UNIVERSIDAD CARLOS III DE MADRID working papers

UC3M Working papers
Economic
13-15
July, 2013

Departamento de Economía

Universidad Carlos III de Madrid
Calle Madrid, 126
28903 Getafe (Spain)
Fax (34) 916249875

Partisan politics: parties, primaries and elections.

Agustin Casas*

This version: July

You can find the most updated version here:
http://sites.google.com/site/agustincasas/research
DRAFT

Abstract

Parties' candidates are chosen by different nomination rules. Recent empirical evidence shows that these rules influence the attributes of the nominees; for instance, open primaries in the U.S. choose more extreme candidates than closed primaries. Despite this evidence, the literature does not provide an explanation of why appealing to a more moderate electorate -as during open primaries- results in more extreme candidates. I build a model that shows that open primaries elect \predictable extremists", while, for instance, party leaders would chose \moderate mavericks". I obtain these results through a model that puts together 3 pieces of partisan politics: affiliation decisions, nomination rules, and an observed endogenous valence, which (together with party membership) signals the candidates' ideologies. Moreover, I investigate the welfare implications of three methods: nomination by the party leader, by closed primaries, and by open primaries. I show the conditions under which nomination by party leaders leads to higher social welfare than nomination by open primaries. Furthermore, I show that higher screening by parties, leads to more ideologically uncertain candidates. In sum, I argue that party affiliation decisions, and endogenous valence play a large role in understanding the effects of nomination rules on the political equilibria.

* Mailto: acasas@u.northwestern.edu . I am indebted to my advisor David Austen-Smith, for his patience and guidance. I also thank to Tim Feddersen, Bard Harstad, and Antoine Loeper. I also benefited from conversations with Steve Callander, Alessandro Pavan, Andre Trindade, Carlo Prato, my fellow grad students, and seminar participants at Northwestern University. All remaining errors are mine.

1 Introduction

As electoral institutions that shape the outcome of elections, partisan nomination rules systematically influence candidates' attributes. Although there is a large body of empirical evidence that examines the relationship between candidates and nomination rules, there are few theoretical papers that address this relationship. In this paper I explain how different nomination rules result in different types of candidates by endogenizing the "median voter" in primary elections. Moreover, political parties are central to the explanation in that they signal candidates' ideology and provide the pool of potential candidates. Thus, I show that open primaries' candidates (where all citizens vote) are expected to be more extremist and less popular (lower valence) than closed primaries' (where only party members vote). Likewise, when the party leader handpicks the candidates, he selects a more moderate and more popular candidate.

Political scientists have long argued for empirical regularities in the U.S. primary systems that select legislative and presidential candidates (Alesina and Rosenthal (1995), Gerber and Morton (1998), Alvarez et al. (1995), Ansolabehere et al. (2007)). Recent evidence shows that (1) open primaries select more extreme candidates than closed ones (King 1998, Kanthak and Morton (2001)), and highlights that (2) primaries induce competition between co-partisans that may be harmful in the general election (Agranov (2009)). Whether there is a nomination method that is superior to any other selection mechanism, remains an open question. However, it is striking that in the U.S. less than a fourth of the states select their candidates through open primaries; and around the globe, only a dozen countries have ever had a primary to select presidential candidates.

In general, any approach to studying the systematic influence of intraparty politics on the equilibrium requires dealing with at least three different, but intertwined, aspects: voters' preferences, the nomination rule itself, and party affiliations. That is why, despite the evidence that partisan nomination methods play a role, few models study their effects on the political equilibria thoroughly (Jackson, Mathevet and Mattes (2007)). Furthermore, there is no formal model that explains both why open primaries' candidates are more extreme on average, and why holding primaries can harm the electoral chances in the general election. Incorporating those three aspects, I investigate the effect of intraparty institutional arrangements on the selection of candidates in a bipartisan democracy. In particular, I study the influence of affiliation and nomination strategies on the candidates' ideology and valence.

Valence is a broad term that has been coined by Stokes (1963) to refer to those issues that cannot be found on the liberal-conservative continuum (Downs (1957)). More generally, valence is modeled as an observable characteristic that affects all voters equally, independent of their ideology. For example, campaigning skills fit the modeling assumptions in this paper, since they influence the citizens' vote without necessarily increasing the voters' utility, *ceteris paribus*. The novelty of my approach is that instead of suggesting valence is exogenous (Groseclose (2001)) or is a partisan investment decision (Ashworth and Bueno de Mesquita (2009)), I argue that all citizens are endowed with a level of valence. Therefore any likable personality traits that may result in electoral returns, such as oratorical abilities, good looks, charisma, and other observable skills fit the model.

In particular I study three games which only differ in their nomination rules. Thus, the institutional comparative statics is focused on the effect of three stylized nomination rules: open primaries (where all citizens vote), closed primaries (where only affiliated citizens vote), and "handpicking" (by the party leader). Each rule defines who is the decisive citizen in choosing the nominee; I call that citizen the "nominator". In open primaries, the nominator is the general median voter; in closed ones, it is the median party member. Regardless of the rule, the nominator is constrained to choose a candidate from the pool of party members, which is an equilibrium outcome of the citizens' affiliation decisions. The decisive citizen is the median voter when the median voter theorem holds: while in the case of closed primaries the theorem holds under very general conditions, the same is not true in open primaries. Nevertheless, I prove that the outcome of all open primaries elections are "median voter like" in the sense that the nominated candidate always coincides with the median voter's perferred choice.

Each individual, described by his private ideology and public valence, optimally decides whether to join a party. The affiliation cost increases as the ideological distance to the party's platform increases; thus, party membership is partly informative of an individual's ideology. Furthermore, it leads to informative party labels, *à* la Snyder and Ting (2002): a candidate from the liberal party is more likely to be more liberal than conservative. Conversely, I assume that the consumption benefit of being a party member increases as the citizen's valence increases, grounded on the intuition that higher valence individuals can extract more rents by joining a party[1].

As a result, there is some degree of substitutability between ideology and valence: citizens that are ideologically far from the party's platform can affiliate if their valence is high enough. On the other hand, those who are close to the platform can affiliate regardless of their valence. Therefore, despite the fact that ideology and valence are independently distributed, voters update their beliefs on each the candidate's ideology by looking at their valence and memberships: high valence candidates are less predictable (higher variance). That is, valence operates as a mean preserving transformation around each party's platform. Thus, I build a spatial model with valence and probabilistic voting where policy motivated politicians enjoy ego rents for winning, but cannot commit to any policy.

I organize the results of the model according to the literature in which they fit. Regarding valence, I show that it is a strategic substitute for platforms' moderation. That is, nomination rules solve the trade off between valence and certainty, such that larger platform differentiation leads to more popular candidates in equilibrium. Intuitively, when parties are ideologically close to each other, their median primary voters are indifferent when choosing a candidate based on policy. Thus, since voters are risk averse, they select a low valence, low variance candidate.

[1]As I explain in the main body of the paper, assuming otherwise results in high valence "alienated" citizens that, although they are very close to one party's platform, they do not affiliate to it.

Regarding the nomination rules, I show that party leaders choose more popular (higher valence) candidates than voters in primaries, and voters in closed primaries nominate higher valence candidates than in open ones. More generally, as party primaries cater to a more moderate electorate, the nominated candidate is lower valence. As before, when the median primary voter is more centrist, he cares less about his candidate winning because he is close to the other party's median voter; hence, he chooses a less popular candidate, who is also less "expensive" in terms of uncertainty. Similarly, when parties are sufficiently moderate and leaders anticipate more democratic nomination arrangements, party leaders choose a more extremist platform; hence, they countervail the nominators' incentives to lower valence, and they reduce their own compromise in policy. In short, open primaries' candidates are (low valence) "predictable extremists"; on the contrary, party leaders' candidates are (high valence) "moderate mavericks". Consequently, I show that primaries are not a self-enforceable institution: when parties can choose a nomination method, and politicians are exclusively office-motivated, they prefer that the party leader handpick the candidate. This new result provides a novel explanation of why so few countries nominate their candidates using primaries. Furthermore, I show that more democratic methods might be less efficient under specific circumstances wherein the median voter might prefer that the leaders pick the nominees.

Regarding the size of parties, although I do not search for an optimal level of screening, I show that, contrary to popular wisdom, more cohesive parties increases the equilibrium uncertainty of the location of candidates. Because of the increasing costs, narrower parties choose higher valence candidates at a lower marginal cost, which causes an overall increase of variance.

Although nomination methods are not central in the explanation of the economic and political performance of countries, they could contribute to understanding some facts. For instance, recent moves toward open primaries in the U.S. (see Serra 2010) could help in our understanding of the phenomenon of party polarization; also, nomination by party leaders in Latin American countries may add to their policy instability. Therefore, partisan nomination methods should not be only seen as exclusively partisan characteristics, but also examined as systemic variables in any given country or state.

1.1 Related literature

Valence

Since Stokes' introduction of valence issues, as opposed to strong unidimensional ideologies, the term has been used to name all those issues that affect voters but cannot be found on the liberal-conservative continuum (Downs 1957). More specifically, political economists have posited that valence refers to an attribute or set of attributes, either of the party or of the candidate, that all voters like equally (in general, without satiation point). For instance, it has been modeled as an unobservable trait that implies commitment to the "correct" policy (Callander and Wilkie (2007), Kartik and McAfee (2007)). Snyder and Ting (2002) and Agranov (2009) explore the ambiguity of politicians[2], which might also be described in relationship to valence. Finally Groseclose (2001), Ashworth and Bueno de Mesquita (2009), and others, model valence as a known characteristic of parties or candidates.

Early work on valence issues has treated it exogenously, finding that candidates with a "valence advantage" choose more centrist platforms (Groseclose (2001)). Similarly, when it is endogenized as a partisan investment decision, the results resemble those of vertical differentiation: less ideological differentiation, higher valence competition (Ashworth and Bueno de Mesquita (2009)).

Parties

Political scientists have developed different models of parties that define their scope and limitations (see Aldrich (1995) for an account on parties and their history in the U.S.). One way or another, parties regulate the competition of their members in a way that limits the behavior of candidates, and affects the political equilibria. However, only recently have intraparty constraints been formally incorporated into the political economy literature. For instance, Hirano et al. (2009) shows how candidates' incentives to cater to certain factions during the primaries affect the post-electoral allocation of resources. Snyder and Ting (2002) emphasize the informative role of party labels as a result of citizens' affiliation decisions. Crutzen et al. (2010), and Serra (2010) study under what conditions it is optimal for a party to nominate candidates through primaries. Ashworth and de Mesquita (2008) investigate the implications of partisan screening on the outcome of elections.

Primaries

Despite this growing trend, and although it has been long acknowledged that party membership and primaries are strongly connected (Berdahl (1949)), few papers study them together. Meirowitz (2005) tackles this issue, and models primaries as devices that allow candidates to learn voters' preferences after observing their membership. The work of Jackson, Mathevet and Mattes (2007) comes closest to the argument of this paper, and provides a model of endogenous parties that specifically studies how the candidates' ideologies depend on the nomination procedure. In particular, it finds that when candidates are nominated by vote (the equivalent to closed primaries in my model), candidates are more moderate than those chosen by the party leader. On the other hand, when they make an allowance for endogenous parties, they show that median outcomes hold in the voting setup.

My work adds to this literature by providing a rationale for the evidence that indicates open primaries choose more extreme candidates. Although open primaries cater to a more moderate electorate, policy-motivated party leaders strategically choose more extreme platforms, closer to their own preferred policy. Therefore, I not only show that open primaries' candidates are more

[2] The literature on ambiguity is broader: Aragones and Neeman (2000) and references thereafter.

extreme on average, but also that their ideology is more predictable (lower variance). Conversely, I call party leaders' candidates moderate mavericks not only because they have more moderate ideologies on average, but also because they have higher ideological variance.

2 The model

There is a continuum of citizens. Each citizen has a private ideology z_i distributed uniformly in the policy space $[-k, k]$, and an observable valence q_i, drawn from a uniform distribution with support $[0, \bar{q}]$, independent of i's ideology. There are also two exogenous political entrepreneurs with private ideologies z_x and z_y in the policy space.

I study three stylized games, which only differ in the way parties nominate their candidates; that is, they differ in their nomination rule r. Before explaining each rule, I introduce the general timing:

- Political entrepreneurs announce platforms simultaneously.
- Citizens affiliate to party X, Y, or none.
- Nominees are chosen simultaneously, from the set of affiliated citizens, and according to the nomination rule r.
- All citizens vote

The political entrepreneurs establish parties X and Y, and announce the political platforms or manifestos taking into consideration the implications of the game form. These platforms are a public declaration of "guidelines and principles" pinpointed in the policy space, that depend on r, and are called x^r and y^r, respectively. Given those platforms citizens decide whether to affiliate to a party, and which one. Then, depending on the nomination rule, each party chooses a candidate from the pool of affiliated citizens. Each nominee is described by his (unobserved) ideology, his valence, and his affiliation: $(\tilde{z}_X^r, q_X^r)$, and $(\tilde{z}_Y^r, q_Y^r)$. Finally, the general election takes place, and the winning candidate implements his own preferred policy.

