Keeping Politicians on Their Toes: Does the Way Parties Select Candidates Matter?

Benoit S Y Crutzen

This Draft: February 2013

Abstract

Accountability is a cornerstone of democracy. Does the way parties select their candidates matter? When should parties favor a selection procedure? To answer these questions, we develop a novel model of elections. Effort provision by politicians increases voter welfare but the power of electoral incentives, the precision of the mapping from effort to electoral success, varies. Parties want to win the election and can select their candidates through an effort-based competition or another, effort-independent, procedure. We identify the optimal selection procedure as a function of the power of electoral incentives and electoral rules (we consider plurality rule, closed-list and open-list proportional representation). We prove that the competitiveness of the selection procedure and the power of electoral incentives, but not the electoral rules themselves, are substitutes. Allowing voters to differ ideologically, we also highlight that the competitiveness and the degree of decentralization of the candidate selection procedure are distinct facets of this procedure.
Abstract

Accountability is a cornerstone of democracy. Does the way parties select their candidates matter? When should parties favor a selection procedure? To answer these questions, we develop a novel model of elections. Effort provision by politicians increases voter welfare but the power of electoral incentives, the precision of the mapping from effort to electoral success, varies. Parties want to win the election and can select their candidates through an effort-based competition or another, effort-independent, procedure. We identify the optimal selection procedure as a function of the power of electoral incentives and electoral rules (we consider plurality rule, closed-list and open-list proportional representation). We prove that the competitiveness of the selection procedure and the power of electoral incentives, but not the electoral rules themselves, are substitutes. Allowing voters to differ ideologically, we also highlight that the competitiveness and the degree of decentralization of the candidate selection procedure are distinct facets of this procedure.

1 Introduction

Holding politicians accountable for their deeds is a cornerstone of democracy. There is by now a vast literature that builds on and studies this assertion, both in political science and in economics. These contributions have focused on different characteristics of government and legislatures to try to explain at least part of the cross-country variation in political moral hazard, typically proxied by indices of political corruption. Important contributions include Myerson (1993), Persson and Tabellini (1999, 2000 and 2003), Persson, Tabellini and Trebbi (2003), Kunicova and Rose-Ackerman (2005), Chang and Golden (2006) and Tavits (2007). This paper contributes to this literature by filling a very important gap: we show that the way parties organize – and especially the degree of competitiveness of their candidate selection procedures – is of first order importance in understanding cross-country differences in effort provision by politicians, or political moral hazard more generally. That the previous literature completely ignored the role parties and their internal organization play in such problems is particularly surprising, given that a large body of comparative and party politics
literature – key contributions include Gallagher and Marsh (1988), Katz and Mair (1994) Norris (1997, 2006), Hazan and Pennings (2001), Hazan and Rahat (2006 and 2010) – does stress that candidate selection procedures differ largely across democracies and are, without any doubt, one of the central loci shaping political competition and the breadth of choice voters have once they go to the electoral booth. To offer just one quote, Norris (2006, p. 89) states that “[T]he process of recruitment to elected and appointed office is widely regarded as one of the most important residual functions of parties, with potential consequences for [...] the accountability of elected members.” It is parties that select candidates, not voters; the electorate has only a somewhat secondary role as it can only, at best, pick its favorite politicians from among a pre-defined set of candidates. The so much acclaimed accountability role of elections may thus be less powerful than what the current common wisdom may have led us to believe.

The second key contribution of the paper is to further our understanding of the determinants of the ways party organize and, in particular, of the adoption of different candidate selection procedures. If several seminal studies have analyzed cross-country differences in candidate selection procedures across countries and their potential consequences – more empirically oriented works include Gallagher and Marsh (1988); Katz and Mair (1994) and Norris (1997); Hazan and Rahat (2010) is the central reference for those wishing to theorize on the subject – only Lundell (2004), Hazan and Voerman (2006) and Shomer (2012), to the best of our knowledge, focus explicitly on the determinants of candidate selection procedures. Yet, these works are exclusively empirical and none examines the role of the competitiveness of the candidate selection procedure. Our theory is thus a first step towards the filling of this crucial gap in the analysis of electoral and political institutions.¹

¹Adams and Merill (2008), Hirano, Snyder and Ting (2009), Castanheira, Crutzen and Sahuguet (2010), Serra (2011) and Snyder and Ting (2011) do offer formal models of selection, but they focus on the selection of a single candidate who will then run for a single executive office under plurality rule only. To the contrary, we focus on the selection of all the electoral candidates a party proposes to voters and our analysis spans different electoral rules.
The competitiveness of the candidate selection procedure in shaping individual political moral hazard – modelled as the choice of individually costly effort by politicians – and thus government policy across electoral rules. We compare optimal individual effort provision by candidates across electoral rules by defining plurality rule as majoritarian electoral competition in single-member-districts (FPTP in what follows) as in the US or the UK and closed-list proportional representation (PR in what follows) as competition between party-specific pre-defined lists of candidates in a single nationwide electoral district, as in Israel or The Netherlands.

The two most important novel feature of our model are the centrality of candidate selection procedures in the game and the power of electoral incentives. We consider two polar candidate selection procedures: a competitive one, in which electoral candidates are the winners of an effort-based tournament within their party; and a non-competitive one, in which the candidates are hand picked by their party following some procedure that is noncompetitive and independent of the effort decisions of the different internal candidates. This paper is thus concerned with the impact the competitiveness of the candidate selection procedure has on the choices of politicians. We leave the study of the impact of other features of the selection procedure to future work.\(^2\)

The power of electoral incentives is the degree to which the individual choices of politicians impact on their electoral appeal, on their electoral score, taking other dimensions and characteristics of the game, such as the selection procedure, the ballot structure and the electoral rule, as given. Electoral incentives are defined as more powerful the more sensitive to the decisions of politicians are the electoral preferences of voters. Then, an important aspect of the power of electoral incentives is the degree of ‘swingability’ of voters: the more reactive voters are to information about the choices of politicians, the closer will be the mapping between the politicians’ performance and their electoral score. Of course, the share of swing voters in the population is not sufficient for electoral incentives to be powerful: voters still need to receive the relevant information. Then, an obvious institution that impacts on

\[^2\text{A more general interpretation of our research is that of the analysis of the conditions under which parties find it optimal to select candidates on the basis of their performance in the dimension(s) voters care about, as opposed to selection on the basis of other criteria that are of lesser or no direct interest to voters.}\]
the power of electoral incentives is the media, which may both increase or decrease electoral incentives, depending on how much and how unbiased is the information they provide voters with. In the model, we assume that such information is available to voters but is subject to noise. The extent of this noise is thus our measure of the power of electoral incentives. Also, notice that the power of electoral incentives refers to all the party candidates, and not to incumbents only. Thus the power of electoral incentives, the ideologic biases of voters and the incumbency advantage are distinct concepts in our analysis.

Turning to our results, we first prove that the competitiveness of candidate selection procedures and the power of electoral incentives are substitutes: when the election offers low-powered incentives to candidates, parties can boost their politicians’ incentives to provide effort by adopting a competitive candidate selection procedure; vice versa, when the election impacts strongly on politicians’ incentives, parties are better off if they rely on a noncompetitive candidate selection procedure. The intuition for this result is as follows: when electoral incentives are high-powered, parties should not add an extra incentive device to ensure politicians exert effort: the electoral game is already providing such an incentive. To the contrary, if politicians know that the mapping from effort provision to election is weak and thus the return to effort provision is low, the electoral contest provides them with little incentive to invest in costly effort. Providing an extra incentive device is then profitable to the party, as politicians are incentivized to exert effort to increase their chances of passing this extra hurdle. As this first theoretical finding holds both within and across electoral rules, a key empirical implication is that one should not expect to find a strong relationship, if there is one at all, between the competitiveness of selection procedures – or, more generally, the extent to which parties select candidates on the basis of their performance along dimensions which are valued by the electorate – and electoral rules.

Next, we prove that the competitiveness of the candidate selection procedure is of first order importance in predicting cross-country differences in effort provision by politicians: if politicians exert more effort under plurality rule than under closed-list proportional representation when the candidate selection procedure is noncompetitive, the opposite holds when
parties find it optimal to let candidate selection procedures be competitive. The intuition for
this result is as follows. Under PR, when parties find it optimal not to rely on a competitive
candidate selection procedure, the negative effect of the closed party list that previous con-
tributions have focused on allows candidates to exert less effort than they would were they
to run for election under FPTP. Yet, when parties find it optimal to select their candidates
through an effort-based competition, parties can rely not only on internal competition to se-
lect the pool of candidates to offer to voters, they can also rely on the competitive allocation
of the different slots on the party list to provide politicians with additional incentives to exert
effort, as the highest slots on the list are associated with a very high chance of election whereas
the lowest slots are associated with a very low probability of making it into the legislature.
Parties can thus use two incentive devices to push candidates to exert effort. Under FPTP,
parties have one incentive device only, as the competitive allocation of slots on the party list is
obviously not available under FPTP: there is only one position available per district. Again,
this finding has important consequences for empirical research. In particular, it implies that
studies focusing on the relationship between political moral hazard – corruption, say – and
political institutions suffer from an omitted explanatory variable problem whenever the way
parties select candidates is not included among the regressors. This may in itself suffice to
explain the lack of clear-cut evidence produced by such empirical exercises, see for example
the conclusion in Persson, Tabellini and Trebbi (2003).

The incentive effect of the party list under PR also turns out to be stronger the larger
is the size of the legislature or the average district magnitude. This fact can be seen as
complementing Myerson (1993) in rationalizing the empirical finding that corruption – to be
interpreted in our model as the opposite of effort provision– and average district magnitude
are negatively correlated under closed-list PR: if Myerson focuses on the role of barriers to
entry, that is, interparty competition, we highlight the role played by the competitiveness of
the candidate selection procedure, that is, intraparty competition.

---

3See for example Persson and Tabellini (1999, 2000) and the references therein.
4In section 6, we show that our results are robust to the extension of the model to allow for the presence
of ideologic biases in the population, so that some districts under FPTP are easier to win than others.
We then consider several extensions. First of all, we study whether and how allowing voters to have an initial ideologic bias in favor of either party modifies our results. This extension allows us to show that the basic results are robust to the introduction of ideologic biases in the electorate. More importantly perhaps, we also show that the degree of competitiveness and the degree of decentralization of the control over the candidate selection procedure become two distinct characteristics of the selection procedure, whereas these two dimensions cannot be disentangled from one another when the polity is homogeneous. Indeed, with ideologically heterogeneous voters, the optimal candidate selection procedure in one district may not be optimal in another, and thus parties may have an incentive to use different selection procedures in different districts, something which is not possible if the party uses a single, one-size-fits-all, candidate selection procedure.\footnote{Instead, when voters all have identical ex-ante preferences, there is a single optimal type of selection procedure, making the issue of the identity of the party official who control the procedure irrelevant (unless different party officials have conflicting interests, something we abstract from in this paper).} To the best of our knowledge, such a distinction has not been identified before in the literature. Yet, such a finding is important for empirical research, as it suggests that one cannot use the existing information on the degree of decentralization of (control rights over) candidate selection to proxy for the degree of competitiveness of this procedure. Further, our findings imply that one should also not expect to find any significant (positive or negative) association between the degree of competitiveness and the degree of decentralization (of control rights over) candidate selection procedures.

In the second key extension of the baseline model, we consider the case of open-list proportional representation, the electoral system used for example in Finland, in which voters have to vote for individual candidates, as opposed to party lists. There, we show that which of open- or closed-list PR is associated to lower effort provision depends again on the competitiveness of the candidate selection procedure. The trade off driving this result is the fact that by moving from closed-list PR to open-list PR, the individual accountability of politicians to voters is enhanced, but the possibility of using strategically the allocation of the different slots on the party list is no longer available to parties, as voters now (have to)
vote for their favorite individual politicians on the party lists and not for the party list as a whole.

