Primary Elections or Smoke-Filled Rooms: A Theory of Party Democratization in Latin America*

Gilles Serra†

December 19, 2006

Abstract

Why do political parties adopt primary elections? This study develops a theory of *party democratization.* It is motivated by the recent proliferation of primary elections in Latin America, which has been well documented but not fully explained. The explanation in this paper consists on making two claims. First that *primaries select more electable but less loyal candidates.* This claim is made precise with a decision-theoretic model whose predictions are consistent with the main empirical findings of the existing empirical literature. And second that *the incentives for party leaders to select electable candidates have increased in the past quarter century.* This claim is based

*Previous versions of this paper were presented at the meeting of the Latin American Studies Association in March 15-18, 2006, as well as political economy seminars at Harvard University and MIT. I gratefully acknowledge financial support by the Institute of Quantitative Social Sciences. I want to thank Christopher Avery, Robert Bates, Barry Burden, Jorge Domínguez, Daniel Gingerich, Nahomi Ichino, Steve Levitsky, John Patty, Kenneth Shepsle and James Snyder for their useful comments on this paper. In particular Kenneth Shepsle, James Snyder, Robert Bates and Christopher Avery have always offered me their insight and support, which not only improved this paper but also made me a better scholar. All remaining errors are my own.

†Ph.D. candidate in Political Economy and Government at Harvard University. Email: serra@fas.harvard.edu

on some historical trends in Latin American politics, namely the increased penetration of television, a more investigative journalism, and a decrease in electoral fraud.

Keywords: primary elections, Latin America, valence, jury theorems

JEL Classification: D02, D72, O54, P48

1 Introduction

Why do political parties adopt primary elections? Political parties across history and around the world have devised many mechanisms to nominate their candidates for public office. At one extreme, parties may use non-democratic methods of nomination such as party elites directly handpicking the candidates in closed-door negotiations (notorious examples being the Democratic Party machine in Chicago during the 1930s or the hegemonic PRI in Mexico throughout the twentieth century). At the other extreme, parties may democratize their nomination process by adopting a primary election where a large number of voters (at least all the party's membership) choose a candidate by majority voting. However, the question of why a party chooses a democratic versus a non-democratic candidate selection process (CSP) remains unanswered.

This study develops a theory of *party democratization*, understood as the adoption of primary elections to nominate a party's candidates.[1] Its regional focus is Latin America which has seen a fast growth in the use of primaries.[2] The theory is developed first by discussing some of the main empirical facts about primary elections in Latin America; second by laying out a decision-theoretic model to explain those facts; and third by linking that theoretical model to some historical trends that predict an increased use of primary elections in that region.

[1] The word "democratization" is not ideal in this context because of its normative connotations. But the literature on candidate selection has not settled on a different term yet (party openness, party inclusiveness etc.). So in this paper I follow the tradition of calling party democratization the adoption of primary elections.

[2] Primary elections have also proliferated in other regions of world. They have multiplied in the United States since the McGovern-Fraser reforms of 1970, and have been used recently in Finland, France, Iceland, Italy, Israel, Palestine, Spain and Taiwan. But as we will see below, the use or primaries by political parties in Latin America is most interesting because it is both widespread and voluntary.

The theory is indeed motivated by an intriguing empirical phenomenon: the proliferation of primary elections in Latin America in the past two decades. The increased frequency of parties adopting primary elections has been well documented: political parties in the region had almost never used primaries to select their candidates before the 1980s but today a large proportion uses them on a regular basis. Indeed, in a comprehensive survey, Alcántara and Freidenberg (2001) report that by the end-2000 as many as 23 out of 73 parties in Latin America had adopted primary elections to select their presidential candidates. Notable examples of presidents who were nominated by a primary election before winning the general election include Argentina's Carlos Menem, Brazil's Ignacio (Lula) da Silva, Chile's Ricardo Lagos, Colombia's César Gaviria, Honduras' Rafael Maduro and Mexico's Felipe Calderón.[3] The majority of those primaries have been *closed* primaries, meaning that the vote is restricted to party members, but the prerequisites to join a party tend to be quite lax resulting in very large party memberships.

The significance of that empirical pattern has been underscored by several academics. For example Alcántara and Espíndola, in a survey of the changes affecting political parties in Latin America since 1980, do not hesitate in saying that:

> "Undoubtedly the most radical institutional change was the adoption of internal or primary elections for the selection of candidates, and particularly the selection of those standing for the Presidency." (Alcántara and Espíndola (2003), 5)

This phenomenon is unexpected, however. As it turns out, parties have democratized their nomination process by voluntary choice. Unlike in the United States where the specific nomination process is mainly mandated by law, political parties in Latin America are largely free to choose their nomination process. The countries where parties have deliberately adopted primaries include Argentina, Chile, Colombia, Costa Rica, Dominican Republic, Ecuador, El Salvador, Mexico, Nicaragua, Paraguay and Venezuela (Freidenberg and Sánchez López (2001), Alcántara (2002)).[4] Parties in those countries often debate heatedly over their choice of a candidate selection

[3] Carey and Polga-Hecimovich counted up to 47 out of 826 presidential candidates nominated through primaries in the period 1978-2004.

[4] In some other countries, parties were forced to adopt primaries by the electoral law rather than by voluntary choice, like Uruguay, but they are still a minority of cases.

process long before they start debating about their choice of the candidate herself. From one election to the next parties go back and forth between primaries and other nomination processes, clearly indicating the strategic nature of CSP choice[5] (Freidenberg (2003)).

And yet the voluntary adoption of primary elections should be considered a puzzle. If a party's leadership has the ability to choose any candidate of its liking, why does it hold primary elections where it cannot fully control the outcome? If party bosses used to decide the nomination in a "smoke-filled room", as Americanists call the old practice of selecting candidates in closed-door negotiations, why did they hand over control to the party's rank and file who might have different and even opposing preferences?

In spite of its significance the question of why parties adopt primary elections remains largely unexplored. In the case of the United Sates, there has been considerable research on the consequences of adopting primaries, but very little on its causes.[6] The origin of primaries in other world regions has also been scarcely researched notwithstanding a few recent contributions.[7] The goal of this paper is to offer a possible explanation.

In a few words, the argument in this paper consists on making two claims. First that *primaries select more electable but less loyal candidates*. This claim is made precise with a decision-theoretic model whose predictions are consistent with the main empirical findings of the existing literature on primary elections in Latin America. And the second claim is that *the incentives for party leaders to select electable candidates have increased in the past quarter century*. This claim is substantiated by pointing out some historical trends in Latin American politics, namely the increased penetration of the media and particularly television, a more investigative journalism, and a decrease in electoral fraud. Taken together, those two claims imply that the incentives for party leaders to adopt primary elections have increased in the past quarter

[5]Levitsky (2003 p.80) quotes an activist of the Partido Justicialista in Argentina complaining that "internal democracy exists only when it serves the interests of the party leadership. When it does not serve the interests of the leadership, there is no internal democracy."

[6]Notable exceptions are the classical accounts in Key (1947), Polsby (1983), Bartels (1988) and more recently Ware (2002). In particular Alan Ware asks a similar question for the United States than I ask for Latin America: Why did party bosses and their "machines" agree to, or refrain from blocking, the adoption of primary elections by state legislatures? Similarly see Adams and Merrill (2006).

[7]Most notably for Latin America, Poiré (2002), Wuhs (2006), Field and Siavelis (2006) and Siavelis and Morgenstern (2006).

century, which would explain the empirical phenomenon just described.

This paper connects three different theoretical literatures. First this is a model of candidate valence (Stokes (1963), Enelow and Hinich (1982), Londregan and Romer (1993), Harrington and Hess (1996), Adams (1999), Ansolabehere and Snyder (2000), Groseclose (2001), Schofield (2004), Schofield and Sened (2005)). Second, it adds to the young but growing formal literature on primary elections (Morton and Williams (2001), Caillaud and Tirole (2002), Owen and Grofman (2004), Meirowitz (2005), Adams and Merrill (2006), Jackson, Mathevet and Mattes (2006), Serra (2006a, 2006b), Snyder and Ting (2006)). In particular it shares some ideas with Adams and Merrill (2006) who independently wrote a model of primary elections where candidates have different valences; from a different approach to mine they find, as I do, that primaries can be expected to increase the valence of the nominated candidate.[8] Jackson et al. (2006) is another model that compares primary elections with handpicked appointments by party leaders, but its focus is on explaining policy divergence, rather than explaining the adoption of primaries. And third, this is also a model of information-aggregation that uses the technology of jury models (Austen-Smith and Banks (1996), Feddersen and Pesendorfer (1998), Duggan and Martinelli (2001), Martinelli (2002)). Thus a theoretical contribution of this paper is to transport the setup and logic of jury theorems to a totally different context. I believe that in this new context, primary elections, the original assumptions in information-aggregation models are even more valid than in jury trials.

The paper is organized as follows. Section 2 summarizes the argument of this paper. Section 3 make a critical review of some of the previous arguments for the proliferation of primary elections. Section 4 lists the main empirical findings from the nascent literature on primary elections in Latin America that my model is set out to explain. Section 5 introduces the model which Section 6 solves. Section 7 derives the main parameter in the model from microfoundations. Section 8, the punch line of the paper, connects the formal model to features of Latin American politics that have evolved in the past three decades in favor of the adoption of primary elections. In the last section we offer a summary and conclusions. And finally the Appendix contains the proofs of theorems and lemmas.

[8] However the focus in Adams and Merrill (2006) is not the choice of primaries versus other CSPs, but the policies adopted by candidates who run in primaries.

2 Some previous arguments

Here is a list of some of the hypotheses that are most often mentioned in academic and journalistic articles to explain the emergence of primary elections in Latin America.[9] The limitations of those hypotheses point to the need of a full-fledged theory.

Party leaders have acquired a genuine preference for democracy: The suggestion is that politicians in Latin America have a strong commitment to democracy and they show it by democratizing their own parties. This argument is most often mentioned by politicians themselves: whenever their party decides to employ a democratic CSP they make large efforts to publicize that decision in order to convince voters of their democratic credentials. The seriousness of this argument is challenged by noticing that the same party often reverts to a traditional back-room haggling in a subsequent election. They do not hold a new press conference though.

Party democratization is consistent with the overall democratization of a country: This line of thinking comes from the observation that party democratization (in particular the use of primaries) seems to have chronologically happened short after the transition from authoritarian to democratic regimes. It assumes that there is a natural link between *government democratization* (the general trend toward fair, transparent and balanced elections to elect candidates for office) and *party democratization* (the use primary elections instead of elite arrangements to select a party's candidate). Government democratization concerns the competition *between* parties while party democratization concerns the competition *within* a party. Note however that those two processes need not be connected. There are examples of democratic parties under undemocratic governments (like the PAN in Mexico), and examples of undemocratic parties under democratic governments (like the orthodox religious parties in Israel). So it is not sufficient to note that government democratization was followed by party democratization: a causal mechanism needs to be proposed, which this paper does.

Primary elections are good for society and society demands them: As mentioned in the conclusion of this paper, there exist normative arguments to wish for a proliferation of primary elections. But a list of benefits does not amount to an explanation. In particular an answer still

[9] The most exhaustive list available of previous explanations comes in Table 1 of Field and Siavelis (2006).

needs to be given to why the party elites accepted and did not block the primaries.

Primaries provide legitimacy and a democratic image to the parties that use them: This argument is based on two assumptions: first that voters pay attention to and care about the details of a party's CSP; and second that a party's image is linked to its use of primaries. Both assumptions are fairly questionable. De Luca et al. (2002) find that the UCR in Argentina has an image of more liberal-democratic and less vertical than the PJ in spite of both being equally likely to hold primary elections. They also find that the media portrays the FG as a "modern" party that does not share the PJ and UCR's "old-school" methods, and yet the FG relies on a back-room CSP instead of primaries much more often than the PJ and UCR (90% of the time). De Luca et al. conclude that "The use of primaries may not be driven as much by external pressures (e.g., the perception of the party's internal functioning by the voters vis-à-vis that of other parties) as by intra-party politics."

Primaries commit the militants and engage the voters: This argument is based on the psychology of voters and militants. Supposedly militants who have been mobilized by a winning precandidate during her primary campaign will "make the candidacy their own" and continue their support during the general election. In addition a voter who voted for a primary candidate would be more likely to vote for her in the general election. These arguments are appealing but do not explain why parties adopt primaries some times and not other times, and why primaries have proliferated in the past two decades.

Party leaders are too weak to choose any other CSP than a primary. This claim comes from the observation that many parties in Latin America fail to have a hierarchical and centralized structure. The implication would be, according to this argument, that party leaders do not really have the ability to make any significant decision about their parties CSP or anything else, and primaries are the natural CSP to be adopted in the absence of any top-down imposition. This view however fails to account for the observation that parties do seem to strategize over their CSP. To explain why parties go back and forth between primaries and other CSPs, we really need a theory of *deliberate* choice of CSP. An additional limitation of this argument is that it does not say why primaries have become more frequent in the past 15 years.

Primary elections preserve party unity: According to this argu-

ment, party leaders rely on primaries when they fear that a traditional elite negotiation will lead to an unsolvable confrontation that might divide the party. This is an appealing argument that has a strategic logic and some anecdotal evidence (although it does not per se explain the *increase* in the number of primaries though time). Indeed the risk is large in Latin America that a party or coalition of parties is split into separate parties (like the PRI in 1988), and the there is also a frequent risk that good candidates will switch to different parties. Carey (2003) believes that primary elections in the 1990s often served as a commitment device for parties (for example in Uruguay) or coalitions of parties (like the Concertación in chile, the Alianza in Argentina) to hold together. Freidenberg (2003) adds Argentina's FREPASO and Ecuador's Izquierda Democrática to the examples of parties that used primaries to avoid a conflict within their leadership. Poiré (2002) documents the series of disastrous defections that followed non-democratic CSPs in the PRI, where the losers of a smoke-filled room nomination cried foul and switched to a different party (and often won the general election). That prompted the PRI to start adopting primaries in the late 1990s. This provides evidence in favor of the argument, but as it has been stated thus far, that argument is incomplete. In particular the question remains of why a losing candidate is less likely to leave the party if she lost after a primary than if she lost after an elite arrangement. What is so special about primaries that they prevent a party fracture? The model in this paper provides an answer, so in fact my theory can be considered as the first step toward a full analysis of parties under the risk of fracture.

3 A new argument: the increased incentive to aggregate information about candidates' electability

In a few words, the argument in this paper is based on two premises: (1) that primaries select more electable but less loyal candidates, and (2) that selecting electable candidates has become increasingly important for party leaders in the past quarter century. Both steps of the argument are elaborated below.