The timing and the nominations rules are also explained by figure (1) below. The nomination rule r can take three values depending on whether the nominee is handpicked by the party leader ($r = l$), or by a majoritarian primary election, which can be a closed primary ($r = c$) or an open one ($r = o$). In each party, for each type of nomination arrangement, there is one citizen who is decisive in the nomination of the candidate. I call that citizen the "nominator", and he is endogenous to the nomination rule: while in the autocratic case the party leader handpicks the candidate, in the democratic ones, the median voter in each primary chooses the candidate.

I abstract from turnout decisions and crossover affiliations: all party members vote in closed primaries, and all citizens vote in open ones (and in the general election). Also, there are no

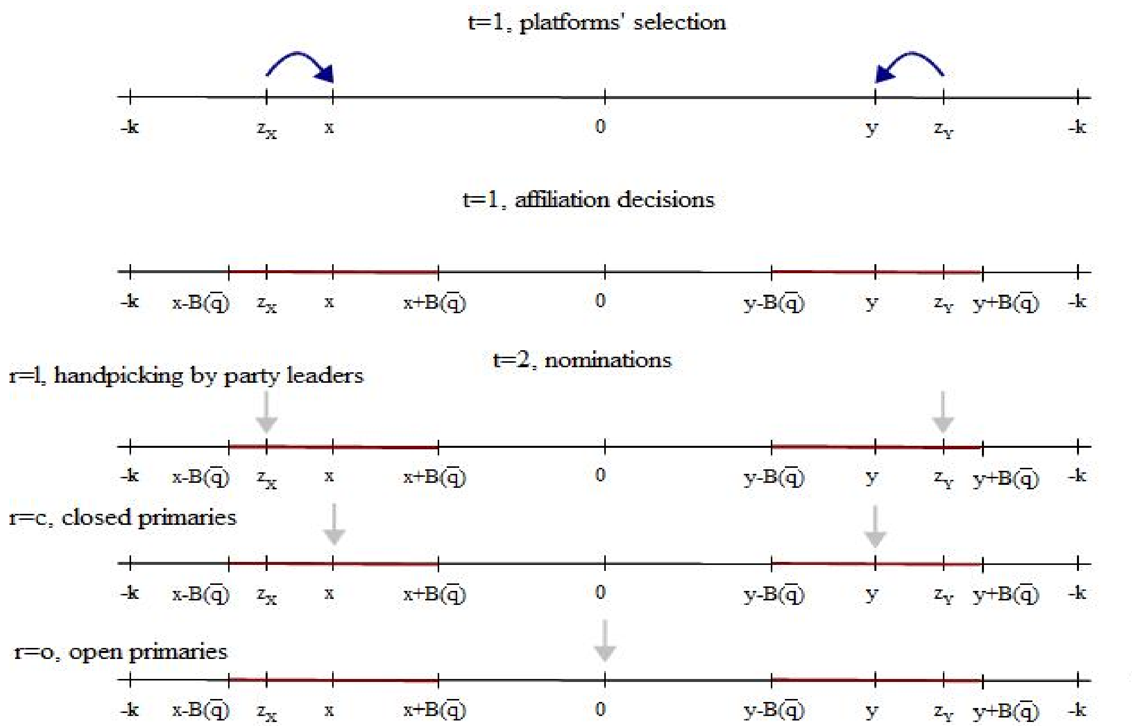

Figure 1: Party formation and nomination stages

crossovers; no citizen that prefers to affiliate to party X will affiliate to party Y in order to change the outcome of the other party's primary election (sincere affiliation and voting).[3] Therefore, the nominator in a closed primary is the median party member. In an open primary, the nominator's ideology is 0, as it coincides with the general median voter. At last, when $r = l$ the nominator's ideology is z_x (the party leader's).

Voters derive utility from their private activity and from the implemented policy. Regarding their private activity, citizens can either become politicians or work in a firm. I define $B(q_i)$ as the payoff of choosing a political career for a citizen with valence q_i, and for simplicity, I fix his opportunity cost (the market return to valence, or wage) to zero.[4] Moreover, I assume that $B(q_i)$ is strictly positive, increasing and convex in q_i, which implies that valence is more rewarded in the political arena, and at an increasing rate.

The cost of affiliation is the ideological distance from an individual's preferred policy to the party's platform; it is the disutility of misrepresenting his ideology. Intuitively, for politicians it

[3] That is, I do not look at the strategic calculus of activism (Aldrich (1983))

[4] More generally, $B(q_i)$ could be defined as the spread between the market and the party's return. Let the gross return to valence be $Priv(q_i)$ in the private sector, and $Pol(q_i)$ in a political party. Then I define $B(q_i) \equiv Priv(q_i) - Pol(q_i)$ as the net return of valence in politics.

is costly to make public statements they do not necessarily agree with, or to vote accordingly to the party's mandate instead of according to their own preference.[5] For easier notation, I drop the superscript r in until necessary. Therefore, the voters net return of joining party X with platform x is

$$B(q_i) - (z_i - x)^2. \tag{1}$$

Similarly for party Y, and 0 if unaffiliated.

Citizens are risk averse, and have quadratic preferences on the implemented policy, which they cannot observe in advance . However, voters update their beliefs on the distribution of candidates' ideology, by looking at their affiliation and valence. Let the $p \in \{X, Y\}$ stand for party X or Y; thus, the expected utility from policy when party p wins is

$$\rho_i(\tilde{z}_p; q_p) = -(z_i - E(\tilde{z}_p|q_p))^2 - V(\tilde{z}_p|q_p) \tag{2}$$

Let $a_p \in \{a_X, a_Y, a_\emptyset\}$ stand for voter affiliation to party p, or none respectively. Putting equations (1) and (2) together, I obtain voter i's expected utility when candidate $(\tilde{z}_X; q_X)$ wins:

$$u_i(a_p, \tilde{z}_X; q_X) = \begin{cases} B(q_i) - (z_i - x)^2 + \rho_i(\tilde{z}_X; q_X) & \text{if i is party X's member} \\ B(q_i) - (z_i - y)^2 + \rho_i(\tilde{z}_X; q_X) & \text{if i is party Y's member} \\ 0 + \rho_i(\tilde{z_x}; q_X) & \text{if i remains unaffiliated} \end{cases} \tag{3}$$

The political entrepreneurs also enjoy ego rents δ if their party wins; therefore, party X's leader utility is

$$u_x(\tilde{z_x}; q_X) = \delta + \rho_x(\tilde{z}_X; q_X),$$

if candidate with q_X wins the election.

Probabilistic voting. All voters have an exogenous bias $\tilde{\alpha}$ for party Y's candidate that parties cannot fully anticipate, but whose distribution depends on the observed valence of the candidates. Thus, i votes for party X if

$$u_i(a_p, \tilde{z}_X; q_X) - u_i(a_p, \tilde{z}_Y; q_Y) > \tilde{\alpha}(q_X, q_Y). \tag{4}$$

I assume that the bias $\tilde{\alpha}$ is distributed uniformly with support $[-\alpha + h(q_X, q_Y), \alpha + h(q_X, q_Y)]$, where $h(q_X, q_Y) = \omega(q_Y - q_X)$. This approach resembles the additive bias model of probabilistic voting (Banks and Duggan (2005)), with the exception that the distribution of the bias depends on a choice variable. This modeling decision intends to emphasize the electoral role of valence: citizens do not necessarily like valence per se, but as a campaigning skill that biases their decision

[5] Consistently with Snyder and Ting (2002) it could also be interpreted as the cost of interacting with members whose average ideology is the party's platform

at the time of voting. For instance, if party X candidate is higher valence than Y's, he is a better campaigner; then the shock $\tilde{\alpha}$ has a negative mean, which means that, in expectation, i votes for party Y, only if q_Y gives him a higher expected policy payoff.

3 Results

In this section I study the consequences of different partisan institutional designs on the nominees; but before getting into the details of each nomination rule, I go over the three common stages of the game. For the rest of the paper, I study the symmetric game with $z_x = -z_y < 0$.

Electoral stage.

As mentioned in the previous section, i votes for party X if equation (4) holds. Thus, it follows that the probability that he votes for party X is

$$\Pr(\text{i votes for X}) = \Pr(\rho_i(\tilde{z}_X; q_X) - \rho_i(\tilde{z}_Y; q_Y) > \tilde{\alpha}(q_X, q_Y)) \equiv P^i.$$

Therefore, given the platforms x and y, and the candidates' valences q_X and q_Y, the probability that party X wins the election is

$$P_x \equiv \frac{1}{2k}\int P^i dz = \frac{1}{2} + \frac{E(\tilde{z}_y|q_y)^2 - E(\tilde{z}_x|q_x)^2 + V(\tilde{z}_Y|q_Y) - V(\tilde{z}_X|q_X) + \omega(q_X - q_Y)}{2\alpha}$$

Nomination stage.

The nomination rule defines who is the decisive citizen in the nomination stage. In particular, for the case of primary elections, throughout the paper the nominator is the median voter. That is, I solve "as if" the median voter theorem holds; nonetheless, in the appendix I show why this approach is appropriate: in closed primaries, the theorem holds under very broad assumptions, and in open primaries, the preferred candidate of the median voter is always nominated.

Therefore, given the platforms x^r and y^r, let the nominator's utility be $u_n(\cdot)$; then he chooses the nominee such that he maximizes his own expected utility. Thus, he solves:

$$\max_{q_X} E(u_n(a_p, \tilde{z}_p; q_p^r)) = P_x u_n(a_p, \tilde{z}_X; q_x^r) + (1 - P_x) u_n(a_p, \tilde{z}_Y; q_y^r).$$

In the Appendix, I provide a discrete version of a game that begins at this stage. Solving that case provides a more straightforward intuition of the results with only two valence levels q_H, and q_L, which could also be interpreted as a two-candidate primary.

Party formation stage.

This stage of the game involves two separate substages: first, the party leaders announce their party platforms, and after observing them, citizens take their affiliation decisions. Either if citizens affiliate to a party, or they stay working for a firm, the decision is irreversible. A direct implication is that party members are potential candidates that can be chosen as nominees of their parties: affiliated citizens are politicians indeed. [6] For simplicity I assume $x < y$.

Affiliations. For any observed pair of platforms, a citizen described by the pair (z_i, q_i) affiliates to party X if

$$E(u_i(a_x, \tilde{z}_p; q_p^r)) \geq E(u_i(a_y, \tilde{z}_p; q_p^r)),$$

and

$$E(u_i(a_x, \tilde{z}_p; q_p^r)) \geq E(u_i(a_\emptyset, \tilde{z}_p; q_p^r)).$$

Thus, party X is fully described by its leader z_x, its platform x^r, and its members

$$M(X) \equiv \left\{ i : -(z_i - x^r)^2 \geq -(z_i - y^r)^2 \text{ and } B(q_i) \geq (z_i - x^r)^2 \right\}.$$

The conditional distribution of ideologies given valence and party affiliation can be easily calculated as a result of the uniform distribution of ideologies. In particular, disregarding for now the bounds of the support of ideologies, the first two central moments are

$$E(\tilde{z}_i | q_i, i \in M(X)) = x^r \tag{5}$$

$$V(\tilde{z}_i | q_i, i \in M(X)) = \frac{B(q_i)}{3} \tag{6}$$

Thus, the expected ideology of any party X's member is the party's platform. Moreover, the median party member in $X = [z_x, x^r, M(X)]$ is a citizen with ideology x^r, the party's platform. Furthermore, a candidates' ideological variance is increasing in his own valence. Therefore, when voters observe a higher valence candidate, they also believe him to be farther from the party's platform (to either side).

Remark 1 *When the cost of affiliation is increasing in the "ideological distance" to the party's platform,* ***party labels*** *are informative.*

Remark 2 *Higher valence candidate are ideologically riskier (less predictable).*

Hence, the main trade off arises. Although parties would like to nominate extremely good campaigners (high valence), their selection comes at the cost of a higher ideological uncertainty. The trade off is solved during the nomination stage by the nominators as explained in detail below.

[6] I also disregard other fair concerns regarding politicians such as those treated in Mattozzi and Merlo (2008), Diermeier et al. (2005), etc.

Furthermore, notice that for any given valence, there is a continuum of citizens; thus the probability of party member i being chosen as the candidate is negligible. Therefore, at the time of deciding their affiliation, citizens do not take into account the probability of being the nominee.

Platforms. At the beginning of the party formation stage, given the nomination method r, the political entrepreneurs announce their own party political platforms or manifestos, anticipating its possible effects on the citizens' affiliation decisions, the probability of winning, and the location of the nominator. Thus, party X's leader maximizes his expected utility:

$$\max_{x^r} P_x(\delta + \rho_x(\tilde{z}_X; q_X)) + (1 - P_x)\rho_x(\tilde{z}_Y; q_Y) \tag{7}$$

From now on, I focus on the symmetric equilibrium, and the case when $x < 0 < y$. Then, I only show party X's results unless stated otherwise, since party Y's problem is conceptually the same. Moreover, as indicated in the remark below, I assume what is necessary to avoid parties hitting the bounds of the support, and parties' overlap. Otherwise, parties could artificially make their brands more informative (reduce variance) by being very extremist or centrist.