The rest of the paper is organized as follows. The next section presents the model. Section three pins down equilibrium effort across the different electoral rules and candidate selection procedures. The next section analyzes how the power of electoral incentives impact on the type of candidate selection procedure parties wish to adopt. Section 5 analyzes the central contribution of this article: we show that candidate selection procedures are of first order importance in determining the choices electoral candidates make. Section 6 discusses our modeling assumptions and some important extensions. The last section concludes and offers some avenues for future research.

2 The Model

Consider a democracy with a continuum of voters of total measure \( L \) and a legislature having \( L \) seats, \( L \) odd. Elections determining the composition of the legislature take place at regular intervals. The electoral rule is either plurality rule in single member districts (with a unit measure of voters in each district) – FPTP in what follows – or proportional representation with closed party lists in a single, nationwide district – PR in what follows. Under FPTP, the candidate who receives the largest number of votes wins the legislative seat assigned to their district. Under PR, the mapping from vote shares to seat shares is such that, for a party to win at least \( m \) legislative seats, its nationwide vote share must be at least equal to \((2m - 1)/2L\), that is, the mapping from vote shares to seat shares offers a premium to diversity over representation. More importantly, this mapping guarantees that, for \( L = 1 \), the game politicians face under FPTP and PR is one and the same, as it should arguably be. That is, for \( L = 1 \), the probability that a party wins the only available seat is the same irrespective of the electoral rule that is in place.

Two parties \( L \) and \( R \) compete for seats in the legislature by proposing their electoral candidates to voters. The party that wins a majority of the legislative seats also wins the executive office. It is common knowledge that parties want their candidates to exert as much
effort as possible, because this maximizes the voters’ welfare, all else equal, as we explain further below.\textsuperscript{6} To each candidate, effort provision is costly and increasingly so; specifically, the cost of effort function of any candidate is given by $C(e) = \frac{1}{2}e^2$. If elected, candidates earn a payoff equal to $V$ – this payoff can be interpreted as the continuation value of the game for a candidate legislator. The objective function of any candidate is thus:

$$Pr(S&E) \cdot V - \frac{1}{2}e^2$$

and $Pr(S&E)$, their probability of being selected by their party and elected by voters is, obviously, a function of, among other things, the candidate selection procedure their party adopts.

Every voter’s indirect utility function is linearly increasing in effort provision by politicians.\textsuperscript{7} Thus, all else equal, voters vote for the politicians who are believed to have exerted highest effort. More precisely, under FPTP, any member of the electorate bases their voting decision on their beliefs about effort provision by the two party candidates running in their district and, under PR, on their beliefs about average effort provision by all the candidates included in the party lists.\textsuperscript{8} Voters base their decisions on beliefs because they do not observe perfectly the choices of the different candidates. The information voter $i$ bases their voting strategy on is subject to an individual-specific noise term $\sigma_i$ that is uniformly distributed over $[\eta - 1, \eta + 1]$, where $\eta$ is a nationwide noise term that is uniformly distributed on $[-\frac{1}{2}, \frac{1}{2}]$ and that impacts on the information set of every voter equally.\textsuperscript{9} $\varepsilon$, the parameter indexing

\textsuperscript{6}We thus assume that the interests of parties and voters are aligned, in the sense that parties want their candidates to focus their efforts on those dimensions that are most valued by voters. This assumption increases simply the starkness of our results. What the actual objective function of parties is remains a largely unanswered question that goes much beyond the scope of this research. Yet, as long as parties also care about voters’ welfare, our results carry through.

\textsuperscript{7}The only purpose of the linearity of the voters’ utility function in effort is to keep the algebra as simple as possible. All our results go through as long as voters value effort provision.

\textsuperscript{8}The sole purpose of this assumption about the information set of voters across electoral rules is to remain in line with the previous literature on the subject, which postulates that accountability through information is higher under FPTP than under PR.

\textsuperscript{9}Readers familiar with probabilistic voting models will have remarked that our incomplete information
the inverse of the variance of \( \eta \) – this variance being equal to \( \frac{1}{12\varepsilon^2} \) – is our measure of the power of electoral incentives. The less noisy are elections, the greater is \( \varepsilon \), the more powerful are electoral incentives.

Turning to the candidate selection procedure, we consider that, irrespective of the type of candidate selection procedure chosen by the party, there are always \( 2L \) politicians who would like to stand for election under their party banner – this assumption is merely to keep the symmetry of the analysis across selection procedures. Parties can choose between a competitive and a non-competitive candidate selection procedure. Under the competitive procedure, the \( 2L \) party candidates – namely the \( l \leq L \) incumbent legislators of the party and the \( 2L - l \) other candidates who wish to be selected by their party – compete inside their party for the right to stand for election at the next electoral contest and the selected candidates are those who (are believed by their party to) have exerted the highest level of effort. Concretely, parties receive signals about their candidates’ effort decision and then rank these in decreasing order of (beliefs relative to) effort provision. Party \( P \) thus believes that its candidate \( i \) has exerted higher effort than its candidate \( j \) if and only if:

\[
e^P_i > e^P_j + \mu_{i,j}
\]

where \( \mu_{i,j} \) is an i.i.d. realization from the Standard Normal distribution \( \mathcal{N}(0,1) \) that impacts on the difference between effort provision by the two candidates \( i \) and \( j \). This type of modelling is standard in the literature dealing with such (principal-agent) problems. See for example Nalebuff and Stiglitz (1983) or, more recently, Moldovanu, Sela and Shi (2007).

Under the non-competitive candidate selection procedure, the party first picks the \( L \) candidates among the pool of \( 2L \) politicians who wish to be selected using a rule that is independent of effort provision by politicians, be them incumbent legislators or not, and then the selected politicians choose how much effort to exert with a view to maximizing their framework can be reinterpreted as a probabilistic voting setup. In that case, the usual interpretation of the shocks is that these represent ideologic swings in voter preferences. The rationale for having two sources of noise is that such a structure allows us to compute the probability that a party wins any seat share under PR, not only a majority. If we were to need to compute the probability that a party wins a majority of the seats only, then one source of noise only would suffice.
chances of being elected by voters.\textsuperscript{10}

The timing of the game is as follows:

1. Parties choose their candidate selection procedure:
   (a) If the procedure is non-competitive, the $L$ electoral candidates are selected;
   (b) If the procedure is competitive, the game moves to following stage.

2. Candidates choose their individual effort level;

3. Parties announce their set of electoral candidates:
   (a) under the non-competitive procedure, this is just the set selected at time 1.(a);
   (b) under the competitive candidate selection procedure, the $L$ electoral candidates are selected based on the party’s beliefs about effort provision.

4. Voters receive some information about candidate effort provision and then cast their ballot.

5. The candidates and the party-in-the-executive are elected, payoffs materialize and the game moves to the next election.

3 Equilibrium Effort Provision Across Candidate Selection Procedures

3.1 Noncompetitive Candidate Selection Procedure

FPTP\textsuperscript{10}In particular, the non-competitive candidate selection procedure includes the set of procedures for which the $L$ incumbent legislators of the party are automatically re-selected, quite obviously. This is, after the one on the information set of voters, the second key assumption that characterizes the existing literature on how electoral rules impact on the voters capacity to hold politicians accountable for their deeds; see for example Persson and Tabellini (2000).
When the candidate selection procedure is noncompetitive, the candidates running under each party banner are chosen by some procedure that is independent of their effort choices. Then, under FPTP, the problem faced by each selected candidate of party \( L \), say, is simply to choose effort so as to maximize their chances of winning the district electoral race. The problem is therefore:

\[
\max_{e^L_d} \Pr \left( Sh^L_d \geq \frac{1}{2} \right) \cdot V - \frac{1}{2} \left( e^L_d \right)^2 \tag{3}
\]

where \( Sh^L_d \) stands for party \( L \)'s vote share in district \( d \) and \( e^L_d \) is effort provision by the candidate of party \( L \) in district \( d \).

Given our modelling assumptions, \( \Pr \left( Sh^L_d \geq \frac{1}{2} \right) \) is given by \( \frac{1}{2} + \varepsilon \left( e^L_d - e^R_d \right) \).\(^{11}\) Then, the first order condition to this problem implies that optimal effort provision, \( e^L_d^* \), is given by:

\[
e^L_d^* = \varepsilon V, \; \forall d = 1, ..., L \tag{4}
\]

Notice that effort is directly proportional to the power of electoral incentives, \( \varepsilon \), and the continuation value of the game, \( V \). This is because, under FPTP, each candidate fully appropriates the returns from extra effort provision. This is the key incentive difference with closed-list PR, as we now show.

**PR**

Under closed-list PR, given that the seats a party wins are filled by its candidates in the order of the party electoral list, the problem each candidate faces depends on his or her position on that list. Focus on party \( L \), say. For the candidate in the \( m \)th position on that party's list, with \( m = 1, 2, ..., L \), the problem is to choose their effort level, \( e^L_m \) optimally, so as to maximize:

\[
\Pr \left( \text{Party } L \text{ wins at least } m \text{ seats} \right) \cdot V - \frac{1}{2} \left( e^L_m \right)^2 . \tag{5}
\]

\(^{11}\)The probability that voter \( i \) in district \( d \) votes for party \( L \) is given by:

\[
\frac{1}{2} + \frac{1}{2} \left( e^L_d - e^R_d - \eta \right) .
\]

Then, by the Law of Large Numbers, averaging over all voters in district \( d \), the vote share of party \( L \) is that district is:

\[
\frac{1}{2} + \frac{1}{2} \left( e^L_d - e^R_d - \eta \right) .
\]

Computing \( \Pr \left( Sh^L_d \geq 1/2 \right) \) yields the expression in the text.
To compute $\Pr (\text{Party } L \text{ wins at least } m \text{ seats})$, remark that the probability (before $\eta$ is realized) that any voter votes for $L$ is equal to:

$$\Pr \left( \frac{\sum_{l=1}^{\ell} e^L_l}{\ell} - \frac{\sum_{l=1}^{\ell} e^R_l}{\ell} - \eta - \sigma_l > 0 \right) = \frac{1}{2} + \frac{1}{2} \left( \frac{\sum_{l=1}^{\ell} e^L_l}{\ell} - \frac{\sum_{l=1}^{\ell} e^R_l}{\ell} - \eta \right).$$

and, integrating over the whole population, the expected population vote share of party $L$ before $\eta$ is realized is equal to

$$\frac{1}{2} + \frac{1}{2} \left( \frac{\sum_{l=1}^{\ell} e^L_l}{\ell} - \frac{\sum_{l=1}^{\ell} e^R_l}{\ell} - \eta \right).$$

Given the mapping from vote shares to seat shares set above, the probability that $L$ wins at least $m$ seats is then given by:

$$\Pr \left( \frac{1}{2} + \frac{1}{2} \left( \frac{\sum_{l=1}^{\ell} e^L_l}{\ell} - \frac{\sum_{l=1}^{\ell} e^R_l}{\ell} - \eta \right) > \frac{2m-1}{2\ell} \right) = \frac{1}{2} + \varepsilon \left( \frac{\sum_{l=1}^{\ell} e^L_l}{\ell} - \frac{\sum_{l=1}^{\ell} e^R_l}{\ell} + \frac{\ell-2m+1}{\ell} \right).$$

