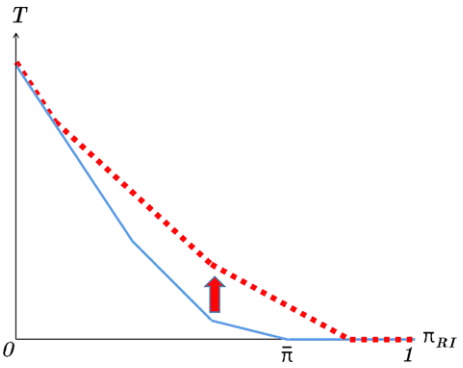

Fig. 7 The effect of increasing the quality of signals q (all things equal) on the likelihood to adopt a primary

primaries, as it could only be obtained by making the realistic assumption that primaries can only reveal information *partially* rather than *fully*.

The last part of the result is more surprising. For low values of π_{RI}, an increase in q will not have any effect on T. The reason is that candidates' performances in the primary would actually being ignored. Primary voters have already made up their minds in favor of an outsider candidates irrespective of her eventual performance in the primary. So increasing or decreasing the amount of information will not alter the nomination decision and consequently will not make primaries more or less attractive.

The fourth, fifth and sixth results broadly indicate that disadvantaged parties are more likely to adopt primaries than advantaged parties. They were all previously found in Serra (2011) so I do not elaborate on them here. Rather I focus on the importance of π_{RI} which is a new contribution.

In particular, the following result departs from previous research as it provides conditions for an insider candidate to avoid a primary challenge. As it turns out, an insider might have a good enough reputation that party leaders will *inevitably* nominate her by not opening the competition to outsiders under *any* circumstance.

Theorem 5 *Suppose the initial expectation that RI is high-skilled, π_{RI}, is such that $\pi_{RI} \in [\overline{\pi}, 1)$. Then the threshold T, which determines the likelihood of primaries, is zero and primaries will never be adopted under any value of the other parameters.*

In other words, the insider's reputation could be so good that leaders will inexorably appoint her. This type or reputation could be enjoyed, for example, by an incumbent who has already won a previous election. Strikingly, a primary election will be eschewed even if primaries reveal a maximum amount of information; even is there is perfect congruence between the elite and the membership of the party; and even if party R has important weaknesses with respect to L. There exists a threshold above which π_{RI} will prevent the use of primary elections for all values of q, X_{RM}, X_{RE}, X_L and π_L.

Hence this result provides an explanation for the empirical observation that many incumbents get re-nominated in their parties without a primary challenge. The rea-

son is that for sufficiently high expectations about the insider candidate's skill, primaries do not bring any advantage at all: both the RAF and the elite are sure to nominate the same candidate. This comes from Lemma 2. Given that primaries do not bring a benefit, any amount of elite-mass incongruence is enough to deter party democratization. S is equal to zero and hence T is equal to zero, which means that any value of d_R is intolerable for party leaders.

8 Conclusions and Discussion

When can an incumbent or any well-known insider feel safe against a challenge for the nomination of a future election? When can he or she be confident that party leaders will directly appoint her rather than holding a competitive primary election? Primary elections are a frequent method used by political parties around the world to select their candidates—and increasingly so. 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.

An implication of those two features is that a primary election will increase the expected valence of the party's nominee. Such benefit has been modeled previously, for example in Adams and Merrill (2008), Serra (2011), Snyder and Ting (2011), and indeed the findings in this paper corroborates some of the findings in that previous literature (for example that primaries are most beneficial to the weakest parties as found by Adams and Merrill (2008), Serra (2011)).

However those models assume that primaries reveal information fully, meaning that candidates' performance in the primary are a perfect forecast of their performance in the general election. In contrast, this paper assumes that primaries only reveal information partially, meaning that candidate's performance in the primary are a noisy and imperfect forecast of their performance in the general election.