Remark 3 *I restrict to (1) moderate values of z_x and z_y (neither close to 0 nor to $|k|$); (2) $B(0)$ and $\bar{q}$ small enough; and (3) k large enough. Then, neither overlapping, nor extreme extremism occurs.*

In what follows I study each game in detail, and in the following section I compare the equilibria, and I study the welfare implications.

3.1 Party leaders handpick the nominees: $r = l$

In this section, I look at the least democratic nomination method. The political entrepreneurs -or party leaders- not only establish the parties, and choose the political platforms or manifests, but they also handpick the nominees of their own parties. This autocratic case is worth studying as it is widespread around the world; for presidential races, only around a dozen countries haver ever had a primary to select their candidates (Mainwaring (1993), Carey and Polga-Hecimovich (2008)).

Nomination stage. Given the platforms x^l and y^l, the party leader with ideology z_x, who is policy and office motivated, chooses the nominee such that he maximizes his utility; he solves:

$$\max_{q_X} P_x(\delta + \rho_x(\tilde{z}_X; q_X)) + (1 - P_x)\rho_x(\tilde{z}_Y; q_Y).$$

Let $\Pi_l(x^l) \equiv \delta + \rho_x(\tilde{z}_X; q_X)) - \rho_x(\tilde{z}_Y; q_Y)$ be the party leaders' winning payoff in terms of policy and ego rents; and let $V(\tilde{z}_X; q_X^l) \equiv V_x(q^l)$. Then, from the F.O.C., the interior equilibrium of this subgame is found by solving

$$\frac{w - V_x'(q^l)}{2\alpha}\Pi_l(x^l) = V_x'(q^l)P_x > 0 \tag{8}$$

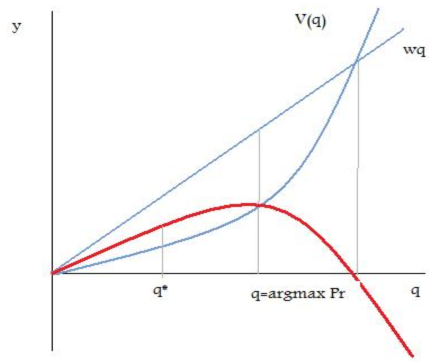

Figure 2: Equilibrium vs. probability of winning maximizing valence

Intuitively, a higher valence candidates is more likely to win, but when $V_x'(q^l) > 0$, it comes at the expense of a higher variance. The nominator dislikes variance as well, therefore he balances the gain in probability with the loss in his policy payoff: that is, the expected net marginal gain of increasing the valence (LHS) is equal to the expected marginal loss (RHS).

Lemma 1 *Let $B(q_i)$ positive, increasing, and convex in q_i. Then there is a unique symmetric equilibrium valence in the nomination subgame, q^l, which is implicitly defined by:*

$$V'(\tilde{z}_X; q^l) \equiv V_x'(q^l) = \omega \frac{\delta + 4z_x x^l}{\alpha + \delta + 4z_x x^l} \tag{9}$$

From equation (6), the variance is positive, increasing, and convex if $B(q_i)$ is. This assures the existence of an interior equilibrium, where the S.O.C. hold. Intuitively, $B(0) > 0$ leads to non-empty parties, $B'(q_i) > 0$ ensures the trade off between valence and closeness to the political platform, and $B''(q_i) > 0$ ensures the existence of an interior equilibrium. These assumptions do not only ensure the unique symmetric interior equilibrium, but also an interesting one. If the returns to valence (and so the variance) were decreasing, we would have corner solution where all parties go for the highest valence candidate. A similarly uninteresting equilibrium arises if the variance is concave, since the net return to variance ($\omega - V'(q)$) would be increasing for all q. Therefore, the convexity of $V(q)$ permits that the probability of voting is maximized for a level of valence q^* such that $V'(q^*) = \omega$, which cannot be achieved by this (or any other) nomination method under this setup (see figure (2)).

Party formation stage. I study a symmetric perfect bayesian equilibrium when $x^l < 0 < y^l$, where the party leader chooses the party's platform such that it maximizes his expected utility (equation (7)). From the F.O.C. I obtain

$$\frac{\partial P_x}{\partial x^l}\Pi_l(x^l) - \frac{\partial V_y}{\partial x^l} = \frac{\partial \Pi_l}{\partial x^l} P_x.$$

The marginal cost for the leader of choosing a more moderate platform is the marginal loss of choosing a policy further to his (RHS). On the other hand, the marginal benefit is the sum of marginal increase in the probability, plus the marginal effect on the equilibrium variance ($\frac{\partial V}{\partial x^l} < 0$). Intuitively, the party leader equates the probability gain of choosing a more centrist platform with the policy loss at the margin, discounting the second order effect of a smaller valence. The following lemma shows the symmetric equilibrium platforms.

Lemma 2 *For k large enough, the political entrepreneurs choose to form parties X and Y with platforms x^l and y^r, such that:*

$$-\frac{x^l}{\alpha}\Pi_l(x^l) - \frac{\partial V_y(q^l)}{\partial x^l} = x^l - z_x \tag{10}$$

Similarly for party Y.

The proof of the lemma can be found in the appendix. The large k ensures that the political entrepreneur does not find optimal to run by himself instead of forming a party. Furthermore, since $x^l < 0$ and $\Pi_l(x^l) = \delta + 4z_x x^l > 0$, the party leader always choose to form a party with a moderate platform, relative to his ideology: $x^l - z_x > 0$.

Intuitively, since I assume that every citizen's ideology is private information and nobody can credibly inform it, the expected ideology of any non-affiliated citizen is $E(z_i) = 0$, and the expected variance is the variance of the distribution, $V(z_i) = \frac{k^2}{3}$. However, when party leaders form a party and announce a platform, independently of any credibility issue, as voters self-screen around that platform, the announcement becomes the median ideology of the party. Therefore, the party leader's gain from establishing parties is caused by (1) signaling candidates' ideologies, and by (2) the possibility of choosing a higher valence nominee. On the other hand, there is a policy compromise; once it is established that party leaders will form parties, it follows that they will do it closer to the median voter, so they increase the likelihood of winning. [7]

Remark 4 *For k large, any unaffiliated candidates' ideology is expected to be $E(z_i) = 0$, and his variance $V(z_i) = \frac{k^2}{3}$. Thus, the party leader finds optimal to form a party.* [8]

3.2 Closed primaries: $r = c$

In this section, I study the game with a primary election where only party members can vote. Exploring this case is specially interesting since this is one of the few institutions in democratic life where qualified voting has endured without having a well studied rationale for it. A salient

[7] Notice that I do not formally let the party leaders, who are not uncertain about their own ideology, to run as candidates from outside the party. However, I look under which conditions they would not find that profitable at all; that is, when parties are "self enforceable"

[8] Both, the remark and its intuition, resemble Snyder and Ting (2002)'s "information rationale" for political parties, but in two dimensions: ideology and valence

difference with the previous case is that the nominator's ideology is now the party's platform x, and he does not have ego rents.

Nomination stage. As in the previous case, the nominator chooses the candidate by maximizing his own expected utility. Therefore we obtain the equilibrium valence, which can be described as a corollary of lemma (1).

Corollary 1 *Under assumptions in lemma* (1)*, when the nominee is chosen by a majoritarian closed primary election, his symmetric equilibrium valence is defined by*

$$V_x'(q^c) = \omega \frac{4(x^c)^2}{\alpha + 4(x^c)^2} \tag{11}$$

Party formation stage. Similarly, when the party leader anticipates a closed primary he chooses the platform x^c, implicitly defined by

$$-\frac{x^c}{\alpha}\Pi_l(x^c) - \frac{\partial V_y(q^c)}{\partial x^c} = x^c - z_x \tag{12}$$

3.3 Open primaries: $r = o$

It has been argued that as open primaries appeal to a more moderate electorate, candidates selected through this institution are expected to be more moderate as well (Gerber and Morton (1998)). However, there are two strategic effects that are disregarded by that argument: as party leaders anticipate the nomination rule, they can accommodate the platform to countervail the open primaries' effect. Also, as a result, the set of potential candidates (affiliated members) changes.

Nomination stage. I have assumed an uniform distribution of ideologies, symmetric around zero; hence the expected median voter's ideology is zero. Therefore, also as a corollary of lemma (1), I state the optimal valence in this game.

Corollary 2 *Under assumptions in lemma* (1)*, when the nominee is chosen by a majoritarian open primary election, his symmetric equilibrium valence is defined by*

$$V_x'(q^0) = 0 \tag{13}$$

Party formation stage. When the party leader anticipates a candidate with the lowest valence, he chooses the platform x^o such that

$$-\frac{x^o}{\alpha}\Pi_l(x^o) = x^o - z_x \tag{14}$$

In the next section I compare the symmetric PBE of each game; that is, I compare the columns in the table below.

	r=l	r=c	r=o
(x^r, y^r)	$x^l, -x^l$	$x^c, -x^c$	$x^o, -x^o$
(q^r, q^r)	q^l, q^l	q^c, q^c	q^o, q^o

4 Analysis

In this section I expand on the results of the previous section, I highlight the results on valence, and I explore the systematic differences across institutions.

4.1 Direct implications

Valence

In contrast to previous research, as parties become less differentiated, the nominator chooses a lower valence candidate - equations (9) and (11). Since platform moderation causes lower valence, these two are regarded as strategic substitutes. This effect does not take place in open primaries, where the nominator always chooses the lowest valence level possible since he is indifferent between parties -see (13).

Previous results in this literature state that higher valence and moderation are complementary. For instance, in Ashworth and Bueno de Mesquita (2009), valence resembles the results of "vertical differentiation" from the industrial organization literature: if a voter is ideologically indifferent between two parties, he would vote for the highest valence candidate with certainty. Moreover, if platforms converge, the candidate with higher valence wins with probability 1. Hence, the less differentiated the parties are, the higher incentives to invest in valence.

Since I introduce the national shocks $\tilde{\alpha}$ (which is not an exclusive shock to valence), convergent parties cannot win with probability one by choosing a higher valence candidate. Since valence is costly, does not guarantee a victory, and parties care about the implemented policy (because parties are described as a set of voters who naturally care about it), then they nominate higher valence candidates to "compensate" for less centrist platforms.

Institutional design

In this section I show the main differences in the candidates expected attributes caused by the use of different partisan institutions to select them. Broadly, the main prediction is that open primaries' candidates are predictable extremists, while candidates handpicked by party leaders are moderate mavericks; with the closed primary being the intermediate case. That is, on average, the location of a candidates chosen by open primaries is more extremist, and their ideological variance is smaller -ceteris paribus- relative to closed primaries' candidates. This result is stated formally below.

Proposition 1 (Institutional comparative statics) *Let $B(q_i)$ be positive, increasing and convex; let $z_x = -z_y < 0$; and let $x^r < 0 < y^r$. Then from the comparison of the symmetric equilibrium across institutions we learn:*

1. *Monotonicity result: given any pair of platforms, more moderate nominators choose lower valence candidates. Thus*
$$q^l > q^c > q^o.$$

2. *Open primaries extremism: when party leaders anticipate an open primary election during the nomination stage methods, they choose more extreme platforms. Thus*
$$x^o < min\{x^c, x^l\}.$$

3. *Closed primaries extremism: let $B(q_i) = aq_i^2 + b$, and assume $\alpha > 2\Pi_z(x^l)$; then*
$$x^c < x^l.$$

And similarly for Y.

The monotonicity result states the main difference across nomination rules when platforms are fixed: roughly, more democratic primaries produce less popular (lower valence) candidates. In order to disentangle what causes the result, notice that, endogenously, the nominator in open primaries is more moderate than the one in closed primaries, who himself is more moderate than the party leader. (This last step is an equilibrium shown in equations (10), (12), and (14).) Then, as the nominator becomes more centrist, his preferences for one party over the other become milder, therefore he cares less about the outcome of the general election with regards to policy. Hence, the nominator chooses a lower valence candidate whose expected location is still the party's platform but whose variance is lower.

That party leaders choose higher valence candidates than closed primaries still holds even if the party leaders are not more extreme than the median party member. Intuitively, if the political entrepreneurs have very large ego rents they will still choose a higher valence nominee; that is, $\frac{\delta}{4} > 4(x^c)^2 - x^l z_x$. Although in equilibrium the party leader is more extreme than the median party member, and the empirical literature partially confirms this hypothesis (Alesina and Rosenthal (1995)[9])

Regarding open primaries extremism, the party leaders' choice of platform has no effect on the nominators' choice of valence, which is $q^o = 0$; therefore this institutional design resonates with a standard probabilistic voting with no valence, and serves as a benchmark case for the following

[9]They show that presidential candidates are on average more extreme than the median party member; assume the elected president to be the leader of the party, then my result holds.

analysis. Policy and office motivated party leaders select their compromise in policy up to the point where they balance the expected loss in probability with the expected gain in utility; disregarding any thought of valence.