Then, the first order condition associated to the problem faced by any candidate of party $L$ yields:

$$e^{L*}_{m} = \frac{\varepsilon}{\ell} V, \quad \forall m = 1, 2, ..., \ell$$

Thus, effort provision is a function of $\frac{V}{\ell}$, and not $V$ as under FPTP. This is because, under closed-list PR, a candidate’s effort provision impacts with weight $1/\ell$ only on how voters perceive the whole list. This is the reduced accountability effect much of the literature on the subject has focused on so far.\(^{12}\)

### 3.2 Competitive Candidate Selection Procedure

**FPTP**

When candidate selection is competitive, candidates first need to win the right to stand for election under their party’s banner. Under FPTP, and in line with the available empirical evidence (see e.g. Galagher and Marsh 1988, Katz and Mair 1994 and also Benedetto and

\(^{12}\)See for example, Persson and Tabellini (2000), Persson, Tabellini and Trebbi (2003) or, more recently, Tavits (2007).
Hix 2007), the intraparty competition is run at the district level.\textsuperscript{13} Also, in each district, two candidates run in the intraparty contest.\textsuperscript{14} Then, ex-ante, at the beginning of the election round (which is the view we take in this paper), the problem for each of the two politicians of party \( L \) is to maximize (with respect to \( e^L_d \)):

\[
\left[ \frac{\Pr (e^L_d - e^L_{dc} > \mu_{d,dc})}{\text{proba internal challenger loses}} \cdot \Pr \left( \frac{S^L_d}{\sqrt{d}} \geq \frac{1}{2} \right) \cdot V - \frac{1}{2} \left( e^L_d \right)^2 \right],
\]

that is:

\[
\Phi \left( e^L_d - e^L_{dc} \right) \left[ \frac{1}{2} + \varepsilon \left( e^L_d - e^R_d \right) \right] V - \frac{1}{2} \left( e^L_d \right)^2,
\]

where \( \Phi \) stands for the cumulative distribution function of the Standard Normal distribution and \( e^R_d \) is effort provision by the representative of party \( R \) in district \( d \).

The first order condition associated to this problem is:

\[
e^L_d = \phi \left( e^L_d - e^L_{dc} \right) \left[ \frac{1}{2} + \varepsilon \left( e^L_d - e^R_d \right) \right] V + \varepsilon \Phi \left( e^L_d - e^L_{dc} \right) V,
\]

where \( \phi (\cdot) \) is the density function of the standard Normal distribution.

Exploiting the fact that, in the symmetric equilibrium, \( e^L_d = e^R_d = e^{L*}_d \) and \( \Phi (0) = 1/2 \), this further reduces to:

\[
e^L_d = \frac{1}{2} V \left[ \varepsilon + \phi (0) \right].
\]

\textbf{PR}

\textsuperscript{13}In some representative democracies, residency requirements limit severely the mobility of candidates across districts. In others, it is the local support base of candidates which makes them unlikely to move too often. The following quote from Benedetto and Hix (2007, p. 777) clearly exemplifies this: “O]nce an MP has been selected in a constituency, the MP is difficult to remove if he or she is supported by his or her local party elite. Contrast this with a national-based, party-list proportional representation system, which gives the national party leadership the power to move a candidate down the party list in the next election and so reduce his or her chances of being reelected.”

\textsuperscript{14}This assumption is not restrictive, as it has been shown that, under FPTP, increasing the pool of candidates beyond two typically does not improve incentives for individual candidates and thus we should expect parties not to let the pool of candidates become too large; see for example Castanheira, Crutzen and Sahuguet (2010)
Under PR, the ex-ante problem each of the $2L$ candidates faces is to maximize with respect to $e_l^L$:

$$\left[ \sum_{m=1}^L \Pr(\text{candidate’s list position is } m) \Pr(\text{Party } L \text{ wins at least } m \text{ seats}) \right] \cdot V - \frac{1}{2} (e_l^L)^2.$$  

(9)

We know from above that:

$$\Pr(\text{party wins at least } m \text{ seats}) = \frac{1}{2} + \varepsilon \left( \frac{\sum_{l=1}^L e_l^L}{L} - \frac{\sum_{l=1}^L e_l^R}{L} + \frac{L - 2m + 1}{L} \right)$$

and, given our assumptions on the way parties select their candidates under the competitive candidate selection procedure,

$$\Pr(\text{candidate } l \text{’s list position is } m) = C_{L1m1} \left( 1 - e_l^L \right)^{L-1},$$

with $C_{(2L-1)_{m-1}} = \frac{(2L-1)!}{(2L-m)! (m-1)!}$, $\Phi$ being shorthand for $\Pr(e_l^L - e_c^L > \mu_c)$ and $e_l^L$ being effort by candidate $l$. The intuition for why this is indeed the probability that candidate $l$ ends up in $m$th position on the list is simply that, to be party $L$’s $m$th candidate, $l$ needs to be perceived as having exerted less effort than $m - 1$ other candidates only. This is captured by the term $\Phi^{2L-m} (1 - \Phi)^{m-1}$. Yet, we also need to consider the total number of combinations of the $2L - 1$ candidates which are consistent with candidate $l$ being beaten by $m - 1$ competitors. The total number of such combinations is $C_{(2L-1)_{m-1}}$.

Then, because, in the symmetric equilibrium, we have that $e_l^{L*} = e_c^{L*} = e_l^{R*}$ and thus $\Phi(e_l^{L*} - e_c^{L*}) = \Phi(0) = 1/2$ and, also, for any $L$, $\sum_{m=1}^L C_{(2L-1)_{m-1}} \cdot \left( \frac{1}{2} \right)^{2L-1} = \frac{1}{2}$, equilibrium effort provision by each candidate is equal to:

$$V \cdot \sum_{m=1}^L \left\{ C_{Lm-1} \left( \frac{1}{2} \right)^{2L-2} (2L - 2m + 1) \phi(0) \left[ \frac{1}{2} + \varepsilon \left( \frac{e_l^{L*} - 2m + 1}{L} \right) \right] + \frac{eV}{2L} \right\}. \quad (11)$$

Having solved for equilibrium effort provision across both the different candidate selection procedures and the different electoral rules, we can now turn to the central questions we wish to focus on.
4 When Should Parties Use the Competitive Candidate Selection Procedure?

If parties wish their candidates to focus on the political dimensions that increase their appeal to voters – in our model, exert as much effort as possible – when should they adopt a candidate selection procedure that is competitive and based on the performance of candidates in these electorally central dimensions? Intuition suggests that such a choice may be optimal when the other existing institutions, such as the general election itself, are not providing politicians with sufficient incentives to focus on such dimensions. Yet, what should parties do when electoral incentives are strong? Should they complement these with a competitive candidate selection procedure, or, rather, should they rely solely on the incentives provided by the election and thus ‘freeride’ on these incentives and select the non-competitive selection procedure? To answer this question, we rely on the following two lemmas.

Lemma 1 Under FPTP, irrespective of the type of candidate selection procedure, individual equilibrium effort is increasing in $\varepsilon$. Further, effort increases faster with $\varepsilon$ when the candidate selection procedure is the noncompetitive one.

Under closed-list PR, equilibrium effort is increasing in $\varepsilon$ when the candidate selection procedure is noncompetitive. When selection is competitive, equilibrium effort provision is increasing in $\varepsilon$ if $L < 4$ but decreasing in $\varepsilon$ if $L \geq 4$.

Proof. See the Appendix. ■

The intuition for this result is as follows. Under FPTP, when selection is noncompetitive, candidates need care about the competition from the representative of the other party only, and effort improves their electoral prospects proportionally to the power of electoral incentives; effort is thus increasing in $\varepsilon$, one-to-one. When selection is competitive, candidates first face the internal competition from a challenger. Their equilibrium probability of making it to the general election is thus $1/2$ only. Hence the equilibrium effect of the power of electoral incentives is still positive, but halved.
Under PR, when selection is non-competitive, candidates also need care only about the general election, and thus effort is increasing in $\varepsilon$, but only in a one-to-$L$ fashion, because of the usual problem team or group mechanics create on the mapping from individual effort to group success when only group performance is available. When selection is competitive, the first effect is still present but, as under PR, its power is halved as the equilibrium probability of making it to the general election is $1/2$ only. Further, when selection is competitive, there is another effect. Indeed, the power of electoral incentives impacts on the party’s equilibrium probability of winning $m$ seats – this probability being $\frac{1}{2} + \varepsilon \left(\frac{L-2m+1}{L}\right)$ – and this effect is positive for the highest positions on the list (for $m < L/2$) but negative for lower list positions (for $m > L/2$). Combining this with the fact that, as the number of competitors each candidate faces, $2L - 1$, rises faster with $L$ than the number of seats with can lead to (re-)election in equilibrium, $L/2$, the candidates’ incentives to provide effort are less powerful the higher is $L$. Then, putting these two effects, we then get that, under PR when selection is competitive, the effect of the power of electoral incentives is positive for low values of $L$ (because the first effect dominates the second) but negative for high values of $L$ (because the second effect is dominates in this case).

Using the above lemma, we can also examine the parties’ incentives to use one candidate select procedure over the other as a function of the power of electoral incentives:

**Lemma 2** Under FPTP, parties find it optimal to use the competitive candidate selection procedure if and only if $\varepsilon \in [0, \phi(0)]$;

Under PR, parties find it optimal to use the competitive candidate selection procedure if and only if $\varepsilon \in [0, \varepsilon^\star]$, with $\varepsilon^\star > \phi(0)$.

**Proof.** See the Appendix.

Armed with the above lemmas, we can state the first main finding of the paper:

**Proposition 3** Under both electoral rules, the power of electoral incentives and the competitiveness of the candidate selection procedure are substitutes for parties: if electoral incentives are weak ($\varepsilon$ is small), parties find optimal to adopt the competitive candidate selection pro-
procedure, whereas the non-competitive candidate selection procedure is optimal for parties if electoral incentives are strong ($\varepsilon$ is large).\footnote{We use the term ‘substitutes’ instead of ‘strategic substitutes’ because the power of electoral incentives is not the result of a game theoretic choice by any of the actors involved in this game.}

**Proof.** Follows directly from Lemma 1.

The basic intuition for this finding is that, when electoral incentives are powerful, candidates already face very strong incentives to exert effort, and adding an extra hurdle on their path to election, by adopting a competitive candidate selection procedure, is suboptimal as it pushes politicians to reduce effort provision as they now think that their odds of being selected and then elected are too small. Vice versa, when electoral incentives are weak, candidates’ incentives to exert effort are low, as effort provision is costly and the return on such an investment is very low – as voters base their voting decision on very noisy information. Then, by adding an extra hurdle at the candidate selection stage, parties can boost their candidates’ incentives to provide effort.

Proposition 3 predicts that parties should have an incentive to adapt strategically the degree of competitiveness of the way they select their candidates to the electoral environment. Yet, remark that there is a fundamental difference between the way parties can achieve this goal across electoral rules. Indeed, under FPTP, parties can make selection competitive by opening up to competition the right to stand for election in any electoral district. Under PR, parties have two levers at their disposal: not only can they open up to competition the right to be included in the electoral party list, but they can also let candidates compete for the best positions on that list. Given that, under closed-list PR, legislative seats are allocated to politicians on the party list following the order in which they appear on the list, politicians understand that being in the top positions on the list implies that they are almost certain of winning a legislative seat, whereas ending up in the bottom of the list only implies that the chances of winning a legislative seat are basically reduced to zero. This second lever is simply not available under FPTP.
Does this imply that we should expect politicians to exert more effort under PR than under FPTP when parties use all the weapons at their disposal? This is the question we turn to now.