Making this realistic assumption led to new insights. The prior reputation of the party insider (the parameter π_{RI}) turns out to play a crucial role in deterring the use of primaries. Primaries are less appealing to party leaders the better the insider candidate is believed to be. In fact, if the party insider has a good enough reputation for winning votes, for example by virtue of being an incumbent who won a previous election, then a primary election will be eschewed altogether. The paper thus provides an explanation for the empirical fact that many incumbents get re-nominated by their parties without a primary challenge.

This new setup also allowed studying the behavior of primary voters more precisely. As expected, primary voters may use the information provided by primary campaigns to select the pre-candidate with a most impressive performance. However, as it turns out they will only do so for moderate expectation about the ability of the insider candidate. If, on the other hand, the insider is believed to be extremely

competent or extremely incompetent, primary voters will actually ignore the contenders' performance in the primary campaigns and vote exclusively according to their preexisting priors. In other words, primary voters will completely disregard the information provided to them.

I finish with a prescriptive note. If we believe that democratization should occur in any representative institution, we should care about when and why political parties become internally democratic. A question for reformers, then, is how to make competitive primary elections more prevalent. This paper provides several suggestions, but the most direct one is to improve the revelation of information during the primary cycle (the parameter q). Political parties and the general public can benefit from improving the design of primaries to test the pre-candidates' campaigning skills thoroughly enough. For example, parties could include more debates, make campaigns longer, and allow tough critiques among contenders. In other words, the more challenging primaries are, the more information they will reveal about the pre-candidates. A recent example is the competition between Hillary Clinton and Barack Obama during the Democratic primary election. Several Democratic supporters complained that the competition between Clinton and Obama was too long and too severe. Those Democrats worried about the possible costs to their party's prospects in the general election. I do not deny that such costs existed: the potential drawbacks of a competitive primary election include division and resentment among the party base, among other possible costs. But this paper points to a benefit that was seldom mentioned during the 2008 primary. Observers claimed that too much information was being revealed about Clinton and Obama—information which could later be misused by the Republicans. My premise, however, is that such information would have been revealed anyway in the course of the general-election campaign. As a consequence, it was beneficial for the Democratic sympathizers to acquire that information beforehand to help them select their nominee wisely. According to this paper, the length and intensity of the primary campaign are not necessarily a curse for the party, but could actually be a blessing.

Appendix with the Proofs

A.1 Proof of Theorem 1

Table 1 here is a particular case of Table 1 in Theorem 1 of Serra (2011).

A.2 Proof of Lemma 1

If there is a primary election, Party R's RAF will vote for the candidate that it believes to have highest probability of being high-skilled. The beliefs it holds about each candidate's skill depend on two pieces of information: its prior beliefs, and

the information acquired throughout the primary campaign. Given that the RAF members are rational, they will update their prior beliefs based on the performances s_{RI} and s_{RO} to form a couple of posterior beliefs about the probabilities that *RI* and *RO* are high-skilled. If the RAF uses Bayes Rule to update its prior beliefs after receiving a given estimate, its posterior beliefs will be given by

$$P(v_{RI} = 1|s_{RI} = low) = \frac{(1-q)\pi_{RI}}{(1-q)\pi_{RI} + q(1-\pi_{RI})}$$

$$P(v_{RI} = 1|s_{RI} = high) = \frac{q\pi_{RI}}{q\pi_{RI} + (1-q)(1-\pi_{RI})}$$

$$P(v_{RO} = 1|s_{RO} = low) = 1 - q$$

$$P(v_{RO} = 1|s_{RO} = high) = q$$

There are four couple of performances (s_{RI}, s_{RO}) that the RAF could observe, which are $(0, 0)$, $(1, 1)$, $(0, 1)$ and $(1, 0)$, I study each of them in turn, along with the decision that the RAF makes upon receiving those couples of estimates.