Primaries select more electable but less loyal candidates: This premise highlights the trade-off that party leaders face in adopting a primary

election. First, the obvious point must be made that party elites benefit from centralized, non-democratic CSPs. Party leaders value having control over the nomination process for several reasons: it allows them to select the candidates closest to their hearts (the hard-core loyalists, or some cronies); it allows them to reward the disciplined members with juicy candidacies and punish the mavericks by blocking their careers; it allows them to reward seniority within the party; and it allows them to enjoy patronage and cronyism on behalf of the winning candidates. On the other hand, party leaders value winning elections because it brings them rents and allows their parties to implement their preferred policies. It is thus clear that party leaders value at least two different attributes in a candidate: loyalty and electability. But the dilemma party leaders face is that the candidates who might provide one type of benefit might not provide the other: the most loyal candidates might not be the most appealing to voters and vice versa. In addition, different CSPs might select different types of candidates. I argue that primary elections are more likely to select the electable type of candidates for three mains reasons: first, the party's rank-and-file membership might be "closer" to voters in the general electorate (closer geographically, and closer in terms of perceptions of candidates). So party leaders who are often removed in central offices might do well in delegating the assessment of a candidates' appeal to the party's foot soldiers. Second, forcing precandidates to run a primary campaign will reveal a lot about their electability: during the primary the party will be able to compare its precandidates' campaigning skills, charisma, and personal records through their initial exposure to journalists and the media. And third, a primary election amounts to a consultation via the ballot box of hundreds if not thousands of primary voters who might in the aggregate have much more information about a candidate than a handful of political bosses. Given these reasons, whenever party leaders feel strongly about winning an election, they will be willing to delegate the nomination decision to the large body of primary voters who will be more likely to nominate the most electable candidate.

This paper is not alone in making the point that primaries are strategically adopted to nominate more electable candidates. In her study of elections in Mexican states, Beer (2003) finds that more participatory CSPs like open primaries lead to the nomination of candidates with local ties and regional sensitivity, who have a stronger appeal to the electorate than "party loyalists" coming from the capital. In an insightful survey of party democratization around the world, Scarrow (2005) argues that beyond the social

benefits of intra-party democracy there are self-interested reasons for parties to adopt primaries, namely that such procedures "may help parties win elections, recruit and select good candidates, and retain popular support". In their survey of Latin American elections, Carey and Polga-Hecimovich (2006) suggest that "primaries may be more effective than elite-driven search processes in identifying candidates with broad popular appeal". Wuhs (2006) argues that "more open selection processes should result in the nomination of more 'electable' candidates who can perform better in the general election." Chand (2001) claimed that primaries would "field more electable candidates because rank-and-file party members were by definition more representative of the state's population than distant party elites." In his field interviews in Mexico, Poiré (2002) found that all the relevant leaders deciding the PRI's CSP shared the view that "the winner of a primary has a better chance of winning the election than whoever turns out to be a loser". One of his general conclusions is that "ultimately, parties do turn to primaries in the hope of winning elections". And he even provides some empirical support: using pre-nomination polling data, he finds that the losers of a primary had on average a lower popularity than the losers of an elite-arrangement. In theoretical work, Adams and Merrill (2006) find that "holding a primary is likely to increase a party's chances of winning the general election, particularly in situations where *valence* issues that involve the candidates' campaigning skills and that are not know prior to the campaign are more salient than policy issues." In studying presidential nominations in the Unites States, Geer (1989) concludes that "in principle, primaries may provide *more* reliable estimates of a candidate's ability to win votes than can party leaders". Even Nelson Polsby, generally a critic of the American primary system, observed that presidential primaries can serve "as means by which politicians inform themselves about the relative popularity of presidential aspirants" (Polsby (1960 p. 617) as cited in Geer (1989 p. 106)).

This view is not shared by everyone however. Some authors argue that primaries are dominated by extreme activists who tend to nominate unrepresentative candidates that are less electable (for the U.S. this view was expressed by Key (1947) and Polsby (1983), for Latin America see Colomer (2002)). However, as we will see below, the empirical evidence in Latin America favors the former point of view over the latter: primaries seem to nominate more electable candidates.

Candidate electability has become increasingly important in the past quarter century: The second step in the argument is connecting the

use of primary elections to other features of Latin American politics that have evolved of the past two decades. My formal model will link the choice of CSP to several variables, and will make a series of empirical predictions of how those variables affect the trade-off between a primary and an elite-arrangement. Three of those variables can be documented to have changed in favor of primaries: (1) the penetration of the media, in particular television, (2) a more investigative journalism, and (3) a decrease in electoral fraud. My model postulates that if those variables shift in a certain direction, primaries will become a better strategic choice for party leaders. And indeed we can document that they have shifted exactly in that direction in the past two decades.

4 Previous empirical findings

Even though primary elections are starting to attract the attention of Latin American scholars, very few empirical papers and even more scarce theoretical papers have been published in that topic yet. However the few empirical papers have already conveyed some findings that a good theory of primary elections should be consistent with. Listed below are some of the main findings of the literature that my formal model is set out to explain.

- **Empirical finding 1: The number of primary elections in Latin America has significantly increased since the 1980s.**

Even though other methods like delegate assemblies or elite anointments are still the most common, the use of primary elections has increased substantially. At the *presidential* level, Carey and Polga-Hecimovich (2006) document that the percentage of candidates nominated by primaries increased from 3% in the 1980s to 4% in the 1990s to 12% in the 2000s, thus increasing fourfold in two decades. The table below illustrates in a different way the increasing frequency of primary elections in presidential primaries in the region. It tracks the elections where one or more candidates where previously nominated through a primary, as a percentage of all the elections (with and without primaries). As can be seen in the table, today a majority of presidential contests are accompanied by at least one primary. At the *congressional*, *gubernatorial* and other levels there does not exist to my knowledge a cross-country database, but there is country-specific evidence that primaries have

also proliferated. For example Wuhs (2006) documents a marked increase in the use of primaries in gubernatorial, Senate and Chamber of Deputies elections by all major parties in Mexico throughout the nineties. De Luca et al. (2002) document the use of primaries for the Chamber of Deputies in Argentina since 1983. It is thus apparent that primaries are quickly becoming a regular feature of Latin American elections.

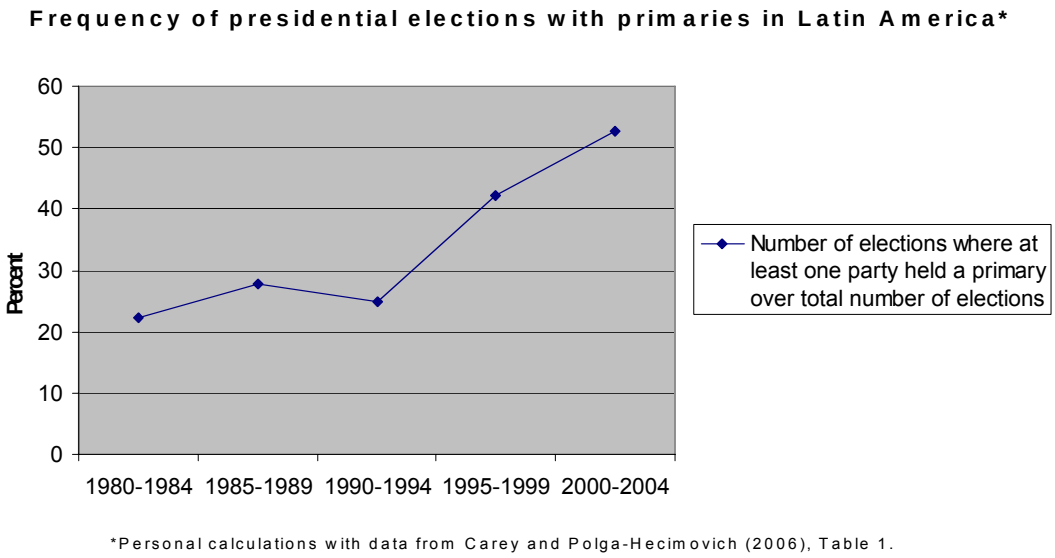


Figure 1:

- **Empirical finding 2: Primaries allow a better performance in the general election compared to other CSPs.**

In a cross-national study of elections in Latin America in the last three decades, Carey and Polga-Hecimovich (2006) find strong evidence of a primary bonus, that is, they find that primary-selected candidates fare better in the general election than those selected by other procedures. Their data set included 826 candidates in 18 countries. They find that all things equal, a candidate nominated by a primary can expect an additional 6 percent points in her voter share during the general election, and this finding is statistically significant and robust to several specifications. These authors also studied the

data from elections in 30 states in Mexico in the period 1994-2000, and find support (though just shy of statistical significance) for the hypothesis that the gubernatorial candidates nominated by a primary have better electoral results than candidates selected by other means. Using different methods, Poiré (2002) studies the same Mexican gubernatorial elections and finds that the PRI did significantly better in those where it used a primary: from his data we can calculate that, controlling for the past vote in each state, the PRI gained 9.3 percent point from using a primary instead of handpicking its candidate.

Consistent with the finding that primaries improve the electoral outcomes of the parties that adopt them, there is evidence that parties are more likely to make use of primaries when they face stronger electoral challenge. The data in Poiré (2002) shows that in the period 1994-2000 the PRI adopted primaries in the states where its electoral advantage over the opposition party PRD was slimmest. Chand (2001) and Beer (2003) reach a similar conclusion.

- **Empirical finding 3: Opposition parties are more likely to hold primaries than parties in the government.**

In an analysis of all the elections for the Chamber of Deputies in Argentina from 1983 to 2001, De Luca et al. (2002) coded the CSP (either a primary or a smoke-filled-room) used by parties to nominate their candidates. Taking the parties' CSP as their dependent variable, they studied the effect of several institutional and partisan features on the probability that a party adopts a direct primary rather than a back-room elite arrangement to select their (list of) candidates. For each district, the authors distinguish between the government party and opposition parties, where government party refers to the party that runs the province that this district belongs to. They find that opposition parties are significantly more likely to hold primaries than government parties (by 18%), and this effect is doubled when the governor of the province is up for reelection (both effects being statistically significant). The authors calculate for example that when the Partido Justicialista is in the opposition it has probability of 0.61 of using primaries to select its congressional candidates for the districts in that province, whereas that probability drops to 0.43 if it holds the governorship of that province, and to 0.21 when its governor is up for reelection. A similar result was found in Mexico by Poiré (2002): the PRI has been more likely to choose a primary in states in which it is not the incumbent.

- **Empirical finding 4: Parties with a smaller membership are less likely to hold primaries.**

In the same study, De Luca et al. (2002) distinguish between the major national parties, and the smaller provincial parties. Two of the main characteristics of the provincial parties are their small membership and their tendency to be dominated by a single person or a small clique. They find that provincial parties are 41% less likely to hold primaries than the national parties. This strong and statistically significant effect leads them to conclude that the smaller the size of a party's membership, the lower the probability that it will hold primaries to choose its candidates for public office. That finding actually carries through to a cross-national level, as shown in Carey and Polga-Hecimovich (2006, Table 3). For the whole Latin American region these authors find that the large parties (those with more than 30% of the vote share) use primaries to nominate 23% of their candidates whereas small parties (those with less than 30% of the vote share) only use primaries to nominate 2% of their candidates.

- **Empirical finding 5: Primaries are more prevalent in countries where politicians are perceived to be more relevant.**

The survey Latinobarómetro, which is an annual survey applied to the citizens of 18 countries in Latin America, includes a question about the relevance of politicians for a country's functioning. The question asks whether a country can function without politicians, which could be interpreted as a measure of how irrelevant politicians are in the eyes of citizens.[10] Alcántara (2002) finds a negative correlation between the number of people answering *yes* and the use of primaries in that country, meaning that the more irrelevant politicians are believed to be, the less likely it is that primaries are adopted in that country. Conversely, the more relevant politicians are believed to be, the more likely it is that primaries are adopted in that country.[11] For example we see at one extreme Uruguay granting the highest relevance to politicians (with only 15% saying that the country could function without

[10] The exact question in Latinobarómetro says "Por lo que Ud. sabe o ha oído, ¿cree que es posible que el país funcione sin políticos?".

[11] In Alcántara's words "parece haber una relación leve negativa entre la creencia de que el país pueda funcionar sin políticos y el uso de las elecciones primarias. Aquellos países en los que es mayor el grado de prescindibilidad de de los políticos son los que menos han adoptado el sistema de primarias (Alcántara [2002], 41)."

them) and having a full adoption of open primaries for all parties. At the other extreme we have Ecuador with people caring the least for politicians (with as much as 40% saying that politicians are irrelevant for the country's functioning), and highly elitist parties that have never adopted primaries.

- **Empirical finding 6: Many parties go back and forth between primaries an other CSPs**

As was mentioned in the introduction, in countries were internal party processes are unregulated, parties are legally able to and they certainly do change their CSP from election to election. Therefore when a party decides to democratize its CSP it is by no means a permanent decision. For example the Partido Justicialista in Argentina used a primary election to select its presidential candidate in 1988, and never again after that (Levitsky(2003)). The Partido Liberal in Colombia experimented with primaries to select its presidential candidates in 1990 and 1994 but went back to a traditional CSP in 1998 (Alcántara (2002)). The PRI in Mexico carefully alternated between primaries and centralized CSPs to nominates its gubernatorial candidates in the period 1994-2000 (Poiré (2002)). If we consider the whole Latin American region in the period 1978-2004, a party that used a primary for a given presidential election only had a 39% probability of using a primary again for the subsequent election (Carey and Polga-Hecimovich (2006, Table 3)).

5 Setup of the Model

We wish to capture the trade-off that the typical Latin American party leader faces in choosing a democratic CSP (e.g. a primary election) versus an undemocratic CSP (e.g. a smoke-filled room). To do so we write the simplest possible model of competition between two parties, one of which still needs to decide on its CSP.[12] In that model there is a unidimensional political spectrum where parties' and voters' preferences are located (Downs (1957)). We

[12]The assumption of only two parties is somewhat restrictive but not entirely unrealistic. Many elections in Latin America are dominated by two major parties or coalitions of parties. Alcántara Sáez [2002] calculated the weighted average of parties in most Latin American countries (giving a smaller weight to smaller parties), finding that many countries had a weighted-average number of parties strictly between two and three throughout the nineties. Those countries included Colombia, Costa Rica, Honduras, Mexico, Paraguay and the Dominican Republic. In any case, extending the model to more than two parties would be feasible and the main results would be mostly unchanged.

will denote policy by x, where $x \in \mathbb{R}$. In this section we describe how we model parties, candidates and voters as well as the election they are involved in.[13]

5.1 Parties

There are two parties denoted by party A and party B. Both parties nominate their candidates at different dates, party A being the first and party B the second in making its nomination.[14] In order to focus on party B's decision, we will assume that party A has already selected a candidate through an elite arrangement and has no other decision to make.[15]

Party B is divided in two groups of people: its leadership (the party elite) and its rank and file (the party members and activists). For convenience whenever I refer to the leaders in party B I will simply write "party B". I will also refer to a member of B's rank and file simply as a "member of B".

We assume that parties A and B have distinct and well defined ideologies that cannot be easily modified, maybe due to a long-held reputation or because their bylaws have rigidly defined their ideology beforehand. Thus A and B have fixed policy platforms denoted by x^A and x^B respectively. Without loss of generality we will assume that $x^B < x^A$, and we normalize both values to $x^B = 0$ and $x^A = 1$. Given that B has a fixed ideology its party leaders cannot strategize over B's policy-platform; they can only strategize over the type of candidate they nominate.

