Party Nomination Strategies in Closed and Flexible List PR*

PRELIMINARY

Peter Buisseret[†] Olle Folke[‡] Carlo Prato[§] Johanna Rickne[¶]

September 18, 2018

Abstract

We develop a model of candidate selection strategies in list proportional representation systems, in which parties allocate candidates of heterogeneous quality (*valence*) to list ranks in order to trade off electoral returns and policy goals. We generate and test two competing theoretical predictions using data on the competence characteristics of the complete universe of Swedish politicians from 1982 to 2014. We reject a *marginal rank* hypothesis that parties should allocate their best candidates to ballot ranks that hang in the electoral balance (the "marginal" ranks). However, we unearth striking evidence in favor of a *top-down rank order* hypothesis, which predicts that parties should place better candidates in higher ballot ranks and that this pattern should be especially pronounced in parties that expect to attain majority status, and thus to appoint politicians to executive posts.

*Financial support from the Swedish Research Council and the Torsten Söderberg Foundation is gratefully acknowledged. We thank Wiola Dziuda and Anthony Fowler for valuable comments and suggestions. We also thank Jonas Ahlerup, Johan Arntyr, Sirus Dehdari, Roza Khoban, and Elin Molin for excellent research assistance.

[†]Harris School of Public Policy, University of Chicago, *Email*: pbuisseret@uchicago.edu

[‡]Department of Government, Uppsala University *Email*: olle.folke@statsvet.uu.se

[§]Department of Political Science, Columbia University, *Email*: cp2928@columbia.edu

[¶]Swedish Institute for Social Research, Stockholm University, *Email*: johanna.rickne@sofi.su.se

1. Introduction

In a representative democracy, elections are the principal mechanism for voters to select skilled and able policymakers. This selection is critically mediated by two distinct forces. First, political parties act as gate-keepers to candidates that aspire to political office. Second, electoral institutions govern how individual votes contribute to a candidate's prospect of election. These two forces, however, do not act in isolation: electoral rules exert their own powerful effects on parties' internal procedures and hierarchies, shaping how parties organize in order to pursue their goals. In this paper, we use theory and data to study how political parties structure candidate selection in closed and flexible list proportional rule ("list PR") electoral systems—the modal electoral rules for electing the world's democratic legislatures.[1]

Lists are the foundation of a party's gate-keeping authority in PR systems. Promoting a candidate to a high rank on the list may virtually guarantee her election; conversely, demoting her to a low rank almost ensures her defeat. We develop and test two competing principles of candidate selection. The first is a *marginal rank* principle, which predicts that purely seat-maximizing parties should place their best politicians in the ballot ranks that are most likely to hang in the balance, i.e., are electorally marginal ranks. The second is a *top-down rank order* principle, which predicts that when political goals are achieved not only by winning additional seats but also through the election of high quality legislative leaders, parties should place their best politicians in the most electorally secure ballot ranks, i.e., the top ranks. Our model predicts that this pattern should be especially pronounced amongst parties (or party coalitions) that enjoy a relatively strong prospect of wielding executive authority.

Our theoretical framework applies to any parliamentary context—local or national—in which voters select politicians to represent them in multi-member districts under closed or flexible list PR election systems. We test our predictions using data on the competence characteristics and list ranks of Swedish politicians in seven municipal electoral cycles. Our empirical analysis decisively rejects the *marginal rank* hypothesis, and uncovers strong support for the *top-down rank order* hypothesis.

Our Approach. In our model, parties have access to a pool of candidates that are differentiated by their quality—i.e., *valence*—a characteristic that voters observe and value. Party leaders rank candidates on a list; all else equal, a higher ranking implies a higher prospect of election for that candidate. Voters observe the party's list and cast their vote for one of

[1] According to the *Database of Political Institutions* (Cruz, Keefer, and Scartascini, 2016), as of 2015 legislative elections in 94 out of 147 democracies employ proportional representation.

the parties. Leaders' choices balance their concerns for a range of outcomes, including their expected share of seats in the assembly and the quality of their post-election legislative team.

We unearth how voters choose between parties as a function of (i) where the parties place their strongest candidates, and (ii) each party's underlying popularity. And, we generate empirically testable predictions about how parties should design lists, depending on the relative importance of seat shares (e.g., majority status) *versus* politician quality in a party's pursuit of policy goals, and whether there are ex-ante sources of electoral imbalance between parties, e.g., a net valence advantage or partisan context that favors one of the parties.

We then test these predictions using the world's most comprehensive and detailed dataset on political selection. Our empirical analysis focuses on Swedish elections at the municipal level, where the local party organizations of the eight established political parties operate in nearly every one of the 290 municipalities. Each municipality replicates the country's national parliamentary democracy, with elections to the council (i.e., legislative assembly) and subsequent appointment of the mayor (i.e., the chief executive) by the largest party in the governing coalition. The parties that have representation in the parliament also have local party branches across all, or most, municipalities, and these local parties enjoy complete autonomy in determining the rank-order of their electoral ballots.

