Ideological extremism and primaries.

Agustin Casas[*]

March 2, 2017

Abstract

This paper is the first one to present a "general equilibrium" model of primaries with endogenous party affiliations. I show that closed primaries (where only affiliated party members can vote) result in more charismatic candidates than in open primaries. That occurs because, in equilibrium, closed-primary voters care more about winning and therefore they are more willing to trade off their ideologically preferred candidate for one who is more likely to win, i.e., a more charismatic one. I also show that under open primaries, the party leaders have higher incentives to choose more extreme platforms. As a consequence, and in line with the most recent empirical evidence, open-primary nominees are more likely to be extremists than closed-primary ones. Finally, I show that, if instead of organizing primaries, party leaders were to handpick the nominees, the candidates would be even more moderate and more charismatic.

1 Introduction

A key issue in democracies is how to select political candidates. Pundits and policymakers agree that partisan nomination rules can systematically influence candidates' attributes and that, in open primaries, candidates are ideologically more moderate. [1] Yet, the empirical evidence on the relationship between candidates and nomination rules does not support this reasoning.[2] Moreover, due to its complexity, there are very few models that address this relationship.

In this paper, I investigate this puzzle with a model that incorporates three central pieces of partisan politics: voters' preferences, nomination rules, and endogenous affiliation decisions. The novelty of combining these three elements allows me to investigate the impact of a political party's governance on the selection of candidates in a bipartisan democracy. Unlike previous literature, my model predicts that candidates nominated in closed primaries are more charismatic and less predictable to voters (in terms of the policy they will implement if elected) than those nominated in open primaries. Additionally, candidates nominated in open primaries are expected to be more extremist (in terms of policy) than those nominated in closed ones. Likewise, when party leaders handpick the candidates, they select more moderate and charismatic candidates than those via primaries.

To understand the intuition behind these results, one must first understand the link between the nomination rules and the selected candidates' ideology and charisma. To elucidate this link, following seminal work on political parties (Aldrich, 1995; Berdahl, 1949b,a), I propose a game with three different but intertwined stages: an affiliation stage, a nomination stage and a general election stage.

Citizens differ in two dimensions, their privately observed ideology and their publicly observed charisma, which are independent of each other. Charisma is an observable individual characteristic that appeals to voters, independently of their preferences, and thus helps a candidate to be elected.

---

[1]See http://www.openprimaries.org/ for a political movement that endorses open primaries, with a reasoning similar to this quote by A. Maldonado, Lieutenant Governor in California (April 29, 2010): "If you want to win a close primary on the Republican side, you have to veer hard to the right, and if you want to win a Democrat primary, you veer hard to the left. In the middle, where you have independents and decline-to-states, guess what they have to do in California? They have to ask for permission of a party to participate in a primary election.".

[2]For instance, Mcghee et al. (2013) and Gerber and Morton (1998), whose contribution I discuss in detail below.

Therefore any likable personality trait that has positive effects on electoral results but is orthogonal to policy would fit my working definition of charisma: oratorical abilities, good looks, and other observable skills.[3] I assume that the candidate who wins the general election implements his ideal policy (i.e., his ideology).

At the **affiliation stage**, after observing the political parties' exogenous platforms, each individual optimally decides whether to join a party. The benefits from affiliation determine the set of party members – who also are the potential candidates – which ultimately shapes the attributes of the party's nominee. 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, à la Snyder and Ting (2002): a candidate from the liberal party is more likely to be more liberal than conservative. Also, I assume that the consumption benefit of being a party member increases as the citizen's charisma increases, grounded on the intuition that higher charisma individuals can extract more rents by joining a party.[4]

An implication of these assumptions is that citizens that are ideologically far from a party's platform can affiliate only if their charisma is high enough. As a result, despite the fact that ideology and charisma are independently distributed, a very charismatic affiliated individual is more likely to be far from the party's platform than one with less charisma. Hence, the ideology of more charismatic candidates is less predictable than the less charismatic ones.

The recent elections in the U.S. may be a good example for the point above. Donald Trump emerged from the Republican primaries as an arguably charismatic politician even though seeming to be very far apart from the party's traditional values. On the other extreme, there was Jeb Bush, who seemed to fit very well the party's values but was less charismatic. Even though Trump's

---

[3]Hamermesh (2006), Berggren et al. (2010) and Lenz and Lawson (2011) show that candidate's beauty has a positive effect on electoral results. However, it is worth emphasizing that the modeling assumptions impose a narrower interpretation of valence than the original one by Stokes (1963). An example of some individual characteristics that do not fit the model because they are intertwined with the policy making, but are sometimes thought of valence would be incumbency (Stone and Simas, 2010), character (a la Callander and Wilkie, 2007 or Kartik and McAfee, 2007) or quality (Caillaud and Tirole, 1999, 2016).

[4]With respect to ideology, the underlying reasoning is that the cost of interacting with other party members is increasing in the disagreement, i.e., in the ideological distance. With respect to charisma, within the scope of the paper, Mattozzi and Merlo (2008) and Berggren et al. (2010) provide some related examples that support the assumption. Beyond the scope of politics, it has been shown that some observables -like beauty- have positive effects on labor market outcomes, independently of whether they are productivity-enhancing, e.g., Mobius and Rosenblat (1966).

implemented policies were arguably unpredictable, as the nominee of the Republican Party he was elected the 45th President of the U.S.[5]

At the **nomination stage**, the institutional comparative statics exercise focuses on three stylized nomination rules, but I allow for a continuum of institutional arrangements as well. Since the median voter theorem holds, a nomination rule defines who selects the nominee; I refer to that citizen as the decisive voter or the "nominator". This decisive voter is the median primary voter: in closed primaries, he is the median party member, while, for instance, in more open primaries, the nominator would be more moderate, closer to the population-wide median voter. 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 novelty of this approach is that the candidates' equilibrium charisma depends on the nomination rule, instead of suggesting that charisma is exogenous (Groseclose, 2001) or a partisan investment decision (Ashworth and de Mesquita, 2009).

For all nomination rules, the risk-averse decisive voter trades off predictability and electability, i.e., the probability of winning. Although the decisive voter wants to choose a charismatic candidate to increase the probability that his party wins the general election, a more charismatic candidate comes at the cost of predictability. The solution to this trade-off depends on the nomination rule: the less open the primary is, the more the decisive voter cares about electability. Hence, the closed-primary decisive voter has larger incentives to nominate a more charismatic candidate.

I model the **electoral stage** as a probabilistic voting model, where voters' behavior is affected by unobserved parameters that we model as a random shock. The mean of this shock is determined by the candidates' charisma: if the left wing candidate is more charismatic, then the shock is more likely to benefit him than the right wing candidate. The winner of the general election will implement his own preferred policy or ideology, a la Osborne and Slivinski (1996) and Besley and Coate (1997).

The main result hence follows: intuitively, a hardcore left wing primary voter would really dislike the right-leaning party to win, thus he would rather nominate a more electable candidate,

---

[5]Please note that this is just a salient example, and that the aim of the paper is, by no means, to explain or fit the 2016 US elections.

even at the cost of high policy uncertainty. Hence, more open primaries result in less charismatic candidates, as the open-primary median voter is relatively more indifferent between parties, in terms of policy.

When I endogenize the parties' platforms, an additional result follows: open primaries lead to candidates with more extreme expected ideologies than closed ones. Before the affiliation stage, I let two party leaders choose the location of the parties such that they maximize their own expected utility. Following from the first result, when the party leaders anticipate open primaries, they expect low-charisma candidates, hence they choose relatively extreme platforms, closer to their own preferred policy, to countervail the loss in expected utility. On the contrary, when the party leader expects closed primaries, he anticipates that the closed-primary median voter will choose a charismatic candidate, hence he can afford a larger policy compromise and choose a more moderate platform. As a result, in line with the empirical literature, candidates nominated in open primaries are (low-charisma) "predictable extremists"; on the contrary, candidates nominated in closed primaries are (high-charisma) "moderate mavericks".

The previous literature on the study of nomination rules in democracies is not abundant. While some scholars argue that primaries have a systematic influence on political equilibria (Alesina and Rosenthal, 1995, Gerber and Morton, 1998, Alvarez et al., 1995, Ansolabehere et al., 2007), others study primaries as a possible equilibrium outcome (Crutzen et al., 2010; Serra, 2011; Aragon, 2013; Hirano et al., 2009). Also, Agranov (2015) and Hummel (2010) argue that holding primaries, which induce intense competition between co-partisans and flip-flopping (change of positions between primaries and elections), may be harmful in the general election.