Party B has two precandidates: a loyal precandidate, L, and a disloyal precandidate, D. The precandidates L and D are identical ex-ante except for one characteristic: if the loyal candidate L is elected, she will bring an

[13] To fix ideas the reader can think of this election as being for president, but the model can also be applied to other executive as well as legislative offices (for legislative elections we could replace the label *candidate* by *list of candidates*).

[14] For Latin America this assumption is more natural than assuming simultaneous nomination dates. Unlike in the United States where most nomination dates are specified in state laws and are usually simultaneous for the Republican and Democratic parties, Latin American electoral laws leave the timing of nominations largely to the parties themselves, typically resulting in sequenced, not simultaneous nominations.

[15] Some interesting analysis wold certainly come out of analyzing party A's decision to use a primary election or not; for example there could be some "contagion" meaning that party A's decision to democratize might provide party B with an incentive to democratize as well. However, since the goal of this paper is to plunge deeply into the *internal* functioning of parties, we leave the analysis of this *external* incentive for future research.

extra benefit to the elite in party B (from patronage, cronyism, or disciplined voting), whereas the disloyal candidate D does not bring B that extra benefit.[16]

Thus the elite in party B cares about winning the election, and conditional on winning the election, it prefers winning with its loyal rather than its disloyal candidate. Concretely, the elite in party B gets a payoff of zero if it loses the election; w if it wins the election with D where w is the party's payoff from being in office; and $w + \ell$ if it wins the election with L where ℓ is the extra payoff from loyalty. w and ℓ are such that $w, \ell \in [0, 1]$.

In addition to its leaders, party B has a rank-and-file membership consisting of n members. For convenience we assume n to be odd. The rank-and-file members in party B have identical preferences to the party leaders, except that they do not receive the payoff ℓ from the patronage or cronyism of the loyal candidate.[17] They only receive w if their party wins and zero if it loses. Given these payoffs, the party members will always want their party to win. As we will prove below, an implication is that whenever they are required to vote between L and D they will always vote for the precandidate they think is most likely to win the general election[18] (we do not allow for abstentions).

One of the two parties, A or B, is the incumbent party previously holding power while the other is in the opposition. Being the incumbent party

[16] We assume that party B is certain about the loyalty (or lack of it) of each precandidate. It is likely however that party leaders are not sure about the level of loyalty that a candidate will exert after she is elected. So an alternative assumption would be that each candidate has an expected (but uncertain) level of loyalty toward the party. All the results of the model would still carry through as long as the expected loyalty of candidate L is higher than the expected loyalty of candidate D.

[17] So we assume that B's rank and file have the same policy preferences as B's party leaders. It could be argued however, that a party's membership would adopt a different policy platform than the leadership would. In the case where the primary voters are more extremist than party leaders, the attractiveness of a primary as a mechanism to improve the odds of winning the election decreases and the results of my model are weakened. But in the case where primary voters are more centrist than party leaders the effectiveness of a primary is reinforced and the results of my model are strengthened. So, in the absence of reliable data about the policy preferences of primary voters in Latin America, it is impossible to say whether the effect of primaries is smaller or larger than predicted in this model.

[18] If party members also cared about nominating the loyal candidate, the results of the model would still go through as long as they did not care *as much* as the party leaders do. The payoff ℓ is meant to capture the difference in the preferences of leaders and rank and file.

is beneficial for winning a subsequent election because the government can engage in electoral fraud to overturn an unfavorable result. If the incumbent party gets less votes than the opposition party, it will commit a fraud that has probability ϕ of succeeding (and thus overturning the result) and probability $1-\phi$ of failing (and thus upholding the result), with $\phi \in (0,1)$. On the other hand, if the incumbent party wins a majority of votes the result is automatically upheld. That is, we assume that the party in power has some resources to tamper with the electoral results in case its share of the vote is not enough to win the election. Such practice has been common in Latin America's history, as seems to have been the case with Noriega in Panama, Fujimori in Peru, Balaguer in the Dominican Republic and the PRI in Mexico.[19]

5.2 Candidates

The candidates only care about winning the election. In particular, they do not have policy preferences of their own: if they are nominated and they win the general election they will implement the policy platform imposed by their parties. In addition to implementing policy x^B, if the candidate L is nominated and wins the general election she will engage in cronyism, patronage and loyal voting in the benefit of B's party leaders. Candidate D would simply implement x^B.

Candidates may differ in their *electability*, where electability refers to a candidate's appeal to voters stemming from personal attributes other than their policy platform or party. In other words, *electability* refers to a candidate's *valence* as introduced by Donald Stokes (1963). In the Latin American context we can think of that valence as the candidate's campaigning abilities, like her charisma, communication skills and reputation for honesty. Conversely, a lack of valence may come from a reputation for corruption and having skeletons in the closet or from having a dull personality. We will denote a candidate's valence by θ, with superscripts $\theta^L, \theta^D, \theta^A, \theta^B$, to denote the valence of candidate L, candidate D, A's nominee and B's nominee respectively.

The valences of candidates L and D are uncertain ex-ante: parties and

[19] Governments have other ways of overturning an unfavorable election than a vote fraud. For example we could interpret ϕ as the probability that a government will successfully impose its candidate with a coup d'état if that candidate loses the election.

voters do not know them for sure before the campaigns.[20] A candidate's valence may be *high*, which we denote by Θ with $\Theta \in (0,1)$, or *low,* which we normalize to 0. Therefore Θ measures the disparity in valence between the most appealing and the least appealing candidates. We interpret Θ as measuring how differentiated the candidates are. In a world where all candidates are perceived to be of uniform "quality" or valence, Θ is small; and in a world where voters perceive large differences among politicians, Θ is large. One implication is that when voters receive more abundant information about the personalities and behavior of candidates, for example through more investigative journalism, we should expect them to be able to better differentiate the candidates, and therefore Θ should increase.

Everyone has a common belief about the uncertainty of L and D's valences. According to that prior belief, a candidate's valence is high with probability π and low with probability $(1-\pi)$. The valence of party A's candidate is also unknown. Given that A's candidate comes from the same pool of politicians everyone has the same prior belief about her valence. Thus θ^L, θ^D and θ^A are thought to be independently and identically distributed. In summary:

$$\begin{aligned} \Pr(\theta^L &= \Theta) = \Pr(\theta^D = \Theta) = \Pr(\theta^A = \Theta) = \pi \\ \Pr(\theta^L &= 0) = \Pr(\theta^D = 0) = \Pr(\theta^A = 0) = (1-\pi) \end{aligned}$$

which imply the following expected electability for each candidate:

$$E\left(\theta^L\right) = E\left(\theta^D\right) = E\left(\theta^A\right) = \pi\Theta$$

We will assume, however, that θ^B and θ^A are perfectly known to voters by the time they decide who to vote for in the *general* election. That is, following Londregan and Romer (1993), we assume that voters will have received enough information during the campaigns, the debates and the media scrutiny to know with certainty the valence of each party's candidate by the time they cast their ballot.

[20] So we assume that party leaders have no information about their precandidates' valences. However it is likely that party leaders actually have some prior idea of their precandidate's appeal to the public, or that they could acquire that information at little cost (for example by commissioning a poll evaluating the precandidates). In that case a primary election would be less attractive than is postulated by my model. But the substance of my results would still hold as long as the primary elections provide better information than leaders previously have (and indeed, a simple poll is unlikely to reveal as much information as a full primary campaign with televised debates and media scrutiny).

5.3 Voters

The electorate cares about the policy implemented by the winning candidate after the election. There is a continuum of voters in the general election, who all have single-peaked, symmetric and linear utility functions over the policy implemented.[21] We will denote voter j's ideal point by x_j. The distribution of voters' ideal points has a median denoted by x_m, but the exact location of that median is uncertain. We call m the median voter, that is the voter (or voters) whose ideal point is x_m. We assume that the median voter's ideal point follows a uniform distribution between x^A and x^B, that is, the median voter in the general electorate could be ideologically located anywhere between party A and party B with equal probability. Given that $x^B = 0$ and $x^A = 1$ this means that $x_m \sim U[0, 1]$.

The voters' evaluation of a candidate depend on that candidate's valence and the weight voters attach to it. The weight that voters place on θ is captured by a parameter γ with $\gamma \in (0, 1)$ .We would expect γ to increase when the media has a deeper penetration, especially television, because that multiplies the effect of a candidate's charisma (through advertisements, interviews, debates etc.).

Then voter j's utility function is

$$U_j(x, \theta) = -|x_j - x| + \gamma\theta \tag{1}$$

There is no abstention and voters never use weakly dominated strategies.

5.4 The nomination technology

Given that B has two possible candidates to choose from (strictly speaking, they are only *pre*-candidates at this stage), its party leaders need a candidate selection process to nominate one of them. We assume that B has the technology to perform one of two possible CSPs.[22]

[21] The assumption of linearity could be relaxed. The utility functions could be concave, convex, or any single-peaked function and the predictions of the model would substantively remain.

[22] A third CSP that is frequently observed is a *delegate assembly*. However in Latin America those delegate assemblies tend to be of one of two types: either they consist of a fair and competitive majority vote, in which case my model captures them as a "primary" where n is small; or they are used to confirm and legitimize a previous decision by the party leaders (Freidenberg (2003)), in which case we can assimilate them to a "smoke-filled room".

1. **Smoke-filled room**: this CSP consists on the party elite simply handpicking the candidate. It is meant to capture an elite arrangement without any consultation to the party membership.

2. **Primary**: this CSP consists on holding a primary election among the n members of B's rank and file, where the candidate winning a majority of votes is nominated.[23]

When B's party leaders decide to hold a primary they are in fact delegating their nomination decision to the party's rank and file. They do so however, knowing that a primary will increase their chances of winning the general election. Indeed, the premise in this section is that a primary has a higher probability than an smoke-filled room of nominating a candidate of high valence. In Section 4 we will show how this premise can actually be derived as a result of party members behavior, rather than just assumed.

But for now we assume that primary elections grant parties an extra probability P of nominating a high-valence candidate. We call P the *primary bonus*, with $P \in (0, 1)$.

$$\Pr\left(\theta^B = \Theta\right)|_{primary} = \Pr\left(\theta^B = \Theta\right)|_{smoke-filled\ room} + P$$

5.5 Timing of the election

Before the election begins, the party leaders and the party members of B perfectly know the values of x^A, x^B, Θ, w, ℓ, ϕ, and n. The ex-ante distributions of x_m, θ^L, θ^D and θ^A are also common knowledge. The timing of the election is the following:

1. B chooses a CSP: smoke-filled room or primary.

2. If B chooses a smoke-filled room, it handpicks one out of its two pre-candidates, L or D.

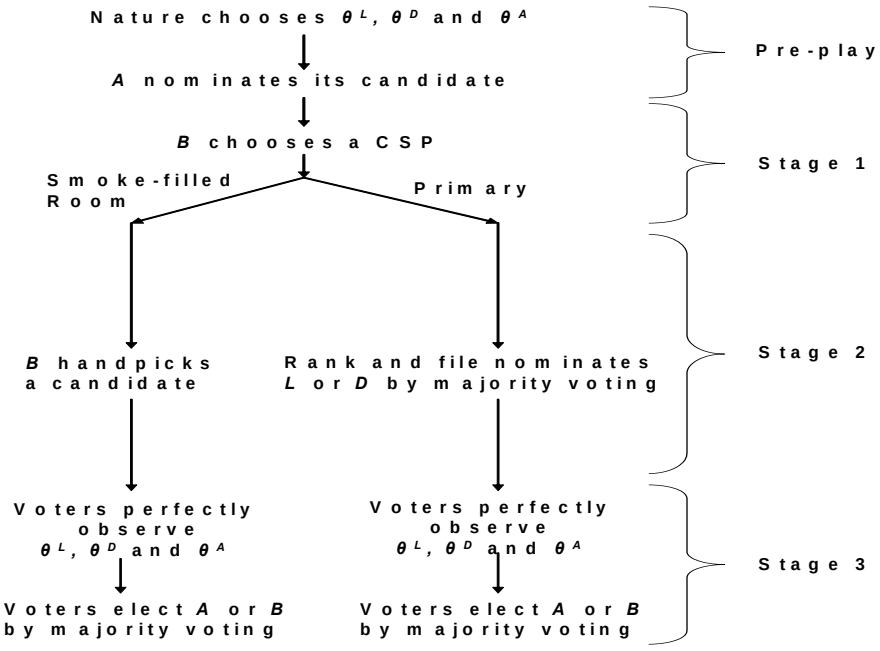
   If B chooses a primary, the n members of B cast their primary vote for L or D. The candidate with most primary votes is nominated.

3. Voters perfectly observe the valences θ^A and θ^B, and vote for one of the two parties, A or B. The party with most votes wins the election.

[23] Given that only the n members of A can vote, this is a model of a *closed* primary rather than an *open* primary. However, n can be assumed to be as large as wanted.

The timing of the election can be seen in more detail in the figure below.

Timing of the election



6 A primary election or a smoke filled room

Being expected-utility maximizers, the party leaders of B will adopt whichever CSP brings them a higher expected payoff. This is exactly what we analyze below by calculating B's payoff under a smoke-filled room, and then its payoff under a primary. We need to solve this game by backward induction and therefore start at the last stage: the general election.

6.1 Third stage: the general election

Given that we are not allowing any abstentions and voters never use weakly dominated strategies, voters will vote as if they were pivotal, that is, they will vote for the party that gives them the highest utility (or flip a coin if they are indifferent between the two).

As a first step in analyzing the voters' behavior, we state the following convenient lemma:

Lemma 1 *Whichever party is preferred by m, the median voter, will win the election.*

That result allows us to ignore all other voters in the general election and just focus on the preferences of the median voter. Thus we can focus on which party is preferred by the median voter m, which in turn depends on where x_m ends up being located in the interval $[0, 1]$.

As a second step we prove that the median voter follows a cut-off rule, meaning that there exists at cut-off policy point $\overline{x}$ such that m will vote for B if $x_m < \overline{x}$, and m will vote for A if $x_m > \overline{x}$ (m would be indifferent if $x_m = \overline{x}$ but that happens with probability zero). This is proved by solving for all the values x_m such that m strictly prefers B to A, which are found by solving the following inequality (remembering that $0 = x^B \leq x_m \leq x^A = 1$)

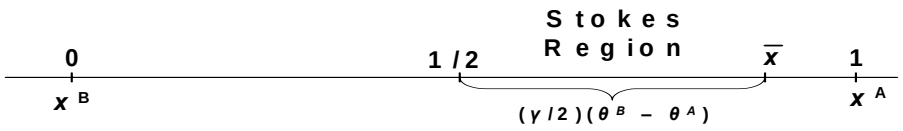
$$\begin{aligned} & U_m(x^A, \theta^A) < U_m(x^B, \theta^B) \\ \Leftrightarrow & -\left|x_m - x^A\right| + \gamma\theta^A < -\left|x_m - x^B\right| + \gamma\theta^B \\ \Leftrightarrow & x_m < \frac{1}{2} + \frac{\gamma}{2}\left(\theta^B - \theta^A\right) \end{aligned}$$

This last inequality implies that m will prefer B to A if and only if $x_m < \overline{x}$ where

$$\overline{x} \equiv \frac{1}{2} + \frac{\gamma}{2}\left(\theta^B - \theta^A\right) \tag{2}$$

Note that if candidates did not have different valences, that is if $\theta^A = \theta^B$, then we would simply have $\overline{x} = \frac{1}{2}$, the midpoint between x^A and x^B just as in a standard median-voter model. However, due to the existence of different valences, we might have the case where the median voter's ideal point is closer to a given party's platform but that party does not win the election because of a relatively low-valence candidate. This is what Groseclose (2001)

called the Stokes region, which we depict in the figure below.