Figure 3: Optimal platforms by nomination rule

In comparison to the previous case, when the leader anticipates a closed primary, he cannot pull the platform so close to his ideal point because now it comes at the cost of a higher variance. First, fix the platform: the closed primary's nominator already chooses a higher valence candidate. Second, let the platform become more extreme: this has two effects that push to an even higher valence; the nominator is now more extreme, and the substitution effect increases ($|\frac{\partial q}{\partial x}|$ is larger). Then, although valence increases the probability of winning -ceteris paribus- it also increases variance in the party leaders' payoff in case of winning, but also in the case of losing (see equation (12)). Summing up, without considering the effect of valence, the party leader pulls the platform as close as possible (as in open primaries), but when he anticipates an effect on valence and variance, he has to choose a more moderate platform. The same reasoning applies to explain $x^o < x^l$. Furthermore, notice that the introduction of valence pushes the symmetric equilibrium to one with less differentiation.

Remark 5 *The equilibrium in standard probabilistic voting models with office and policy motivated politicians is more moderate with the introduction of valence as purely campaigning skills.*

Whether $x^c < x^l$ depends on the party leaders' relative payoff of winning. Intuitively, if the gains of winning are really high, he would prioritize a more moderate platform (which comes only at an individual policy cost), over a higher valence-higher variance candidate. In the proposition, I look at the opposite case: $\alpha > \Pi_z(x^l)$. When the party leader himself handpicks the nominee, he is already solving the trade off with the pair (q^l, x^l) he likes the most. However, when the nominee is chosen by voting in a closed primary, given a platform, the party leader expects a lower valence candidate. Then when, he does not care a lot about winning ($\alpha > \Pi_z(x^l)$) the party leader will move to a more extreme platform in order to increase valence. Moreover, the strategic substitution effect is decreasing in $\Pi_z(x^l)$, thus the adjustment in valence is higher for low $\Pi_z(x^l)$; $|\frac{\partial q}{\partial x}|$ is decreasing in $\Pi_r(x^{r'})$.

An alternative and also intuitive explanation is the following: when the noise in the economy is large ($\alpha > \Pi_z(x^l)$), and the party leader does not control the choice of the nominee in the second period (primaries), he does not find as profitable to invest in a moderate platform so he pulls it to a point where he likes it better. A different approach is to think of the ex-ante moderation of the party: if the party leader is already moderate (low $\Pi_z(x^l)$), not only he does not lose much by reducing his compromise in policy, but also he compensates it by a larger adjustment of valence.

For a complete proof of the proposition, please refer to the appendix.

Welfare analysis

In this section I study whether the social welfare in the society, defined as the general median voter's utility, changes under the different institutional settings. Loosely, since the median voter chooses the nominees for both parties in the case of open primaries, I investigate under what circumstances he prefers a different nomination rule, where he delegates the choice on somebody else.

Intuitively, social welfare when the median voter is the nominator (Ω^o) can be larger than when the party leaders handpick the nominees (Ω^l), only if the loss in variance is larger than the gain in policy. Recall that when the political entrepreneurs anticipate an open primary, they choose a more extreme platform. Therefore $\Omega^o \leq \Omega^l$ if and only if

$$u_0(x^o, 0) = -(x^o)^2 \leq -(x^l)^2 - V_x(q^l) = u_0(x^l, q^l) \tag{15}$$

Lemma 3 *Let $B(q_i) = aq_i^2 + b$, and $\alpha > \Pi_l(x^l)^2 + \alpha\Pi_l(x^l)$. Then, exists $\bar{m} = m(\alpha, \delta, b; x^o)$ such that for all $\frac{\omega}{a} > \bar{m}$, the following holds: $\Omega^o \leq \Omega^l$*

A large $\frac{\omega}{a}$ is a sufficient condition that indicates that the role of valence in voting is large relative to its role in affiliation decisions. That is, despite campaigning skills do matter when citizens vote, its role in extracting higher rents from the party -and as a substitute of ideological distance- is rather limited. In such case, although the variance increases fast with valence, the substitution effect increases even more rapidly; that means that as the party leader chooses a more centrist platform, he will reduce the valence by much. Thus, under the nomination rule $r = l$, not only the platforms are more centrist, but also the overall variance does not increase so much with respect to the open case. Intuitively, decreasing a is similar to increase party's screening (as for the same valence, an affiliated citizen has to be closer to the platform); and I show in the extensions that the equilibrium valence and variance increase with partisan screening. Regarding the remaining assumptions, as in proposition (1), the one on $B(q_i)$ simplifies the analysis; and the assumption on α is necessary, otherwise for low α's, the welfare is always the highest under open primaries. The proof of the lemma can be found in the appendix.

From the lemma above, it follows that there is also a $\bar{\bar{m}}$, such that for any $\frac{\omega}{a}$ larger than it, then $\Omega^o \leq \Omega^c$. Under the same assumptions of the lemma above: $\bar{\bar{m}} < \bar{m}$, thus for $\frac{\omega}{a} \in [\bar{\bar{m}}, \bar{m}]$ the median voter prefers closed primaries to open ones, although not necessarily prefers $r = l$ to $r = c$. In general, whether $\Omega^l \leq \Omega^c$ is ambiguous.

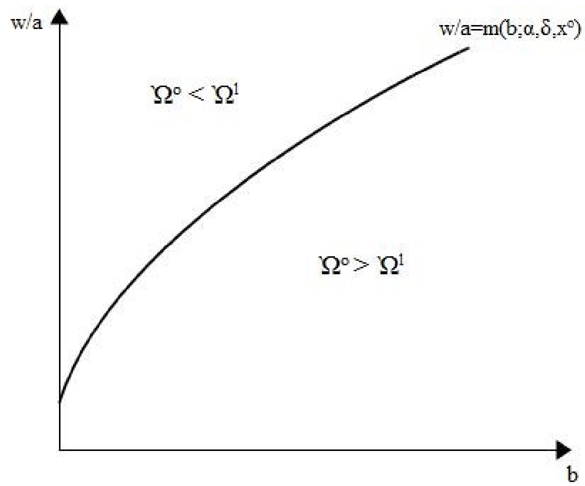

Figure 4: Welfare analysis

4.2 Comparative statics

The institutional comparative statics have shown that nomination rules matter essentially through two mechanisms. First, they change the location of the median voter, and, consequently, the set of potential candidates. Second, as the nomination rules solve the trade off between valence and variance differently, they also change the incentives of the party leader in the first stage of the game. Furthermore, in general, the regular comparative statics exercises also have to take into account the nomination methods.

Recall that the open primaries case serve as a the no valence benchmark, where the first stage choices do not affect the choice of valence. Therefore, the static and dynamic effects are null.

Remark 6 (Open primaries) *In open primaries, the optimal valence is always $q^o = 0$, therefore, as the party leaders' ego rents or the systemic uncertainty α change, q^o remains at the minimum.*

On the other hand, due to the substitution effect between valence and platforms moderation ($\frac{\partial q}{\partial x} \leq 0$), the case of of closed primaries and handpicking turn out to be more interesting. During the nomination stage, a decrease in α, or an increase in ω result in a higher valence q^r, for $r \in \{c, l\}$. Also, as δ increases, q^l increases. The static effect (disregarding the effects on the following stages) during the party formation stage, would be the same as before: higher δ, and lower α would push platforms to the center. However, this compromise by the party leaders would indirectly reduce the equilibrium valence. Therefore, they have to take into account the overall effects on valence as well

$$\frac{dq}{d\delta} = \frac{\partial q}{\partial \delta} + \frac{\partial q}{\partial x}\frac{\partial x}{\partial \delta},$$

and,

$$\frac{dq}{d\alpha} = \frac{\partial q}{\partial \alpha} + \frac{\partial q}{\partial x}\frac{\partial x}{\partial \alpha}.$$

For the remainder of the section I study the effect of the systemic noise and ego rents on the equilibrium valence and platforms, for each case; and let $B(q_i) = aq_i^2 + b$. Thus,

Lemma 4 (Systemic noise: $r \in \{l, c\}$) .

- ***Extreme parties:*** *$\frac{dq}{d\alpha} < 0$. Let $\frac{\omega}{\alpha}$ be large. If $\Pi_r \geq 2\alpha$, then as the systemic uncertainty α increases, the party leaders choose more moderate platforms, and so the equilibrium valence q^r decreases.*

- ***Moderate parties.*** *If $\Pi_r \leq 2\alpha$, then as the systemic uncertainty α increases, the party leaders choose more extreme platforms. The effect on the equilibrium valence q^r is ambiguous.*

As mentioned above, the direct static effect of a larger α would be to decrease both valence, and platforms (more extreme). However, in this game those two decisions may be taken by different players. The party leader's response to increased uncertainty has to take into account the "nominator's" response to it, plus the indirect effect. At the nomination subgame, every nominator decreases the valence as α goes up. In order to increase his overall utility, the party leader could choose a less moderate platform to compensate the future reduction in valence, or he could do the opposite and choose a more centrist platform (which increases the probability of winning). It turns out that the more the nominator cares about winning (for $2\alpha \leq \Pi_r$), the substitution effect between platforms and valence gets milder. Therefore, the party leader finds optimal to choose a more moderate platform to compensate for the future reduction in valence. On the other hand, when $2\alpha > \Pi_r$, the party leader chooses a more extreme platform, and the overall effect is ambiguous.

Regarding the effect of δ, I study each case separately.

Lemma 5 (Ego rents, $r = l$) .

- ***Extreme parties:*** *$\frac{dq}{d\delta} > 0$.Let $\frac{\omega}{\alpha}$ be large. If $2\Pi_l \geq \alpha$, then as the party leaders' ego rents increase, they choose more extreme platforms, and so the equilibrium valence q^l increases.*

- ***Moderate parties.*** *If $2\Pi_l \leq \alpha$, then as the party leaders' ego rents increase, they choose more moderate platforms. The effect on the equilibrium valence q^r is ambiguous.*

The intuition is straightforward. When the returns to winning the elections are high, increasing them makes parties more eager to win. Since ex-ante all voters are equally affected by valence, and the underlying noise (α) is low, the party would like to improve their electoral chances choosing a higher valence candidate. Therefore, both the indirect effect (more extreme platform) and direct effect, increase valence as ego rents increase.

On the other hand, when there are closed primaries, and the party leaders' ego rents go up, they always increase their compromise in policy by choosing more moderate platforms. However, they move the platforms up to a point where still the overall effect on valence is positive. That is, when δ increases, the probability of winning is increased both by a higher valence and by more moderate platforms.

Lemma 6 (Ego rents, $r = c$) *When δ goes up, party leaders choose more moderate platforms, and valence increases unambiguously.*

4.3 Other implications

Democratization of nomination rules

In this section I briefly study why parties do not nominate their candidates by more democratic methods more often. As mentioned above, the current bias against primaries in general, and open primaries in particular is intriguing. While political scientists and world leaders advocate for clean and popular elections as a necessary condition to choose democratic representatives [10], arguably one of the most important and common institutions to all democracies such as political parties are still halfway in the democratization path. [11]

When political parties' selection methods are not regulated exogenously, it is party leaders (or the elite) who can choose to democratize their internal procedures. However, when the platforms are fixed (i.e. during the nomination stage), party leaders would only delegate the decision to a median primary voter if it does not make a difference: trivially, if voters do not care about valence ($\omega = 0$), or if there is no party differentiation ($x = 0 = y$). Otherwise, when the preferred party leaders' candidate is different to the preferred median voters' candidate, the former would never democratize the nomination rules.

This argument could lead to the conclusion that moderate parties have more open nomination methods. However, this claim is true only if the party platforms are fixed. Otherwise, when party leaders anticipate a more open nomination method, they choose a more extremist platform (see Proposition (1) above). Instead of addressing the effect of platforms on the democratization of nomination rules, in the following remark I highlight the effect of nomination methods on the polarization of parties.

Remark 7 *In the short run, when the platforms are fixed, more democratic nomination rules can be observed if there is platforms convergence. However, in the long run, more democratic nomination rules are observed in extremist parties.*

[10]The two most widely quoted criteria from Dahl (1989) are (1) effective participation and (2) voting equality at the decisive stage

[11]"Genuine democratic elections serve to resolve peacefully the competition for political power within a country and thus are central to the maintenance of peace and stability.", from the Declaration of Principles for International Election Observation, endorsed by the UN and various organizations such as "The Carter Center".

Primaries harm the probability of winning

A related but different topic is whether holding primaries harms the likelihood of winning elections. By the hypothesis of "divisive primaries", candidates nominated through primaries have a smaller probability of winning when they are more intense contests (see Agranov (2009) for references and a theoretical explanation). [12]

I do not account for the intensity of primaries, but I briefly address the issue of whether primaries reduce the probability of winning. It is worth emphasizing that no nomination method studied here maximizes the probability of winning (see figure (2)). Moreover, fix party Y's choices to (y, q), then party X's probability of winning is highest when the party leader is the nominator, and lowest when voting in open primary does.