5 Comparative Politics with Optimal Candidate Selection Procedures

Now that we know that the power of electoral incentives and the competitiveness of the candidate selection procedure are substitutes under both FPTP and PR, we can turn to the paper’s key comparative politics question: can the competitiveness of the candidate selection procedure explain cross-country differences in political moral hazard? To answer this question, we proceed in two steps. First, we show that the strength of the incentives coming from the competitive candidate selection procedure under PR is unambiguously increasing in average district magnitude.\(^{16}\) Then, we show that which of the two electoral rules we should expect to be associated with higher effort provision by politicians turns out to depend crucially on the competitiveness of the candidate selection procedure parties adopt.

Analyzing the effect of average district magnitude on effort provision, we have:

**Proposition 4** When it is optimal for parties to adopt a noncompetitive candidate selection procedure, equilibrium effort provision under closed-list proportional representation is strictly decreasing in average district magnitude. To the contrary, when electoral incentives are weak and thus it is optimal for parties to adopt a competitive candidate selection procedure, equilibrium effort provision under closed-list proportional representation is strictly increasing in average district magnitude.

**Proof.** See the Appendix.

Three forces are at play when average district magnitude increases: first, irrespective of one’s position on the party list, the impact of any candidate’s effort choice on the fortunes of their party shrinks as the size of the party list increases, and this depresses incentives to

\(^{16}\)This question is obviously irrelevant under FPTP as there district size is unitary by assumption.
exert effort. This is the usual group or team incentives dilution effect. Second, given one’s (absolute) position on the party list, an increase in $\mathcal{L}$ increases mechanically one’s chances of (re-)election, thus depressing incentives for effort provision.\textsuperscript{17} Third, as $\mathcal{L}$ grows larger, so does the number of available legislative seats, that is, the size of the pool of available prizes, and this boosts the probability of being elected if positioned in the very first slots on the list, all else equal; this final effect boosts the candidate’s incentives to provide effort, so as to be positioned near the top of the party list and thus secure re-election, whenever effort impacts on a candidate’s list positioning.

What the above proposition states is that which combination of the three forces dominates depends on the type of candidate selection procedure parties find optimal to adopt. When it is optimal to have the noncompetitive selection procedure, so that effort and one’s position on the electoral lists are unrelated, the third force disappears, making effort a decreasing function of $\mathcal{L}$: given a candidate’s position on the list, the impact of their effort on the party’s fortunes is decreasing in $\mathcal{L}$. When the optimal selection procedure is the competitive one, the third effect is the dominating one, as securing a high position on the list is what matters for candidates.

The second part of the proposition also yields a prediction that is complementary to that of Myerson (1993). Myerson showed that larger district should be associated with higher effort provision (or less corruption) because, all else equal, larger districts are associated with lower entry barriers, that is, higher interparty competition and thus voters should find it easier to punish politicians for corrupt or slacking behavior by casting their vote in favor of another, ideologically close but less corrupt candidate. The above proposition stresses that district size matters too for the incentive effects intraparty competition has on effort provision by politicians: larger districts lead to higher effort provision by politicians if the candidate selection procedure is competitive. To say it another way, our model suggests that average district magnitude, the power of electoral incentives and the type of candidate selection procedures interact together in their impact on effort decisions by politicians.

\textsuperscript{17}Indeed, the equilibrium probability of (re-)election if a candidate is second, say, on the party list is much higher when $\mathcal{L}$ is equal to 15, say, than when $\mathcal{L}$ is equal to 2.
We can now state the central result of the paper:

**Proposition 5**  *Comparing optimal effort provision by individual politicians across electoral rules:*

1) when electoral incentives are high-powered, so that it is optimal for parties to adopt a noncompetitive candidate selection procedure, effort provision is lower under closed-list proportional representation than under first-past-the-post; conversely,

2) when electoral incentives are low-powered, so that it is optimal for parties to adopt a competitive candidate selection procedure, effort provision is higher under closed-list proportional representation than under first-past-the-post.

**Proof.** See the Appendix. ■

The intuition for this result is as follows. When selection is not competitive, effort provision under FPTP is greater than that under PR as, under PR, the effect of the presence of a list implies that, on average, politicians have less incentives to provide effort: each candidate does not appropriate fully the returns to their effort as what voters compare is the party lists. Yet, when selection is competitive, under PR, parties can not only choose – through an effort-based competition – the set of candidates who will be allowed to stand for election, they can also use the allocations of the different slots on the party electoral list to provide politicians with additional incentives to exert effort, simply by letting the allocation of these slots be competitive and a function of the different candidates’ effort decisions. This second intraparty incentive device is simply not available under FPTP. What proposition 5 highlights is that, when elections do not provide strong incentives to exert effort, increasing the size of the pool of candidates and adding an extra effort-based hurdle (competition for the best slots on the party list) has a strong effect on every candidate’s choice. Given that, under FPTP, the best a party can do is increase the size of the pool of candidates only, it should come as no surprise that PR is associated to higher effort provision than FPTP when selection is competitive.

The following figure illustrates the proposition. It plots equilibrium effort provision under FPTP (the upward sloping black line is individual effort under FPTP with the competitive
candidate selection procedure, the dashed curve refers to effort under FPTP with the noncompetitive procedure) and under PR when the selection procedure is competitive for different values of average district magnitude – the two directional arrows show how the relationship between the power of electoral incentives and effort rotates as we let average district magnitude increase; the three curves correspond to district magnitudes equal to 3 (upward sloping line starting at roughly 0.375), 5 (slightly downward sloping line) and 11 (downward sloping line) – as a function of the power of electoral incentives – the greater is the value on the horizontal axis, the more powerful are electoral incentives. The vertical dashed line splits the space into the two subspaces for which the optimal candidate selection procedure under FPTP is either the competitive (solid upward sloping black line to the left of the dashed line) or the noncompetitive one (dashed upward sloping black line to the right of the vertical dashed line).\textsuperscript{18} Notice finally that the figure highlights that, to the left of the threshold value of the power of electoral incentives for which parties find it optimal to switch between the two types of candidate selection procedures under FPTP, namely for $\varepsilon = \phi (0) \simeq 0.39894$, effort under PR is higher than that under FPTP when selection is competitive.\textsuperscript{19}

\textsuperscript{18}We focus on the incentives of parties under FPTP to split the space into two subspaces as the proof of lemma 2 shows that for any value of the average district magnitude, parties have an incentive to choose the competitive selection procedure for a larger set of parameter values under PR. The threshold that determines the switch in procedures under FPTP is thus the binding one for our comparative politics exercise.

\textsuperscript{19}We did not represent optimal effort under PR when the candidate selection procedure is noncompetitive as it is obvious that it is lower than optimal effort under FPTP: $\forall \mathcal{L} \geq 1$, $\varepsilon V / \mathcal{L} \leq \varepsilon V$.  

22
6 Discussion and Extensions

6.1 Candidate- versus Leadership selection

The model we developed focuses on candidate selection only. We do recognize that leadership selection is also key in determining the success of a party in elections, either directly, when the party leader runs for election in a district or on his or her party list, or indirectly, for example through possible coattail effects, as in Zudenkova (2011). Yet, as vividly argued by, among others, Rahat and Kenig (2011) and Samuels and Shugart (2010), candidate and leadership selection differ in some essential features that imply that it is more useful to analyze these two selection procedures separately, not jointly. Hirano, Snyder and Ting (2009), Castanheira, Crutzen and Sahuguet (2010), or Hortala-Vallve and Mueller (2011) are formal models that can be described as analyzing leadership selection, but under FPTP only.

6.2 Incumbency Advantage

If, for example, we think of incumbency advantage as being a feature of a candidate such that any opposition must be perceived by voters as being superior to this candidate by some discrete margin for the election results to lead to the ousting of the incumbent – the margin
being a measure of the incumbency advantage,—then the effect on the results of our analysis above are somewhat indeterminate under FPTP: incumbents may face less incentives to exert effort, but non-incumbents should face stronger incentives to exert effort, as they now know that they need not only to win, but also to win by a discrete margin.20

Similarly, under PR, the presence of an incumbency effect should imply that incumbents and more senior politicians more generally find it easier to reach the top spots on the party list. Again this implies that incumbents can be expected to exert less effort while non-incumbents will exert more effort.

All in all, unless we have a good reason to expect incumbency advantage to be stronger under one of the two electoral rules, this feature of the political game should not impact much our comparative politics results.

6.3 Ideology and Safe Seats

Our baseline model does not allow for ideology to play a key role in the game: the distributions of the individual-specific and nationwide noise terms are both symmetric around 0. Thus, ex-ante, no voter has a particular preference for a party over another. To extend our model to a world with voters who on average prefer one party over the other, suppose now that voter preferences are also impacted by an individually-specific ideologic bias $b_i$ in favor of either party. These individual ideologic biases $b_i$ are distributed symmetrically (according to some distribution $G$) around 0 in the whole nation. Yet, if the electoral rule is FPTP we allow the distribution of individual biases in any district $d$ to be symmetric around some $b_d \geq 0$, so that our model does not restrict every district to have an unbiased view of effort provision by the two party candidates but, rather, allows some districts to favor (all else equal) $L$ whereas other prefer (all else equal) $R$. Given that, nationwide, the average ideologic bias is 0, we obviously impose that the average of all district-specific biases is equal to 0.

When voter preferences include these ideologic biases, the problem facing each electoral

---

20The idea that beating an incumbent requires the challenger to win by some margin may also rationalize the evidence on the lack of observed competition in many races in which incumbents stand. See for example, Ansolabehere, Hansen, Hirano and Snyder (2006).
candidate is not symmetric anymore under FPTP, as a candidate running in a district in which their party is a likely winner faces an easier task whereas the representative from the opposition faces a tougher electoral challenge. This implies that the algebra becomes more involved under FPTP. To keep things tractable, when solving for equilibrium effort under the different scenarios below, we restrict attention to the case in which \( L = 3 \). Then, straightforward computations show that the probability that a random voter in any district \( d \) votes for the candidate of party \( L \) is equal to \( \frac{1}{2} + \frac{1}{2} (e^L_d - e^R_d - \eta - b_d) \), which implies, integrating over all voters in district \( d \), the the expected vote share of party \( L \) is equal to \( \frac{1}{2} + \frac{1}{2} (e^L_d - e^R_d - \eta - b_d) \) where, remember, \( b_d \leq 0 \) is the average ideologic bias of voters in district \( d \). The probability that party \( L \) wins the election in district \( d \), \( \Pr (Sh^L_d > \frac{1}{2}) \), is then equal to \( \frac{1}{2} + \varepsilon (e^L_d - e^R_d - b_d) \). Similarly, the probability that party \( R \) wins the election in that same district is given by \( \frac{1}{2} + \varepsilon (e^R_d - e^L_d + b_d) \).

If the candidate selection procedure is noncompetitive, the problem faced by the candidate of party \( P, P = L, R \) is thus:

\[
\max_{e_d^P} \Pr (Sh^P_d > \frac{1}{2}) V - \frac{1}{2} (e_d^P)^2, \tag{12}
\]

and it is immediate to check that the first order condition associated to this problem is the same as that in the main text, implying that optimal effort is affected neither by the organizational choices of the other party nor by the presence of the ideologic bias and is still equal to (4), namely \( e^*_d^P = \varepsilon V \). The intuition for this result is that, given our distributional assumptions, the ideologic bias does not affect the marginal incentives in favor of or against effort provision, and thus the optimal choice of every candidate is unaffected by the safety of the seat.