Note that for the allowed values of Θ and γ we have that $\overline{x} \in (0,1)$. Intuitively this condition implies that the effect of valence is never so large that it will determine by itself the winner of the election irrespective of the location of the median voter.

Thus the outcome will be uncertain in spite of the fact that the parties' platforms are fixed and known in advance, and that the candidates' valences are fully revealed by the time voters cast their ballot. Party B cannot calculate the exact payoff from its actions, but must calculate *expected* payoffs based on its probability of winning or losing. Given the uniform distribution of the median voter's ideal point, and remembering that we have normalized the parties platforms to $x^B = 0$ and $x^A = 1$, the probability that B wins the election will be $\overline{x}$ and the probability that A wins the election will be $1 - \overline{x}$. In sum:

$$\text{Probability that } B \text{ wins} = \overline{x} \tag{3}$$

$$\text{Probability that } A \text{ wins} = 1 - \overline{x} \tag{4}$$

6.2 Second stage: the nomination

Under a smoke-filled room regime, B will always handpick L over D. This is because at this stage candidates L and D have identical expected valences. The only difference in case of winning the election, is that L will bring B an extra payoff from loyalty that D will not. So for any positive value of ℓ, B's choice is straightforward: always handpick L. This implies that

$$\Pr(\theta^B = \Theta)|_{smoke-filled\ room} = \Pr(\theta^L = \Theta) = \pi$$

So the expected valence of B's candidate will be the same as the expected valence of L:

$$E(\theta^B)|_{smoke-filled\ room} = E(\theta^L) = \pi\Theta$$

which allows us to calculate the expected cut-off point for the median voter's decision between A and B.

$$E\left(\overline{x}\right)|_{smoke-filled\ room} = E\left(\frac{1}{2} + \frac{\gamma}{2}\left(\theta^B - \theta^A\right)\right) = \frac{1}{2}$$

using the fact that A and B's candidate come from the same pool and therefore have the same expected valences.

Under a primary election regime, B delegates its nomination decision to its rank-and-file members. The primary bonus P implies that

$$\Pr(\theta^B = \Theta)|_{primary} = \Pr(\theta^B = \Theta)|_{smoke-filled\ room} + P = \pi + P$$

which allows to calculate the expected valence of B's candidate after a primary election:

$$\begin{aligned} E(\theta^B)|_{primary} &= \Pr(\theta^B = \Theta)|_{primary}(\Theta) + \Pr(\theta^B = 0)|_{primary}(0) \\ &= (\pi + P)\,\Theta \end{aligned}$$

And this allows us to calculate the expected cut-off point for the median voter's decision between A and B.

$$\begin{aligned} E\left(\overline{x}\right)|_{primary} &= E\left(\frac{1}{2} + \frac{\gamma}{2}\left(\theta^B - \theta^A\right)\right) \\ &= \frac{1}{2} + \frac{\gamma}{2}\left(E\left(\theta^B\right)|_{primary} - E\left(\theta^A\right)\right) \\ &= \frac{1}{2} + \frac{\gamma}{2}P\Theta \end{aligned}$$

6.3 First stage: payoffs for the party elite

The probability that B wins a majority of votes in the election is given by $E\left(\overline{x}\right)$. With that result B can calculate its expected payoff from choosing a smoke-filled room or a primary as its CSP. But we must distinguish whether B is the incumbent party or the opposition party.

6.3.1 B is the incumbent party

If B is the incumbent party and it loses the election it will attempt a fraud against the opposition party A that has probability ϕ of overturning A's victory. B will therefore get elected to office and get the corresponding payoffs under two scenarios: if it wins a majority of votes, or if it does not win majority of votes but its fraud attempt is successful. This allows B to calculate its expected payoff from choosing a smoke-filled room or a primary.

Under a smoke-filled room given that B will always handpick L, its expected payoff is equal to its probability of winning times its benefit from winning with L, plus its probability of losing times its payoff from losing.

$$\begin{aligned} EU^B|_{smoke-filled\ room} &= E(\overline{x})(w+\ell) + (1 - E(\overline{x}))\left[\phi(w+\ell) + (1-\phi)(0)\right] \\ &= \left(\frac{1}{2} + \frac{1}{2}\phi\right)(w+\ell) \end{aligned} \tag{5}$$

B can also calculate its expected payoff from choosing a primary election as its CSP. B's payoff from winning is equal to w, plus ℓ if and only if candidate L gets nominated. Given that the distributions of both L and D's valence are identical, they both have ex-ante the same one-half probability of being nominated. Therefore B's expected payoff from winning office is $w + \frac{\ell}{2}$. It can therefore be calculated that

$$\begin{aligned} EU^B|_{primary} &= E(\overline{x})\left(w + \frac{\ell}{2}\right) + (1 - E(\overline{x}))\left[\phi(w + \frac{\ell}{2}) + (1-\phi)(0)\right] \\ &= \left(\frac{1}{2} + \frac{1}{2}\phi + \frac{\gamma}{2}P\Theta(1-\phi)\right)\left(w + \frac{\ell}{2}\right) \end{aligned} \tag{6}$$

In comparing the equations (5) and (6) we can readily see the trade-off that party B faces in choosing a primary over a smoke-filled room. Looking at the second equation we see that adopting a primary election instead of a smoke-filled room reduces B's expect payoff from loyalty from ℓ to $\frac{\ell}{2}$, but it increases its probability of winning by $\frac{\gamma}{2}P\Theta(1-\phi)$ due to a higher expected candidate valence. In other words *in choosing a primary, party B is trading off some loyalty for some valence.*

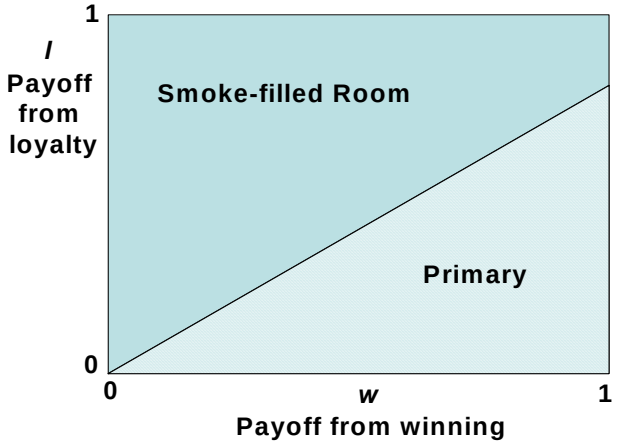
So when does party B decide to adopt a primary over a smoke-filled room? Comparing the payoffs of each different CSP gives us the following result.

Theorem 1 *When B is the incumbent party, it will strictly prefer a primary*

(strictly prefer a smoke-filled room) (be indifferent) if and only if

$$\ell < (>)\,(=)\,\frac{2\gamma P\Theta\,(1-\phi)}{1+\phi-\gamma P\Theta\,(1-\phi)}w$$

These preferences can be depicted in the graph below. The map of all possible values of w and ℓ is divided in two regions by the straight line from the origin whose slope is $\frac{2\gamma P\Theta(1-\phi)}{1+\phi-\gamma P\Theta(1-\phi)}$. Note that whenever party B cares much about winning per se relative to winning with its loyal candidate, that is whenever w is large relative to ℓ, it will choose a primary election. Given that w and ℓ can vary significantly from election to election, we would expect party B's choice of CSP to vary from election to election as well, which explains the Empirical Finding 6.



The area below the line contains all the values of w and ℓ such that B prefers a primary over a smoke-filled room. We will call that area the *likelihood that B chooses a primary.* Conversely we call the *likelihood that B chooses a smoke-filled room* the area above the line.

To answer the questions posed by this paper, what matters is how these likelihoods change with the main parameters of the model γ, Θ, P, π and ϕ. The following theorem describes those changes.

Theorem 2 *When B is the incumbent party, the likelihood that it will choose a primary strictly increases when*

1. *P increases*

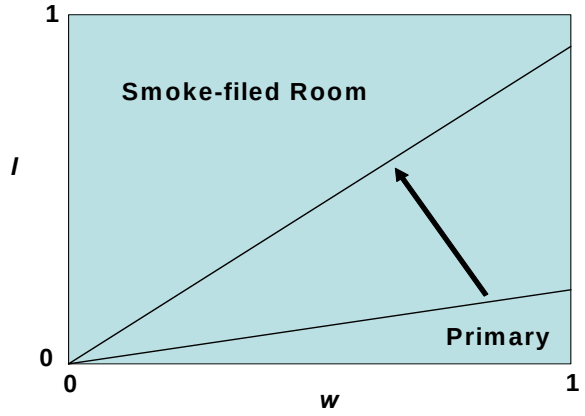
2. γ *increases*

3. Θ *increases*

4. ϕ *decreases*

These comparative statics are depicted in the figure below. The effects of P, γ and Θ are quite intuitive. If primaries are increasingly effective in their purpose of nominating a high-valence candidate (an increase in P), if valence looms larger in the voters' decision (an increase in γ), or if the difference in valences among candidates increases (an increase in Θ), adopting a primary election will be more attractive. Why does a primary become more attractive when ϕ decreases, or conversely why does a smoke-filled room become more attractive when ϕ decreases? The reason comes from the benefit to party B if it receives less votes than the opposition. In that case B will lose the vote but will have probability ϕ of imposing its candidate through an electoral fraud. But if B will impose its candidate, it prefers to have nominated its loyal rather than its disloyal precandidate. Given the probability of nominating its loyal candidate is larger with a smoke-filled room, the ex-ante benefit of a fraud is larger after a smoke-filled room than after a primary. Thus increasing ϕ makes a smoke-filled room more attractive than a primary.

Effect of increasing P, γ, Θ or decreasing ϕ if B is the incumbent
Effect of increasing P, γ or Θ if B is in the opposition



6.3.2 B is the opposition party

If B is the opposition party and it wins a majority of votes it will suffer a fraud attempt by the incumbent party A that has probability ϕ of overturning B's victory. On the other hand if B does not win a majority of votes, its defeat will be respected. These two scenarios allow B to calculate its expected payoff from choosing a smoke-filled room or a primary.

Under a smoke-filled room given that B will always handpick L, its expected payoff is equal to its probability of winning times its benefit from winning with L, plus its probability of losing times its payoff from losing.

$$\begin{aligned} EU^B|_{smoke-filled\ room} &= E(\overline{x})\,(1-\phi)\,(w+\ell) \\ &= \left(\frac{1}{2}-\frac{1}{2}\phi\right)(w+\ell) \end{aligned} \tag{7}$$

B can also calculate its expected payoff from choosing a primary election as its CSP remembering that B's expected payoff from winning office is $w+\frac{\ell}{2}$. It can therefore be calculated that

$$\begin{aligned} EU^B|_{primary} &= E(\overline{x})\,(1-\phi)\left(w+\frac{\ell}{2}\right) \\ &= \left(\frac{1}{2}-\frac{1}{2}\phi+\frac{\gamma}{2}P\Theta\,(1-\phi)\right)\left(w+\frac{\ell}{2}\right) \end{aligned} \tag{8}$$

Equations (7) and (8) allow us to state the following theorem.

Theorem 3 *When B is the opposition party, it will strictly prefer a primary (strictly prefer a smoke-filled room) (be indifferent) if and only if*

$$\ell < (>)\,(=)\,\frac{2\gamma P\Theta}{1-\gamma P\Theta}w$$

How does B's likelihood of adopting a primary change with the main parameters of the model γ, Θ, P, π and ϕ? The following theorem provides the answer.

Theorem 4 *When B is the opposition party, the likelihood that it will choose a primary strictly increases when*

1. *P increases*

2. γ increases

3. Θ increases

B's likelihood of adopting a primary is unaffected by an increase or decrease in ϕ.

These comparative statics are depicted in the graph above. The effects of P, γ and Θ are the same as when B is the incumbent party. But the effect of ϕ is different, perhaps surprisingly so. Why does a change in ϕ not affect B's choice at all? The reason is that a change in ϕ changes the expected utilities from a smoke-filled room and a primary *in the same proportion*, thus leaving the trade-off between them intact.[24] Intuitively when the probability of a successful fraud increases, the opposition party sees its expected utility decrease, but that decrease does not depend on whether it nominated its loyal or disloyal candidate and therefore its preference between a primary and a smoke-filled room remains unaffected.

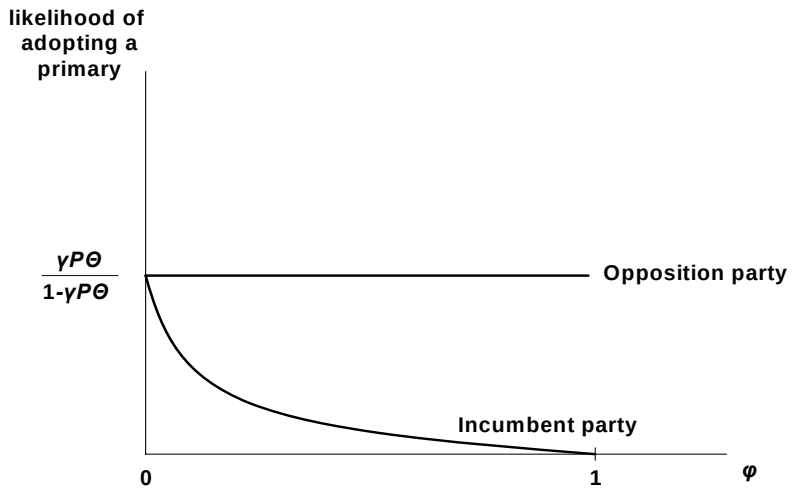
6.3.3 Comparing an incumbent and an opposition party

Comparing Theorems 1 and 3 allows us to determine whether an incumbent party or an opposition party is more likely to adopt a primary.

Theorem 5 *The likelihood that B will choose a primary is strictly larger if it is the opposition than if it is the incumbent party.*

In turn, comparing Theorems 2 and 4 allows to determine the relative frequency with which we would expect to observe primary elections with respect to smoke-filled rooms in a given polity when the probability of a successful fraud is reduced. Concretely the theorems tell us that when ϕ decreases we we should expect an incumbent party to be more inclined to adopt a primary while we should expect an opposition party to remain equally inclined to adopt a primary. Such a pattern is depicted in the figure below.