- If the RAF observes $s_{RI} = low$ and $s_{RO} = low$:

The RAF will vote for *RI* if $P(v_{RO} = 1|s_{RO} = low) < P(v_{RI} = 1|s_{RI} = low)$ which is equivalent (after some algebra) to $\frac{1}{2} < \pi_{RI}$. Then, given my indifference assumption, the RAF will vote for *RO* if $\pi_{RI} < \frac{1}{2}$, will vote for *RI* if $\frac{1}{2} < \pi_{RI}$, and will randomize equally if $\pi_{RI} = \frac{1}{2}$.

- If the RAF observes $s_{RI} = high$ and $s_{RO} = high$:

The RAF will vote for *RI* if $P(v_{RO} = 1|s_{RO} = high) < P(v_{RI} = 1|s_{RI} = high)$ which is equivalent (after some algebra) to $\frac{1}{2} < \pi_{RI}$. Then, given my indifference assumption, the RAF will vote for *RO* if $\pi_{RI} < \frac{1}{2}$, will vote for *RI* if $\frac{1}{2} < \pi_{RI}$, and will randomize equally if $\pi_{RI} = \frac{1}{2}$.

- If the RAF observes $s_{RI} = low$ and $s_{RO} = high$:

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

- If the RAF observes $s_{RI} = high$ and $s_{RO} = low$:

The RAF will vote for *RO* (in other words, disregard the candidates' performance) if $P(v_{RO} = 1|s_{RO} = low) < P(v_{RI} = 1|s_{RI} = high)$ which is equivalent (after some algebra, and noting that $1 - 2q + 2q^2 > 0$) to $\pi_{RI} < \frac{(1-q)^2}{1-2q+2q^2}$. Then, given

Table A.1 The primary vote as a function of the signals

	$s_{RI} = low$ $s_{RO} = low$	$s_{RI} = high$ $s_{RO} = high$	$s_{RI} = low$ $s_{RO} = high$	$s_{RI} = high$ $s_{RO} = low$
if $\pi_{RI} \in (0, \underline{\pi}]$	Vote for *RO*	Vote for *RO*	Vote for *RO*	Vote for *RO*
if $\pi_{RI} \in (\underline{\pi}, \frac{1}{2})$	Vote for *RO*	Vote for *RO*	Vote for *RO*	Vote for *RI*
if $\pi_{RI} = \frac{1}{2}$	Randomize	Randomize	Vote for *RO*	Vote for *RI*
if $\pi_{RI} \in (\frac{1}{2}, \overline{\pi})$	Vote for *RI*	Vote for *RI*	Vote for *RO*	Vote for *RI*
if $\pi_{RI} \in [\overline{\pi}, 1)$	Vote for *RI*	Vote for *RI*	Vote for *RI*	Vote for *RI*

my indifference assumption (and noting that $\frac{(1-q)^2}{1-2q+2q^2} < \frac{1}{2}$), the RAF will vote for *RO* if and only $\pi_{RI} \leq \underline{\pi}$, with $\underline{\pi} \equiv \frac{(1-q)^2}{1-2q+2q^2}$.

Table A.1 summarizes these results. Which is what the lemma claims.

A.3 Proof of Theorem 2

This conclusion comes directly from two observations: (1) With an elite selection, the party will directly appoint *RI*, and thus $P(v_R = V | m_R = elite) = \pi_{RI}$. And (2) with a primary election the probability of nominating a high-skilled candidate will increase by S by definition, such that $P(v_R = V | m_R = primary) = \pi_{RI} + S$.

A.4 Proof of Lemma 2

I start by calculating the exact value of S. All its properties are derived from this value. We can use the RAF's behavior described in the previous lemma. For that, I first need to calculate $P(v_R = V | primary)$. We can do so by noting that

$$P(v_R = V | primary) = \sum_{v_{RI}, v_{RO}} \sum_{s_{RI}, s_{RO}} P(v_R = V | primary, s_{RI}, s_{RO}; v_{RI}, v_{RO}) \cdot P(s_{RI}, s_{RO} | v_{RI}, v_{RO}) \cdot P(v_{RI}, v_{RO})$$

which uses the definition of conditional probability twice.