The empirical evidence is inconclusive with respect to differences between closed and open primaries. However, some authors have shown that primaries' openness may lead to more extreme candidates than closed ones (for instance, Gerber and Morton, 1998).[6] The most recent evidence, Mcghee et al. (2013), shows that "...the openness of a primary election system has little to no effect

---

[6]It is worth clarifying that Gerber and Morton (1998) conclude that in closed primaries the median voter is more extreme than in more open primaries, and they claim that the nominated candidates should also be more moderate in the more open primaries. Even though they conclude that closed primaries lead to more extreme candidates than semi-closed ones, Tables 3 and 4 in their paper unequivocally show that open primaries result in more extreme (winning) candidates than (i) semi-closed primaries, for any slicing of the data, and (ii) closed primaries, when the focus is on the Republican party.

on the ideological positions of the politicians it elects. **Our estimates of effects are rarely robust, and when they are, they are generally the opposite of the ones that are expected – more open primaries electing legislators who are more extreme.**" (page 2, bold is mine). Moreover, despite the public debate and this empirical evidence on partisan nomination methods playing a role, few models study their effects on the political equilibria thoroughly (Jackson et al., 2007; Cho and Kang, 2014; Amorós et al., 2016; Grofman et al., 2016). Furthermore, there is no formal model that explains simultaneously why open primaries' candidates may be more extreme on average, and why holding primaries can harm the electoral chances in the general election.

The work of Jackson et al. (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, they find 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. The work of Snyder and Ting (2011) is also similar as they build a model with primaries and valence, which also provides an argument for why ideological extremism is associated with more democratic nomination methods. However, since they do not analyze different types of primaries, they do not explicitly incorporate an affiliation stage, which is specially relevant to distinguish closed primaries from open ones.

The paper is organized as follows: in Section 2, I introduce the bare bone of the model and I thoroughly describe the players' roles at each stage of the game. In Section 3, I describe the equilibrium. In Section 4, I discuss the institutional comparative statics with exogenous and endogenous platforms. Furthermore, I show that under different specifications for the affiliation stage, equilibrium charisma is always lower in (pure) open primaries. In Section 5, I extend the game to introduce an informal "nomination rule": handpicking by party leaders, among other extensions. Lastly, Section 6 concludes.

2 The model

There are two parties $p \in \{L, R\}$ with platforms $\pi_p \in \{l, r\}$ in the policy space $X = [-\bar{x}, \bar{x}]$. There is a continuum of citizens, who are characterized by their ideology $x_i \in X$, which is private information, and their *charisma* or valence $c_i \in C = [0, \bar{c}]$, publicly observed. Ideology and charisma are independently drawn from distributions with full support on X and C, respectively. For tractability, ideology is assumed to be uniformly distributed and the population-wide median voter's ideology is 0.[7]

Any citizen in this economy is fully described by the pair (x_i, c_i) and their party affiliation $a_i \in \{\emptyset, L, R\}$. Let $\delta_i \equiv \delta(x_i, \pi_p)$ be the distance between the ideology of a citizen i and the party's platform.[8] Let x be the ideology or implemented policy by the winner of the general election. Then, the utility of a citizen can be written in terms of (a_i, x_i, c_i, x) as follows:

$$U_i(a_i; x_i, c_i, x) = B(\delta_i, c_i | a_i) - (x_i - x)^2 \tag{1}$$

The **first term**, $B(\delta_i, c_i | a_i)$, captures the benefits of affiliation, which depend on the citizen's ideological position relative to the party's platform, δ_i, and on the charisma of the individual. Regarding these benefits of being an (active) party member, I assume the following:

Assumption 1 *The benefits of affiliation are (i) decreasing in the citizens' ideological distance to the party's platform, $B_\delta \leq 0$, (ii) increasing in charisma $B_c \geq 0$, and (iii) additive separable in charisma and ideological distance, so $B_{\delta c} = B_{c\delta} = 0$. Last, (iv) citizens without charisma have negative affiliation benefits, $B(\delta_i, 0) < 0$, and the opportunity cost of affiliation is zero, $B(\delta_i, c_i | a_i = \emptyset) = 0$*

Hence all citizens i will affiliate to party p if and only if they prefer party p to p' and to being unaffiliated. Therefore, the set of party members of p, A_p, is also the set of pre-candidates at the

---

[7]The current setup is a low information environment regarding ideological positions. Consistently with the current empirical evidence (Snyder and Ting, 2002 and references therein), voters cannot distinguish between conservative and liberal candidates **within** a party, but they can use party labels to distinguish between candidates position overall. Thus, in the main model, I make the assumption that candidates' ideology is private information. In the extensions I relax this assumption and I obtain similar results.

[8]$\delta(x_i, \pi_p)$ is a distance function in $\mathbb{R}$, symmetric around π_p.

nomination stage, which is the following:

$$A_p = \{i : B(\delta_i, c_i | a_i = p) \geq B(\delta_i, c_i | a_i = p') \;\; and \;\; B(\delta_i, c_i | p) \geq B(\delta_i, c_i | a_i = \emptyset) = 0\}. \tag{2}$$

The **second term** in Equation 1, $-(x_i - x)^2$, corresponds to the policy payoff, which I call u_i. Let x, c and a indicate the ideology, charisma and affiliation of the winning candidate. Voters are risk averse, so they care about the expected quadratic distance to the winning candidates' ideology. This expected policy payoff, $u_i^e(c, a)$, is

$$E(u_i | c, a) = u_i^e(c, a) = -E(x_i - x | c, a)^2 = -(x_i - E(x | c, a))^2 - V(x | c, a). \tag{3}$$

Equation 3 indicates that the voters care about the expected implemented policy (or ideology), $E(x | c, a)$, and they also care about the variance, $V(x | c, a)$, which I call the "ideological unpredictability" of the candidate or policy.

I consider a range of different **nomination rules**, with a special focus on (pure) closed and (pure) open primaries, which I explain in detail in Subsection 3.2. Among all the affiliated members, only one is chosen as the party nominee following nomination rule n, hence a nomination rule is a function $n : A_p \to A_p$. That is, a nomination rule determines who is the decisive voter in the primary. I call $d_{p,n}$ the decisive voter of party p under nomination rule n. Finally, the nominee from party p is then described by his (unobserved) ideology and charisma: (x_p, c_p).

Let $P_L(x_L, c_L, x_R, c_R)$ be the probability that the candidate from party L wins, hence we can define $EU_i(x_L, c_L, x_R, c_R)$, i's expected utility, as follows:

$$P_L(x_L, c_L, x_R, c_R) E(u_i | c_L, a_L) + (1 - P_L(x_L, c_L, x_R, c_R)) E(u_i | c_R, a_R) + B(\delta_i, c_i | a_i). \tag{4}$$

Furthermore, when parties overlap or hit the bounds of the support, the ideological variance depends not only on the equilibrium affiliations but also on the platforms. Hence, in order to avoid

"artificial" decreases of the variance I impose the following constraints on l and r.

Assumption 2 *Let $(l, r) \in X^2$ satisfy that*

   i. $B(0 - \pi_p, \bar{c}) < 0$,

  ii. $\exists (i, j) : x_i < l, x_j > r$ *and* $(a_i, a_j) = (\emptyset, \emptyset)$.

Condition (i) above implies that parties are sufficiently extreme (far from 0) such that the population-wide median voter does not want to affiliate to any party even if he has the largest possible charisma. And Condition (ii) implies that parties are sufficiently moderate such that there are some extreme voters who do not want to affiliate to any party.[9]

The **timing** of the game is as follows: at $t = 1$ citizens observe the location of the political parties and decide whether to affiliate to either party or none. Then, at $t = 2$, among the pool of affiliated citizens A_p, each party has to choose a nominee following the nomination rule n (explained in detail in next subsection). The chosen candidates will compete in a general election against each other. Lastly, at $t = 3$, the political campaigns and the general election take place, and afterwards, the winning candidate implements his preferred policy (a la Osborne and Slivinski (1996), Besley and Coate (1997)).

3 Equilibrium Analysis

Equation 1 shows that the affiliation benefits are independent of the winner of the election and the implemented policy, which allows for studying the affiliation and voting decisions separately. Hence, for simplicity, I summarize first the main characteristics of the affiliation stage, before solving the full game by backward induction.