[24] Note that this result is quite robust: it does not depend on the linearity of voters utilities, nor the additivity of w and ℓ, nor the fact that the the payoff from losing is zero, nor having a uniform distribution of the median voter. The fact that the likelihood of adopting a primary is unaffected by ϕ would remain if voters utilities were convex or concave, if w and ℓ multiplied each other, if the payoff from losing was a non-zero constant, or if we assumed a more general pdf for the median voter's ideal point.



6.4 Empirical predictions

The results in this section give us a series of empirical predictions. The interpretations we suggested for the different parameters naturally translate the results in Theorems 1 through 5 in the following claims.

	Primaries are expected to be more likely when:
Prediction 1	Primaries get better at selecting the most electable candidates
Prediction 2	Voters increase the importance they place on candidates attributes
Prediction 3	The media, in particular TV, increases its penetration
Prediction 4	Journalism becomes more investigative
Prediction 5	Patronage to party elites and cronyism decrease
Prediction 6	Rewards to the party from being in office, e.g. corruption increase
Prediction 7	Electoral fraud decreases
Prediction 8	We consider opposition parties rather than incumbent parties
Prediction 9	Parties care more about winning elections

We can connect some of these predictions to the empirical findings that we listed in Section 3: in particular our model predicts the Empirical Finding 3 (with Prediction 8) and also predict the Empirical Finding 5 (with Prediction 2).

7 Primaries as an information-aggregation mechanism

7.1 Motivation

We have thus far assumed that primaries have an extra probability P with respect to a smoke-filled room of nominating a precandidate with high-valence. We called P the primary bonus and assumed it was larger than zero. In this section we illustrate where P might come from, based on the behavior of individual party members and the mechanism through which they reach a decision in the primary election. In formal-theoretic parlance, we will derive P from "micro-foundations".

We postulate that primary elections can serve as an information-aggregation mechanism to reveal the actual valences of each candidate. Primaries can achieve this goal in two steps. The first is forcing candidates to display their valence inside their party through a primary campaign: the party members will be able to form an opinion on how electable a candidate is by observing her performance in the primary campaign, e.g. during the televised debates and the stories uncovered during the initial media scrutiny. In other words the candidates are forced to send an informative *signal* that will partially reveal their "type". The second step is aggregating all the bits of information contained in the evaluations that thousands or millions of party members have independently made about each candidate. It is the *majority-voting rule* of the primary election that allows aggregating all that information into a well-informed group decision.

7.2 The Bayesian game

We now describe in detail what happens during a primary election, if there is one. Each member i of B receives a couple of private signals s_i^L and s_i^D about L's and D's valence, respectively. Each signal can take one of two values: Θ or 0. Those signals come from the party members' perception of the candidates' actual electability. We assume however that those signals are noisy and do not fully reveal the true valences θ^L and θ^D. We assume that the signals s_i^L and s_i^D are mutually independent of each other, and are also

mutually independent of the signals received by other party members.[25] All signals are identically distributed, and have probability q of being "correct" with $q \in \left(\frac{1}{2}, 1\right)$. That is, we have

$$\begin{aligned} \Pr\left(s_i^L = 0 | \theta^L = 0\right) &= \Pr\left(s_i^L = \Theta | \theta^L = \Theta\right) = q \\ \Pr\left(s_i^L = 0 | \theta^L = \Theta\right) &= \Pr\left(s_i^L = \Theta | \theta^L = 0\right) = 1 - q \end{aligned}$$

and a similar set of equations for s_i^D.

The decision that each party member needs to make is who to vote for in the primary. A *voting strategy* for member i is a mapping $v_i : \left\{s_i^L, s_i^D\right\} \to \{L, D\}$ describing whether i votes for L or votes for D after observing her signals s_i^L and s_i^D (i may also randomize between L and D).

There are three types of voting strategies that are particularly interesting to study because of their intuitive appeal. Following Austen-Smith and Banks (1996), we defined them as: *informative voting*, which consists of voting for the candidate whose signal was highest; *sincere voting*, which consists of voting for the precandidate that would give the highest expected utility; and *rational voting*, which consists of all party members' strategies forming an equilibrium.

Potentially those three types of strategies could lead to different decision and different outcomes for the same set of signals (as has been illustrated for example in Austen-Smith and Banks (1996), Feddersen and Pesendorfer (1998), Duggan and Martinelli (2001)). Fortunately we can prove that in this context those three different strategies coincide. But before stating the lemma with that result, let us define those strategies precisely.

Definition 1 *A voting strategy v_i is called* informative *if*

$$v_i\left(s_i^L, s_i^D\right) = \begin{cases} L \text{ if } s_i^D < s_i^L \\ D \text{ if } s_i^L < s_i^D \\ \text{Randomize equally if } s_i^D = s_i^L \end{cases}$$

In order to define sincere voting we need to introduce a notation for the party members' payoffs, assuming that they all are expected-utility maximizers. We will call $U_i(\theta)$ the utility that a member i of party B derives when her party nominates a precandidate of valence θ.

[25] In this context assuming *independence* is not enough, we need *mutual independence* (Casella and Berger (2002 p. 26)).

Definition 2 *A voting strategy v_i is called* sincere *if*

$$v_i\left(s_i^L, s_i^D\right) = \begin{cases} L \text{ if } E\left(U_i\left(\theta^D\right)|s_i^L, s_i^D\right) < E\left(U_i\left(\theta^L\right)|s_i^L, s_i^D\right) \\ D \text{ if } E\left(U_i\left(\theta^L\right)|s_i^L, s_i^D\right) < E\left(U_i\left(\theta^D\right)|s_i^L, s_i^D\right) \\ \text{Randomize equally if } E\left(U_i\left(\theta^L\right)|s_i^L, s_i^D\right) = E\left(U_i\left(\theta^D\right)|s_i^L, s_i^D\right) \end{cases}$$

In order to define *rational voting*, note that the other two types of strategies, informative and sincere, do not take into account the strategies of the other party members. For the case where party members think about the possible strategies that other party members might be adopting, the most appropriate prediction is that all the strategies will form an equilibrium. Given that the party members have private information, where their signals can be interpreted as their "types", they are playing a Bayesian game. We thus need to derive a Bayesian Nash equilibrium, which requires us to introduce some additional notation. We call v_{-i} the profile of voting strategies adopted by all party members other than i. And we call s_{-i}^L and s_{-i}^D the vectors of signals received by all party members other than i about L's valence and D's valence respectively.

Definition 3 *A profile of voting strategies $v_1^*, ..., v_n^*$ is called* rational *if it forms a Bayesian Nash equilibrium, that is, if for each member i and for each set of signals s_i^L and s_i^D that i might receive, the voting strategy $v_i^*\left(s_i^L, s_i^D\right)$ must solve*

$$\max_{v_i} E\left(U_i\left(\theta\right)|v_i, v_{-i}^*, s_i^L, s_i^D\right)$$

We can now state a useful result that allows us to determine the behavior of party members during this primary election.[26]

Theorem 6 *In this primary election, informative voting is equivalent to sincere voting and it is rational.*

So we can safely assume that party members in this election vote informatively: it is an equilibrium for them to do so.

[26] This lemma can be considered the equivalent in this context to Theorem 1 in Austen-Smith and Banks (1996).

7.3 The information aggregation

The previous result allows us compute the possible outcomes of this primary. What we are interested in calculating is the probability that a primary election will result in the nomination of a high-valence candidate, $\Pr(\theta^B = \Theta)|_{primary}$. We can separate the possible situations that B can face in three cases: (1) both precandidates have low valence, $\theta^L = \theta^D = 0$; (2) both precandidates have high valence, $\theta^L = \theta^D = \Theta$; and (3) one candidate has low valence and the other has high valence, $\theta^L \neq \theta^D$. Cases (1) and (2) are trivial: the probability that B nominates a high-valence candidate are 0 and 1 respectively.

The interesting case is (3), where the valence of the candidate nominated is uncertain. In that case the outcome will come from the behavior of party members when they vote in the primary elections, as well as from the way the primary election processes that behavior into a nomination. Given Theorem 6 we know that party members will vote informatively, that is, they will "vote their signal", which allows us to derive the following result.

Lemma 2 *When one precandidate has low valence and the other has high valence, the probability that a given party member i votes for the high valence precandidate is q.*

Now that we know how the individual voters will vote, we need to know how the primary election will aggregate those votes into a nomination. From Lemma 2 we know that each party member has probability q of voting for the high-valence candidate. Therefore the number of votes that the high-valence precandidate receive in the primary election follows a binomial distribution $B(n, q)$. Furthermore, given that the nomination is decided by majority voting, the probability that the high-valence precandidate is nominated is equal to the probability that she receives more than half of the n votes. The following result follows from that observation.

Lemma 3 *When one precandidate has low valence and the other has high valence, the probability that the high valence precandidate is nominated in the primary election is equal to* $\sum_{k=\frac{n+1}{2}}^{n} \left(\begin{array}{c} n \\ k \end{array} \right) q^k (1-q)^{n-k}$

With this result we can calculate $\Pr(\theta^B = \Theta)|_{primary}$ and compare it to $\Pr(\theta^B = \Theta)|_{smoke-filled\ room}$ to finally calculate P. In other words, we will

be able to know if the primary bonus P which we defined as $P \equiv \Pr(\theta^B = \Theta)|_{primary} - \Pr(\theta^B = \Theta)|_{smoke-filled\ room}$ is positive.

Theorem 7 *The probability that B nominates a high-valence candidate with a primary election is given by* $\Pr(\theta^B = \Theta)|_{primary} = 2\pi(1-\pi)\sum_{k=\frac{n+1}{2}}^{n} f_n(k) + \pi^2$. *The primary bonus is equal to*

$$P = \pi(1-\pi)\left(2\sum_{k=\frac{n+1}{2}}^{n}\binom{n}{k}q^k(1-q)^{n-k} - 1\right)$$

which is such that $P \in \left(0, \frac{1}{4}\right)$.

The previous theorem confirms that the mechanism described above is indeed effective at increasing the expected valence of the party's nominee. Remembering that valence is useful for winning votes in the general election, the previous lemma has an immediate consequence for the odds that B will win the election.

Corollary 1 *Party B has a larger probability of winning a majority of votes when its CSP is a primary rather than a smoke-filled room.*

Note that this corollary directly predicts the Empirical Finding 2.

7.4 Comparative statics

We are now interested in what affects the primary bonus, i.e., what makes a primary election more or less effective at nominating a high-valence candidate. We would like to know how P varies with its main parameters q, n and π.

Remember from Theorems 2 and 4 that the likelihood that B will choose a primary is strictly increasing with P, and therefore any parameter that increases P will also increase the likelihood that B will adopt a primary. The following theorem tell us how the parameters must change to have exactly that effect.

Theorem 8 *The primary bonus P, and therefore the likelihood that B adopts a primary, strictly increases when*

1. q increases

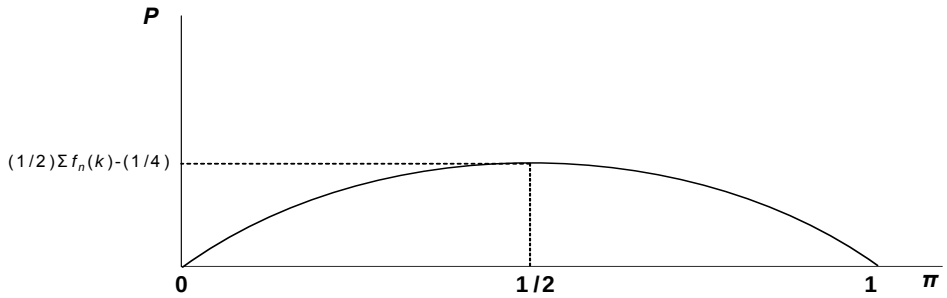
2. n increases

3. The variance of θ^L and θ^D increases

4. π gets closer to $\frac{1}{2}$

The effects of q and n are quite intuitive: both are parameters that improve the effectiveness of a primary as an information aggregation mechanism. By increasing q, the signals received by party members about the valence of candidates is more accurate; in other words the *quality of the information* conveyed by the primaries increases. By increasing n there is a larger number of party members receiving and processing that information; in other words the *quantity of information* aggregated by the primaries increases.

But the effects of π and the variance of skills might seem surprising, and require more elaboration. They point to the fact that primaries are more useful when the pool of candidates is more heterogeneous. This is better visualized with the figure below which graphs the non-monotonic relationship between π and P. For low values of π the pool of candidates is very homogenous (almost all of them having a low valence), and choosing a smoke-filled room is very likely to result in a candidate of the same valence as the one resulting from a primary. This reduces the benefit of a primary relative to a smoke-filled room. The same can be said for large values of π – but the cost of adopting a primary remains the same. In contrast for values of π close to $\frac{1}{2}$ the respective valences of L and D are very unpredictable and that is when an information-aggregation mechanism such as a primary is the most beneficial. Therefore primaries have the most to offer to party leaders when the pool of candidates is the most heterogeneous, or equivalently, when the variance of valences θ^L and θ^D is the largest. On the other hand, when the pool of candidates is homogeneous, as smoke-filled room does almost as good a job as a primary in selecting the most electable candidates – and thus the trade-off of choosing a primary is not as attractive.



7.5 Empirical predictions

The model in this section gives us a series of empirical predictions. The interpretations we suggested for the different parameters in this paper allow us translate the results in Theorems 6, 7 and Corollary 1 in the following claims.

Primaries are expected to be more likely when:

Prediction 10 The quality of information conveyed in primary campaigns is larger

Prediction 11 The number of party members voting in the primary increases

Prediction 12 The heterogeneity of candidate valences increases

We can connect some of these predictions to the empirical finding in the previous literature that we listed in 3. Empirical Finding 2 is consistent with Prediction 9, and Empirical Finding 4 is consistent with Prediction 11.

8 Historical trends in Latin American politics

We now use the theory developed in the previous sections to answer the question that motivated this paper: why have primary elections proliferated in Latin America in the past two decades? We answer by linking the likelihood that a party adopts a primary versus a smoke-filled room according to

the previous model, to some features of Latin American elections that have shown a consistent pattern in the last three decades. This paper is not the place for a cross-country analysis of how those variables correlate with the adoption of primary elections, an endeavor that is left for future research; but the aggregate trends in the following variables are indicative of the possible causes for the proliferation of primaries in the region.

1. **The media has increased its penetration, especially television:** An increasing number of households have access to radio and television broadcasting in Latin America and the Caribbean: in the period 1980-1996 the number of TV receivers per one thousand inhabitants increased from 97 to 204, and radio receivers from 259 to 413 (Fox and Waisbord (2002)). Modern campaigns now include televised debates, interviews and advertisements that have shaped the kind of candidates that win elections (Adler (1993), de Lima (1993), Skidmore (1993), Zuleta-Puceiro (1993)).[27] The implication is that the weight voters place in the personal attributes of candidates like their charisma, communication skills and physical appearance has been magnified. This translates in an increase in γ in our model

2. **Journalism has become more investigative**: In the past decades the media in Latin America has become more free and more professional, thus investigating political candidates more aggressively and thoroughly (Waisbord (2000), Lawson (2002)). The implication is that voters revive more profound and more abundant information about the candidates. As a consequence both their valence-increasing behavior (such a as church attendance, charity donations etc.) and their valence-decreasing behavior (corruption, scandals) will surface, which allows candidates to be differentiated. This translates in an increase in Θ in our model.