Each summand in that expression is straightforward to calculate. $P(v_{RI}, v_{RO})$ depends only on the prior probabilities that v_{RI} and v_{RO} are high-skilled, which are π_{RI} for the insider and $\frac{1}{2}$ for the outsider. $P(s_{RI}, s_{RO} | v_{RI}, v_{RO})$ depends only on the accuracy of the signals, which is q. And $P(v_R = V | primary; s_{RI}, s_{RO}; v_{RI}, v_{RO})$ depends on how the RAF will vote given the candidates' performances, which I just computed in the table above. Multiplying and adding those probabilities is easy but

too long to develop here (the detailed calculations are reported in previous versions of this paper). With the appropriate algebra we find that

$$P(v_R = V|primary) = \begin{cases} \frac{1}{2} & \text{if } \pi_{RI} \in (0, \underline{\pi}] \\ \pi_{RI}q^2 + q - \frac{1}{2}q^2 - \pi_{RI}q + \frac{1}{2}\pi_{RI} & \text{if } \pi_{RI} \in (\underline{\pi}, \frac{1}{2}) \\ \frac{1}{2}q + \frac{1}{4} & \text{if } \pi_{RI} = \frac{1}{2} \\ \pi_{RI}q - \pi_{RI}q^2 + \frac{1}{2}q^2 + \frac{1}{2}\pi_{RI} & \text{if } \pi_{RI} \in (\frac{1}{2}, \overline{\pi}) \\ \pi_{RI} & \text{if } \pi_{RI} \in [\overline{\pi}, 1) \end{cases}$$

I can now calculate the value of interest, S. The values above are used to calculate $S \equiv P(v_R = V|primary) - P(v_R = V|leadership)$, remembering that $P(v_R = V|leadership) = \pi_{RI}$. With some algebra and noting the continuity of S at $\pi_{RI} = \underline{\pi}$, $\pi_{RI} = \frac{1}{2}$ and $\pi_{RI} = \overline{\pi}$, we find that

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

which are the values we were looking for.

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

A.5 Proof of Lemma 3

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

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

A.6 Proof of Lemma 4

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

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

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

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

A.7 Proof of Lemma 5

See the proof of Lemma 1 in Serra (2011).

A.8 Proof of Theorem 3

See the proof of Theorem 2 in Serra (2011).

A.9 Proof of Theorem 4

For points 1, 2, 5, 6 and 7, see the proof of points 1, 2, 6, 7 and 8 of Theorem 4 in Serra (2011), respectively.

To study the effect of q (point 3 in the theorem), we note that it only has an indirect effect on T through its effect on S. I proved in Lemma 5 that q has a strictly positive effect on S whenever for $\pi_{RI} \in [\underline{\pi}, \overline{\pi}]$. And I have proved (in point 2 of the theorem) that S has a strictly positive effect on T. Therefore, combining both

partial derivatives, I prove that q has a strictly positive effect on T whenever for $\pi_{RI} \in [\underline{\pi}, \overline{\pi}]$.

To study the effect of π_{RI} we must note that it has two effects on T: a direct effect, and an indirect effect through its effect on S. In total, we have that $\frac{dT}{d\pi_{RI}} = \frac{\partial T}{\partial \pi_{RI}} + \frac{\partial T}{\partial S}\frac{\partial S}{\partial \pi_{RI}}$. It is easy to calculate that $\frac{\partial T}{\partial \pi_{RI}} = \frac{-S[X_{RE}(1-\pi_L)-X_L\pi_L]}{(1-\pi_L)(\pi_{RI}+S)^2}$ which is strictly negative. On the other hand I just calculated that $\frac{\partial T}{\partial S}$ is strictly positive, and we know from Lemma 4 that $\frac{\partial S}{\partial \pi_{RI}}$ is non-positive. We therefore have that $\frac{dT}{d\pi_{RI}} < 0$ and T is strictly decreasing with π_{RI}.