All citizens can choose whether to affiliate to either party or none. The affiliations result in two sets of party members A_L and A_R, defined in Equation 2. Besides the parties' platforms,

---

[9]Notice that, given $B(\delta_i, c_i)$, Assumption 2 is a condition on the platforms. However, keeping the platforms fix, it could be reinterpreted as a condition on $B(\delta_i, c_i)$, for instance, a very negative B_δ.

voters only observe the candidates' affiliation and charisma. Hence, citizens update their beliefs on the candidates' ideology through party labels and charisma. Under Assumptions (1) and (2), and taking into account that ideology is distributed uniformly, all citizens drawn from the pool of members, A_p, have the following expected ideology

$$E(x_i|c_i, i \in A_p) = E(x_i|i \in A_p) = \pi_p, \tag{5}$$

and the following ideological variance or unpredictability

$$V(x_i|c_i, i \in A_p) = V(c_i). \tag{6}$$

Thus, affiliation decisions are critical to generate an **endogenous cost of nominating more charismatic candidates**: more charismatic members have the same expected ideology but they are less predictable. During the nomination stage, charisma "produces" mean-preserving transformations of the candidate's perceived ideology. This endogenous cost of charisma is driven by the opposite effect of δ_i and c_i on the benefits of affiliation. Charisma and ideological "closeness" to the party's platform are (weak) complements: a citizen more distant to the party's platform must also be very charismatic to become a member. Hence, very charismatic party members are more likely to be far from the party's platform than non-charismatic members, i.e., $\frac{\partial V(c_i)}{\partial c_i} > 0$.

As a result, voters' preferences over expected policies (Equation 3) can be re-written in terms of induced preferences over the candidates' charisma and affiliation, summarized in c_p. For simplicity, since these induced preferences resemble the expected policy payoff, we write them as $u_i^e(c_p)$. Thus, we can re-write Equation 4, the expected utility, as follows:

$$EU_i(c_L, c_R) = P_L(c_L, c_R)u_i^e(c_L) + (1 - P_L(c_L, c_R))u_i^e(c_R) + B(\delta_i, c_i|a_i) \tag{7}$$

Therefore, for every pair of symmetric platforms (l, r), parties (A_L, A_R) and nomination rule n, in a Nash equilibrium it must be the case that voters in the primary election choose the candidate

who maximizes their expected utility.

Definition 1 *Let $l = -r < 0$, $(c_L^*, c_R^*) \in C^2$ is an electoral equilibrium if for all $y \in C$ there is a majority $\mathcal{M}$ of voters in the primary such that*

$$EU_i(c_p^*, c_{p'}^*) \geq EU_i(y, c_{p'}^*),$$

for all $i \in \mathcal{M}$.

That is, given the individual behavior in the general election, we can obtain the probability of winning for any pair of candidates, $P_L(c_L, c_R)$, and use that information to calculate the expected utility, and obtain the Condorcet winners in each primary, (c_L^*, c_R^*).

In the following two sections I solve the game by backward induction. First, I describe the behavior of voters during the general election $(t = 3)$. Second, I explain in detail the different types of primaries and how they affect the incentives of the median voter in a primary $(t = 2)$. Then, I characterize the equilibria and explain the comparative statics exercises.

3.1   General election $(t = 3)$

The last stage of the game is the general election. Once they are nominated, let the candidates from each party engage in persuasive campaigns that may not necessarily affect the voters' utility but influence their behavior.[10] Since the effect of these attributes on the voters' behavior cannot be fully predicted by external observers, I model it as a random shock, as in the "probabilistic voting" literature. Given the two candidates, with charisma c_L and c_R, citizens will vote for the one who delivers them higher expected utility, relative to an unobserved "national shock".[11] Hence a voter i votes for party L if

$$u_i^e(c_L) - u_i^e(c_R) > \tilde{\alpha}(c_L, c_R) \tag{8}$$

---

[10]For instance, charisma affects the votes that a candidate gets but is unrelated to policy (see Berggren et al. (2010), Hamermesh (2006), Lenz and Lawson (2011) and Lawson et al. (2010) for some examples and empirical evidence).

[11]In the Appendix, I show that results do not change when we account for an individual "shock" as well.

Where the mean of the "shock" is determined by the relative charismatic advantage,

$$\tilde{\alpha} \sim U\left[-\alpha + \omega \times (c_R - c_L); \alpha + \omega \times (c_R - c_L)\right], \tag{9}$$

a la Banks and Duggan (2005); and ω is a parameter that captures the marginal effect of a charismatic advantage on voters' behavior. Hence, from Equations 3 and 8, I obtain the probability that i votes for L:

$$\text{Pr(i votes for L)} = \text{Pr}\left(u_i^e(c_L) - u_i^e(c_R) > \tilde{\alpha}(c_L, c_R)\right) \equiv P^i.$$

Therefore, using Equations 5 and 6, given the platforms and the candidates' charisma, the probability that party L wins the election is obtained by integrating P^i over the distribution of ideologies $\left(\int P^i dF(x)\right)$:

$$P_L = \frac{1}{2} + \frac{r^2 - l^2 + V(c_R) - V(c_L) + \omega \times (c_L - c_R)}{2\alpha}. \tag{10}$$

Some remarks are in order: first, Equation 10, is the probability that a candidate (the one from party L in this case) wins the general election, i.e., the **electability** of a given candidate.[12] Since primary voters care about the policy eventually implemented by the winner, they care about the ideology of their party nominee, but also about his electability. Second, notice that the affiliation benefits do not enter the probability of winning, which depends positively on the candidate's relative expected ideological moderation, but negatively on the uncertainty of the implemented policy. Ceteris paribus, more centrist candidates are more "electable", less predictable candidates (i.e. mavericks) are less electable, and more charismatic ones are more electable (through its direct effect ω).

3.2 Nomination rules $(t = 2)$

Each party holds a primary election where there is a decisive voter whose location depends on the nomination rule n. Let $d_{p,n}$ be the ideology of party p's decisive voter under nomination rule n. That is, $d_{p,n}$ is the individual who nominates the party's candidate to compete in the

---

[12]The full derivation of the equation can be found in the Appendix.

general election.[13,14] From now on, I refer to moderate or extreme ideologies taking into account the distance to the population-wide median voter, located at zero.

For instance, in pure closed primaries ($n = closed$), only affiliated party members vote, hence the decisive voter is the median party member $d_{p,closed} = med(A_p)$. In open primaries, for all possible variations of the rule discussed in the literature (pure open, semi-open, semi-closed) the decisive voter is more moderate than the closed primaries' median voter, $|d_{p,n'}| < |med(A_p)|$, for n' being some open primary.[15]

In either institutional setting, the decisive voter $d_{p,n}$ nominates the candidate who maximizes his expected utility but since ideology is unobservable, the candidates can be fully described by their charisma and party affiliation (c_p). Thus, the decisive voter chooses the charisma of the party's nominee:

$$c_p^* \in argmax_{c_i \in C, \forall i \in A_p} EU_{d_{p,n}}(c_p). \tag{11}$$

Since the affiliation benefits do not depend on the political outcome, the expected utility can be expressed in terms of the probabilities of winning and the expected policy payoff enjoyed by the decisive voter. Let Π_i be i's relative policy gain when the candidate from party L, instead of R, wins the general election.[16] The decisive voters' maximization problem can be re-written in a way that highlights the trade-off between electability and expected policy:

$$max_{c_L} P_L(c_L, c_R) \Pi_{d_{L,n}}(c_L, c_R), \tag{12}$$

Notice that in any symmetric equilibrium, the decisive voter would always prefer his own party to the rival one. The decisive voter's idelogy is to the left of the population-wide median voter

---

[13]In order to avoid the proliferation of variables, I refer to $d_{p,n}$ as both the decisive voter and the decisive voters's ideology.

[14]In the appendix I show that the median voter theorem holds quite generally.

[15]More specifically, $d_{L,open} < med(A_L)$ and $d_{R,open} > med(A_R)$.

[16]Remember that the decisive voter's expected policy payoff is

$$u_{d_{p,n}}^e(c_p) = EU_{d_{p,n}}(c_p).$$

Hence, the expected policy gain is:

$$\Pi_{d_{L,n}} \equiv \left[u_{d_{p,n}}^e(c_L) - u_{d_{p,n}}^e(c_R) \right].$$

(included), then it holds that $\Pi_{d_{L,n}}(c_L, c_R) \geq 0$, for all (c_L, c_R).