3. **Elections have become more fair and transparent**: In the past three decades, as authoritarianism has given way to democratization, Latin American countries have seen an increase in the competitiveness

[27] A similar trend happened in the United States before the expansion of presidential primary elections. In the 1950s and 1960s the mass media of communication became increasingly important actors in the nominating process, particularly television. This favored those candidates with popular appeal among voters, who might otherwise have been ignored by party leaders (Bartels 1988 pp. 17-18, 24).

of their elections. Reforms have been undertaken which have created electoral institutions and legislation. And the international community has paid more attention to elections with NGOs such as the Carter Center. That has decreased the incidence of electoral fraud (Hartlyn (1998), Eisenstadt (2004)) which translates in a decrease in ϕ in our model.

All these trends in the past three decades have shifted the party leaders' trade-off in choosing a CSP in favor of a primary over an elite arrangement. Indeed according to the theory in this paper, these historical trends have shifted some parameters exactly in the direction that increases the likelihood that a party will adopt a primary, which allows us to state the following predictions:

	We should expect primaries to:
Prediction 13	Have become more frequent in the past three decades
Prediction 14	Become even more frequent in the coming decades

Note that Prediction 13 explains the Empirical Finding 1, which is the phenomenon that motivated this paper: the proliferation of primary elections in the past two decades.

9 Conclusions

We have elaborated a formal model that predicts some of the features observed in Latin American primary elections in the recent decades. Most importantly, it provides an explanation for the endogenous adoption of primary elections by traditionally non-democratic party leaders. We have characterized endogenous party democratization as a puzzle: why would party leaders voluntarily forgo their control of the CSP by adopting a primary election that they cannot control? The explanation focuses on the simple argument that primary elections serve to reveal the information about the valences of different candidates. This benefit from primary elections has become more attractive with the changes of Latin American politics in the last two decades.

The comparative statics of our model suggest why party democratization followed government democratization in Latin America: as the ability to commit fraud decreased it became more important for incumbent parties to have electable candidates. That created incentives for party leaders to select good campaigners instead of loyal cronies. Primary elections provided the

information-revelation mechanism to insure the highest valence candidate got the nomination. Thus we see how the democratization of elections *between* parties led to the democratization of nominations *within* parties.

We therefore provide a formal-theoretic account, the first one that I am aware of, for the proliferation of primary elections in Latin America during the past quarter century. However, several extensions are in order. Future research should account for increasingly complex situations, such as candidates switching parties, and the possibility that loser precandidates form a third party.

There is a normative debate about the benefits and costs of adopting primaries.[28] The model developed in this paper contributes to this debate by calling attention to two social benefits of primaries: an increased valence of the candidates, understood as a candidate's appeal to voters; and a larger independence of candidates vis-à-vis their party leaders which should translate in reduced patronage, cronyism and partisan voting.

By pointing to these normative benefits, the paper also addresses an important policy debate that is taking place throughout Latin America: Should candidate selection be regulated by the state, as it is for example in Costa Rica, or should it be left to parties themselves?[29] If regulation is decided, should primaries be made compulsory, as they are for example in Uruguay?[30] The social benefits underscored by this paper provide an argument in favor of making primaries compulsory for political parties.

The significance of this question cannot be understated. Political parties are key institutions in the consolidation of democracy: they have come to be seen, especially in developing countries, as guarantors of free elections against a rise of authoritarianism. Given the manifest fragility of democratic institutions in Latin America, a good comprehension of when parties become democratic is much desired and long overdue.

[28] For an overview of these normative arguments see Scarrow (2005).

[29] For example, in Mexico a prominent left-wing ideologue proposed that the Federal Electoral Institute should organize the parties' selection of candidates (*El Universal* (2003)).

[30] Such is the proposal of the new president of Ecuador, Rafael Correa, who wants to call a referendum to require parties to hold primaries (*The Economist* (2006)).

10 Appendix

10.1 Proof of Lemma 1

Proof of Lemma 1. We must show that there exists a median voter whose vote is decisive in determining which party will win. This is shown by considering two arbitrary voters i and j, such that $x_i < x_j$. By manipulating the equations above it is easy to prove that if voter j prefers party B to party A then so does voter i, and if i prefers A to B then so does j. What this implies for the median voter m is that if m prefers B to A then so do all the voters to her left – which amount to at least 50%-plus-one voters – thus allowing B to win the election. And if m prefers A to B then so do all the voters to her right – which amount to at least 50% plus one voters – thus allowing A to win the election. Therefore whichever party is preferred by the median voter m will win the election. If m was located exactly at the mid-point between A and B then she wold be indifferent between A and B, but that happens with probability zero. ■

10.2 Proof of Theorem 1

Proof of Theorem 1. B will strictly prefer a primary over a smoke-filled room whenever its expected payoff from adopting the former exceeds that of adopting the latter. Using the Equations (5) and (6), B will adopt a primary whenever

$$\left(\frac{1}{2}+\frac{1}{2}\phi\right)(w+\ell) < \left(\frac{1}{2}+\frac{1}{2}\phi+\frac{\gamma}{2}\Theta P\left(1-\phi\right)\right)\left(w+\frac{l}{2}\right)$$

Solving for ℓ in the above inequality we obtain:

$$\ell < \frac{2\gamma P\Theta\left(1-\phi\right)}{1+\phi-\gamma P\Theta\left(1-\phi\right)}w$$

and similar equations can be obtained when B strictly prefers a smoke-filled room or is indifferent. The line defined by $\ell = \frac{2\gamma P\Theta(1-\phi)}{1+\phi-\gamma P\Theta(1-\phi)}w$ is the *indifference curve* between a primary and A smoke-filled room. ■

10.3 Proof of Theorem 2

Proof of Theorem 2. First note that the area below the straight line is strictly increasing with the slope of the line, therefore to show that the likelihood that B will choose a primary increases we only need to show that the slope of the line increases. We start by noting that this slope is always strictly positive which comes from the fact that all the terms in the numerator are strictly positive, and because $\gamma\Theta P(1-\phi) < 1 < 1+\phi$ which implies that the denominator is strictly positive.

To prove the effect of P we need to differentiate the value of the slope verify that the sign of the differential is positive. Indeed we have

$$\frac{\partial\left(\frac{2\gamma P\Theta(1-\phi)}{1+\phi-\gamma P\Theta(1-\phi)}\right)}{\partial P} = \frac{\left[2\gamma\Theta(1-\phi)\right]\left[1+\phi-\gamma P\Theta(1-\phi)\right]+\left[2\gamma P\Theta(1-\phi)\right]\left[\gamma\Theta(1-\phi)\right]}{\left[1+\phi-\gamma P\Theta(1-\phi)\right]^2}$$

where all the bracketed terms are strictly positive, and therefore the whole expression is strictly positive.

We obtain very similar expression by differentiating the slope with respect to γ and Θ which give strictly positive differentials. Similarly straightforward calculations show that the differential of the slope with respect to ϕ is strictly negative. ■

10.4 Proof of Theorem 3

Proof of Theorem 3. B will strictly prefer a primary over a smoke-filled room whenever its expected payoff from adopting the former exceeds that of adopting the latter. Using the Equations (7) and (8), B will adopt a primary whenever

$$\left(\frac{1}{2}-\frac{1}{2}\phi\right)(w+\ell) < \left(\frac{1}{2}-\frac{1}{2}\phi+\frac{\gamma}{2}P\Theta(1-\phi)\right)\left(w+\frac{\ell}{2}\right)$$

Solving for ℓ in the above inequality we obtain:

$$\ell < \frac{2\gamma P\Theta}{1-\gamma P\Theta}w$$

and similar equations can be obtained when B strictly prefers a smoke-filled room or is indifferent. ■

10.5 Proof of Theorem 4

Proof of Theorem 4. The effects P, Θ and γ are calculated exactly as for Theorem 2. Straightforward differentiation proves that the derivatives of the slope with respect to those variables are strictly positive.

To prove ϕ has no effect note that ϕ does not appear in the slope's expression: it dropped out of the comparison of expected utilities. Therefore the differentiation of the slope with respect to ϕ gives zero. ■

10.6 Proof of Theorem 5

Proof of Theorem 5. We must prove that the slope of the indifference curve is larger when B is the opposition than when it is the incumbent party. That is we must prove that

$$\begin{aligned} \frac{2\gamma P\Theta}{1-\gamma P\Theta} &< \frac{2\gamma P\Theta\left(1-\phi\right)}{1+\phi-\gamma P\Theta\left(1-\phi\right)} \\ \left(2\gamma P\Theta\right)\left(1+\phi-\gamma P\Theta\left(1-\phi\right)\right) &< \left(2\gamma P\Theta\left(1-\phi\right)\right)\left(1-\gamma P\Theta\right) \\ -\phi &< \phi \end{aligned}$$

which immediately follows with some algebra form the fact that $0<\phi$. ■

10.7 Proof of Theorem 6

Proof of Theorem 6. We will start by proving the first part of the theorem, that sincere and informative voting are equivalent. The first step is proving that i votes for L with certainty under sincere voting if and only if she would also vote for L with certainty under informative voting, that is, we want to prove that $E\left(U_i\left(\theta^D\right)|s_i^L,s_i^D\right)<E\left(U_i\left(\theta^L\right)|s_i^L,s_i^D\right)\Longleftrightarrow s_i^D<s_i^L$. For that we need to state the following four remarks. Remark 1 comes directly form the definition of $\overline{x}$ and expected payoffs of party members. Remarks 2 and 3 are straightforward applications of Bayes rule to the posterior beliefs of party members after receiving their signals. Remark 4 comes from easy algebra.

Remark 1 $U_i\left(\theta\right)=\begin{cases}\left[\left(\frac{1}{2}+\frac{\gamma}{2}\left(\theta-\theta^A\right)\right)\left(1-\phi\right)+\phi\right]w & \textit{if } B \textit{ is the incumbent party}\\ \left(\frac{1}{2}+\frac{\gamma}{2}\left(\theta-\theta^A\right)\right)\left(1-\phi\right)w & \textit{if } B \textit{ is the opposition party}\end{cases}$

Remark 2 $P\left(\theta^L=\Theta|s_i^L=0\right)=P\left(\theta^D=\Theta|s_i^D=0\right)=\frac{\pi(1-q)}{\pi(1-q)+(1-\pi)q}$

Remark 3 $P\left(\theta^L = \Theta | s_i^L = \Theta\right) = P\left(\theta^D = \Theta | s_i^D = \Theta\right) = \frac{\pi q}{\pi q + (1-\pi)(1-q)}$

Remark 4 $\frac{\pi(1-q)}{\pi(1-q)+(1-\pi)q} < \frac{\pi q}{\pi q+(1-\pi)(1-q)}$ *for* $q > \frac{1}{2}$

These results allow us to derive the following equivalences:

$$\begin{aligned}
&E\left(U_i\left(\theta^D\right)|s_i^L, s_i^D\right) < E\left(U_i\left(\theta^L\right)|s_i^L, s_i^D\right) \\
\Leftrightarrow\ &E\left(\theta^D | s_i^L, s_i^D\right) < E\left(\theta^L | s_i^L, s_i^D\right) \text{ given Remark 1} \\
\Leftrightarrow\ &E\left(\theta^D | s_i^D\right) < E\left(\theta^L | s_i^L\right) \text{ given mutual independence of the signals} \\
\Leftrightarrow\ &P\left(\theta^D = \Theta | s_i^D\right) < P\left(\theta^L = \Theta | s_i^L\right) \\
\Leftrightarrow\ &s_i^D < s_i^L \text{ from Remarks 2, 3 and 4}
\end{aligned}$$

Following identical steps we can prove that $E\left(U_i\left(\theta^D\right)|s_i^L, s_i^D\right) > E\left(U_i\left(\theta^L\right)|s_i^L, s_i^D\right) \Longleftrightarrow s_i^D > s_i^L$ and that $E\left(U_i\left(\theta^D\right)|s_i^L, s_i^D\right) = E\left(U_i\left(\theta^L\right)|s_i^L, s_i^D\right) \Longleftrightarrow s_i^D = s_i^L$. So the conditions that determine sincere voting and informative voting are equivalent.

We now prove the second part of the theorem, that informative voting is rational. That is, we need to prove that it is a Bayesian Nash equilibrium for all party members to vote informatively. To do so, we assume that all party members other than i are voting informatively we, and we prove that it is optimal for i to vote informatively. We call v_{-i}^* the profile of strategies of all party members other than i when they are voting informatively.

We also introduce the following notation: we define n_{-1}^L the number of votes that party members other than i cast for L, and n_{-1}^D the number of votes that party members other than i cast for D. Not that having $n_{-1}^L = n_{-1}^D$ means that i's vote is *pivotal*.

We will prove now that, assuming v_{-i}^*, party member i has a strictly lower expected utility from voting for D than voting for L if and only if she received a strictly lower signal for D than for L. That is, we want to show that $E\left(U_i\left(\theta\right)|D, v_{-i}^*, s_i^L, s_i^D\right) < E\left(U_i\left(\theta\right)|L, v_{-i}^*, s_i^L, s_i^D\right) \Longleftrightarrow s_i^D < s_i^L$. The the proof will use the following remarks. Remarks 5 and 6 come from the law of total probability, and form the fact that i's vote cannot influence the outcome when $n_{-1}^L < n_{-1}^D$ or $n_{-1}^L > n_{-1}^D$. Remark 7 is true because the realizations of θ^L and θ^D do not depend on the strategies that party members decided to adopt. Remark 8 comes from the fact that party members are voting informatively, the prior distributions of θ^L and θ^D are identical, and the prior distribution of signals s_i^L and s_i^D are also identical. Remark 9 is

derived directly form the distribution of signals. Remark 10 does not require explanation.