In the same line, it is fair to ask what is the best nomination method for an opposition party, when the incumbent's attributes (valence and ideology) are perfectly known. It turns out that for any observable pair (ϵ_y, q_y), party X maximizes the probability of winning when the party boss hand-picks the candidate. These facts provide additional intuition for the observed scarcity of open primaries in particular, and primaries in general, to nominate candidates. That is why, I pose, primaries are rare and unstable events when they are not mandatory (Hirano et al 2009).

5 Extensions

In this section I modify some details of the original setup in order to answer new questions that emerge from some features of the model. In particular, I study the effect of different affiliations costs on the equilibrium. Then, I briefly address the robustness of my results to the perfect observability of valence (instead of ideology).

5.1 Partisan screening

In order to study the sensitivity of the equilibrium valence to the affiliation costs I assume that the platforms are fixed, with $|z_x| \geq |x|$ and $|z_y| \geq |y|$, and I introduce partisan screening s into the net returns to affiliation:

$$B(q_i) - c(z_i)s.$$

For any pair (ϵ_i, q_i), a larger s makes affiliation more costly. Therefore, for a given valence q, higher screening costs leads to smaller uncertainty on the location of the candidate. On the contrary, lower affiliation costs makes for a wider party ideologically (less cohesive), and the party label becomes less informative. [13] In this setup, a more informative party label can lead the nominator to choose

[12] The empirical literature still struggles with this hypothesis since some studies find either no effect (Alvarez et al. (1995), for the U.S.) or a positive effect (Carey and Polga-Hecimovich (2008) for Latin America).

[13] If we think parties as a "tent", as it is commonly mentioned, lower screening costs enlarge that tent.

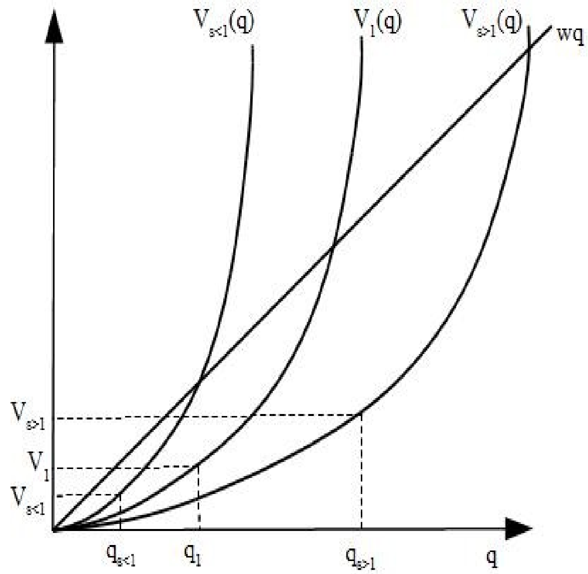

Figure 5: Optimal valence for different screening levels vs benchmark s=1

higher valence candidates at the same cost, so his incentives change, and therefore the equilibrium changes.

Lemma 7 *Let $q^r_P(s)$ be the equilibrium valence in party P, given the nomination rule r, and the screening cost s. The optimal valence $q^r_P(s)$ increases monotonically in s.*

On the other hand, party leaders also accommodate the party's platform to the affiliation costs. In particular, more informative party labels (higher s) reduces the uncertainty of any given candidate, which permits the leader to bring the platform closer to his preferred policy. For a formal proof of this intuition, and the previous lemma, see the appendix.

A conjecture: The case of California

In June 2010, California has changed its nomination rule for candidates to the U.S. Congress by moving toward more open primaries [14]. As argued above, despite that the new nomination rule in California does not fit any of my previously studied stylized rules, if platforms stay fixed in the short time, opening up primaries should produce lower valence candidates than the previous Californian candidates. Moreover, political scientists have stated that the change of rules could also affect candidates' entry (see Kanthak and Morton (2001) for a summary of this literature). In particular, open primaries could lead to entry of more candidates, which within my model can be

[14] To a system where the top two candidates with more votes in the primaries are the ones that run in the general election, independently of their party affiliation

re-stated as lower screening costs, "wider" parties, and less informative labels. Consequently, lower valence candidates will arise not only as a direct result of the open primaries but also because of a lower s. Thus, assuming that still one candidate from each party makes it to the top two vote getters, extrapolating the results of the model, and adding screening as explained, the following prediction is in place.

Conjecture 1 *In the short term, Californian candidates' will be of more predictable, and lower valence on average.*

Intuitively, larger candidate entry, and the loosening of party ties dilutes the party brands. Hence, voters select candidates whose ideology is easily predicted, instead of high-valence mavericks. It is worth noting that in the medium and long run, as platforms can be changed, two effects take place. As stated in the main body of the paper, the parties will become more extremist ($|x^o| > |x^c|$) due to the change in the nomination rule and candidates' entry. However, this platforms' movement also counterweights (partially or fully) the reduction in valence.lutes the party brands. Hence, voters select candidates whose ideology is easily predicted, instead of high-valence mavericks. It is worth noting that in the medium and long run, as platforms can be changed, two effects take place. As stated in the main body of the paper, the parties will become more extremist ($|x^o| > |x^c|$) due to the change in the nomination rule and candidates' entry. However, this platforms' movement also counterweights (partially or fully) the reduction in valence.

Optimal screening

An interesting question naturally arises: what is the optimal level of screening? From the lemma above, and $q \in [0, \bar{q}]$, with $\bar{q} < \infty$, choosing an infinitely large screening cost would ensure that you can choose the highest valence candidate at zero cost. Therefore, the main trade-off disappears. [15]

5.2 Observed ideologies

Throughout the paper I have assumed that valence is common knowledge, while ideologies are private (as in Snyder and Ting (2002). This is a sensible assumption, and it is consistent with the remaining modeling choices. However, in order to show that the main result is robust to this assumption, I provide a brief intuition on how the mechanisms in the model would change if the ideology of citizens is observed.

Suppose that z_i is observed, and that q_i is private knowledge, and independent from the ideologies. The set of affiliated voters does not change, and the size and composition of parties is the same as before. Moreover, among the affiliated citizens, the correlation between ideology and valence still holds. Therefore, when voters observe the ideology and the affiliation of the candidates, they

[15] However, if s could only be chosen up to $\bar{s}$, then the gain from a larger s may depend on whether the optimal q is above or below $\bar{q}$

update their beliefs on the conditional density $q_i|z^P$. They would still be risk averse on ideologies, but now there's no incomplete information on that dimension.

Lemma 8 *Assume $B(q)$ as before. For α large enough, there is large enough $\frac{\omega}{a}$, such that the candidates' valence is decreasing in nominator's ideology. That is, more moderate nominators choose candidates whose valence is smaller, on average.*

Intuitively, when $\frac{\omega}{a}$ is large two effects are taking place. First, as ω increases, the value of signaling a large valence increases, relative to choosing a more moderate policy; therefore, more extreme nominators chose candidates whose preferred policy is closer to them, and therefore have higher valence. Second, as a decreases, ideology becomes a stronger signal of valence; therefore as the nominator is more extreme, he cares more about winning, thus he chooses to depart more from the party's platform, in order to signal a higher valence. A proof of the lemma can be found in the appendix. Also, as with the main model, higher uncertainty leads the nominator to choose smaller valence levels. In this case, that allows him to pick a candidate whose ideology is closer to him.

On the contrary, suppose $\lim \frac{\omega}{a} \longrightarrow 0$. When valence does not play a role, the nominator only have incentives to choose a moderate policy to increase the probability of winning. Thus, more moderate nominators are willing to compromise to more moderate policies. When choosing a more moderate policy also has the benefit of signaling a higher valence, depending on the strength of the signal, more moderate nominators may decide to compromise more (when the signal is weak), or less (when the signal is strong). Moreover, if α is too small, all nominators find more profitable to choose more moderate policies.

At last, notice that if high valence party members scale up positions within the party more easily than low valence ones, then both voters and party members could learn the candidates' valence. This is one more reason, why in the general setup it makes more sense to assume that valence and not ideologies are observed.

5.3 Cross-over voting

In this section, I simultaneously tackle two caveats of my approach. First, not all primaries fit my open-closed setup; and not all citizens vote in primaries as assumed. By making an allowance for mixed types, between the two stylized electoral rules discussed in the paper, the median primary voter's location varies according to the institutional details.

Regarding the institutional design, it should be pointed that in the U.S. there are open and closed primaries, but there are a few intermediate cases as well. Following McGhee (2010) classification, there are two more types of partisan primaries: semi-closed, and semi-open. Broadly, in the former voters can declare their affiliation on election day. In the latter, they are not required of an affiliation, but they choose in which primary they are voting; that is, they cannot participate in both.

Taking into account the mixed cases helps to get around the assumptions of full turnout and sincere voting, which despite being demanding (specially in open primaries), they are useful to capture the idea that more open primaries appeal to a more moderate electorate. The main difference among these mixed types is the level of crossover voting and turnout. For instance, right-leaned citizens may turn out to vote in the left party's primaries, either to help nominating a candidate they like better ("sincere crossover") or to help nominating a candidate with less chances ("strategic crossover"). Therefore, within our stylized characterization of primaries, the amount and type of crossover affect the location of the median voter in a primary election.

The introduction of a flexible "intermediate" case allows to account for differences in turnout and crossover voting in a simple way: changing the location and affiliation of the nominator. Let m_p be the ideology in any primary where the nominator is a member, and o_p when he is not. For instance, assume that higher sincere crossover voting moves the median voter in the primary closer to the general median voter, hence the nominator could either be a more moderate party member than the median member ($|m_p| > |m'_p|$), or a more moderate non-member ($|m_p| > |o_p|$). Similarly, higher levels of "strategic" crossover would produce a change "as if" the median voter were more extremist. Intuitively, higher sincere crossover voting leads to less popular, lower valence, more certain candidates. Let δ_m be stand for party member's ego rents when his party wins, then the following lemma formalizes the intuition.

Lemma 9 *For $\delta_m > 4x(o_x - m_x)$ any party member always chooses a more popular candidate than a non-member.*

A party member with ideology $m_x \leq o_x$ always chooses a more popular candidate than a non member. When $m_x \geq o_x$, for sufficiently high ego rents, and sufficiently centrist parties (x close to 0), the statement also holds. Intuitively, when the nominator is a non-member with extreme preferences he cares more about winning and so he will lean toward a more popular candidate than a member, unless the party is too centrist and/or the ego-rents are too high.

In sum, I posit that when the institutional design favors strategic cross-over instead of sincere cross-over, the nominator is more extreme, and therefore the nominees are higher valence. Whether semi-closed or semi-open primaries incentivize voters to behave in one way or another, depends on the details of the rule, which cannot be easily captured just by the taxonomic description of primaries.

6 Conclusions

I presented a model where party membership and valence play an important role in understanding how nomination rules influence the choice of candidates. In particular, I have shown a rationale for a piece of unexplained empirical evidence: open primaries' nominees are more extremist and more predictable than closed primaries' nominees. Similarly, if party leaders were to handpick the candidates, they choose moderate mavericks, that is, less predictable, lower valence, more moderate candidates. Moreover, when the campaigning effect of valence is high relative to its effect on membership, social welfare under open primaries is strictly smaller than under an autocratic nomination method as handpicking by the party leader. That is, under those conditions, before the game begins, the median voter in open primaries would rather delegate his choice to the party leader.

These results also speak to other branches of the literature, and provide novel results. I have shown that valence, as campaigning skills that only affect the chances of winning, decreases with platform moderation. Moreover, I provide one more explanation for the increases polarization of parties in the U.S. that is not supported by a change in preferences. That is, the current trend of opening up primaries leads to higher party polarization. Also, it is not necessarily true that parties with more democratic internal structures (such as more open primaries) are more moderate. All the contrary, when platforms are flexible, open primaries lead to more extreme parties. Last, I also studied the effect of screening costs on the equilibrium valence of candidates: increasing the screening costs makes for tighter parties that leads to nominate higher valence candidates whose ideology is less predictable.

To sum up, the main predictions from the model are that (1) candidates chosen by more democratic nomination rules are expected to be more extremist and more predictable; (2) candidates chosen by less democratic methods are more moderate but they are less predictable (higher variance); (3) nomination rules have an effect on the overall economy, as contribute to the explanation of party polarization, democratization, policy instability, and the popularity of candidates. Therefore, as rules that shape not only shape electoral outcomes, nomination rules should also be examined as systemic variables in any given country or state.