Figure 1 summarizes the differences between the nomination rules above, and incorporates the timing of the model. In panel (a), at $t = 1$, voters observe the parties' platforms ($l(n)$ and $r(n)$) and decide whether they affiliate or not. To exemplify the point, the thicker line indicates affiliated members with a maximum given level of charisma, say $\bar{c}$. At $t = 2$, out of those affiliated members, in closed primaries, the median party member also becomes the median voter in closed primaries, or the nominator (indicated with an arrow in panel (b)). However, in more open primaries, the nominator or decisive voter is more moderate than the closed-primary median voter, as shown by the arrows in panel (c).
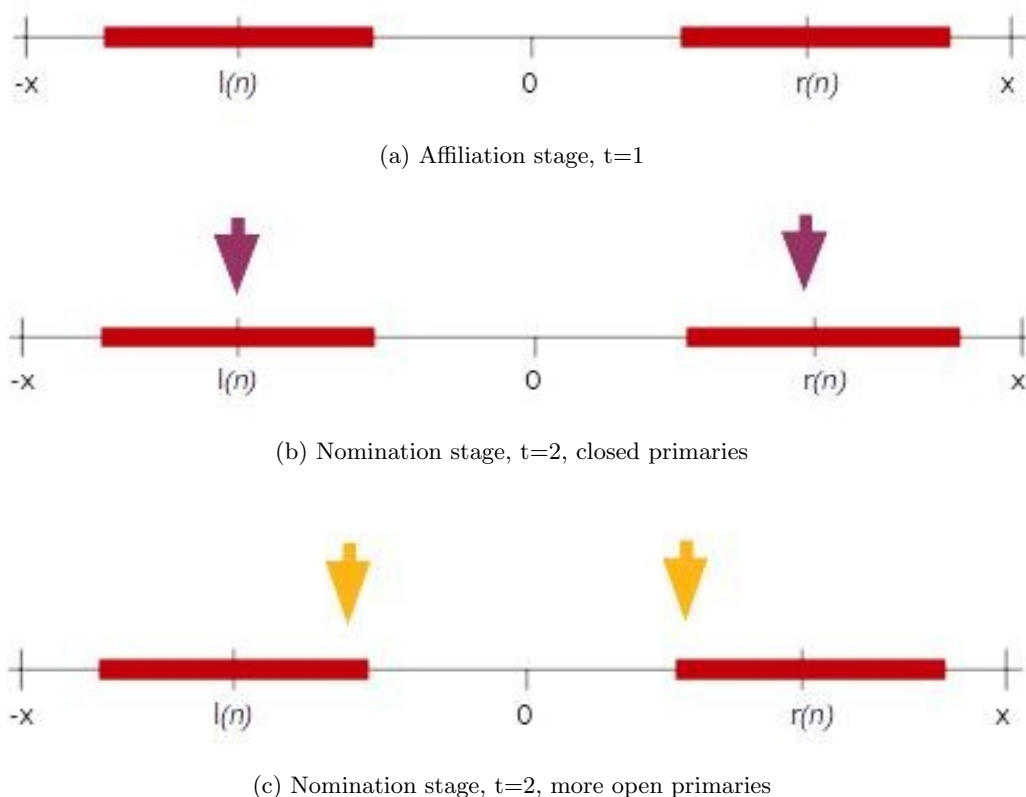

(a) Affiliation stage, t=1

(b) Nomination stage, t=2, closed primaries

(c) Nomination stage, t=2, more open primaries

Figure 1: Party formation and nomination stages (for a given level of charisma)

14

3.2.1 The decisive voter

Assumption 1 determines the nominator's trade-off: charisma affects the probability of winning positively through ω and negatively through $V(x_i|c_i, i \in A_p)$ but decreases his policy payoff.

Assumption 2 allows for sufficiently moderate parties that nonetheless are not too close, which limits the presence of cross-over voting (voters from the opposite party voting for the least electable candidate) and provides a cleaner interpretation of the results. That is, the median voter theorem holds, and therefore the decisive voter in a primary – the nominator – is the median primary voter.[17]

Lemma 1 (Median voter theorem) *Let $C = [0, \bar{c}]$ such that $\bar{c} = argmaxP_p < \infty$ and $\frac{\partial^2 V(x_i|c)}{\partial c_p^2} \geq 0$, hence under assumptions 1 and 2 the median voter theorem holds.*

Since all potential candidates' expected ideology is identical, the voters' expected utility only depends on the candidates' charisma. Hence, the problem becomes unidimensional, and the proof of Lemma 1 (in the Appendix) is not cumbersome. The two conditions stated in the lemma ensure that the preferences are single-peaked. The restriction on the choice set of charisma ensures that charisma monotonically increases the probability of winning: $V'(\bar{c}) = \omega$ and $\omega - V'(c) \geq 0$ for all $c \in C$. Even if there were voters who wanted to harm the electoral prospects of the party, this restriction would never bind since none of the voters would like to choose a level of charisma above $\bar{c}$.[18] The condition of the variance holds when the affiliation benefits increase at a sufficiently fast rate with charisma. Assumption 1 implies that the affiliation benefits are additive separable, hence, as an example consider $B(\delta_i, c_i) = c_i^\alpha - |x_i - x|^\beta$ with $(\alpha, \beta) \in \mathcal{R}_+^2$. In this case, the conditions are met for a large range of functions: for all $0 < \beta < 2\alpha$, the variance is convex in c_i.[19]

---

[17]Moreover, even if I allowed for cross-over, median-voter-like results would hold, as discussed below and shown in the appendix.

[18]If those voters exist, it would be optimal for them to choose a lower level of charisma and, therefore, lower levels of uncertainty (in case that party actually wins).

[19]More generally, let $B(\delta_i, c_i) = b(c_i) - |x_i - x|^\beta$. The variance is convex in c_i if and only if

$$(\frac{2}{\beta} - 1)(b_c)^2 + b_{cc}b(c) > 0,$$

which always holds for $0 < \beta < 2$ and $b_{cc} > 0$, or for $0 < \beta < 2$ and $b_{cc} < 0$ with b_{cc} large (i.e., close to 0).

4 Results

4.1 Equilibrium with exogenous platforms.

In this section, I investigate the incentives of voters in primaries across different institutional arrangements. This setup serves as the foundation for the next section, in which I show that more open primaries may result in more extreme candidates.

The main institutional comparative statics exercise consists in understanding and explaining the effect of the rules under which primaries are organized. At a hypothetical level, we could think of a continuum of institutions between pure closed primaries and pure open primaries, described only by the location of the decisive voter. In what follows, we say that a nomination rule n is more open than n' if the decisive voter in n is more moderate (closer to the population-wide median voter) than the decisive voter in n'.

In this model, all candidates at a primary have the same expected ideology (Equation 5). So, from Equation 12, the decisive voter or nominator will choose the charisma that optimally trades off his desire to lower the candidate's unpredictability and increase his electability. As shown in the following proposition, the more extreme the nominator is, the more he is willing to choose a more electable candidate at the expense of predictability.

Proposition 1 (Institutional comparative statics) *When the median voter theorem holds, the more open a primary is, the less charismatic the resulting candidate.*

Proposition 1 (proof below) provides an answer to the main question of this subsection: how do different rules solve the trade-off between charisma and policy certainty. The driving force behind the result is that the relative policy payoff, $\Pi_{d_{p,n}}$, increases as the median voter is more extreme because the decisive voter has more to lose in terms of expected implemented policy. Intuitively, a hardcore left-wing voter would really dislike the right-leaning party to win, thus he would rather nominate a more electable candidate, even at the cost of more policy uncertainty. Hence, more open primaries result in less charismatic candidates, as the primary's median voter is relatively more indifferent between parties, in terms of policy. [20]

---

[20]Both in the proof and the intuition, Assumption 2 plays an important simplifying role, but it is not a necessary

This intuition can be more easily understood in the following extreme case: in pure open primaries the population-wide median voter is indifferent between candidates in terms of the expected implemented policy, hence he will always choose the candidate with the lowest charisma possible. While he does not gain anything for making one candidate more likely to win, he would have to pay a cost for choosing a higher-charisma candidate: policy uncertainty.

The corollary below emerges from the intuitive explanation of Proposition 1: any increase in party polarization (i.e. larger $|r - l|$) hurts the median voter in any primary, but at the same time provides him with larger incentives to choose a more charismatic candidate. To fix ideas, it is useful to look at the equilibrium charisma, implicitly defined as follows:[21]

$$V'(c_L^*) = \frac{\partial V(c_L^*)}{\partial c_L} = \frac{\omega \Pi_{d_L}}{\alpha + \Pi_{d_L}}. \tag{13}$$

When parties are more polarized, the decisive voter's relative policy payoff Π_{d_L} increases (except in the pure open primary). Thus, his "willingness" to invest in the party's candidate also increases, hence he nominates a more charismatic and more electable individual.

Corollary 1 *In all symmetric equilibria, more party polarization leads to more charismatic candidates (and more policy uncertainty).*

As a consequence, more party polarization causes higher policy uncertainty due to the nomination of candidates with more charisma. This result is different to the standard results in valence models, where they find that more polarization comes with lower valence (i.e. Groseclose (2001), Ashworth and de Mesquita (2009)). There are two reasons for these contrasting findings: as I mentioned above, the interpretation of valence and charisma are different, since charisma has an endogenous cost suffered by all citizens; and, the cost of nominating a more charismatic candidate is borne by all voters, as they all equally dislike policy uncertainty.

Proof of Proposition 1 and Corollary 1.

In the proof of Lemma 1 in the Appendix, it has been shown that the objective function is

---

one.