To the contrary, when the candidate selection is competitive in both parties, the problem faced by each candidate is equal to:

\[
\max_{e_d^P} \Pr (e_d^P > e^P_{dc} + \mu_{d,dc}) \Pr (Sh^P_d > \frac{1}{2}) V - \frac{1}{2} (e_d^P)^2 \tag{13}
\]

In this case, the safety of the seat matters and we can thus not invoke symmetry anymore to simplify the first order conditions. Indeed, the first order conditions pinning down
equilibrium effort are now party-specific:

\[
\begin{align*}
  e_{Ld}^* &= \phi \left( e_{Ld}^* - e_{Rd}^* \right) \left[ \frac{1}{2} + \varepsilon \left( e_{Ld}^* - e_{Rd}^* - b_d \right) \right] V + \varepsilon \Phi \left( e_{Ld}^* - e_{Rd}^* \right) V \\
  e_{Rd}^* &= \phi \left( e_{Rd}^* - e_{Ld}^* \right) \left[ \frac{1}{2} + \varepsilon \left( e_{Rd}^* - e_{Ld}^* + b_d \right) \right] V + \varepsilon \Phi \left( e_{Rd}^* - e_{Ld}^* \right) V
\end{align*}
\]

Of course, within each party, the problem is still symmetric, implying that \( \phi \left( e_{d}^{Pd} - e_{d}^{Pc} \right) = \phi \left(0\right) = \frac{1}{\sqrt{2\pi}} \) and \( \Phi \left( e_{d}^{Rd} - e_{d}^{Rc} \right) = \Phi \left(0\right) = 1/2 \). The first order conditions above then simplify in equilibrium to:

\[
\begin{align*}
  e_{Ld}^* &= \phi \left(0\right) \left[ \frac{1}{2} + \varepsilon \left( e_{Ld}^* - e_{Rd}^* - b_d \right) \right] V + \frac{\varepsilon}{2} V \\
  e_{Rd}^* &= \phi \left(0\right) \left[ \frac{1}{2} + \varepsilon \left( e_{Rd}^* - e_{Ld}^* + b_d \right) \right] V + \frac{\varepsilon}{2} V
\end{align*}
\]

Solving the above system yields:

\[
\begin{align*}
  e_{Ld}^* &= \frac{\phi \left(0\right) + \varepsilon}{2} V + \frac{2\phi \left(0\right) \varepsilon V}{4\phi \left(0\right) \varepsilon - 2} b_d \\
  e_{Rd}^* &= \frac{\phi \left(0\right) + \varepsilon}{2} V - \frac{2\phi \left(0\right) \varepsilon V}{4\phi \left(0\right) \varepsilon - 2} b_d
\end{align*}
\]  

(14)

Yet, one of the two parties (or even both!) may be better off going for the non-competitive selection procedure, if \( \varepsilon V \) is greater than the relevant effort level in (14). Then, if \( R \), say, adopts the noncompetitive selection procedure but \( L \) sticks to the competitive one, equilibrium effort provision by the two candidates of party \( L \) is given by

\[
\begin{align*}
  e_{Ld}^* &= \phi \left(0\right) \left[ \frac{1}{2} + \varepsilon \left( e_{Ld}^* - e_{Rd}^* - b_d \right) \right] V + \frac{\varepsilon}{2} V + \frac{\varepsilon}{2} e_{Ld}^* = V \frac{\phi \left(0\right) \left( \varepsilon \left( b_d + \varepsilon V \right) - \frac{1}{2} \varepsilon \right) - \frac{1}{2} \varepsilon}{\phi \left(0\right) \varepsilon - 1}.
\end{align*}
\]  

(15)

Under PR, suppose for the sake of simplicity that each party enjoys a nationwide group of unconditional supporters such that each party is certain to win one legislative seat and that the last, centrist and contested, seat goes to party \( L \), say, with probability \( \frac{1}{2} + \varepsilon \left( \frac{1}{3} \sum_{i=1}^{3} e_{d_i}^{L} - \frac{1}{3} \sum_{i=1}^{3} e_{d_i}^{R} \right) \).

Then, the problem each candidate of party \( L \) faces when candidate selection is competitive
is given by:\textsuperscript{21}
\[
\max \sum_{m=1}^{3} \Pr (\text{candidate } l \text{ is elected if in } m\text{th position on party list})
\]
\[
= \Pr (\text{first}) \cdot V + \Pr (\text{second}) \left[ \frac{1}{2} + \varepsilon \left( \frac{\sum_{l=1}^{3} e_{l}^{L}}{3} - \frac{\sum_{l=1}^{3} e_{l}^{R}}{3} \right) \right] \cdot V - \frac{1}{2} (e_{1}^{L})^{2}
\]
as the probability that \( L \) wins the first, left-wing seat is 1 and the probability that \( L \) wins the last, right-wing seat is 0. Also, applying (10) to the current case,

\[
\Pr (\text{first on the list}) = \left[ \Pr (e_{1}^{L} - e_{c}^{L} > \mu_{c}) \right]^{5}
\]
and

\[
\Pr (\text{second on the list}) = 5 \cdot \left[ \Pr (e_{1}^{L} - e_{c}^{L} > \mu_{c}) \right]^{4} \left[ 1 - \Pr (e_{1}^{L} - e_{c}^{L} > \mu_{c}) \right].
\]

The first order condition associated to this problem boils down to:

\[
e_{1}^{L*} = \frac{V}{32} \left( 25 \phi (0) + \frac{5}{3} \varepsilon \right). \tag{16}
\]

Given the above results, as the following plot shows, allowing for the polity to be ideologically heterogeneous does not invalidate our key results: for low values of \( \varepsilon \) effort is typically higher under PR, whereas the opposite holds for large values of \( \varepsilon \). The figure graphs equilibrium effort under the different scenarios considered above: the flatter curve starting at an effort level of 0.37 is optimal effort under PR (when selection is competitive, equation (16), as in the graph \( \varepsilon \leq 1 \) and then the optimal selection procedure is the competitive one, see lemma 2 and its proof for details). The other curves refer to FPTP: the solid piece-wise linear curve is optimal effort under FPTP when districts are ideologically homogeneous. The dashed curves which start at an effort level of 0.2 and form a trumpet are effort when \( b_{d} > 0 \) and both parties rely on the competitive selection procedure (equation (14); optimal effort is given by these dashed curves until these cut the piecewise linear curve, when it is optimal to switch to the noncompetitive procedure under FPTP). Optimal effort under PR when

\textsuperscript{21}With the above ideologic biases, it is obvious that when the candidate selection is noncompetitive, average effort by candidates under PR is lower than that of candidates under FPTP: only candidate 2 under PR will exert some effort, as the first candidate on the list is sure to be elected (and thus does not need to exert any effort) and the last is sure not to be (and thus there is no point in exerting any effort).
selection is competitive is thus higher than under FPTP under all scenarios provided \( \varepsilon \) is small.\(^{22}\)

If our comparative politics result of proposition 5 goes through when we allow for ide-ology, our findings do highlight one very important point for scholars of candidate selection procedures: as effort provision now varies with the ideologic bias of the relevant voters and the competitiveness of the candidate selection procedure of both parties, parties will have an incentive to use the competitive candidate selection procedure in some districts only, at least if the way they are organized allows them to do so. Thus, our analysis highlights that one should not expect there to exist an obvious, one-to-one relationship between the degree of decentralization of the candidate selection procedure and its average, party-wide degree of competitiveness. In a nutshell, the degree of decentralization and the competitiveness of the candidate selection procedure are two distinct facets of this process that do not necessarily map monotonically into each other. Electoral score maximizing parties, like the ones considered in this research, should thus be expected to organize in such a way that the degree of competitiveness of the candidate selection procedure is tailored to the ideologic characteristics of the districts, but decentralization of candidate selection at the district level is just

\(^{22}\)We set \( V = 1 \) and \( b = 0.1 \). Other parameter values yield a similar plot.
one way to achieve this. A centralized mechanism which takes district-specific characteristics into account can also achieve this goal.

6.4 Open-List PR

Some countries, like Finland, rely on open-list PR, in which candidates are listed in random or alphabetical order on the party list and voters are free to (or even must) pick their favorite individual candidates from these lists. How does this system compare to the previous two we analyzed in terms of effort provision by candidates? At first sight, open-list PR combines the best features of both FPTP and PR. Yet, a move from closed- to open-list PR does imply for the party that the competitive allocation of slots on the list does not belong to its incentive tool kit anymore.

To be consistent with the received wisdom on voter behavior under open-list PR, we assume that voters have enough information to be able to compare candidates individually.\textsuperscript{23} Also, each party candidate is selected by a local selection committee.\textsuperscript{24} Essentially, we thus model intraparty candidate selection under open-list PR as that under FPTP but let inter-party candidate selection at the general election be based on the nationwide comparison of all individual performances of each electoral candidate of the two parties, rather than on a district-by-district comparison of the performance of the district-specific candidate of each party as under FPTP.

If the candidate selection procedure is competitive, conditional on being selected by one’s local selection committee, all a candidate needs to secure in order to win a legislative seat is to be ranked by voters among the top \(L\) political candidates available in the race \textit{overall}, that is, \textit{across parties}.

Specifically, when the candidate selection procedure is competitive, the problem for any

\textsuperscript{23}If voters cannot rank candidates individually then the very purpose of open-list PR is largely lost.
\textsuperscript{24}Candidate selection is indeed very decentralized under open-list PR. See e.g. Gallagher and Marsh (1988), Katz and Mair (1994), Lundell (2004) and Shomer (2012).
candidate is to set their effort $e^L_l$ so as to maximize:

$$\Phi \left( e^L_l - e^L_c \right) \cdot \left\{ \sum_{m=1}^\ell C \left( \frac{2\mathcal{L} - 1}{m - 1} \right) (P)^{2\mathcal{L} - m} (1 - P)^{m-1} \right\} V - \frac{1}{2} \left( e^L_l \right)^2$$

with $P = \frac{1}{2} + \varepsilon \left( e^L_l - e^P_l \right)$, where $e^P_l$ stands for effort provision by any other candidate who is running in the general election. $\Phi \left( e^L_l - e^L_c \right)$ represents the probability of a candidate $l$ being selected by their own local selection committee. The second term on the left hand side of the problem is the probability of being elected by voters.

The associated first order condition boils down in equilibrium to, as all candidates exert the same level of effort and thus $P = \Phi (0) = 1/2$:

$$e^L_{l^*} = V \cdot \left[ \frac{\phi (0) + \varepsilon (2\mathcal{L} + 1)}{2} \right] - 2V \cdot \varepsilon \sum_{m=1}^\ell \left\{ C \left( \frac{2\mathcal{L} - 1}{m - 1} \right) \left( \frac{1}{2} \right)^{2\mathcal{L} - 1} m \right\}$$

When the candidate selection procedure is noncompetitive, the problem for each selected candidate under open-list PR is slightly simpler (as the candidate selection stage does not rely on effort provision) and given by:

$$\left\{ \sum_{m=1}^\ell C \left( \frac{2\mathcal{L} - 1}{m - 1} \right) (P)^{2\mathcal{L} - m} (1 - P)^{m-1} \right\} V - \frac{1}{2} \left( e^L_l \right)^2.$$

The associated first-order condition boils down to:

$$e^L_{l^*} = V \cdot \sum_{m=1}^\ell \left\{ C \left( \frac{2\mathcal{L} - 1}{m - 1} \right) \left( \frac{1}{2} \right)^{2\mathcal{L} - 1} (2\mathcal{L} - 2m + 1) \varepsilon \right\}.$$

How do the above findings impact on proposition 3? Not much. Indeed, intuition suggests that closed-list PR may lead to higher effort provision than open-list PR when electoral incentives are quite weak as, in that case, the strategic use of competition for the best positions on the party list under closed-list PR may push candidates to exert more effort. What if the candidate selection procedure is non-competitive? Then, politicians face only the electoral hurdle. Here, because the competition for a seat in the legislature is more individually-based under open-list than under closed-list PR, intuition suggests effort provision should be higher under open-list PR. The following proposition confirms that the above intuitions are indeed correct:
Proposition 6  Comparing equilibrium effort provision across electoral rules:

1) if parties do not adopt a competitive candidate selection procedure, closed-list PR is the electoral rule that is associated with the lowest level of effort provision;

2) if parties adopt a competitive candidate selection procedure,
2.a) FPTP is the electoral rule that is associated with the lowest level of effort provision.
2.b) provided electoral incentives are weak enough, open-list PR is associated to lower effort provision than closed-list PR.