Remark 5 $E\left(U_i\left(\theta\right)|D, v_{-i}^*, s_i^L, s_i^D\right)$
$= E\left(U_i\left(\theta^D\right)|v_{-i}^*, s_i^L, s_i^D, n_{-1}^L < n_{-1}^D\right) P\left(n_{-1}^L < n_{-1}^D|v_{-i}^*, s_i^L, s_i^D\right)$
$+E\left(U_i\left(\theta^L\right)|v_{-i}^*, s_i^L, s_i^D, n_{-1}^L > n_{-1}^D\right) P\left(n_{-1}^L > n_{-1}^D|v_{-i}^*, s_i^L, s_i^D\right)$
$+E\left(U_i\left(\theta^D\right)|v_{-i}^*, s_i^L, s_i^D, n_{-1}^L = n_{-1}^D\right) P\left(n_{-1}^L = n_{-1}^D|v_{-i}^*, s_i^L, s_i^D\right)$

Remark 6 $E\left(U_i\left(\theta\right)|L, v_{-i}^*, s_i^L, s_i^D\right)$
$= E\left(U_i\left(\theta^D\right)|v_{-i}^*, s_i^L, s_i^D, n_{-1}^L < n_{-1}^D\right) P\left(n_{-1}^L < n_{-1}^D|v_{-i}^*, s_i^L, s_i^D\right)$
$+E\left(U_i\left(\theta^L\right)|v_{-i}^*, s_i^L, s_i^D, n_{-1}^L > n_{-1}^D\right) P\left(n_{-1}^L > n_{-1}^D|v_{-i}^*, s_i^L, s_i^D\right)$
$+E\left(U_i\left(\theta^L\right)|v_{-i}^*, s_i^L, s_i^D, n_{-1}^L = n_{-1}^D\right) P\left(n_{-1}^L = n_{-1}^D|v_{-i}^*, s_i^L, s_i^D\right)$

Remark 7 $P\left(\theta^L = \Theta|v_{-i}^*\right) = P\left(\theta^D = \Theta|v_{-i}^*\right)$ *and* $P\left(\theta^L = 0|v_{-i}^*\right) = P\left(\theta^D = 0|v_{-i}^*\right)$

Remark 8 $P\left(n_{-1}^L = n_{-1}^D|v_{-i}^*, \theta^L = \Theta, \theta^D = 0\right) = P\left(n_{-1}^L = n_{-1}^D|v_{-i}^*, \theta^L = 0, \theta^D = \Theta\right)$

Remark 9 $P\left(s_i^L = 0, s_i^D = 0|\theta^L = 0, \theta^D = \Theta\right) = P\left(s_i^L = 0, s_i^D = 0|\theta^L = \Theta, \theta^D = 0\right) = q\left(1-q\right)$
$P\left(s_i^L = \Theta, s_i^D = \Theta|\theta^L = 0, \theta^D = \Theta\right) = P\left(s_i^L = \Theta, s_i^D = \Theta|\theta^L = \Theta, \theta^D = 0\right) = q\left(1-q\right)$
$P\left(s_i^L = \Theta, s_i^D = 0|\theta^L = 0, \theta^D = \Theta\right) = P\left(s_i^L = 0, s_i^D = \Theta|\theta^L = \Theta, \theta^D = 0\right) = \left(1-q\right)^2$
$P\left(s_i^L = 0, s_i^D = \Theta|\theta^L = 0, \theta^D = \Theta\right) = P\left(s_i^L = \Theta, s_i^D = 0|\theta^L = \Theta, \theta^D = 0\right) = q^2$

Remark 10 $\left(1-q\right)^2 < q\left(1-q\right) < q^2$ *for* $q > \frac{1}{2}$

$$E\left(U_i\left(\theta\right)|D, v^*_{-i}, s^L_i, s^D_i\right) < E\left(U_i\left(\theta\right)|L, v^*_{-i}, s^L_i, s^D_i\right)$$
$$\Leftrightarrow E\left(U_i\left(\theta^D\right)|v^*_{-i}, s^L_i, s^D_i, n^L_{-1} = n^D_{-1}\right) < E\left(U_i\left(\theta^L\right)|v^*_{-i}, s^L_i, s^D_i, n^L_{-1} = n^D_{-1}\right)$$
from comparing Remarks 5 and 6,
$$\Leftrightarrow E\left(\theta^D|v^*_{-i}, s^L_i, s^D_i, n^L_{-1} = n^D_{-1}\right) < E\left(\theta^L|v^*_{-i}, s^L_i, s^D_i, n^L_{-1} = n^D_{-1}\right) \text{ given Remark 1}$$
$$\Leftrightarrow P\left(\theta^D = \Theta|v^*_{-i}, s^L_i, s^D_i, n^L_{-1} = n^D_{-1}\right) < P\left(\theta^L = \Theta|v^*_{-i}, s^L_i, s^D_i, n^L_{-1} = n^D_{-1}\right)$$
$$\Leftrightarrow \frac{P\left(s^L_i, s^D_i, n^L_{-1} = n^D_{-1}|v^*_{-i}, \theta^D = \Theta\right) P\left(\theta^D = \Theta|v^*_{-i}\right)}{P\left(s^L_i, s^D_i, n^L_{-1} = n^D_{-1}|v^*_{-i}\right)}$$
$$< \frac{P\left(s^L_i, s^D_i, n^L_{-1} = n^D_{-1}|v^*_{-i}, \theta^L = \Theta\right) P\left(\theta^L = \Theta|v^*_{-i}\right)}{P\left(s^L_i, s^D_i, n^L_{-1} = n^D_{-1}|v^*_{-i}\right)}$$
using Bayes rule,
$$\Leftrightarrow P\left(s^L_i, s^D_i, n^L_{-1} = n^D_{-1}|v^*_{-i}, \theta^D = \Theta\right) < P\left(s^L_i, s^D_i, n^L_{-1} = n^D_{-1}|v^*_{-i}, \theta^L = \Theta\right)$$
given the symmetry between θ^D and θ^L,
$$\Leftrightarrow P\left(s^L_i, s^D_i|n^L_{-1} = n^D_{-1}, v^*_{-i}, \theta^D = \Theta\right) P\left(n^L_{-1} = n^D_{-1}|v^*_{-i}, \theta^D = \Theta\right)$$
$$< P\left(s^L_i, s^D_i|n^L_{-1} = n^D_{-1}, v^*_{-i}, \theta^L = \Theta\right) P\left(n^L_{-1} = n^D_{-1}|v^*_{-i}, \theta^L = \Theta\right)$$
by definition of conditional probability,
$$\Leftrightarrow P\left(s^L_i, s^D_i|\theta^L = 0, \theta^D = \Theta\right) P\left(n^L_{-1} = n^D_{-1}|v^*_{-i}, \theta^L = 0, \theta^D = \Theta\right) P\left(\theta^L = 0|v^*_{-i}\right)$$
$$+P\left(s^L_i, s^D_i|\theta^L = \Theta, \theta^D = \Theta\right) P\left(n^L_{-1} = n^D_{-1}|v^*_{-i}, \theta^L = \Theta, \theta^D = \Theta\right) P\left(\theta^L = \Theta|v^*_{-i}\right)$$
$$< P\left(s^L_i, s^D_i|\theta^L = \Theta, \theta^D = 0\right) P\left(n^L_{-1} = n^D_{-1}|v^*_{-i}, \theta^L = \Theta, \theta^D = 0\right) P\left(\theta^D = 0|v^*_{-i}\right)$$
$$+P\left(s^L_i, s^D_i|\theta^L = \Theta, \theta^D = \Theta\right) P\left(n^L_{-1} = n^D_{-1}|v^*_{-i}, \theta^L = \Theta, \theta^D = \Theta\right) P\left(\theta^D = \Theta|v^*_{-i}\right)$$
by the law of total probability,
$$\Leftrightarrow P\left(s^L_i, s^D_i|\theta^L = 0, \theta^D = \Theta\right) < P\left(s^L_i, s^D_i|\theta^L = \Theta, \theta^D = 0\right)$$
using Remarks 7 and 8,
$$\Leftrightarrow s^D_i < s^L_i \text{ from Remarks 9 and 10}$$

Following identical steps we can prove that $E\left(U_i\left(\theta\right)|D, v^*_{-i}, s^L_i, s^D_i\right) > E\left(U_i\left(\theta\right)|L, v^*_{-i}, s^L_i, s^D_i\right) \iff s^D_i > s^L_i$ and that $E\left(U_i\left(\theta\right)|D, v^*_{-i}, s^L_i, s^D_i\right) = E\left(U_i\left(\theta\right)|L, v^*_{-i}, s^L_i, s^D_i\right) \iff s^D_i = s^L_i$. So given v^*_{-i}, for any pair of signals s^D_i and s^L_i it is optimal for i to vote informatively. Therefore informatively voting by all party members is a Bayesian Nash equilibrium. ■

10.8 Proof of Lemma 2

Proof of Lemma 2. We need to calculate the probability that whenever there is a low-valence and a high-valence precandidate, an individual party member i will vote for the latter. Without loss of generality assume that $\theta^L = 0$ and $\theta^D = \Theta$ and let us calculate the probability that i votes for D.

$$\begin{aligned}
\Pr(i \text{ votes for } D|\theta^L &= 0, \theta^D = \Theta) = \\
&\Pr\left(s_i^L = 0 \text{ and } s_i^D = \Theta|\theta^L = 0, \theta^D = \Theta\right) \\
&+\frac{1}{2}\Pr\left(s_i^L = 0 \text{ and } s_i^D = 0|\theta^L = 0, \theta^D = \Theta\right) \\
&+\frac{1}{2}\Pr\left(s_i^L = \Theta \text{ and } s_i^D = \Theta|\theta^L = 0, \theta^D = \Theta\right) \\
&= (q)(q) + \frac{1}{2}(q)(1-q) + \frac{1}{2}(1-q)(q) \\
&= q
\end{aligned}$$

■

10.9 Proof of Lemma 3

Proof of Lemma 3. We need to calculate the probability that whenever there is a low-valence and a high-valence precandidate, the latter will be nominated in the primary. Without loss of generality let us consider the case where $\theta^L = 0$ and $\theta^D = \Theta$ and let us calculate the probability that D is nominated. We will denote by k the number of party members who vote for the high-valence candidate D in the primary election, and we will denote by $f_n(k)$ its corresponding probability distribution function. We know from Lemma 2 that the probability that any single party member votes for D is q. Given that all the party members' votes are independent (because they only depend on their signals which are themselves independent), we have that k, the number of votes for the candidate of valence Θ, follows a binomial distribution $B(n, q)$. Therefore

$$f_n(k) = \left(\begin{array}{c} n \\ k \end{array}\right) q^k (1-q)^{n-k}$$

We know that D will be nominated if she gets a majority of votes. Given that n is odd this implies $k \geq \frac{n+1}{2}$. So the probability that the high-valence candidate D is nominated is equal to the sum of probabilities $f_n(k)$ for all values of k larger than or equal to $\frac{n+1}{2}$, that is, $\sum_{k=\frac{n+1}{2}}^{n} f_n(k)$. ■

10.10 Proof of Theorem 7

Proof of Theorem 7. We can now calculate the probability that party B will nominate a high-valence candidate by majority voting in the primary. If $\theta^L = 0$ and $\theta^D = 0$ that probability is obviously zero. If $\theta^L = \Theta$ and $\theta^D = \Theta$ that probability is obviously one. If $\theta^L = 0$ and $\theta^D = \Theta$, or $\theta^L = \Theta$ and $\theta^D = 0$, the probability depends is given by Lemma 3.

The table below summarizes these calculations:

Probability that B nominates a candidate of valence Θ:

	$\theta^D = 0$	$\theta^D = \Theta$
$\theta^L = 0$	0	$\sum_{k=\frac{n+1}{2}}^{n} f_n(k)$
$\theta^L = \Theta$	$\sum_{k=\frac{n+1}{2}}^{n} f_n(k)$	1

Looking at this table of conditional probabilities, we can calculate the unconditional probability that a candidate of high valence is nominated, by noting that each of the four cases in the box has the following probabilities of arising:

$$\begin{aligned}
\Pr\left(\theta^L = 0, \theta^D = 0\right) &= (1-\pi)^2 \\
\Pr\left(\theta^L = 0, \theta^D = \Theta\right) &= \pi(1-\pi) \\
\Pr\left(\theta^L = \Theta, \theta^D = 0\right) &= \pi(1-\pi) \\
\Pr\left(\theta^L = \Theta, \theta^D = \Theta\right) &= \pi^2
\end{aligned}$$

which implies the following result:

$$\Pr(\theta^B = \Theta)|_{primary} = 2\pi(1-\pi)\sum_{k=\frac{n+1}{2}}^{n} f_n(k) + \pi^2$$

Remembering that $\Pr(\theta^B = \Theta)|_{smoke-filled\ room} = \pi$ we can calculate P to be

$$\begin{aligned}
P &= 2\pi(1-\pi)\sum_{k=\frac{n+1}{2}}^{n} f_n(k) + \pi^2 - \pi \\
&= \pi(1-\pi)\left(2\sum_{k=\frac{n+1}{2}}^{n} f_n(k) - 1\right)
\end{aligned}$$

which is exactly what the theorem says.

We now want to verify that $P \in \left(0, \frac{1}{4}\right)$. Proving that $P < \frac{1}{4}$ can be done easily by using the fact that $f_n(k)$ is a pdf, and therefore $\sum_{k=\frac{n+1}{2}}^{n} f_n(k) < 1$.

Proving that $P > 0$ hinges on proving that $\sum_{k=\frac{n+1}{2}}^{n} f_n(k) > \frac{1}{2}$, which in turn can easily be proved using the binomial theorem and the fact that $\left(\begin{array}{c} n \\ k \end{array} \right) = \left(\begin{array}{c} n \\ n-k \end{array} \right)$. ■

10.11 Proof of Corollary 1

Proof of Corollary 1. Remember that the probability that B wins a majority of votes is $E(\overline{x})$. For a smoke-filled room this is equal to $E(\overline{x})|_{smoke-filled\ room} = \frac{1}{2}$, whereas for a primary it is equal to $E(\overline{x})|_{primary} = \frac{1}{2} + \frac{\gamma}{2} P\Theta$. Given that we just proved for Theorem 6 that $P > 0$, we immediately see that $E(\overline{x})|_{smoke-filled\ room} < E(\overline{x})|_{primary}$. ■

10.12 Proof of Theorem 8

Proof of Theorem 8. To prove that P is strictly increasing in q we need to prove that $\frac{\partial P}{\partial q}$ is strictly positive. Form the formula for P we see that $\frac{\partial P}{\partial q} = \frac{\partial}{\partial q} \sum_{k=\frac{n+1}{2}}^{n} f_n(k)$. To prove that this is positive we need the following remarks. Remark 11 comes from differentiating the expression $\sum_{k=0}^{\frac{n-1}{2}} f_n(k) + \sum_{k=\frac{n+1}{2}}^{n} f_n(k)$ and noting that the result must be equal to zero because $\sum_{k=0}^{n} f_n(k) = 1$ for any q. Remarks 12, 13 and 14 come from straightforward algebra.