[21]That is, the solution to Equation 11, derived in the appendix as Equation 20.

strictly concave, and therefore there is an unique interior equilibrium. Hence, I need to show that $\frac{\partial c_p^*}{\partial x_{d_L}}$ has the right sign; i.e., negative for $p = L$ and positive for $p = R$. Proving it for one of the two cases is enough, so for consistency, I show it for $p = L$. Using the implicit function theorem, it is enough to show that the cross derivative of the objective function is negative. That is, the partial derivative of the FOC, Equation 20 (in the proof of Lemma 1 in the Appendix), with respect to x_{d_L} must be negative.

$$sign(\frac{\partial c_L^*}{\partial x_d}) = sign(\frac{FOC}{\partial x_d})$$

$$= sign\left(\frac{\omega - \frac{\partial V(c_L)}{\partial c_L}}{2\alpha} \frac{\partial \Pi_i(c_L, c_R)}{\partial x_{d_L}}\right)$$

$$= sign\left(\frac{\omega - \frac{\partial V(c_L)}{\partial c_L}}{2\alpha} 2(l - r)\right).$$

In the equation, $\omega - \frac{\partial V(c_L)}{\partial c_L} > 0$ follows from $c < \bar{c}$, and $l - r < 0$ from the assumption that $l < 0 < r$. Hence $\frac{\partial c_L^*}{\partial x_{d_L}} < 0$.

Following the same reasoning, the sign of the derivative of Equation 20 w.r.to the parties' platforms will determine the sign of $\frac{\partial c_p^*}{\partial \pi_p}$. First notice that the corollary holds in the symmetric case where $r = -l$. Hence,

$$sign(\frac{\partial c_L^*}{\partial l})|_{r=-l} = sign(\frac{FOC}{\partial l})|_{r=-l}$$

$$= sign\left(\frac{\omega - \frac{\partial V(c_L)}{\partial c_L}}{2\alpha} \frac{\partial(4x_{d_L}l)}{\partial l} - V'\frac{\partial P_L}{\partial l}\right)$$

$$= sign\left(\frac{\omega - \frac{\partial V(c_L)}{\partial c_L}}{\alpha} 2x_{d_L} - 0\right).$$

And since the decisive voter in a left primary has $x_{d_L} \leq 0$, hence $\frac{\partial c_L^*}{\partial l}|_{r=-l}$ is negative. Therefore, charisma and policy uncertainty increase with polarization. ∎

In sum, acknowledging that the median voter in a closed primary is more likely to have less moderate preferences than a median voter in an open primary, results in interesting comparative statics: closed primaries are more likely to nominate charismatic mavericks, while open primaries are more likely to nominate predictable candidates. Hence, since political platforms are fixed in the short run, a sudden change in the nomination rules would be followed by a change in the equilibrium charisma and the policy uncertainty of their nominees. It would be interesting to investigate, how

parties' location (i.e. the platforms) changes with changes in the nomination rules, so in the next subsection I introduce endogenous platforms in my model.

4.2 Equilibrium with endogenous platforms.

The affiliation stage defines not only the set of party members, but also the potential candidates from which the primaries' voters are going to choose the nominee. Thus, the parties' platforms, which precede and shape the affiliation stage, play a fundamental role. More importantly, as it is usually the case in reality, at the time the platforms are chosen, party leaders already know the nomination rules: in general, these rules are not decided by the political parties but are superseded by other regulations. For instance, in the United States, the primaries for state legislatures are subject to different rules depending on the state and/or the National Party regulation (see Mcghee et al. (2013), Shor and McCarty (2011) and Serra (2010, 2011)). In Argentina, the electoral law regulates the primaries, while France and Spain are two notable exceptions, for which the nomination rules depend entirely on the parties' internal structure.

I model two political entrepreneurs or party leaders with symmetric ideologies, $z_r = -z_l$, who choose the location of the parties $l(n)$ and $r(n)$ taking into account the nomination rule n and anticipating the effects of their choice on the development of the game. These party leaders have full information except for the individual ideologies of the voters. Specifically, the party leaders choose the party platforms taking into account the primary voters response at the nomination stage, specially the nominator's response. Also, they anticipate its effects on the probabilities of winning in the general election stage. Without loss of generality, party L's leader chooses $l^*(n) \in X$ such that it maximizes his expected utility

$$EU_{z_L}(l(n), r(n)) = P_L u^e_{z_L}(c_L) + (1 - P_L)u^e_{z_L}(c_R), \qquad (14)$$

where P_L, c_L and c_R are also functions of the platforms $l(n), r(n)$.

Party leaders' trade-off can be explained through a direct and an indirect channel: a moderate platform (relative to the leader's ideal point) decreases the leader's utility $(u^e_{z_L})$ but increases the probability of his party winning (P_L). I refer to this effect as the "platform effect" or the direct

channel. In pure open primaries, this is the only effect of platforms on the party leader's expected utility, and it determines the location of the party.

However, in closed primaries there is also an "indirect effect": platforms affect the location of the median voter, who chooses the equilibrium charisma. Hence, from Proposition 1, more moderate platforms imply that a more moderate decisive voter nominates a less charismatic candidate. Therefore, the leader's utility ($u^e_{z_L}$) increases due to greater predictability, but the probability of winning (P_L) decreases. I call this effect the "charisma effect" or the indirect channel.

While the location of the party in pure open primaries is solely determined by the direct channel, in closed primaries the indirect channel also plays a role. The platforms would be more moderate due to the latter channel only if the charisma effect is positive, that is, if there is a net gain of choosing more moderate platforms.

It turns out that, due to the convexity of the variance ($\frac{\partial^2 V}{\partial c^2} > 0$), the probability of winning increases at a decreasing rate with charisma. Thus, the positive effects on the party leaders utility ($u^e_{z_L}$) always dominate the negative effects on the probability of winning (P_L). Therefore, under closed primaries the party leader chooses more moderate platforms.

In other words, the marginal effect on the probability of winning ($\frac{\omega - V'}{2\alpha}$) is always smaller than the effect on the utility (V') when the ideological variance increases rapidly with charisma (for large V').

The variance may increase with charisma through two different paths: the benefits of affiliation and the nominator's best response to changes in platforms. First, when the benefits of affiliation increase with charisma (as it follows from Assumption 1), charismatic citizens who are far from the party are more attracted to it. Hence, the larger the returns to charisma (B_c), the noisier the signal of charisma (i.e., larger variance for the same level of charisma). Second, a question remains: how much would the decisive voter in a primary adjust his choice of charisma as a response to a change in platforms.[22] These two mechanisms are summarized in $\frac{\partial V(c^n_L)}{\partial l(n)} = \frac{\partial V(c^n_L)}{\partial c^n_L} \frac{\partial c^n_L}{\partial l(n)}$, which I call the median voters' sensitivity to platform changes. Formally,

---

[22]Remember that there is no adjustment in pure open primaries.

Proposition 2 *If the closed primary's decisive voter is relatively sensitive, i.e.,*

$$0 \geq \frac{\partial V(c_L^{open})}{\partial l(open)} > \frac{\partial V(c_L^{closed})}{\partial l(closed)},$$

then the parties' platforms are more extreme in relatively open primaries (with respect to the closed ones).

When the closed-primary median voters respond "too much" to platforms (i.e. when they increase charisma extensively due to more polarization), the party leaders choose moderate platforms to avoid the nomination of extreme mavericks,i.e., ideologically extreme and unpredictable nominees. On the same lines, when the open-primary decisive voter is relatively insensitive to the platforms, the party leaders can afford choosing extreme platforms without running the risk of policy uncertainty, in the symmetric equilibrium. As shown in the proof of Proposition 2, the exact choice of the location of the platforms perfectly balances these incentives: if platforms were more extreme than the equilibrium, the decisive voters would be reducing the expected utility of the party leaders by nominating candidates who are too unpredictable. Conversely, if platforms were more moderate, they would be nominating unappealing candidates (low-charisma).[23]

As an illustration, in the next corollary I focus on the extreme case of pure-open primaries, in which the equilibrium charisma is unaffected by the choice of the platforms and therefore the variance does not change with the platforms either. Hence, the charisma or indirect channel is shut down.

Corollary 2 *Pure open primaries lead to the most extreme candidates.*

From Proposition 1 we know that the population-wide median voter, the decisive voter in a pure open primary, has no incentives to choose a charismatic candidate because he is indifferent in terms of policy between the parties ($x_d = 0$). In terms of Proposition 2, he is not sensitive to

---

[23]A complicating but realistic feature of my modeling assumptions is that all stages are intertwined, in a systematic way. For instance, the extent to which a decisive voter is sensitive to platforms depends on the affiliation benefits, $B(\delta_i, c_i)$, and the effect of charisma on electability, ω. Intuitively, if the payoff to charisma is very large at the affiliation stage, hence party members can be more distant to the platforms, and therefore policy uncertainty increases very fast with charisma. In that case, for a given platform and ideal point, a voter is more likely be "sensitive".

changes in platforms, $\frac{\partial V(c_L)}{\partial l} = 0$. Hence, he would choose the lowest charisma possible in both primaries ($c_L = c_R = 0$), regardless of the parties location. Therefore, the party leaders can exploit the platforms channel to increase their expected utility without affecting the candidates' charisma: they choose more extreme platforms.