The above proposition also provides a theoretical rejoinder for the partially conflicting empirical findings of Kunicova and Rose-Ackermann (2005) and Chang and Golden (2006): which of closed- or open-list PR is associated with less corruption (to be interpreted here as the inverse of effort provision) depends on which type of candidate selection procedure parties adopt. Thus, cross-country empirical exercises on the causes of corruption may yield conflicting results if proxies for the competitiveness of the candidate selection procedure are omitted from the analysis.

7 Conclusion

This paper proposed a formal model to study how differences in the competitiveness of candidate selection procedures map into different chosen levels of effort by candidate legislators. The main result of the analysis was to show that the way parties organize their candidate selection procedure is of first order importance in explaining effort provision by candidates.

Further, the model highlighted that the power of electoral incentives, but not the electoral rule itself, and the competitiveness of the candidate selection procedure are substitutes. Finally, extending the model to allow for voter preferences to be ideologic allowed us to highlight that the degree of competitiveness and the degree of decentralization of candidate selection procedures are two distinct facets of these processes.

Where do we go from here? An obvious avenue for further research is the collection of data on both candidate selection procedures and the candidates parties propose to the electorate. Shomer (2009, 2012) are recent empirical studies that follow this route.
Another upshot of the present paper is that formal model of politics should investigate more deeply the role parties and their internal structures play in shaping the decisions taken by the different players in the game. To say it another way, to further our understanding of politics and the consequences it has on the formation of government policy, we need to go even further than what has been done so far in terms of micro-analysis of political games, by moving beyond games between politicians and voters only, in favor of games in which parties are active players who make strategic decisions too. If this paper focused on the candidate selection procedure, other dimensions of the way parties organize are of first order importance too. Obvious such dimensions include the way parties select their ministers and other executive office holders: should the current leader of the party be free to pick their favorite party members, or should this selection be more democratic and inclusive? What consequences do these choices have? One recent study offering an answer to some of these questions is Dewan and Hortala-Vallve (2011).

Finally, if our theory suggests that large districts may make intraparty incentive effects more powerful under PR, they may also make the electorate face an embarrassment of riches in terms of the number of candidates or parties to choose from. The optimal solution may then be to adopt medium-sized districts, as argued recently by Carey and Hix (2011).

References


Comparative Political Studies 40: 755-781.


8 Formal Proofs

8.1 Proof of Lemma 1

The first part follows directly from equations (4) and (8). To prove the second part of Lemma 1 we compute explicitly equilibrium effort for $L = 2$, 3 and 4 to prove that equilibrium effort is increasing in $\varepsilon$ for $L = 2$, 3 but decreasing in $\varepsilon$ for $L = 4$ and then prove that $\varepsilon$ impacts increasingly negatively on effort as we let $L$ take larger and larger values.

From (11), for $L = 2$, equilibrium effort is equal to $V$ times:

\[
\frac{3\phi(0)}{4} + \frac{\varepsilon}{4}
\]

and thus effort is obviously increasing in $\varepsilon$.

For $L = 3$, equilibrium effort is equal to $V$ times:

\[
\frac{15}{16}\phi(0) + \varepsilon \left( \frac{1}{6} - \frac{5\phi(0)}{24} \right) \simeq 0.37401 + 0.0835\varepsilon
\]

and thus effort is still increasing in $\varepsilon$.

For $L = 4$, equilibrium effort is equal to $V$ times:

\[
\frac{35}{32}\phi(0) + \varepsilon \left( \frac{1}{8} - \frac{112}{256}\phi(0) \right) \simeq 0.43634 - 0.0495\varepsilon
\]

and thus effort is now decreasing in $\varepsilon$.

To prove that equilibrium effort provision is decreasing in $\varepsilon$ for any $L \geq 4$ it suffices to show that the derivative of equilibrium effort – equation (11) – with respect to $\varepsilon$ is decreasing in $L$ for any $L \geq 4$.

\[25\] Note to the Editor and the referees: some of the proofs below could be streamlined. We kept all the steps so as to make the checking of the proofs easier.
This derivative is given by:

\[
V \sum_{m=1}^{\mathcal{L}} \left\{ C \left( \frac{2\mathcal{L} - m - 1}{2} \right) \left( \frac{1}{2} \right)^{2\mathcal{L} - 2} \phi(0) \left( 2\mathcal{L} - 2m + 1 \right) \left( \frac{\mathcal{L} - 2m + 1}{\mathcal{L}} \right) \right\} + \frac{V}{2\mathcal{L}}.
\]

For \( \mathcal{L} + 1 \), this derivative is equal to

\[
V \sum_{m=1}^{\mathcal{L}+1} \left\{ C \left( \frac{2\mathcal{L} + 1 - m - 1}{2} \right) \left( \frac{1}{2} \right)^{2\mathcal{L}} \phi(0) \left( 2\mathcal{L} - 2m + 3 \right) \left( \frac{\mathcal{L} - 2m + 2}{\mathcal{L} + 1} \right) \right\} + \frac{V}{2(\mathcal{L} + 1)}.
\]

The difference between the second and the first expression above is equal to \( V \) times:

\[
\sum_{m=1}^{\mathcal{L}+1} \left\{ C \left( \frac{2\mathcal{L} + 1 - m - 1}{2} \right) \left( \frac{1}{2} \right)^{2\mathcal{L}} \phi(0) \left( 2\mathcal{L} - 6m - \frac{10}{\mathcal{L} + 1} m + \frac{4}{\mathcal{L} + 1} m^2 + \frac{6}{\mathcal{L} + 1} + 7 \right) \right\}
- \sum_{m=1}^{\mathcal{L}} \left\{ C \left( \frac{2\mathcal{L} - m - 1}{2} \right) \left( \frac{1}{2} \right)^{2\mathcal{L}-2} \phi(0) \left( 2\mathcal{L} - 6m - \frac{4}{\mathcal{L} m} + \frac{4}{\mathcal{L} m^2} + \frac{1}{\mathcal{L}} + 3 \right) \right\}
- \frac{1}{2\mathcal{L} (\mathcal{L} + 1)}
\]

We need to show that this difference is negative for any \( \mathcal{L} \geq 4 \).

Using Blaise Pascal’s Triangle for combinations, we have that

\[
C \left( \frac{2\mathcal{L} + 1}{m - 1} \right) = C \left( \frac{2\mathcal{L}}{m - 1} \right) + C \left( \frac{2\mathcal{L}}{m - 2} \right)
= C \left( \frac{2\mathcal{L} - 1}{m - 1} \right) + 2C \left( \frac{2\mathcal{L} - 1}{m - 2} \right) + C \left( \frac{2\mathcal{L} - 1}{m - 3} \right).
\]

Then

\[
\sum_{m=1}^{\mathcal{L}+1} \left\{ C \left( \frac{2\mathcal{L} + 1}{m - 1} \right) \left( \frac{1}{2} \right)^{2\mathcal{L}} \phi(0) \left( 2\mathcal{L} - 6m - \frac{10}{\mathcal{L} + 1} m + \frac{4}{\mathcal{L} + 1} m^2 + \frac{6}{\mathcal{L} + 1} + 7 \right) \right\}
\]

is equal to:

\[
\phi(0) \left( \frac{1}{2} \right)^{2\mathcal{L}-2} \sum_{m=1}^{\mathcal{L}+1} C \left( \frac{2\mathcal{L} - 1}{m - 1} \right) \left( 2\mathcal{L} - 6m - \frac{10}{\mathcal{L} + 1} m + \frac{4}{\mathcal{L} + 1} m^2 + \frac{6}{\mathcal{L} + 1} + 7 \right)
+ \phi(0) \left( \frac{1}{2} \right)^{2\mathcal{L}-2} \sum_{m=1}^{\mathcal{L}+1} C \left( \frac{2\mathcal{L} - 1}{m - 2} \right) \left( 2\mathcal{L} - 6m - \frac{10}{\mathcal{L} + 1} m + \frac{4}{\mathcal{L} + 1} m^2 + \frac{6}{\mathcal{L} + 1} + 7 \right)
+ \phi(0) \left( \frac{1}{2} \right)^{2\mathcal{L}-2} \sum_{m=1}^{\mathcal{L}+1} C \left( \frac{2\mathcal{L} - 1}{m - 3} \right) \left( 2\mathcal{L} - 6m - \frac{10}{\mathcal{L} + 1} m + \frac{4}{\mathcal{L} + 1} m^2 + \frac{6}{\mathcal{L} + 1} + 7 \right)
\]
Rewriting the above as a sum in terms for $m$ going from 1 to $L$ instead of $L + 1$, we have:

\[
- \frac{\phi(0)}{4} \left( \frac{1}{2} \right)^{2L-2} C \left( \frac{2L-1}{L} \right) \frac{5L - 1}{L + 1} C \left( \frac{2L-1}{L - 1} \right) - \frac{3\phi(0)}{4} \left( \frac{1}{2} \right)^{2L-2} C \left( \frac{2L-1}{L - 1} \right) \\
+ \frac{\phi(0)}{4} \left( \frac{1}{2} \right)^{2L-2} \sum_{m=1}^{L} C \left( \frac{2L-1}{m-1} \right) \left( 2L - 6m - \frac{10}{L + 1} m + \frac{4}{L + 1} m^2 + \frac{6}{L + 1} + 3 \right) + \phi(0) \\
+ \phi(0) \left( \frac{1}{2} \right)^{2L-2} \sum_{m=1}^{L} C \left( \frac{2L-1}{m-1} \right) \left( 2L - 6m - \frac{10}{L + 1} m + \frac{4}{L + 1} m^2 + \frac{6}{L + 1} + 3 \right) \\
+ \phi(0) \left( \frac{1}{2} \right)^{2L-2} \sum_{m=1}^{L} C \left( \frac{2L-1}{m-1} \right) \left( \frac{4m - 3}{L + 1} \right) + 2\phi(0) \\
+ \phi(0) \left( \frac{1}{2} \right)^{2L-2} \sum_{m=1}^{L} C \left( \frac{2L-1}{m-1} \right) \left( \frac{4m - 1}{L + 1} - \frac{9}{4} \right) + \phi(0)
\]

which simplifies to

\[
- \frac{\phi(0)}{4} \left( \frac{1}{2} \right)^{2L-2} C \left( \frac{2L-1}{L} \right) \frac{5L - 1}{L + 1} \\
+ \frac{\phi(0)}{4} \left( \frac{1}{2} \right)^{2L-2} \sum_{m=1}^{L} C \left( \frac{2L-1}{m-1} \right) \left( 2L - 6m - \frac{10}{L + 1} m + \frac{4}{L + 1} m^2 + \frac{6}{L + 1} + 3 \right) \\
+ \phi(0) \left( \frac{1}{2} \right)^{2L-2} \sum_{m=1}^{L} C \left( \frac{2L-1}{m-1} \right) \left( \frac{4m - 3}{L + 1} \right) \\
+ \phi(0) \left( \frac{1}{2} \right)^{2L-2} \sum_{m=1}^{L} C \left( \frac{2L-1}{m-1} \right) \left( \frac{4m - 1}{L + 1} \right) \\
- \frac{5}{4} \phi(0) - \frac{3\phi(0)}{4} \left( \frac{1}{2} \right)^{2L-2} C \left( \frac{2L-1}{L - 1} \right).
\]