Remark 11 $\frac{\partial}{\partial q} \sum_{k=\frac{n+1}{2}}^{n} f_n(k) > 0 \Longleftrightarrow \frac{\partial}{\partial q} \sum_{k=0}^{\frac{n-1}{2}} f_n(k) < \frac{\partial}{\partial q} \sum_{k=\frac{n+1}{2}}^{n} f_n(k)$

Remark 12 *For any integer* $k \leq \frac{n-1}{2}$ *we have* $1 \leq n - 2k$

Remark 13 *For any* $q > \frac{1}{2}$ *we have* $\frac{q}{1-q} > 1$

Remark 14 *For any integer* $k \leq \frac{n-1}{2}$ *we have* $\frac{k-nq}{n-k-nq} < 1$

This allows us to derive the following equivalences

$$\frac{\partial}{\partial q}\sum_{k=\frac{n+1}{2}}^{n} f_n(k) > 0$$

$$\Leftrightarrow \frac{\partial}{\partial q}\sum_{k=0}^{\frac{n-1}{2}} \binom{n}{k} q^k(1-q)^{n-k} < \frac{\partial}{\partial q}\sum_{k=\frac{n+1}{2}}^{n} \binom{n}{k} q^k(1-q)^{n-k} \text{ from Remark 11}$$

$$\Leftrightarrow \sum_{k=0}^{\frac{n-1}{2}} \binom{n}{k} q^{k-1}(1-q)^{n-k-1}(k-nq) < \sum_{k=\frac{n+1}{2}}^{n} \binom{n}{k} q^{k-1}(1-q)^{n-k-1}(k-nq)$$

$$\Leftrightarrow \sum_{k=0}^{\frac{n-1}{2}} \binom{n}{k} q^{k-1}(1-q)^{n-k-1}(k-nq) < \sum_{j=0}^{\frac{n-1}{2}} \binom{n}{j} q^{n-j-1}(1-q)^{j-1}(n-j-nq)$$

by substituting k by $n-j$ in the right-hand side

$$\Leftrightarrow \sum_{k=0}^{\frac{n-1}{2}} \binom{n}{k} \left[q^{k-1}(1-q)^{n-k-1}(k-nq) - q^{n-k-1}(1-q)^{k-1}(n-k-nq)\right] < 0$$

To prove this last inequality it is sufficient to prove that the bracketed term is always strictly negative.

$$q^{k-1}(1-q)^{n-k-1}(k-nq) - q^{n-k-1}(1-q)^{k-1}(n-k-nq) < 0$$

$$\Leftrightarrow \frac{k-nq}{n-k-nq} < \frac{q^{n-k-1}(1-q)^{k-1}}{q^{k-1}(1-q)^{n-k-1}}$$

$$\Leftrightarrow \frac{k-nq}{n-k-nq} < \frac{q^{n-2k}}{(1-q)^{n-2k}}$$

$$\Leftrightarrow \frac{k-nq}{n-k-nq} < \left(\frac{q}{(1-q)}\right)^{n-2k}$$

which is true given Remarks 12, 13 and 14.

To prove that P is strictly increasing in n we note that n only affects P through $\sum_{k=\frac{n+1}{2}}^{n} f_n(k)$. So what we need to prove is that $\sum_{k=\frac{n+1}{2}}^{n} f_n(k)$ is strictly increasing with n. But that is a direct consequence of q begin larger than $\frac{1}{2}$ and $f_n(k)$ representing a binomial distribution $B(n,q)$. To see that consider the random variable $\frac{k}{n}$, which represents the percentage of the vote going to the high-valence candidate. The remarks below are derived from the expected value and the variance of k inherent to the binomial distribution.

Remark 15 $E\left(\frac{k}{n}\right) = q$

Remark 16 $Var\left(\frac{k}{n}\right) = \frac{q(1-q)}{n}$

Note form Remark 16 that the variance of $\frac{k}{n}$ is decreasing with n. An implication is that as n increases the distribution of $\frac{k}{n}$ will have a higher concentration of mass around its expected value, which is q. And given that $q > \frac{1}{2}$ this implies that there will be a higher concentration of mass above $\frac{1}{2}$ and the probability that $\frac{k}{n}$ is larger than $\frac{1}{2}$ will be larger. That means that $\Pr\left(\frac{k}{n} > \frac{1}{2}\right)$ increases with n.

But note that

$$\sum_{k=\frac{n+1}{2}}^{n} f_n(k) = \Pr\left(k > \frac{n}{2}\right) = \Pr\left(\frac{k}{n} > \frac{1}{2}\right)$$

which implies that $\sum_{k=\frac{n+1}{2}}^{n} f_n(k)$ increases with n, and therefore P is also increasing with n.

To study the effect of π note that

$$\frac{\partial P}{\partial \pi} = \left(2\sum_{k=\frac{n+1}{2}}^{n} f_n(k) - 1\right)(1 - 2\pi)$$

which is strictly positive for $\pi < \frac{1}{2}$ and strictly negative for $\pi > \frac{1}{2}$. Therefore P reaches a peak at $\pi = \frac{1}{2}$.

To study how the variance of the distribution of candidate valences θ^L affect P, note that given the distribution of θ^L this variance can be calculated to be

$$Var\left(\theta^L\right) = \Theta^2 \pi (1 - \pi)$$

With that result P can be rewritten as

$$P = \frac{Var\left(\theta^L\right)}{\Theta^2}\left(2\sum_{k=\frac{n+1}{2}}^{n} f_n(k) - 1\right)$$

and therefore P is strictly increasing with $Var\left(\theta^L\right)$. Exactly the same can be said for $Var\left(\theta^D\right)$. ■

11 References

References

[1] Adams, James (1999), "Policy Divergence in Multicandidate Probabilistic Spatial Voting", *Public Choice*, 100: 103-122.

[2] Adams, James and Samuel Merrill, III (2006), "The Spatial Strategies of Candidates Facing Both a Primary and a General Election: When Do Primary Elections Help Parties Win Office?", paper presented at the MPSA 64th Annual Meeting.

[3] Adler, Ilya (1993), "The Mexican Case: The Media in the 1988 Presidential Election", in *Television, Politics, and the Transition to Democracy in Latin America*, Thomas E. Skidmore ed., The John Hopkins University Press, Baltimore, Maryland.

[4] Alcántara Sáez, Manuel (2002), "Experimentos de Democracia Interna: Las Primarias de Partidos en América Latina," The Helen Kellogg Institute for International Studies, University of Notre Dame, Working Paper #293.

[5] Alcántara Sáez, Manuel and Roberto Espíndola (2003), "Political Parties in Latin America: A Tentative Typology", manuscript.

[6] Alcántara Sáez, Manuel and Flavia Freidenberg (eds.) (2001), *Partidos Políticos de América Latina,* Ediciones de la Universidad de Salamanca, Salamanca, Spain.

[7] Ansolabehere, Stephen and James M. Snyder Jr., (2000), "Valence Politics and Equilibrium in Spatial Election Models", *Public Choice*, Vol. 103, no. 3-4, June 2000, pp. 327-36

[8] Austen-Smith, David and Jeffrey S. Banks (1996), "Information Aggregation, Rationality, and the Condorcet Jury Theorem", *The American Political Science Review*, Vol. 90, No. 1, (Mar. 1996), pp. 34-45.

[9] Beer, Caroline C. (2003), *Electoral Competition and Institutional Change in Mexico*, University of Notre Dame Press, Notre Dame, Indiana.

[10] Caillaud, Bernard and Jean Tirole (2002), "Parties as Political Intermediaries", Quarterly Journal of Economics, Vol. 117, no. 4, November 2002, pp. 1453-89.

[11] Carey, John M. (2003), "Presidentialism and Representative Institutions," in *Constructing Democratic Governance in Latin America,* Eds. Jorge I. Domínguez and Michael Shifter, The John Hopkins University Press, Baltimore, Maryland.

[12] Carey, John M. and John Polga-Hecimovich (2006), "Primary Elections and Candidate Strength in Latin America", *Journal of Politics*, Vol. 68, No. 3,August 2006, pp. 530-543

[13] Casella, George and Roger L. Berger (2002), *Statistical Inference, Second Edition,* Duxbury, Pacific Grove, California.

[14] Chand, Vikram K. (2001), *Mexico's Political Awakening*, University of Notre Dame Press, Notre Dame, Indiana.

[15] Chen, Kong-Pin and Sheng-Zhang Yang (2002), "Strategic Voting in Open Primaries", *Public Choice*, 112: 1-30.

[16] Colomer, Josep M. (2002), "Las Elecciones Primarias Presidenciales en America Latina y sus Consecuencias Políticas", in Marcelo Cavarozzi and Juan Manuel Abal Medina (eds.), *El asedio a la Política. Los partidos politicos en America Latina en la era neoliberal.* Buenos Aires: Gedisa, 2002.

[17] de Lima, Venicio A. (1993), "Brazilian Television in the 1989 Presidential Campaign: Constructing a President", in *Television, Politics, and the Transition to Democracy in Latin America*, Thomas E. Skidmore ed., The John Hopkins University Press, Baltimore, Maryland.

[18] de Luca, Miguel, Mark P. Jones and María Inés Tula (2002), "Back Rooms or Ballot Boxes? Candidate Nomination in Argentina", *Comparative Political Studies*, Vol.. 35, No 4, 413-436.

[19] Díaz-Cayeros, Alberto (1997), "Federalism and Veto Players: Equilibrium in Local Political Ambition in Mexico", paper presented at the Annual Meeting of the American Political Science Association, Washington, DC, August 28-31.

[20] Downs, Anthony (1957), *An Economic Theory of Democracy*, Harper and Brothers Publishers, New York.

[21] Duggan, John and César Martinelli (2001), "A Bayesian Model of Voting in Juries", *Games and Economic Behavior*, 37, 259-294

[22] Eisenstadt, Todd A. (2004), *Courting Democracy in Mexico: Party Strategies and Electoral Institutions*, Cambridge University Press, New York.

[23] *El Universal* (2003), "Urge Frenar Debacle de Partidos", August 25, 2003.

[24] Enelow, James M. and Melvin J. Hinich (1982), "Nonspatial Candidate Characteristics and Electoral Competition", *The Journal of Politics*, Vol. 44, No. 1, Feb., 115-130.

[25] Field, Bonnie N. and Peter Siavelis (2006), "Internal Party Rules and Candidate Selection in Democratizing Countries", paper presented at the 2006 Meeting of the Latin American Studies Association.

[26] Feddersen, Timothy and Wolfgang Pesendorfer (1998), "Convicting the Innocent: The Inferiority of Unanimous Jury Verdicts under Strategic Voting," American Political Science Review, vol. 92, no. 1, March 1998, pp. 23-35

[27] Freidenberg, Flavia (2003), *Selección de Candidatos y Democracia Interna en los Partidos de América Latina*, Biblioteca de la Reforma Política Nº 1, Lima

[28] Fox, Elizabeth and Silvio Waisbord (2002), "Latin Politics, Global Media", in *Latin Politics, Global Media,* Elizabeth Fox and Silvio Waisbord eds., University of Texas Press, Austin.

[29] Geer, John G. (1989), *Nominating Presidents: An Evaluation of Voters and Primaries*, Greenwod Press, Westport, Connecticut.

[30] Groseclose, Tim (2001), "A Model of Candidate Location When One Candidate Has a Valence Advantage", *American Journal of Political Science*, Vol. 45, No. 4, 862-886

[31] Harrington, Joseph E. Jr., and Gregory D. Hess (1996), "A Spatial Theory of Positive and Negative Campaigning", *Games and Economic Behavior*, 17: 209-229.

[32] Hartlyn, Jonathan, *The Struggle for Democratic Politics in the Dominican Republic*, University of North Carolina Press, Chapel Hill.

[33] Jackson, Matthew O., Laurent Mathevet and Kyle Mattes (2006), "Nomination Processes and Policy Outcomes", typescript

[34] Key, V.O. (1947), *Politics, Parties and Pressure Groups,* New York: Thomas Y. Crowell Company.

[35] Lawson, Chappell H. (2002), *Building the Fourth Estate: Democratization and the Rise of a Free Press in Mexico*, University of California Press, Berkeley, Los Angeles, London.

[36] Levistky, Steven (2003), *Transforming Labor-Based Parties in Latin America: Aregentint Peronism in Comoparative Perspective*, Cambridge University Press, New York, NY.

[37] Londregan, John and Thomas Romer (1993), "Polarization, Incumbency, and the Personal Vote", in *Political Economy: Institutions, Competition and Representation*, W.A. Barnett, M.J. Hihicn, and N.J. Schofield (eds), Cambridge University Press, Cambridge, UK.

[38] Martinelli, César (2002), "Convergence Results for Unanimous Voting", *Journal of Economic Theory*, 105, 278-297.

[39] Meirowitz, Adam (2005), "Informational Party Primaries and Strategic Ambiguity", *Journal of Theoretical Politics,* 17(1): 107-136.

[40] Morton, Rebecca B. and Kenneth C. Williams (2001), *Learning by Voting*, University of Michigan Press, Ann Arbor.

[41] Owen, Guillermo and Bernard Grofman (2004), "Two-Stage Competition in Two-Party Contests: Persistent Divergence of Party Positions", forthcoming in *Social Choice and Welfare.*

[42] Poiré Romero, Alejandro Alfonso (2002), "Bounded Ambitions. Party Nominations, Discipline and Defection: Mexico's PRI in comparative

Perspective", Unpublished Ph.D. thesis, Cambirdige, Massachusetts, Harvard University.

[43] Polsby, Nelson W. (1960), "Decision-Making at the National Conventions", *Western Political Quarterly*, 13:609-619

[44] Polsby, Nelson W. (1983), *Consequences of Party Reform*, New York: Oxford University Press.

[45] Scarrow, Susan (2005), "Implementing Intra-party Democracy", in *Political Parties and Democracy in Theoretical and Practical Perspective,* National Democratic Institute for International Affairs, research paper series.

[46] Schofield, Norman (2004), "Equilibrium in the Spatial 'Valence' Model of Politics", *Journal of Theoretical Politics*, 16 (4): 447-481.

[47] Schofield, Norman and Itai Sened (2005), "Multiparty Competition in Israel, 1988-96", *British Journal of Political Science*, 35, 635-663.

[48] Serra, Gilles (2006a), "Primary Divergence: the Effects of Primary Elections on Candidate Strategies in the Downsian Model", Harvard University mimeo.

[49] Serra, Gilles (2006b), "Primary Elections with Costly Entry of Candidates: Does Divergence Prevail?", Harvard University mimeo.

[50] Skidmore, Thomas E. (1993), "Politics and the Media in a Democratizing Latin America", in *Television, Politics, and the Transition to Democracy in Latin America*, Thomas E. Skidmore ed., The John Hopkins University Press, Baltimore, Maryland.

[51] Siavelis, Peter M. and Scott Morgenstern (2006), *Pathways to Power: Political Recruitment and Candidate Selection in Latin America*, Unpublished Manuscript, http://www.wfu.edu/politics/conference/paper.htm

[52] Snyder, James M. and Michael M. Ting (2005), "Electoral Control and Selection with Parties and Primaries", typescript.

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

[54] *The Economist* (2006), "A Mandate for Change", December 2, 2006.

[55] Waisbord, Silvio (2000), *Watchdog Journalism in South America: News, Accountability and Democracy*, Columbia University Press, New York.

[56] Ware, Alan (2002), *The American Direct Primary,* Cambridge University Press, Cambridge.

[57] Wuhs, Steven T. (2006), "Democratization and the Dynamics of Candidate Selection Rule Change in Mexico, 1991-2003", *Mexican Studies/Estudios Mexicanos* Vol. 22, Issue 1, Winter 2006, pp. 33-55.

[58] Zuleta-Puceiro, Enrique (1993), "The Argentine Case: Television in the 1989 Presidential Election", in *Television, Politics, and the Transition to Democracy in Latin America*, Thomas E. Skidmore ed., The John Hopkins University Press, Baltimore, Maryland.