Proof of Proposition 2.

I prove Proposition 2 in two parts. First, I rewrite the equilibrium charisma, which party leaders must anticipate when they choose their platforms, and then I show their choice.

Equation 20, derived in the proof of Lemma 1, can be re-written, taking into account the different nomination rules n:

$$V'(c_L^{n,*}) = \frac{\partial V(x_L|c_L^{n,*})}{\partial c_L^n} = \frac{\omega \Pi_{d_L}^n}{2\alpha P_L + \Pi_{d_L}^n}.$$

Hence, for a given nomination rule n (the same for both parties as assumed throughout the paper) the party leaders choose platforms that maximize their expected utility:

$$max_{l(n)} P_L \left[u_{z_L}^e(l(n), c_L|n) - u_{z_R}^e(r(n), c_R|n) \right] + u_{z_R}^e(c_R, r(n)|n)$$

And, in agreement with footnote 16, since

$$\Pi_{d_L,n} \equiv \left[u_{d_{p,n}}^e(c_L) - u_{d_{p,n}}^e(c_R) \right],$$

the maximization problem can be re-written in terms of the expected policy gain of the party leader with ideology z_L when the nomination rule n is used, $\Pi_{z_L}^n$,

$$max_{l(n)} P_L \Pi_{z_L}^n(l(n), r(n), c_L, c_R) + u_{z_R}^e(r(n), c_R|n),$$

Hence, the F.O.C. is:

$$0 = \frac{-2l(n) + \frac{\partial V_R^n}{\partial l(n)} - \frac{\partial V_L^n}{\partial l(n)}}{2\alpha} \Pi_{z_L}^n + P_L 2(z_L - l(n)) - \frac{\partial V_R^n}{\partial l(n)}$$

And notice that in the fully symmetric equilibrium with $z_L = -z_R$, it is the case that $l(n) = -r(n)$ and $\frac{\partial V_L^n}{\partial l(n)} = -\frac{\partial V_R^n}{\partial r(n)}$, so

$$\frac{dV_R^n}{dl(n)} = \frac{\partial V_R^n}{\partial r(n)} \frac{\partial r(n)}{\partial l(n)} = -\frac{\partial V_R^n}{\partial r(n)} = \frac{\partial V_L^n}{\partial l(n)},$$

22

hence, from the FOC at the symmetric equilibrium, we obtain

$$0 = -\frac{l(n)}{\alpha}\Pi_{z_L}^n + (z_L - l(n)) - \frac{\partial V_L^n}{\partial l(n)} \tag{15}$$

Rearranging, for two nomination rules, n and n', it must be the case that

$$\frac{l(n)}{\alpha}\Pi_{z_L}^n + l(n) + \frac{\partial V_L^n}{\partial l(n)} = \frac{l(n')}{\alpha}\Pi_{z_L}^{n'} + l(n') + \frac{\partial V_L^{n'}}{\partial l(n')}. \tag{16}$$

Suppose n' is a more open primary, i.e. the decisive voter is more moderate. I want to prove that

$$\text{for } \frac{\partial V_L^{n'}}{\partial l(n')} = \frac{\partial V(\tilde{x}_L^{n'}|c_L^{n'})}{\partial l(n')} > \frac{\partial V(\tilde{x}_L^n|c_L^n)}{\partial l(n)} = \frac{\partial V_L^n}{\partial l(n)}, \text{ then } l(n') < l(n) < 0,$$

i.e., that when the decisive voter in the comparatively closed primary is relatively sensitive to changes in the platform, more open primaries lead to extreme platforms. To prove it by contradiction, suppose $0 > l(n') > l(n)$. For $0 \geq \frac{\partial V_L^{n'}}{\partial l(n')} > \frac{\partial V_L^n}{\partial l(n)}$, equation 16 holds only if $l(n') + l(n')\Pi^{n'} < l(n) + l(n)\Pi^n$. Yet Corollary 1 implies $\Pi^{n'} < \Pi^n$, and so $l(n')\Pi^{n'} > l(n)\Pi^n$. Hence, by contradiction, it must be the case that $l(n') < l(n)$. $\blacksquare$

5 Further discussion

In this section, I discuss a widely used method of nominations - discretionary handpicking by a party leader- and its implications for intra-party democratization. Then, I explain the counterintuitive result that handpicking delivers higher welfare than any other nomination rule. Lastly, I briefly discuss why all the results here would hold under the assumption that ideology is observed instead of charisma.

5.1 Intraparty democratization.

Although I do not endogenize the choice of the nomination rules, the model discussed in the previous section sheds some light on this issue.[24] In particular, it is possible to extend the results to the case of the party leader handpicking the nominee from his party. In this case the decisive party

---

[24]See Aragon (2013) or Serra (2011), for related literature.

member, who is also the party leader, is more extreme than the decisive voter in the closed primary. Therefore, he will nominate an even more charismatic individual. Similarly, when the platforms are endogenous, the party leader is not delegating the choice of the nominee, therefore he can afford to choose a candidate that is more electable in every dimension: besides choosing a more charismatic one, he can also choose more moderate platforms, in comparison to a primary. Therefore, rather than a new result this is just a straightforward corollary from Proposition 2:

Corollary 3 (Informal nomination rules: handpicking) *Under the conditions of Proposition 2, it follows that party leaders choose more moderate platforms when they handpick the nominee theirselves than when they expect any other formal nomination rule. Similarly, party leaders nominate more charismatic candidates than any other primary rule.*

While political scientists and world leaders advocate for clean and popular elections as a necessary condition to choose democratic representatives, political parties are still halfway in the democratization path.[25] The corollary above is useful to understand this current bias against primaries in general, and open primaries in particular.

When political parties' selection methods are not regulated exogenously, party leaders (or the party elite) 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 the decisivevoter in the primary if they do not experience a loss in their expected utility: trivially, that is the case if voters do not care about charisma ($\omega = 0$), or if there is no party differentiation ($l = 0 = r$). Else, when the party leaders' preferred candidate is different from the decisive voters' preferred 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 (2) above). Instead of addressing the effect of platforms on the democratization of

---

[25]The two most widely quoted criteria for a democracy are due to Dahl (1989): (1) effective participation and (2) voting equality at the decisive stage. Furthermore, "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".

nomination rules, in the following remark I highlight the effect of nomination methods on the polarization of parties.

Remark 1 *In the short run, when the platforms are fixed, more democratic nomination rules can be observed if there is platform convergence. However, in the long run, more democratic nomination rules are observed in extremist parties.*

A related but different topic is whether holding primaries harms the likelihood of winning an election: according to the hypothesis of "divisive primaries", candidates nominated through primaries have a smaller probability of winning when the primaries are more intense and contested (see Agranov (2015) for references and a theoretical explanation). [26] Although in the paper, I do not have a measure of intensity, I can briefly address the effect of primaries on the probability of winning, within the limits of the model. Since in the symmetric case, for any nomination rule n, both parties have the same probability of winning, it is fair to ask what the best nomination method is for an opposition party, taking the incumbent's attributes (charisma and ideology) as known. A party maximizes its probability of winning when choosing a more centrist and more charismatic candidate, which is the case when the party leader handpicks the nominee, as stated in Corollary 3. 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)).

5.2   Welfare

In this section I study to what extent the social welfare in this economy, defined as the population-wide median voter's utility, changes with the institutional setting. In particular, since the decisive voter in a pure open primary coincides with the population-wide median, he chooses the nominees for both parties. I investigate under what circumstances he prefers a different nomination rule, by which he would "delegate" the choice to somebody else. Intuitively, when the population-wide median voter is the open primary's decisive voter, i.e. with ideology at 0, social welfare in pure

---

[26]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).

open primaries Ω^o can be larger than when the party leaders handpick the nominees, Ω^h, only if the loss in variance is larger than the gain in policy. From Proposition 2, when party leaders anticipate an open primary, they choose a more extreme platform. Let the utility of the population-wide median voter (with preferred policy equal to 0), be u_0. Therefore, focusing on the left-wing party, $\Omega^o \leq \Omega^h$ if and only if

$$\Omega^o \equiv u_0\left(l(o), c_L^o = 0\right) = -\left(l(o)\right)^2 \leq -\left(l(h)\right)^2 - V\left(c_L^h\right) = u_0\left(l(h), c_L^h\right) \equiv \Omega^h \tag{17}$$

Rearranging the inequality, social welfare with the least democratic nomination method (discretionary handpicking) might be larger when the policy unpredictability with handpicking is smaller than a distance between the platforms. Simply put,

$$V\left(c_L^h\right) \leq \left(l(o)\right)^2 - \left(l(h)\right)^2$$

As explained in Proposition 2, whether the decisive party member (in this case, the party leader) is "sensitive" to changes in the platforms, relative to the platforms distance, is what will determine which nomination method maximizes the population-wide median voter's utility. Using the equation that defines implicitly the optimal platforms (equation 15, in the Appendix), I can rewrite the condition above as,

$$\frac{4z_L}{\alpha}V\left(c_L^h\right) \geq (l(h) - l(o)) + \frac{\partial V\left(c_L^h\right)}{\partial l(h)}.$$

While the left-hand side is negative, the right-hand side will only be negative if $|\frac{\partial V\left(c_L^h\right)}{\partial l(h)}|$ is very large, since $l(h) - l(o)$ is negative in equilibrium,i.e. , only when the party leader adjusts his choice of charisma by much, due to a more moderate platform. In that case, the policy uncertainty would be too large and the general median voter will prefer open primaries. Otherwise, the party leader handpicks better candidates in terms of social welfare.