Then, the difference between the above and

\[
\sum_{m=1}^{L} \left\{ C \left( \frac{2L-1}{m-1} \right) \left( \frac{1}{2} \right)^{2L-2} \phi(0) \left( 2L - 6m - \frac{4}{L} m + \frac{4}{L} m^2 + \frac{1}{L} + 3 \right) \right\} + \frac{1}{2L(L + 1)}
\]

\footnote{Given that for any function $f(m)$ and $k > 1$, $\sum_{m=1}^{L} C \left( \frac{2L-1}{m-k} \right) f(m)$ is equal to $\sum_{m=1}^{L} C \left( \frac{2L-1}{m-k} \right) f(m + 1)$, to $\sum_{m=1}^{L} C \left( \frac{2L-1}{m-k} \right) f(m + 2)$ and so on.}
is equal to
\[- \frac{\phi(0)}{4} \left( \frac{\frac{1}{2}2^L - 2}{L-1} \right) \frac{5L - 1}{L+1} \]
\[- \frac{1}{L(L+1)} \frac{\phi(0)}{4} \left( \frac{\frac{1}{2}2^L - 2}{L-1} \right) \sum_{m=1}^{L} C \left( \frac{2^L - 1}{m-1} \right) \left( -2Lm - 4m - L + 4m^2 + 1 \right) \]
\[- \frac{5}{4} \frac{\phi(0)}{4} - 3\frac{\phi(0)}{4} \left( \frac{\frac{1}{2}2^L - 2}{L-1} \right) - \frac{1}{2L(L+1)}. \]

Further, using the following two identities involving combinations:
\[ \sum_{m=0}^{2L-1} \frac{C}{m} \left( \frac{2L - 1}{m-1} \right) = 2^{2L-1}, \]
\[ \sum_{m=0}^{2L-1} \frac{C}{m} \left( \frac{2L - 1}{m-1} \right) m = (2L - 1)2^{2L-2}, \]
and the fact that, for \( L > 4 \), we have that
\[ \sum_{m=1}^{L} \frac{C}{m} \left( \frac{2L - 1}{m-1} \right) m^2 > (L - 1) \sum_{m=1}^{L} \frac{C}{m} \left( \frac{2L - 1}{m-1} \right) m, \]
and the fact that, exploiting the symmetry property of the values combinations take on, we have:
\[ \sum_{m=1}^{L} \frac{C}{m} \left( \frac{2L - 1}{m-1} \right) \left( \frac{1}{2} \right)^{2L-2} = 1, \]
\[ \sum_{m=1}^{L} \frac{C}{m} \left( \frac{2L - 1}{m-1} \right) \left( \frac{1}{2} \right)^{2L-2} m < 1 \sum_{m=1}^{2L-1} \frac{C}{m} \left( \frac{2L - 1}{m-1} \right) \left( \frac{1}{2} \right)^{2L-2} m \]
\[ < \frac{1}{2} \sum_{m=0}^{2L-1} \frac{C}{m} \left( \frac{2L - 1}{m} \right) \left( \frac{1}{2} \right)^{2L-2} m = L - \frac{1}{2}, \]
the difference we need to evaluate is smaller than
\[- \frac{\phi(0)}{4} \left( \frac{\frac{1}{2}2^L - 2}{L-1} \right) \frac{5L - 1}{L+1} \]
\[+ \frac{2L - 4}{L(L+1)} \frac{\phi(0)}{4} \left( \frac{\frac{1}{2}2^L - 2}{L-1} \right) + \frac{\frac{1}{2}2^L + 1}{L(L+1)} \phi(0) \]
\[- \frac{1}{L(L+1)} \frac{\phi(0)}{4} (L - 1) \left( \frac{\frac{1}{2}2^L - 2}{L-1} \right) \]
\[- \frac{1}{L(L+1)} \frac{\phi(0)}{4} \left( \frac{\frac{1}{2}2^L - 2}{L-1} \right) \frac{5L - 1}{L+1} \]
\[- \frac{5}{4} \frac{\phi(0)}{4} - 3\frac{\phi(0)}{4} \left( \frac{\frac{1}{2}2^L - 2}{L-1} \right) - \frac{1}{2L(L+1)}. \]
which is clearly negative as
\[
\phi(0) \left[ \frac{2\mathcal{L} + 4}{\mathcal{L}(\mathcal{L}+1)} \left( \mathcal{L} - \frac{1}{2} \right) + \frac{\mathcal{L}}{\mathcal{L}(\mathcal{L}+1)} - \frac{4}{\mathcal{L}(\mathcal{L}+1)} (\mathcal{L} - 1) \left( \mathcal{L} - \frac{1}{2} \right) - \frac{1}{\mathcal{L}(\mathcal{L}+1)} \right] \\
= -\frac{\phi(0)}{\mathcal{L}^2 + \mathcal{L}} (2\mathcal{L}^2 - 10\mathcal{L} + 5)
\]
is negative for any \( \mathcal{L} \geq 5 \). Thus effort is decreasing in \( \varepsilon \) for any \( \mathcal{L} > 3 \). QED

8.2 Proof of Lemma 2

We build on lemma 1.

Under FPTP, comparing (4) and (8) implies that parties find it in their interest to adopt the competitive candidate selection procedure if and only if \( \varepsilon < \phi(0) = 1/\sqrt{2\pi} \simeq 0.39894 \).

Thus, the competitiveness of the selection procedure and the power of electoral incentives are substitutes for the party.

Under PR, we rely on explicit computations for \( \mathcal{L} = 2, 3 \) and 4 and then exploit the fact that equilibrium effort is decreasing in \( \varepsilon \) for \( \mathcal{L} > 3 \) to prove that there always exists a threshold value of \( \varepsilon \) below which parties find it optimal to choose the competitive candidate selection procedure. To close the proof, we finish by proving that for \( \varepsilon = \phi(0) \), parties still find it optimal to use the competitive candidate selection procedure so that the cutoff for \( \varepsilon \) above which parties switch to the noncompetitive candidate selection procedure under PR is always higher than \( \phi(0) \).

From (11), for \( \mathcal{L} = 2 \), equilibrium effort is equal to \( V \) times:
\[
\frac{3\phi(0)}{4} + \frac{\varepsilon}{4}
\]
and thus parties find it optimal to adopt the competitive candidate selection procedure if and only if
\[
\frac{3\phi(0)}{4} + \frac{\varepsilon}{4} > \frac{\varepsilon}{2} \Leftrightarrow \varepsilon < 3\phi(0) \simeq 1.1968.
\]

For \( \mathcal{L} = 3 \), equilibrium effort is equal to \( V \) times:
\[
\frac{15}{16}\phi(0) + \varepsilon \left( \frac{1}{6} - \frac{5\phi(0)}{24} \right)
\]
and thus parties find it optimal to adopt the competitive candidate selection procedure if and only if
\[ \frac{15}{16} \phi (0) + \varepsilon \left( \frac{1}{8} - \frac{5\phi (0)}{24} \right) > \varepsilon \Leftrightarrow \varepsilon < \frac{15}{16} \left( \frac{\phi (0)}{1 + \frac{5\phi (0)}{24}} \right) \approx 1.4974 \]

For \( \mathcal{L} = 4 \), equilibrium effort is equal to \( V \) times:
\[ \frac{35}{32} \phi (0) + \varepsilon \left( \frac{1}{8} - \frac{112 \phi (0)}{256} \right) \]
and parties find it optimal to adopt the competitive candidate selection procedure if and only if
\[ \frac{35}{32} \phi (0) + \varepsilon \left( \frac{1}{8} - \frac{112 \phi (0)}{256} \right) > \varepsilon \Leftrightarrow \varepsilon < \frac{35}{32} \frac{1}{8} + \frac{112 \phi (0)}{256} \approx 1.4567 \]

From lemma 1, we know that the derivative of (11) with respect to \( \varepsilon \) is decreasing in \( \mathcal{L} \) for any \( \mathcal{L} \geq 4 \). We now show that, at \( \varepsilon = \phi (0) \), parties still want to use the competitive selection procedure, that is, we show that effort under PR is higher when parties choose the competitive candidate selection procedure. This is true if and only if:
\[
\sum_{m=1}^{\mathcal{L}} \left\{ C(2\mathcal{L} - m - 1) \left( \frac{1}{2} \right)^{2\mathcal{L}-2} (2\mathcal{L} - 2m + 1) \phi (0) \left[ \frac{1}{2} + \varepsilon \left( \frac{\mathcal{L}-2m+1}{\mathcal{L}} \right) \right] \right\} + \frac{\phi (0)}{2\mathcal{L}} > \frac{\phi (0)}{\mathcal{L}}
\]
\[
\Leftrightarrow \sum_{m=1}^{\mathcal{L}} \left\{ C(2\mathcal{L} - m - 1) \left( \frac{1}{2} \right)^{2\mathcal{L}-2} \left[ (2\mathcal{L} - 2m + 1) \left\{ \frac{\phi (0)}{2} + (\phi (0))^2 + \frac{(\phi (0))^2}{\mathcal{L}} - \frac{2m}{\mathcal{L}} (2\mathcal{L} - 2m + 1) (\phi (0))^2 \right\} \right] \right\} - \frac{\phi (0)}{2\mathcal{L}} > 0
\]

As \( \phi (0) < 1/2 \) and \( 2m \leq 2\mathcal{L} \), replacing \( m \) by \( \mathcal{L} \), we have:
\[
\sum_{m=1}^{\mathcal{L}} C\left( \frac{2\mathcal{L} - 1}{m - 1} \right) \left( \frac{1}{2} \right)^{2\mathcal{L}-2} \frac{2m}{\mathcal{L}} (2\mathcal{L} - 2m + 1) (\phi (0))^2
\]
\[
< \sum_{m=1}^{\mathcal{L}} C\left( \frac{2\mathcal{L} - 1}{m - 1} \right) \left( \frac{1}{2} \right)^{2\mathcal{L}-2} 2 (2\mathcal{L} - 2m + 1) (\phi (0))^2
\]
then
\[
\sum_{m=1}^{\mathcal{L}} \left\{ C(2\mathcal{L} - m - 1) \left( \frac{1}{2} \right)^{2\mathcal{L}-2} \left[ \frac{2\mathcal{L}-2m+1}{2} \left[ \phi (0) + 2 (\phi (0))^2 + 2 \left( \frac{\phi (0))^2}{\mathcal{L}} \right] - \frac{2m}{\mathcal{L}} (2\mathcal{L} - 2m + 1) (\phi (0))^2 \right] \right\} - \frac{\phi (0)}{2\mathcal{L}}
\]
\[
> \sum_{m=1}^{\mathcal{L}} \left\{ C(2\mathcal{L} - m - 1) \left( \frac{1}{2} \right)^{2\mathcal{L}-2} \left[ \frac{2\mathcal{L}-2m+1}{2} \left[ \phi (0) + 2 (\phi (0))^2 + 2 \left( \frac{\phi (0))^2}{\mathcal{L}} \right] - 2 (2\mathcal{L} - 2m + 1) (\phi (0))^2 \right] \right\} - \frac{\phi (0)}{2\mathcal{L}}
\]
\[
= \sum_{m=1}^{\mathcal{L}} \left\{ C(2\mathcal{L} - m - 1) \left( \frac{1}{2} \right)^{2\mathcal{L}-2} \left[ \frac{2\mathcal{L}-2m+1}{2} \left[ \phi (0) - 2 (\phi (0))^2 + 2 \left( \frac{\phi (0))^2}{\mathcal{L}} \right] \right] \right\} - \frac{\phi (0)}{2\mathcal{L}}
\]
\[
= \sum_{m=1}^{\mathcal{L}} \left\{ C(2\mathcal{L} - m - 1) \left( \frac{1}{2} \right)^{2\mathcal{L}-2} \left[ \frac{2\mathcal{L}-2m}{2} \left[ \phi (0) - 2 (\phi (0))^2 + 2 \left( \frac{\phi (0))^2}{\mathcal{L}} \right] \right] \right\} + \frac{\phi (0)}{2} \left( 1 - \frac{1}{\mathcal{L}} \right) + (\phi (0))^2 \left( \frac{1}{\mathcal{L}} - 1 \right).
\]
The first term,
\[
\sum_{m=1}^{\mathcal{L}} \left\{ C \left( \frac{2\mathcal{L} - 1}{m - 1} \right) \left( \frac{1}{2} \right)^{2\mathcal{L}-2} \left[ \frac{1}{2} (2\mathcal{L} - 2m) \left( \phi (0) - 2 (\phi (0))^2 + 2 \frac{(\phi (0))^2}{\mathcal{L}} \right) \right] \right\},
\]
is always positive. The sum of the two terms making up the rest of the expression above is also always positive: it takes on value 0 for \( \mathcal{L} = 1 \) and the derivative of that sum with respect to \( \mathcal{L} \) is equal to \( \frac{\phi (0)}{2 \mathcal{L}^2} - (\phi (0))^2 \frac{1}{4 \mathcal{L}}, \) which is strictly positive as \( \frac{\phi (0)}{2 \mathcal{L}^2} > (\phi (0))^2. \)