7 Appendix

7.1 Bounded probability

Lemma 10 *Let $q^* \equiv argmax P_x$, then for $\alpha \geq \omega q^* - V(q^*) + V(0)$ the probability of winning is bounded between 0 and 1.*

Proof. The probability of winning for party X in the symmetric equilibrium can be written as

$$P_X \equiv \frac{\alpha + \omega(q_x - q_y) + V(q_y) - V(q_x)}{2\alpha}.$$

Consider $q^* \equiv argmax P_x$; no voter will choose a candidate whose valence is larger than q^*. Since the benefit of higher valence, $\omega q - V(q)$ is convex in q, for any probability of winning with $q > q^*$, there is a lower valence candidate that leaves the same probability of winning at a lower individual cost (lower variance). Thus, we can focus on $q \in [0, q^*]$. For all values of $(q_x, q_y) \in [0, q^*]^2$, the probability above is bounded between 0 and 1 when the following conditions hold: First, let take the biggest possible value of P_X, that is, when X chooses q^*, and Y, 0:

$$P_X < \frac{\alpha + \omega q^* - V(q^*) + V(0)}{2\alpha} < 1$$

And, the smallest, when the opposite is true:

$$P_X > \frac{\alpha - \omega q^* + V(q^*) - V(0)}{2\alpha} > 0.$$

Both conditions boil down to

$$\alpha \geq \omega q^* - V(q^*) + V(0)$$

■

7.2 Median voter theorem

Closed primaries

Lemma 11 *Let $q^* \equiv argmax P_x$, then for $\alpha \geq \omega q^* + V(q^*)$ the median voter theorem holds*

I need to proof that the expected utility of voter i with ideology $z_i \leq 0$ (for X primaries) - $EU_i(q_x, q_y)$ - is concave in q_x, in the compact set $[0, q^*]$ for all q_y in the same set. If the $EU_i(q_x, q_y)$, is concave, then it is single peaked. **Proof.** Let $\bar{q} \equiv \lambda q' + (1 - \lambda)q''$; then $EU_i(q_x, q_y)$ is strictly concave in q_x if for all $(q', q'') \in [0, q^*]^2$, and for all $\lambda \in [0, 1]$, the following holds:

$$EU_i(\bar{q}, q_y) > \lambda EU_i(q', q_y) + (1 - \lambda)EU_i(q'', q_y) \tag{16}$$

Let the LHS be

$$\frac{\alpha + \omega(\bar{q} - q_y) + V(q_y) - V(\bar{q})}{2\alpha}(4z_i x + V(q_y) - V(\bar{q})) - (z_i + x)^2 - V(q_y),$$

and in the RHS, we can write

$$\lambda EU_i(q', q_y) \equiv \lambda[\frac{\alpha + \omega(q' - q_y) + V(q_y) - V(q')}{2\alpha}(4z_i x + V(q_y) - V(q')) - (z_i + x)^2 - V(q_y)],$$

and

$$(1-\lambda)EU_i(q'', q_y) \equiv (1-\lambda)\frac{\alpha + \omega(q'' - q_y) + V(q_y) - V(q'')}{2\alpha}(4z_i x + V(q_y) - V(q'')) - (z_i + x)^2 - V(q_y)$$

Thus, adding up the last two lines we get:

$$\frac{\alpha + \omega(\bar{q} - q_y) + V(q_y) - \lambda V(q') - (1-\lambda)V(q'')}{2\alpha}[4z_i x + V(q_y)] - \lambda V(q')P(q', q_y) - (1-\lambda)V(q'')P(q'', q_y)$$

We can re-write the equation (16) as

$$[4z_i x + V(q_y)][\lambda V(q') + (1-\lambda)V(q'') - V(\bar{q})] > V(\bar{q})P(\bar{q}, q_y) - \lambda V(q')P(q', q_y) - (1-\lambda)V(q'')P(q'', q_y)$$

Since $V(q)$ is convex in q, then $V(\bar{q}) < \lambda V(q') + (1-\lambda)V(q'')$, so the left hand side is always positive for $x < 0$ and $z_i < 0$. Thus, if the function

$$V(q)P(q, q_y) = V(q)\frac{\alpha + \omega(q - q_y) + V(q_y) - V(q)}{2\alpha}$$

is convex in q, then the RHS is always negative. Its second derivative is

$$\frac{\partial^2 V(q)P(q, q_y)}{\partial q^2} = V''(q)\frac{\alpha + \omega(q - q_y) + V(q_y) - 2V(q)}{2\alpha} + 2V'(q)\frac{\omega - V'(q)}{2\alpha}.$$

Since $V(q)$ is positive and convex, and $\omega - V'(q) \geq 0$, a sufficient condition for $\frac{\partial^2 V(q)P(q,q_y)}{\partial q^2} \geq 0$ is

$$\alpha \geq \omega q^* + V(q^*).$$

■

Notice that the minimum condition for $\alpha + \omega(q - q_y) + V(q_y) - V(q) > 0$ is

$$\alpha > V(q^*)$$

if $\omega q^* - 2V(q^*) > -2V(0)$, otherwise the sufficient and necessary condition is

$$\alpha \geq \omega q^* - V(q^*) + 2V(0).$$

However, to avoid making more assumptions on $V(q)$ we just impose an additional constraint on the distribution of $\tilde{\alpha}$.

Open primaries

The median voter theorem does not necessarily hold in open primaries because the voters doing crossover (i.e. that prefer Y but vote in X primaries) may be indifferent between voting for a low valence candidate with a low probability of winning, and for another one with such a high valence that makes him less likely to win because of his high unpredictability (variance). However, the equilibrium in open primaries resembles a median voter equilibrium since the winning candidate would be one that maximizes the general median voter's utility.

For this proof, I only look at the open primaries in party X to prove that $q_x = q_y = 0$ is the only equilibrium. First, I show that for any $q_y > 0$ all voters to the left of the general median voter would rather vote for any candidate with $q \leq q_y$[16] than for a candidate with no valence. Then, I show that all voters to the right of the median voter (including him) would never vote for a candidate with $q \geq q_y$. Lastly, I show that if there is a voter to the right of the general median voter that prefers a candidate with $q < q_y$, then that candidate wins the election. Thus, with a similar reasoning, in Y primaries a candidate with valence smaller than q would win, and so in X's primaries they would like to vote for a candidate with an even smaller valence, and so on, until $q_x = q_y = 0$, which I prove is an equilibrium.

Define the interim utility of voter i - $EU_i(q_X, q_Y)$- as the expected utility of the voter when X's candidate has a valence of q_X, and Y candidate's valence is q_Y. For instance, if $EU_i(q = 0, 0) > EU_i(q > 0, 0)$, then voter i votes for the candidate with no valence, and otherwise. Below is the full proof, where I assumed that $q \in [0, argmaxP_x]$

Lemma 12 *For all voters with $z_i < 0$, and for all q: $V(q) < 4z_i x + V(q_y)$, then*

$$EU_i(q, q_y) > EU_i(0, q_y)$$

Proof. Let $P(q, q_y) - P(0, q_y) \equiv \Delta P \geq 0$. The equation above implies

$$[4z_i x + V(q_y)][P(q, q_y) - P(0, q_y)] > V(q)P(q, q_y) - V(0)P(0, q_y)$$
$$\Leftrightarrow$$
$$[4z_i x + V(q_y)]\Delta P > V(q)\Delta P > V(q)P(q, q_y) - V(0)P(0, q_y)$$
$$\Leftrightarrow$$
$$4z_i x + V(q_y) > V(q)$$
$$\Leftrightarrow$$
$$q \leq V^{-1}(4z_i x + V(q_y)) \Rightarrow q \leq q_y$$

■

Therefore, the question is what do voters with $z_i \geq 0$ vote for? If they all prefer to choose a zero-valence-candidate, then that candidate wins. On the other hand, they may prefer to vote for the high valence candidate, as that increases the candidate's variance as well. If they would rather

[16]Actually, the condition is less restrictive:

$$q \leq V^{-1}(4z_i x + V(q_y))$$

choose the high valence candidate, then the high valence candidate wins.

Lemma 13 *For all voters with $z_i \geq 0$, and for all $q \geq q_y$, then*

$$EU_i(q, q_y) < EU_i(0, q_y)$$

Proof. As above, the equation in the lemma, implies

$$[4z_i x + V(q_y)]\Delta P < V(q)P(q, q_y) - V(0)P(0, q_y)$$
$$\Leftrightarrow$$
$$4z_i x + V(q_y) < V(q)$$

Which always holds under the conditions in the lemma. ■

Lemma 14 *For all voters with $z_i \geq 0$, then $q < q_y$ is an equilibrium where*

$$EU_i(q, q_y) \geq EU_i(0, q_y)$$

only if $q_x = q_y = 0$.

When $q < q_y$, the condition in (17) may hold for a large set of parameters. However, it cannot hold in equilibrium for $q \neq 0$ as I prove below, by contradiction.

Proof. Suppose exists $z_j > 0$ such that j votes for $q < q_y < \bar{q}$. Then, all voters in X primaries with $z_i \leq 0$ and j vote for the candidate with $q_x = q$. But then in Y primaries, exists j', with $-z_{j'} > 0$ such that he votes for $q < q_y < \bar{q}$. Then all voters in Y primaries with $z_i \geq 0$ and j' vote for candidate with $q_y = q$. But then the cycle begins again, and allways a candidate with lower valence would be selected until $q_x = q_y = 0$. It remains to be proved that the pair $(0, 0)$ is an equilibrium. Thus, I investigate under which conditions the best response of a voter with $z_i \geq 0$ is to vote for a candidate with zero valence when $q_y = 0$.

$$EU_i(0,0) > EU_i(q_{>0},0)$$
$$\Leftrightarrow$$
$$P_x(0,0)(u_i(x;0,0) - u_i(y;0,0)) - (z_i - y)^2 - V(0) > P_X(q,0)(u_i(x;q,0) - u_i(y;q,0)) - (z_i - y)^2 - V(0)$$
$$\Leftrightarrow$$
$$0.5(u_i(x;0,0) - u_i(y;0,0)) > \frac{\alpha + \omega q - V(q)}{2\alpha}(u_i(x;q,0) - u_i(y;q,0))$$
$$0.5(-(z_i - x)^2 + (z_i - y)^2) > \frac{\alpha + \omega q - V(q)}{2\alpha}(2z_i(x - y) - V(q))$$
$$\Leftrightarrow$$
$$4z_i x > \frac{\omega q - V(q)}{\alpha}(4z_i x - V(q)) + 4z_i x - V(q)$$
$$\Leftrightarrow$$
$$V(q) > \frac{\omega q - V(q)}{\alpha}(4z_i x - V(q)) \quad (17)$$

Which always holds, since the LHS is positive, and the RHS is always negative for $\omega q - V(q) > 0$ and $z_i > 0$. Therefore, i's best response in open primaries, when $z_i > 0$ is:

$$BR_i(q_Y = 0) = 0, \ \forall \ z_i > 0 \ \text{ in X's open primaries}$$

■

7.3 The nomination subgame with discrete valence

Studying the discrete case serves two purposes. First, it illustrates the main trade-off and results in a transparent and plain setting. Second, assuming that valence can take only two values can also be interpreted as if there are two candidates with different valence in each primary election. This allows to fairly situate my results into the literature of primaries, where most of the models assume two exogenous candidates with different valence and ideology (Meirowitz (2005), Hummel (2010)).

Say that there are two valence levels $q_H > q_L$, and the probability of winning increases with high valence candidates:

$$\omega q_H - V(q_H) > \omega q_L - V(q_L).$$

Let the platforms be symmetric, and the probability of X winning when the party chooses a low valence candidate be $P_x(q_L)$, as defined in the main body of the paper. Let Π_i stand for the net payoff of winning to voter i with ideology z_i, and let Π_l stand for the net payoff of winning to the party leader, with ideology z_x. The median voter in the primaries with ideology z_i and ego rents

δ_i will choose a lower valence candidate than the party leader (z_x, δ), for all $q_y \in \{q_L, q_H\}$ if and only if

$$P_x(q_L)\Pi_i(q_L) > P_x(q_H)\Pi_i(q_H),$$

and

$$P_x(q_L)\Pi_l(q_L) < P_x(q_H)\Pi_l(q_H),$$

Let the uncertainty α has to be large enough such that the probability of X winning is bounded by 1, hence $\alpha \geq \omega(q_H - q_L) + V_L - V_H$. Let $x = -y < 0$ be the parties' platforms; and let V_y stand for party Y's choice, and V_{-y} for the complement, i.e. if Y's choice is V_H, then $V_{-y} = V_L$.

Lemma 15 *Let $V_H \equiv V(q_H)$, and $V_L \equiv V(q_L)$, then the two conditions above are satisfied if and only if*

$$\Pi_i \equiv \delta_i + 4xz_i < \frac{(V_H - V_L)\alpha}{\omega(q_H - q_L) + V_L - V_H} + V_{-y} - V_y < \delta + 4z_x x \equiv \Pi_l. \tag{18}$$

As explained in the main body of the paper, parties create a correlation between valence and variance among their members. I am interested in the case when parties attract high valence individuals far from their platform, then that correlation is positive; and, consequently, the ideology of a high valence candidate is more uncertain than a low valence candidate.