5.3 Observed ideologies

Throughout the paper I have assumed that charisma is observed but ideologies are private information but, in order to show that the main result is robust, here I provide a brief intuition on how the mechanisms in the model would change if the ideology of citizens were observed.

Suppose that x_i is observed, and that c_i is private knowledge, and independent from the ideologies. Since the affiliation decisions are not affected, the size and composition of parties is the same as before. Moreover, among the party members, the ex-post relationship between ideology and charisma still holds. Therefore, when voters observe the ideology and the affiliation of the candidates (i.e. x_P), they update their beliefs on the conditional density $c_i | x^P$. They would still be risk averse on ideologies, but now there is no incomplete information on that dimension. Suppose that x_i is observed, and that c_i is private knowledge, and independent from the ideologies. Since the affiliation decisions are not affected, the size and composition of parties is the same as before. Moreover, among the party members, the ex-post relationship between ideology and charisma still holds. Therefore, when voters observe the ideology and the affiliation of the candidates (i.e. x_P), they update their beliefs on the conditional density $c_i | x^P$. They would still be risk averse on ideologies, but now there is no incomplete information on that dimension.

Lemma 2 *Assume $B(\delta_i, c_i | p) = -(x_i - \pi_p)^2 + ac_i^2 + t$, where $a > 0$ and t some constant. For α large enough, there is a large enough $\frac{\omega}{a}$, such that the candidates' valence is decreasing in nominator's ideology. That is, more open primaries leads to nominating less charismatic candidates.*

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 choose candidates whose preferred policy is closer to them, and 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, and 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.

At last, notice that if more charismatic party members were able to scale up positions within

the party more easily than less charismatic ones, then both, voters and party members, could learn the candidates' charisma. This is one more reason why in the general setup I take the modeling choice that charisma, instead of ideologies, is observed.

Proof of Lemma (2)   To show that the necessary monotonicity also holds if the ideologies are observed instead of charisma, it is enough to show that: $\frac{\partial x_L}{\partial x_{L,n}^d} < 0$. First, I show how the affiliation decisions change as I change the previous assumption; second, I show that the result still holds.

If a voter (x_i, c_i) affiliates to party L, then $b(c_i) \geq (x_i - l(n))^2$. Let x_L be the observed candidates' ideology; and let $|x_L| < |l|$, and the same for R. Then

$$E(c_i | x_L) = \int_{c_i \geq b^{-1}(x_l - l)^2} c_i dF(c).$$

Assume $b(c_i) = ac_i^2 + t$, and let $d_{L,n}$ be the nominator's ideology from party L with ideology $x_{L,n}^d$; in a symmetric game he maximizes

$$\frac{\alpha + x_R^2 - x_L^2 + \omega \left[E(c_R | x_R) - E(c_L | x_L) \right]}{2\alpha} \left(4x_L x_{L,n}^d \right)$$

The lemma (2) states that for a large enough α there exists a large ω/a such that the equilibrium x_L is decreasing in $x_{L,n}^d$. 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

$$E' \equiv \frac{\partial E(c_L | x_L)}{\partial x_L} = \frac{(x_L - x_{L,n}^d)}{2a \sqrt{\frac{(x_L - x_{L,n}^d)^2 - t}{a}}} > 0.$$

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

$$((x_R - x_{L,n}^d)^2 - (x_L - x_{L,n}^d)^2)(\frac{\omega}{2\alpha} E' - \frac{x_L}{\alpha}) - 2(x_L - x_{L,n}^d)P_L = 0.$$

The S.O.C always holds:

$$\frac{\partial FOC}{\partial x_L} = -\frac{t}{2a(\frac{(x_L - x_{L,n}^d)^2 - t}{a})^{3/2}} = -\frac{t}{2a(\cdot)^{3/2}} < 0.$$

Then, by the implicit function theorem, $\frac{\partial x_L}{\partial x_{L,n}^d} < 0$ if

$$\frac{2x_L}{\alpha} (\frac{\omega}{2a} \frac{(x_L - x_{L,n}^d)}{(\cdot)^{1/2}} - 2x_l) + 1 < 0,$$

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

$$\alpha > 4x_L^2$$

■

6 Conclusion

As mentioned in the Introduction, pundits and policy-makers in the US support open primaries. In France, for the first time in history the two most important parties – the Socialist Party and the Republican Party – had open primaries to select the nominees for the 2017 Presidential race. In Spain and Netherlands, the political parties began to use primaries in the last decade. In Latin America, most regimes have had at least one primary to nominate their presidential and their legislative candidates.

Due to these recent moves toward (open) primaries worldwide, it has become more important to understand their effects. In particular, the consequences over a countrys policy-making and ideological polarization, I claim, will partly depend on the types of primaries and the role played by charisma, during the affiliation stage and/or the electoral stage.

If charisma does not play a role during the electoral stage, then under all types of primaries the voters would choose a non-charismatic nominee. If charisma does not play a role during the affiliation stage, then charisma and ideology would be orthogonal in equilibrium and voters would pick an equally charismatic nominee under any type of primary (at no cost, they would nominate a candidate with the maximum possible charisma).

The most recent empirical literature that looks for a causal effect of primaries on candidates' ideology in the US is in line with my results. In a model where party membership and charisma play an important role in the understanding of how nomination rules influence the choice of candidates, 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. Hence, contrary to the conventional wisdom, when both parties nominate candidates with open primaries, my model predicts higher party polarization in the long term. Similarly, if party leaders were to handpick the candidates, they would choose candidates who are more moderate and predictable than candidates nominated in primaries.

As shown, any general equilibrium model of primaries should include endogenous party mem-

bership because it affects the pool of potential candidates and the voters' beliefs on the types of candidates. The interplay between the nomination rules and the parties' organization also affects the equilibrium candidates, and therefore, the implemented policies.

This paper highlights the mechanisms by which charisma may interact with the partys rules to induce more polarization in a society. Hence, the choice of a nomination rule should be decided after careful examination of its consequences, beyond the electoral arena.

7   Appendix

7.1   Probabilistic voting: Bounds and Derivation

Lemma 3 *Let $c^* \equiv argmaxP_L$, then for $\alpha \geq \omega c^* - V(c^*) + V(0)$ the probability of winning is bounded between 0 and 1.*

Proof.

The probability of winning for party L in the symmetric equilibrium can be written as

$$P_L \equiv \frac{\alpha + \omega(c_L - c_R) + V(c_R) - V(c_L)}{2\alpha}.$$

Consider $c^* \equiv argmaxP_L$; no voter will choose a candidate whose valence is larger than c^*. Since the benefit of higher valence, $\omega q - V(q)$ is convex in c, for any probability of winning with $c > c^*$, 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 $c \in [0, c^*]$. For all values of $(c_L, c_R) \in [0, c^*]^2$, the probability above is bounded between 0 and 1 when the following conditions hold: First, let take the biggest possible value of P_L, that is, when L chooses c^*, and R, 0:

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

And, the smallest, when the opposite is true:

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

Both conditions boil down to

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

∎

Derivation of equation 10.