A.10 Proof of Theorem 5

Note from Lemma 2 that $S = 0$ when $\pi_{RI} \in [\overline{\pi}, 1)$. And remember that $T \equiv \frac{S[X_{RE}(1-\pi_L)-X_L\pi_L]}{(1-\pi_L)(\pi_{RI}+S)}$. Hence, when $\pi_{RI} \in [\overline{\pi}, 1)$ we have that $T = 0$ for any value of the other parameters.

References

Adams J, Merrill S III (2008) Candidate and party strategies in two-stage elections beginning with a primary. Am J Polit Sci 52(2):344–359

Adams J, Merrill S III, Stone WJ, Simas EN (2009) When candidates value good government: a spatial model with applications to congressional elections. Manuscript

Ashworth S, de Mesquita EB (2009) Elections with platform and valence competition. Games Econ Behav 67(1):191–216

Austen-Smith D, Banks JS (1996) Information aggregation, rationality, and the Condorcet jury theorem. Am Polit Sci Rev 90(1):34–45. 1996

Caillaud B, Tirole J (2002) Parties as political intermediaries. Q J Econ 117(4):1453–1489

Callander S (2008) Political motivations. Rev Econ Stud 75:671–697

Calvert RL (1985) Robustness of the multidimensional voting model: candidate motivations, uncertainty, and convergence. Am J Polit Sci 29:69–95

Carrillo J, Castanheira M (2008) Information and strategic political polarisation. Econ J 118(530):845–874

Castanheira M, Crutzen BSY, Sahuguet N (2010) Party organization and electoral competition. J Law Econ Organ 26(2):212–242

Cho S-J, Kang I (2008) Open primaries and crossover voting. Paper presented at the annual national conference of the Midwest Political Science Association

Condorcet M de [1785](1994) Essai sur l'application de l'analyse a la probabilite des decisions rendues a la pluralite des voix. McLean I and Hewitt F (trans) Paris

Ezra M (2001) The benefits and burdens of congressional primary elections. In: Galderisi PF, Ezra M, Lyons M (eds) Congressional primaries and the politics of representation. Rowman & Littlefield, New York

Feddersen T, Pesendorfer W (1998) Convicting the innocent: the inferiority of unanimous jury verdicts under strategic voting. Am Polit Sci Rev 92(1):23–35. 1998

Jackson MO, Mathevet L, Mattes K (2007) Nomination processes and policy outcomes. Q J Polit Sci 2(1):67–94

Kselman D (2012) Median-activists or median-voters? A spatial model with intra-party competition. Manuscript, Juan March Institute

Meirowitz A (2005) Informational party primaries and strategic ambiguity. J Theor Polit 17(1):107–136
Meirowitz A (2008) Electoral contests, incumbency advantages, and campaign finance. J Polit 70(03):681–699
Owen G, Grofman B (2006) Two-stage electoral competition in two-party contests: persistent divergence of party positions. Soc Choice Welf 26(3):547. Heidelberg
Roemer JE (2001) Political competition: theory and applications. Harvard University Press, Cambridge
Schofield N (2007) The mean voter theorem: necessary and sufficient conditions for convergent equilibrium. Rev Econ Stud 74:965–980
Schofield N, Claassen C, Ozdemir U, Schnidman E, Zakharov AV (2008) Party positioning under proportional representation and plurality rule. Center in Political Economy, Washington University. Typescript
Schofield N, Sened I (2005) Multiparty competition in Israel, 1988–1996. Br J Polit Sci 35:635–663
Serra G (2010) Polarization of what? A model of elections with endogenous valence. J Polit 72(2):426–437
Serra G (2011) Why primaries? The party's trade-off between policy and valence. J Theor Polit 23(1):21–51
Snyder JM Jr, Ting MM (2011) Electoral selection with parties and primaries. Am J Polit Sci 55(4):781–795
Wittman D (1973) Parties as utility maximizers. Am Polit Sci Rev 67:490–498