However, before introducing this case, I briefly digress to study what would happen in the other one: if the return to valence $B(q)$ is decreasing in valence, then the correlation is negative. In that case, voters expect that high-valence candidates are also closer to the party platform (more certainty). Then, there is no trade off, and when voters are risk averse the parties always choose the highest valence candidate. In the toy model above, when the variance is decreasing in valence, $V_H < V_L$, the middle term in equation (18) is always negative, and so the left hand of the equation does not hold. So under any nomination rule, the chosen candidates are high-valence.

On the other hand, when the variance increases with valence, $V_H > V_L$, a trade-off between valence and variance arises. Then, parties do not always nominate the highest valence party member because voters learn that his preferred policy is more uncertain. In the model, notice that as the cost of higher valence $(V_H - V_L)$ increases, the middle term in (18) increases, and so the scope for nominating higher valence candidates is reduced. On the other hand, as the uncertainty (α) decreases, the "randomness" of the electoral process diminishes and so it makes sense for parties to invest in higher valence members. In the equation above, the left part becomes binding for a low enough value of α.

It is also interesting that valence would decrease monotonically (in the continuous case) with the ideology of the "nominator". In this simple setup, as you increase $z_i < 0$, the left hand side becomes smaller, and so the scope for choosing higher valence candidates is reduced. Intuitively, as the parties are less differentiated, the gain in utility through policy when your party wins is reduced because the two parties have more similar policies. Therefore, the returns to investing in high valence are lower.

Thus, the more centrist the parties are (and nominators), the more likely the candidate is low valence. This intuition also drives the main result: since, in equilibrium party leaders are always more extreme than the party mean[17], they always choose more popular candidates than primaries. That is, for standard values of $V_H - V_L$, the lemma is satisfied for $|z_x| > |z_i|$.

Then, I sum up the results in the following proposition.

Proposition 2 (The discrete nomination subgame) .

1. *Party leaders choose more popular candidates than closed primaries(i.e. more centrist nominators are more likely to select a low valence candidate)*

2. *Less ideological differentiation between parties leads to lower valence candidates (i.e. valence and platforms are substitutes)*

3. *Larger ego rents (δ), and/or less uncertainty (α) lead to higher valence candidates*

4. *As the returns to valence increase (ω), and its cost decrease ($V_H - V_L$), parties are more likely to chose higher valence candidates*[18].

All the facts above still hold in the more general continuous case, which allows for richer conclusions.

7.4 Probabilistic voting

There is a continuum of voters whose utility depend on the implemented policy and their affiliation decisions. Let the pair of the unobserved ideology, and observed valence -$(\tilde{z}_i, qi)$- identify each voter, with $z_i \sim U[-k, k]$, $q_i \sim U[0, \bar{q}]$, and $z_i \perp q_i$. To keep it short, let $u_i(\cdot, X)$ be the utility if party X's candidate wins, for all affiliation decisions. All voters have an exogenous preference for party Y over X, which can be decomposed in an individual preference $\beta_i \sim U[-\beta, \beta]$, and a general bias

$$\tilde{\alpha} \sim U[-\alpha + h(q_Y - q_X), \alpha + h(q_Y - q_X)],$$

which could be described as idiosyncratic and national shocks, unobserved by the parties; the national shock's distribution depends on the valence -campaigning skills- of the candidates. Also notice that in the main body of the paper I do not introduce the idiosyncratic shock: I do it here to highlight that, under this setup, it does not affect the probability of winning.

[17] I show this when I solve the party formation game

[18] The derivative of the middle term with respect to $V_H - V_L$ is

$$\frac{\alpha\omega(q_H - q_L)}{\omega(q_H - q_L) - (V_H - V_L)} \pm 1$$

which is always positive.

Thus, a voter i votes for party X if:

$$u_i(\cdot, X) - u_i(\cdot, Y) > \beta_i + \tilde{\alpha},$$

hence

$$\begin{aligned}
\Pr(u_i(\cdot, X) - u_i(\cdot, Y) \geq \beta_i + \tilde{\alpha}) &= \Pr(u_i(\cdot, X) - u_i(\cdot, Y) - \tilde{\alpha} \geq \beta_i) \\
&= \frac{u_i(\cdot, X) - u_i(\cdot, Y) - \tilde{\alpha} + \beta}{2\beta} \\
&= \frac{2z_i(x-y) + \omega(q_x - q_y) - (V_x - V_y) + y^2 - x^2 - \tilde{\alpha}}{2\beta} + \frac{1}{2} \\
&\equiv P^i(u_i(\cdot, X), u_i(\cdot, Y); \tilde{\alpha}) \equiv P^i \qquad (19)
\end{aligned}$$

Notice that parties are all of the same size since they have the same recruiting technology, and the uniform distribution of ideologies. Thus, integrating over the affiliated and non-affiliated voters does not change any of the following results. Hence, given the realization of the national shock, the share of people that votes for party x is:

$$\begin{aligned}
\int P^i \frac{1}{2k} dz_i &= \frac{\omega(q_x - q_y) - (V_x - V_y) + y^2 - x^2 - \tilde{\alpha}}{2\beta} + \frac{1}{2} \\
&\equiv S_x \qquad (20)
\end{aligned}$$

And lastly, the probability that party x wins is:

$$\Pr(S_x > \frac{1}{2}) = \frac{\omega(q_x - q_y) - (V_x - V_y) + y^2 - x^2}{2\alpha} + \frac{1}{2} \equiv P_x$$

7.5 Analysis

Optimal valence

The results stated in Proposition (1) come from solving the maximization problems at the nomination and party formation stages. I proceed to solve each problem here. Regarding the nomination stage.

Proof. First, let a nominator with utility $u_n(\cdot)$ solve his expected utility, for any given pair (x, y)

$$\max_{q_X} E(u_n(a_p, \tilde{z}_p; q_p^r)) = P_x u_n(a_p, \tilde{z}_X; q_x^r) + (1 - P_x) u_n(a_p, \tilde{z}_Y; q_y^r).$$

Then, remember that for any nomination rule r, I define $\Pi_r \equiv u_n(\cdot; q_X) - u_n(\cdot; q_Y)$, and

$$\begin{aligned} 0 =& P_x'(u_n(\cdot; q_X) - u_n(\cdot; q_Y)) + u_n'(\cdot; q_X)P_x + u_n'(\cdot; q_Y)(1 - P_x) \\ 0 =& \frac{\omega - V'(q_X)}{2\alpha}\Pi_r - V'(q_X)P_x \\ V'(q_X) =& \omega\frac{\Pi_r}{\Pi_r + 2\alpha P_x} \end{aligned}$$

And the SOC hold strictly: $-V''(q_X)(\frac{\Pi_r}{2\alpha} + P_X) < 0$. Thus, when $x = -y$:

$$V'(q_X) = \omega\frac{\Pi_r}{\Pi_r + \alpha},$$

there is a unique symmetric equilibrium, shown in equations (9), (11), and (13). ■

Optimal platforms

Regarding the party formation stage.

Proof. Second, during the party formation stage, the party leader chooses the platform such as he maximizes his expected utility, thus he solves equation (7), by which I obtain:

$$0 = [\frac{-x}{\alpha} + \frac{1}{2\alpha}(\frac{dV_Y}{dx} - \frac{dV_X}{dx})]\Pi_l - P_X 2(x - z_x) - P_x\frac{dV_X}{dx} - (1 - P_x)\frac{dV_Y}{dx} \tag{21}$$

And notice that in the symmetric equilibrium with $x < 0 < y$, $\frac{\partial V_x}{\partial x} = -\frac{\partial V_y}{\partial y}$, so

$$\frac{dV_Y}{dx} = \frac{\partial V_Y}{\partial y}\frac{\partial y}{\partial x} = -\frac{\partial V_Y}{\partial y} = \frac{\partial V_X}{\partial x} \tag{22}$$

Hence, in the symmetric equilibrium, the equation (21) becomes:

$$x - z_x = \frac{-x}{\alpha}\Pi_l - \frac{dV_X}{dx}.$$

■

Before moving to the proof of the Proposition itself, I prove lemma (10) and the following remark, by which party leaders find optimal to form a party for large enough k.

Why do leaders form parties? Here I prove that a political entrepreneur $(z_x, \bar{q})$ with the maximum possible valence forms a party when he faces (z_y, q_i). Then it follows that party leaders with lower valence will form parties as well. Recall that if we would let a non affiliated citizen to be a candidate, the voters' beliefs on their policy preference and uncertainty would be: $E(z_p) = 0$, and $V(z_p) = \sigma = \frac{k^2}{3}$. Notice that this proof holds for any pair (z_x, z_y), and not only for the symmetric case.

Proof. Let $P_x(e_x|e_y)$ be the probability that X wins, given that z_y has chosen $e_y \in \{0, Y\}$, where 0 means that he is running alone, while $e_y = Y$ means that he formed a party. With this notation, the sketch of the proof is straightforward: I want to show that the best response of X to Y's strategy is to form a party. So, if

$$P(X|Y)\Pi(X|Y) > P(0|Y)\Pi(0|Y),$$

and

$$P(X|0)\Pi(X|0) > P(0|0)\Pi(0|0),$$

then forming the party is a dominant strategy for both entrepreneurs, due to symmetry. Then it follows that

$$\frac{\delta + (V_y - V_x) + (y - z_x)^2 - (x - z_x)^2}{\delta + V_y + (y - z_x)^2} > 1 + \frac{y^2 + V_y - \sigma^2 + (\bar{q} - q_y)}{\alpha},$$

and

$$1 + \frac{-x^2 - V_y + \sigma^2 - (\bar{q} - q_x)}{\alpha} > \frac{\delta + (z_y - z_x)^2}{\delta - V_x + (z_y - z_x)^2 - (x - z_x)^2}.$$

And there exists finite σ such that they hold. ■

7.5.1 Proposition (1)

The **monotonicity result** in the proposition is straightforward from equations (9), (11), (13):
Proof. Since $V'(q^r) \geq 0$ for all r, then if $V'(q) > V'(q')$, then $q > q'$. Then, the last inequality below follows from z_x, x^l, x^c negative,

$$\omega \frac{\delta + 4z_x x^l}{\alpha + \delta + 4z_x x^l} > \omega \frac{4(x^c)^2}{\alpha + 4(x^c)^2} > 0$$

while the first inequality follows from $z_x < \frac{4(x^c)^2 - \delta}{4x^l}$. Thus for any given platform $x^l = x^c = x$, the inequality always holds since $z_x < x$ in equilibrium. ■

On the other hand, I prove that $x^o < x^c < x^l < 0 < y^l < y^c < y^o$ by the absurd.

Recall that

$$-\frac{x^o}{\alpha}\Pi_l(x^o) = x^o - z_x \quad \text{, and} \quad -\frac{x^r}{\alpha}\Pi_l(x^r) - \frac{\partial V_y(q^r)}{\partial x^r} = x^r - z_x.$$

Then, the **extremism results** are proved.
Proof. First, suppose that $x^o \geq x^r$ for $r \in \{c, l\}$. Then,

$$x^o - z_x \geq x^r - z_x \quad \text{, and} \quad -\frac{xo}{\alpha}\Pi_r(x^o) \leq \frac{xr}{\alpha}\Pi_r(x^r)$$

Therefore, it has to be the case that

$$-\frac{\partial V_y(q^r)}{\partial x^r} < 0,$$

which is absurd. Then, $x^o < x^r$. ■

With the same reasoning as above, let $B(q_i) = aq_i^2 + b$, I prove that for $\alpha > \Pi_r(x^l$, then $x^l > x^c$.