There is a continuum of voters whose utility depend on the implemented policy and their

affiliation decisions. Let the pair (x_i, c_i) identify each voter, with $x_i \sim U[-x, x]$, $c_i \sim F$ with support $C = [0, \bar{c}]$, and $x_i \perp c_i$. To keep it short, let $u_i(\cdot, L)$ be the utility if party L's candidate wins, for all affiliation decisions. All voters have an exogenous preference for party R over L, which can be decomposed in an individual preference $\beta_i \sim U[-\beta, \beta]$, and a general bias

$$\tilde{\alpha} \sim U[-\alpha + \omega(c_R - c_L), \alpha + \omega(c_R - c_L)],$$

which could be described as idiosyncratic and national shocks, unobserved by the parties; the national shock's distribution depends on the candidates' relative charismatic advantage. 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. Thus, a voter i votes for party p, say L, if:

$$u_i(\cdot, L) - u_i(\cdot, R) > \beta_i + \tilde{\alpha},$$

hence

$$
\begin{aligned}
\Pr(u_i(\cdot, L) - u_i(\cdot, R) \geq \beta_i + \tilde{\alpha}) &= \Pr(u_i(\cdot, L) - u_i(\cdot, R) - \tilde{\alpha} \geq \beta_i) \\
&= \frac{u_i(\cdot, L) - u_i(\cdot, R) - \tilde{\alpha} + \beta}{2\beta} \\
&= \frac{2x_i(l - r) + \omega(c_L - c_R) - (V(c_L) - V(c_R) + r^2 - l^2 - \tilde{\alpha}}{2\beta} + \frac{1}{2} \\
&\equiv P^i(u_i(\cdot, L), u_i(\cdot, R); \tilde{\alpha}) \equiv P^i \qquad (18)
\end{aligned}
$$

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 L is:

$$\int P^i \frac{1}{2x} dx_i = \frac{\omega(c_L - c_R) - (V(c_L) - V(c_R)) + r^2 - l^2 - \tilde{\alpha}}{2\beta} + \frac{1}{2} \equiv S_L \qquad (19)$$

And lastly, the probability that party L wins is:

$$\Pr(S_L > \frac{1}{2}) = \frac{\omega(c_L - c_R) - (V(c_L) - V(c_R)) + r^2 - l^2}{2\alpha} + \frac{1}{2} \equiv P_L$$

∎

7.2   Other proofs.

Proof of lemma 1.

For an interior equilibrium, from the first derivative of the expected utility in equation 12, I

obtain

$$\frac{\partial P_L}{\partial c_L}\Pi_i(c_L, c_R) + P_L\frac{\partial \Pi_i(c_L, c_R)}{\partial c_L} = 0$$

$$\frac{\omega - \frac{\partial V(c_L)}{\partial c_L}}{2\alpha}\Pi_i(c_L, c_R) - P_L\frac{\partial V(c_L)}{\partial c_L} = 0 \tag{20}$$

Before moving into the second derivative, first notice that if $\frac{\partial V(c_L)}{\partial c_L}$ was negative, hence the equation above cannot be satisfied with equality, hence there would be a corner solution: the upper bound of C, with $\bar{c} < \infty$.

Hence, the second derivative can be written as

$$\frac{\partial FOC}{\partial c_L} = 2\frac{\partial P_L}{\partial c_L}\frac{\partial \Pi_i(c_L, c_R)}{\partial c_L} + P_L\frac{\partial^2 \Pi_i(c_L, c_R)}{\partial c_L^2} + \Pi_i\frac{\partial^2 P_L(c_L, c_R)}{\partial c_L^2}$$

$$= -2\frac{\omega - \frac{\partial V(c_L)}{\partial c_L}}{2\alpha}\frac{\partial V(c_L)}{\partial c_L} - P_L\frac{\partial^2 V(c_L)}{\partial c_L^2} - \frac{\Pi_i}{2\alpha}\frac{\partial^2 V(c_L)}{\partial c_L^2},$$

which, for convenience, I write as

$$\frac{\partial FOC}{\partial c_L} = -\frac{\omega - V'}{\alpha}V' - V''(P_L + \frac{\Pi_i}{2\alpha}),$$

and satisfies the SOC if and only if

$$V'' \geq \frac{-2(\omega - V')V'}{2\alpha P_L + \Pi_i}$$

In particular, for all $V'' \geq 0$ the SOC is satisfied. Notice that if $V'' < 0$, hence there is no interior c_L that maximizes the probability of winning, i.e., everybody would choose $c^* = \bar{c}$. And if $V' < 0$, everybody would choose $c^* = 0$. $\blacksquare$

References

Agranov, M. A. (2015). Flip-Flopping, Intense Primaries and the Selection of Candidates. (March), 1–48.

Aldrich, J. (1995). *Why Parties? The Origin and Transformation of Party Political Parties in America* (1 ed.). Chicago: University of Chicago Press.

Alesina, A. and H. Rosenthal (1995). *Partisan Politics, Divided Government, and the Economy.* Cambridge, UK: Cambridge University Press.

Alvarez, R. M., D. T. Canon, and P. Sellers (1995). The Impact of Primaries on General Election Outcomes in the US House and Senate. *Working Papers*.

Amorós, P., M. S. Puy, and R. Martínez (2016, apr). Closed primaries versus top-two primaries. *Public Choice 167*(1-2), 21–35.

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

Aragon, F. (2013). Political parties, candidate selection, and quality of government. *The BE Journal of Economic Analysis & Policy 13*(2), 783–810.

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

Banks, J. S. and J. Duggan (2005). Probabilistic Voting in the Spatial Model of Elections: The Theory of Office-motivated Candidates. *Social Choice and Strategic Decisions: Essays in Honor of Jeffrey S. Banks*, 15–56.

Berdahl, C. a. (1949a). Presidential Selection and Democratic Government. *The Journal of Politics 11*(01), 14.

Berdahl, C. A. (1949b). Some Notes on Party Membership in Congress, I. *The American Political Science Review 43*(2), 309–321.

Berggren, N., H. Jordahl, and P. Poutvaara (2010, feb). The looks of a winner: Beauty and electoral success. *Journal of Public Economics 94*(1-2), 8–15.

Besley, T. and S. Coate (1997). An Economic Model of Representative Democracy. *The Quarterly Journal of Economics 112*(1), 85–114.

Caillaud, B. and J. Tirole (1999). Party governance and ideological bias. *European Economic Review*.

Caillaud, B. and J. Tirole (2016). Parties as Political Intermediaries. *Quarterly Journal of Economics 117*(4), 1453–1489.

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

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

Cho, S.-j. S. and I. Kang (2014, may). Open primaries and crossover voting. *Journal of Theoretical Politics 27*(3), 351–379.

Crutzen, B. S. Y., M. Castanheira, and N. Sahuguet (2010). Party organization and electoral competition. *Journal of Law, Economics, and Organization 26*(2), 212–242.

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

Grofman, B., O. Troumpounis, and D. Xefteris (2016). Electoral competition with primaries and quality asymmetries.

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

Hamermesh, D. S. (2006, dec). Changing looks and changing discrimination: The beauty of economists. *Economics Letters 93*(3), 405–412.

Hirano, S., O. Kemahlioglu, and R. Weitz-Shapiro (2009, jan). Why Primaries in Latin American Presidential Elections? *The Journal of Politics 71*(01), 339.

Hummel, P. (2010). Flip-flopping from primaries to general elections. *Journal of Public Economics 94*(11-12), 1020–1027.

Jackson, M. O., L. Mathevet, and K. Mattes (2007). Nomination Processes and Policy Outcomes. *Quarterly Journal of Political Science 2*(1), 67–92.

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

Lawson, C., G. S. Lenz, A. Baker, and M. Myers (2010). Looking Like a Winner: Candidate Appearance and Electoral Success in New Democracies. *World Politics 62*(04), 561–593.

Lenz, G. S. and C. Lawson (2011). Looking the part: Television leads less informed citizens to vote based on candidates' appearance. *American Journal of Political Science 55*(3), 574–589.

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

Mcghee, E., S. Masket, B. Shor, S. Rogers, and N. Mccarty (2013). A Primary Cause of Partisanship? Nomination Systems and Legislator Ideology. *American Journal of Political Science 58*(2), 337–351.

Mobius, M. M. and T. S. Rosenblat (1966). Why Beauty Matters. *Science (New York, N.Y.) 96*(1), 222–235.

Osborne, M. J. and a. Slivinski (1996). A model of political competition with citizen-candidates. *Quarterly Journal of Economics 111*, 65–96.

Serra, G. (2010). Polarization of What? A Model of Elections with Endogenous Valence. *The Journal of Politics 72*(02), 426.

Serra, G. (2011). Why primaries? The party's tradeoff between policy and valence. *Journal of Theoretical Politics 23*(1), 21–51.

Shor, B. and N. McCarty (2011). The Ideological Mapping of American Legislatures. *American Political Science Review 105*(03), 530–551.

Snyder, J. M. and M. M. Ting (2002). An Informational Rationale for Political Parties. *American Journal of Political Science 46*(1), 90–110.

Snyder, J. M. and M. M. Ting (2011). Electoral Selection with Parties and Primaries. *American Journal of Political Science 55*(4), 782–796.

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

Stone, W. J. and E. N. Simas (2010). Candidate Valence and Ideological Positions in U.S. House Elections. *American Journal of Political Science 54*(2), 371–388.