Thus, we can thus conclude that for any value of \( \varepsilon \) between 0 and \( \phi (0) \), parties will choose the competitive candidate selection procedure under both electoral rules. Further, under PR, the value of \( \varepsilon \) for which parties switch selection procedures is strictly greater than \( \phi (0) \).

To conclude, given that \( \frac{\partial e}{\partial \varepsilon} \) is strictly positive under PR when candidate selection is noncompetitive – as is obvious from (6) – we have that, under PR, when \( \varepsilon \) is small enough, parties find it optimal to select the competitive selection procedure, whereas when \( \varepsilon \) is large enough, parties find it optimal to select the noncompetitive procedure.

This concludes the proof of lemma 2.

**8.3 Proof of Proposition 4**

That equilibrium effort provision is decreasing in \( \mathcal{L} \) when the candidate selection procedure is noncompetitive is obvious as equilibrium effort is given by (6).

To prove the second part of proposition 4, we first prove that equilibrium effort when candidate selection is competitive under PR is increasing in \( \mathcal{L} \) if \( \varepsilon \) tends to 0. Letting \( \varepsilon \) tend to zero, equilibrium effort tends to

\[
V \sum_{m=1}^{\mathcal{L}} \left\{ C \left( \frac{2\mathcal{L} - 1}{m - 1} \right) \left( \frac{1}{2} \right)^{2\mathcal{L}-1} (2\mathcal{L} - 2m + 1) \phi (0) \right\}
\]

Denote the above expression as \( e (\mathcal{L}) \). We now prove that \( e (\mathcal{L} + 1) - e (\mathcal{L}) \) is strictly positive for any \( \mathcal{L} \). \( e (\mathcal{L} + 1) - e (\mathcal{L}) \) is given by:

\[
V \sum_{m=1}^{\mathcal{L}+1} \left\{ C \left( \frac{2\mathcal{L} + 1}{m - 1} \right) \left( \frac{1}{2} \right)^{2\mathcal{L}+1} (2\mathcal{L} - 2m + 3) \phi (0) \right\}
\]

\[
- V \sum_{m=1}^{\mathcal{L}} \left\{ C \left( \frac{2\mathcal{L} - 1}{m - 1} \right) \left( \frac{1}{2} \right)^{2\mathcal{L}-1} (2\mathcal{L} - 2m + 1) \phi (0) \right\}.
\]
Again, using Blaise Pascal’s Triangle for combinations, we have that $C^{(2\ell+1)}$ is equal to

$$C^{(2\ell-1 \atop m-1)} + 2C^{(2\ell-1 \atop m-2)} + C^{(2\ell-1 \atop m-3)}.$$ 

Then the difference to evaluate is equal to:

$$V\phi(0) \left( \frac{1}{2} \right)^{2\ell+1} \sum_{m=1}^{\ell+1} \left\{ C^{(2\ell-1 \atop m-1)}(2\ell-2m+3) \right\}$$

$$- V \sum_{m=1}^{\ell} \left\{ C^{(2\ell-1 \atop m-1)} \left( \frac{1}{2} \right)^{2\ell-1}(2\ell-2m+1)\phi(0) \right\},$$

namely:

$$\frac{V\phi(0)}{4} \left( \frac{1}{2} \right)^{2\ell-1} \sum_{m=1}^{\ell+1} \left\{ C^{(2\ell-1 \atop m-1)}(2\ell-2m+3) \right\}$$

$$+ \frac{V\phi(0)}{2} \left( \frac{1}{2} \right)^{2\ell-1} \sum_{m=1}^{\ell+1} \left\{ C^{(2\ell-1 \atop m-2)}(2\ell-2m+3) \right\}$$

$$+ \frac{V\phi(0)}{4} \left( \frac{1}{2} \right)^{2\ell-1} \sum_{m=1}^{\ell+1} \left\{ C^{(2\ell-1 \atop m-3)}(2\ell-2m+3) \right\}$$

$$- V \sum_{m=1}^{\ell} \left\{ C^{(2\ell-1 \atop m-1)} \left( \frac{1}{2} \right)^{2\ell-1}(2\ell-2m+1)\phi(0) \right\}$$

$$= \frac{V\phi(0)}{4} \left( \frac{1}{2} \right)^{2\ell-1} C^{(2\ell-1 \atop \ell)} + \frac{V\phi(0)}{2} \left( \frac{1}{2} \right)^{2\ell-1} \sum_{m=1}^{\ell} C^{(2\ell-1 \atop m-1)}(2\ell-2m+1)$$

$$+ \frac{V\phi(0)}{4} \left( \frac{1}{2} \right)^{2\ell-1} \sum_{m=1}^{\ell} C^{(2\ell-1 \atop m-2)}(2\ell-2m+1) + \frac{V\phi(0)}{2} \left( \frac{1}{2} \right)^{2\ell-1} \sum_{m=1}^{\ell} C^{(2\ell-1 \atop m-3)}(2\ell-2m+1)$$

$$- V \sum_{m=1}^{\ell} \left\{ C^{(2\ell-1 \atop m-1)} \left( \frac{1}{2} \right)^{2\ell-1}(2\ell-2m+1)\phi(0) \right\}$$

$$= \frac{V\phi(0)}{4} \left( \frac{1}{2} \right)^{2\ell-1} C^{(2\ell-1 \atop \ell)} + \frac{V\phi(0)}{4} \left( \frac{1}{2} \right)^{2\ell-1} C^{(2\ell-1 \atop \ell-1)}$$

which is strictly positive.

Thus, when $\varepsilon$ tends to zero, parties find it in their interest to adopt the competitive candidate selection procedure and the more so the larger is average district magnitude, as
equilibrium effort provision by their candidates is strictly increasing in $L$. By continuity, this result holds too for small but strictly positive values of $\varepsilon$, that is, the result holds too when the power of electoral incentives is weak but strictly positive.

### 8.4 Proof of Proposition 5

The proof of the first part of proposition 5 is immediate by comparing directly (4) and (6).

To prove the second part of the proposition, we know from lemma 2 that for $\varepsilon$ close to 0, parties find it optimal to adopt the competitive candidate selection procedure under PR, just like under FPTP. Let $\varepsilon$ tend to zero, then optimal effort provision under the competitive candidate selection procedure under PR tends to

$$\sum_{m=1}^{L} \left\{ C \left( \frac{2L-1}{m-1} \right) \left( \frac{1}{2} \right)^{2L-1} (2L - 2m + 1) \phi (0) \right\}$$

$$= L\phi (0) - 2\phi (0) \sum_{m=1}^{L} C \left( \frac{2L-1}{m-1} \right) \left( \frac{1}{2} \right)^{2L-1} m + \frac{\phi (0)}{2}$$

$$> L\phi (0) - \phi (0) \left( L - \frac{1}{2} \right) + \frac{\phi (0)}{2}$$

$$= \phi (0)$$

where we made use of the fact that\(^{27}\)

$$2 \sum_{m=1}^{L} C \left( \frac{2L-1}{m-1} \right) \left( \frac{1}{2} \right)^{2L-1} m < \frac{1}{2} \sum_{m=1}^{2L-1} C \left( \frac{2L-1}{m-1} \right) \left( \frac{1}{2} \right)^{2L-2} m$$

$$< \frac{1}{2} \sum_{m=0}^{2L-1} C \left( \frac{2L-1}{m} \right) \left( \frac{1}{2} \right)^{2L-2} m = L - \frac{1}{2}.$$

As optimal effort under FPTP with the competitive selection procedure tends to $\phi (0) / 2$ when $\varepsilon$ tends to 0, effort under PR is greater than under FPTP. By the continuity of both optimal effort functions, this results holds too for small but strictly positive values of $\varepsilon$.

This completes the proof of proposition 5.

### 8.5 Proof of proposition 6

When selection is competitive, equilibrium effort under open-list PR is given by (18). Yet, we have that:

$$\sum_{m=1}^{L} C \left( \frac{2L-1}{m-1} \right) \left( \frac{1}{2} \right)^{2L-1} m < \sum_{m=1}^{L} C \left( \frac{2L-1}{m-1} \right) \left( \frac{1}{2} \right)^{2L-1} L = \frac{L}{2}.$$

\(^{27}\)Notice that the sum in the second term runs up to $2L - 1$ and not $L$ only.
Then, equilibrium effort under open-list PR is always strictly greater than:

\[ V \cdot \frac{\phi(0) + \varepsilon(2L+1)}{2} - V \varepsilon L = V \cdot \frac{\phi(0) + \varepsilon}{2} \]  

(21)

and thus open-list PR does always better than FPTP when parties adopt a competitive candidate selection procedure.

Comparing optimal effort under open-list PR to that under closed-list PR, (11), if the power of electoral incentives is low enough, that is, if \( \varepsilon \) is low enough, effort under closed-list PR is higher than that under open-list PR. Indeed, let \( \varepsilon \) tend to 0, then effort under open-list PR tends to \( \phi(0)/2 \), which is lower than that under closed-list PR, which tends to:

\[
\sum_{m=1}^{L} C \left( \frac{2L-1}{m-1} \right) \left( \frac{1}{2} \right)^{2L-1} (2L - 2m + 1) \phi(0)
\]

\[
= \frac{\phi(0)}{2} + \phi(0) \left[ 2 \sum_{m=1}^{L} C \left( \frac{2L-1}{m-1} \right) \left( \frac{1}{2} \right)^{2L-1} - 2 \sum_{m=1}^{L} C \left( \frac{2L-1}{m-1} \right) \left( \frac{1}{2} \right)^{2L-1} \right].
\]

When selection is noncompetitive, equilibrium effort is given by (20), which is always strictly greater than effort provision under closed-list PR \( \varepsilon V/L \) and also than effort provision under FPTP, provided \( L \) is large enough, as \( e^{L^*} \) is increasing in \( L \).

This completes the proof of proposition 6.