Proof. First, suppose $x^l \leq x^c$. Then, it has to be the case that

$$-\frac{\partial V_y(q^c)}{\partial x^c} > -\frac{\partial V_y(q^l)}{\partial x^l}.$$

Then,

$$\frac{\Pi_c}{(\Pi_c + \alpha)^3} > \frac{\Pi_l}{(\Pi_l + \alpha)^3}\frac{z}{x^c},$$

which is absurd because when $0 > x^c \geq x^l$, then $\Pi_l \geq \Pi_c$, and both sides of the inequality are increasing in Π_r for $\alpha > 2\Pi_r$. Thus, $x^c < x^l$. ■

Notice that $\frac{\partial \frac{\Pi_r}{(\Pi_r+\alpha)^3}}{\partial \Pi_r} = \frac{\alpha - 2\Pi_r}{(\Pi_r+\alpha)^4}$

7.5.2 Welfare analysis: proof of lemma (3)

Proof. First, subtract equation (14) from (10). Then

$$[(x^l - x^o)(\alpha + \delta) + \alpha\frac{\partial V_y}{\partial x^l}]\frac{1}{-4z_x\alpha} = (x^l)^2 - (x^o)^2.$$

Second,

$$\Omega^o \equiv u_0(x^o, 0) = -(x^o)^2 \leq -(x^l)^2 - V_x(q^l) = u_0(x^l, q^l) \equiv \Omega^l$$

$$\Leftrightarrow$$

$$(x^l)^2 - (x^o)^2 \leq -V_x(q^l)$$

$$\Leftrightarrow$$

$$[(x^l - x^o)(\alpha + \delta) + \alpha\frac{\partial V_y}{\partial x^l}]\frac{1}{-4z_x\alpha} \leq -V_x(q^l)$$

$$\Leftrightarrow$$

$$(x^l - x^o)(\alpha + \delta)\frac{1}{-4z_x\alpha} \leq -V_x(q^l) + \frac{\alpha}{4z_x\alpha}\frac{\partial V_y}{\partial x^l}$$

Third, let $B(q_i) = aq_i^2 + b$, and assume symmetry. Then I can calculate $V_x(q^l)$ and its derivative in the RHS:

$$RHS = -[\frac{3\omega^2}{2a}(\frac{\Pi_l(x^l)}{\Pi_l(x^l) + \alpha})^2 + \frac{b}{3}] + \frac{\alpha}{4z_x\alpha}\frac{3\omega^2}{2a}\frac{\Pi_l(x^l)}{(\Pi_l(x^l) + \alpha)^3}\alpha 4z_x.$$

Notice that the RHS is positive if and only if

$$\Pi_l(x^l) - \frac{\alpha}{\Pi_l(x^l) + \alpha} < 0,$$

which implies $\alpha > \Pi_l(x^l)$, when $0.5 \leq \Pi_l(x^l) \leq 1$. And notice that the LHS is positive and smaller than:

$$-x^o(\alpha + \delta)\frac{1}{-4z_x}.$$

Therefore, given x^o, b, and δ. Assume large α as above, then the RHS is increasing in $\frac{\omega}{a}$. Thus, exists $\bar{m} \equiv m(\alpha, \delta, b; x^o)$:

$$LHS \leq -x^o(\alpha + \delta)\frac{1}{-4z_x} \leq RHS(\bar{m}).$$

■

7.5.3 Comparative statics

Let $B(q_i) = aq_i^2 + b$. Then, for reference,

$$V(q^r) = \frac{3}{2}\frac{\omega^2}{a}(\frac{\Pi_r}{\Pi_r + \alpha})^2 + \frac{b}{3},$$

and

$$q^r = \frac{3}{2}\frac{\omega}{a}\frac{\Pi_r}{\Pi_r + \alpha}.$$

First, I show the proof of (4). **Proof.**

- Let the parties be extreme: $\max_r \Pi_r > 2\alpha > 0.5\alpha$. Then, since $\frac{\partial q}{\partial \alpha} < 0$, and $\frac{\partial q}{\partial x} < 0$, when $\frac{\partial x}{\partial \alpha} > 0$, the total effect $\frac{dq}{d\alpha}$ is unambiguous. Therefore, I look under which conditions this is the case, using the implicit function theorem. Let $R(x, \frac{\omega}{\alpha}) \equiv -\frac{\partial FOC}{\partial x} > 0$. Then, in equation (10), for $r = l$, and equation (12), for $r = c$, I take derivatives with respect to α and I obtain:

$$sign(\frac{x\Pi_r}{\alpha^2} - \frac{\partial^2 V(q^r)}{\partial x \partial \alpha}) = sign(\frac{\partial x}{\partial \alpha}),$$

which is positive if $\frac{\partial^2 V(q^r)}{\partial x \partial \alpha}$ is negative and large. Let $\frac{\omega}{\alpha}$ is large; then, when $\Pi_l > 2\alpha$

$$\frac{dq^l}{d\alpha} = \frac{3\omega}{2a}\frac{1}{(\Pi_l + \alpha)^2}\{-1 + \frac{\alpha 4z_x}{R(x, \frac{\omega}{\alpha})}\Pi_l[\frac{x}{\alpha} + \frac{3\omega^2}{2a}\frac{4\alpha z_x(\alpha - 2\Pi_l)}{(\Pi_l + \alpha)^4}]\} < 0 \tag{23}$$

And, when $\Pi_c > 2\alpha$

$$\frac{dq^c}{d\alpha} = \frac{3\omega}{2a}\frac{1}{(\Pi_c + \alpha)^2}\{-1 + \frac{\alpha 4z_x}{R(x, \frac{\omega}{\alpha})}[\Pi_l\frac{x}{\alpha} + \Pi_c\frac{3\omega^2}{2a}\frac{4\alpha x(\alpha - 2\Pi_c)}{(\Pi_c + \alpha)^4}]\} < 0$$

- Let the parties be moderate: $\min_r \Pi_r < 0.5\alpha < 2\alpha$. Then, $\frac{\partial x^r}{\partial \alpha} < 0$, and the total effect is ambiguous.

■

Second, I show the proof of lemma (7.5.3). The proof of lemma (7.5.3), as mentioned below, resembles the one above closely. lemma(4)

Proof. Here, I prove the lemmas (7.5.3), and (7.5.3).

- **Ego rents,** $r = c$. Notice that in the equation below, $\alpha R(\cdot) = \Pi_c + 4xz_x + \alpha + \frac{\partial^2 V}{\partial x^2}$, where $\frac{\partial^2 V}{\partial x^2} \geq 0$. Thus, $\alpha R(\cdot) > \Pi_c$. Hence,

$$\frac{dq^c}{d\delta} = \frac{3\omega}{2a}\frac{\alpha}{(\Pi_c + \alpha)^2}[1 - \frac{\Pi_c}{\alpha R(x, \frac{\omega}{\alpha})}] > 0$$

Where $R(x, \frac{\omega}{\alpha}) = -\frac{\partial FOC}{\partial x}$, which is increasing in $\frac{\omega}{\alpha}$, and decreasing in x, is always smaller than Π_c.

- **Ego rents,** $r = l$. This proof follows step by step the proof of lemma (4).

■

7.6 Extensions

7.6.1 Partisan screening: Lemma (7), and $x(s)$

Proof. Let $\tilde{B}(q) \equiv \frac{B(q)}{s}$, then $\tilde{V}(q) \equiv \frac{V(q)}{s}$, and so the optimal valence can be written as

$$\tilde{V'}(q) = \omega\frac{\delta + 4n_x^r x}{\alpha} + \delta + 4n_x^r x.$$

Therefore

$$\tilde{V'}(q) = \frac{\tilde{B}'(q)}{3} = \frac{V(q)}{s},$$

and it follows that $q_P^r(s)$ increases with the screening costs. ■

The proof of $x(s) < 0$ increasing monotonically in s follows from equation (7).

Proof. Notice that using the implicit function theorem, as in proposition (1) and assuming linear costs, then

$$-\frac{\partial^2 V^r}{\partial x \partial s} = -\frac{\partial V^r}{\partial x} > 0$$

. ■

7.6.2 Observed ideologies: lemma (8)

To show that the monotonicity result in proposition (1) also holds if the ideologies are observed instead of valence, it is enough to show that: $\frac{\partial c_x}{\partial n_x} < 0$. First, I show how does the affiliation decisions change as I change the previous assumption; second, I show that the result still holds.

If a voter $(\tilde{z}_i, q_i)$ affiliates to party P, then $B(q_i) > (z_i - x_p)^2$. Let c_p be the observed candidates' ideology; and let $|c_X| < |x|$, and the same for Y. Then

$$E(q_i|c_x) = \int_{q_i \geq B^{-1}(z_x - x)^2} q_i dF(q).$$

Assume $B(q_i) = aq_i^2 + b$, and let n_x be the nominator's ideology from party X with ideology, in a symmetric game; then he maximizes

$$\frac{\alpha + c_y^2 - c_x^2 + \omega[E(q_y|c_y) - E(q_x|c_x)]}{2\alpha}(\delta_n + 4c_x n_x)$$

The lemma (8) states that for large enough α there exists a large ω/a such that the equilibrium c_x is decreasing in n_x. Thus, the main result of monotonicity holds: more moderate nominators choose lower valence candidates. Using the implicit function theorem, and for a large enough α, the lemma holds. Let

$$E' \equiv \frac{\partial E(q_x|c_x)}{\partial c_x} = \frac{(c_x - n_x)}{2a\sqrt{\frac{(c_x-n_x)^2-b}{a}}} > 0.$$

Proof. For $c_x > x$, in the F.O.C. the ideology of the candidate is implicitly defined by

$$(\delta_n + (c_Y - n_x)^2 - (c_X - n_x)^2)(\frac{\omega}{2\alpha}E' - \frac{c_x}{\alpha}) - 2(c_X - n_x)P_x = 0.$$

The S.O.C always holds:

$$\frac{\partial FOC}{\partial c_x} = -\frac{b}{2a(\frac{(c_X-n_x)^2-b}{a})^{3/2}} = -\frac{b}{2a(\cdot)^{3/2}} < 0.$$

Then, by the implicit function theorem, $\frac{\partial c_x}{\partial n_x} < 0$ if

$$\frac{2c_x}{\alpha}(\frac{\omega}{2a}\frac{(c_X - n_x)}{(\cdot)^{1/2}} - 2c_x) + 1 < 0,$$

and since $0 > c_x > x$, the inequality above holds for large enough $\frac{\omega}{a}$, given that α is not to small; that is,

$$\alpha > 4c_x^2$$

■

References

Agranov, M. (2009). Flip-flopping: contrasting effects of intense primaries and general elections on selection of candidates.

Aldrich, J. (1983). A downsian spatial model with party activism. *The American Political Science Review 77*(4), 974–990.

Aldrich, J. (1995). Why parties?: The origin and transformation of political parties in america.

Alesina, A. and H. Rosenthal (1995). Partisan politics. *Divided Government, and the Economy (Cambridge University press, Cambridge, UK).*

Alvarez, R., D. Canon, and P. Sellers (1995). The impact of primaries on general election outcomes in the us house and senate. *Working Papers*.

Ansolabehere, S., J. Hansen, S. Hirano, and J. Snyder Jr (2007). The incumbency advantage in us primary elections. *Electoral Studies 26*(3), 660–668.

Aragones, E. and Z. Neeman (2000). Strategic ambiguity in electoral competition. *Journal of Theoretical Politics 12*(2), 183.

Ashworth, S. and E. Bueno de Mesquita (2009). Elections with platform and valence competition. *Games and Economic Behavior 67*(1), 191–216.

Ashworth, S. and E. de Mesquita (2008). Informative party labels with institutional and electoral variation. *Journal of Theoretical Politics 20*(3), 251.

Banks, J. and J. Duggan (2005). Probabilistic voting in the spatial model of elections: The theory of office-motivated candidates. *Social Choice and Strategic Decisions*, 15–56.

Berdahl, C. (1949). Some notes on party membership in congress, i. *The American Political Science Review 43*(2), 309–321.

Callander, S. and S. Wilkie (2007). Lies, damned lies, and political campaigns. *Games and Economic Behavior 60*(2), 262–286.

Carey, J. and J. Polga-Hecimovich (2008). Primary elections and candidate strength in latin america. *The Journal of Politics 68*(03), 530–543.

Crutzen, B., M. Castanheira, and N. Sahuguet (2010). Party organization and electoral competition. *Journal of Law, Economics, and Organization 26*(2), 212.

Diermeier, D., M. Keane, and A. Merlo (2005). A political economy model of congressional careers. *The American Economic Review 95*(1), 347–373.

Downs, A. (1957). An economic theory of political action in a democracy. *The Journal of Political Economy 65*(2), 135–150.

Gerber, E. and R. Morton (1998). Primary election systems and representation. *Journal of Law Economics and Organization 14*(2), 304–324.

Groseclose, T. (2001). A model of candidate location when one candidate has a valence advantage. *American Journal of Political Science 45*(4), 862–886.

Hirano, S., J. Snyder, and M. Ting (2009). Distributive politics with primaries. *The Journal of Politics 71*(04), 1467–1480.

Hummel, P. (2010). On the nature of equilibria in a downsian model with candidate valence. *Games and Economic Behavior*.

Kanthak, K. and R. Morton (2001). The effects of electoral rules on congressional primaries. *Congressional primaries and the politics of representation*, 11631.

Kartik, N. and R. McAfee (2007). Signaling character in electoral competition. *The American Economic Review 97*(3), 852–870.

Mainwaring, S. (1993). Presidentialism, multipartism, and democracy: The difficult combination. *Comparative Political Studies 26*(2), 198–228.

Mattozzi, A. and A. Merlo (2008). Political careers or career politicians? *Journal of Public Economics 92*(3-4), 597–608.

McGhee, E. (2010). Open primaries. *San Francisco: Public Policy Institute of California*.

Meirowitz, A. (2005). Informational party primaries and strategic ambiguity. *Journal of Theoretical Politics 17*(1), 107.

Serra, G. (2010). Why primaries? the partys tradeoff between policy and valence. *Forthcoming, Journal of Theoretical Politics*.

Snyder, J. M. and M. M. Ting (2002). An informational rationale for political parties. *American Journal of Political Science 46*(1), 90–110.

Stokes, D. (1963). Spatial models of party competition. *The American Political Science Review 57*(2), 368–377.