Parties are required to report their ballots to the electoral authorities with the (mandatory) personal identification (ID) code of every candidate in every list rank. ID codes for all nominated politicians on all electoral ballots are linked to other administrative registers held by the Swedish government that include a host of demographic information, whether the individual secured election, and any executive posts he or she may have held. We use the information in these registers to develop four distinct measures of candidate quality.

In each of seven elections (1991 to 2010) we therefore observe the complete quality-rank ordering of candidates for over 2,000 local parties. Our variables that measure candidate quality come from high-quality administrative data; they are never self-reported, and have very few missing values. The size of this dataset creates high statistical precision in observing how parties organize ballots. It also allows sample splits, for example, by the electoral imbalance in favor of one party in the municipality, and thus the competitiveness of local elections.

Our formal model focuses on purely closed list contexts. However, Buisseret and Prato (2018) show that when list flexibility is not too large, outcomes are observationally equivalent to a fully closed list. This prediction is reflected in our data: in our period of study, about 1% of all municipal councilors secured election purely because of preference votes, i.e., they would not have been elected solely on the basis of their ranking. Nonetheless, a 1998 reform that transitioned from *closed* to *flexible* list PR offers us an additional window into individual voting behavior, since we are able to observe the distribution of preferences votes *within* a party's list.

Predictions and Findings. We begin with a benchmark model in which party leaders care solely about designing their lists to maximize their share of seats. This could reflects an institutional setting in which the winning party or coalition enjoys considerable agenda-setting powers and executive appointments. Thus, we ask how parties should design their lists in order to maximize their appeal to voters. We derive a *marginal rank hypothesis*, which predicts that parties should place their very best candidates in the ballot ranks whose prospects of election hang most in the balance.

For example, if a party typically wins between four and six out of ten possible seats, its fifth-ranked candidate's prospect of election is in greater contention that its first-ranked candidate (who is almost sure to win election) and its tenth-ranked candidate (who is almost sure to lose). The marginal rank hypothesis predicts that vote-maximizing parties should place their strongest candidates in the ranks whose election hangs in the balance; by doing so, they bolster the appeal of supporting the party's list. Our theoretical model identifies contexts—i.e., subsets of our data—in which the clearest empirical patterns should be visible.

We then extend our model to admit the possibility that—like voters—political parties also place a premium on the skills and experience of their elected politicians, and that these skills are an essential input to the achievement of a party's legislative goals. We allow for the possibility that legislative parties view majority control and the human capital of their legislative teams as *complements*. In that environment, we derive a competing *top-down rank order* hypothesis: parties should prioritize their very best candidates in higher ranks on their lists. And, even more importantly, it is electorally advantaged parties with a strong expectation of enjoying majority control that derive the greatest relative value from choosing lists that place their best candidates in better ranks.

We confront our competing theories with data. We use historical election results in each of our municipalities to construct predicted probabilities, for each electoral ballot, that the candidate on each of the list ranks wins a seat. We then use these predicted probabilities to identify the 'marginal' ranks on each electoral ballot in every local election. We develop four alternative measures of a politician's quality, i.e., valence, and relate these measures to a politician's position on his or her party's electoral ballot.

We decisively reject the marginal rank hypothesis, but find strong support for the *top-down rank order* hypothesis. In particular, the electoral ballots that generate the strongest positive relationship between ballot rank and our measures of quality are parties that have a history of executive control—in particular, appointing the chief executive.

Our Contribution. While considerable attention has been devoted to the effects of electoral rules on *inter*-party competition—e.g., the number of parties (Duverger, 1959)—we focus on their consequences for *intra*-party organization. Like Eguia (2011), Krasa and Polborn (2018), and Dewan and Hortala-Vallve (2013), our theoretical framework explicitly acknowledges that parties are *teams* of politicians, both in elections and in post-election assemblies. Our model allows us to assess different views of how both voters and party leaders value these teams, as well as the individual politicians that constitute these teams. It also explores how skills and experience complement institutional features such as the relative value of majority control in a legislative assembly. More broadly, our framework asks: how do the qualities of individual politicians in a team contribute to a party's electoral and legislative *capacity*?

Our paper studies party nomination processes and their implications for political selection. These issues are widely studied in single-member contexts—notably, the trade-off between democratic versus elite-driven candidate selection—both theoretically (e.g., Caillaud and Tirole, 2002 and Crutzen, Castanheira, and Sahuguet, 2010) and empirically (e.g., Hirano and Snyder, 2014). Our paper addresses this question both theoretically and empirically in a multi-member settings. It therefore contributes to a growing literature on political selection in list PR systems, whose focus has included: the link between personal votes and political promotions in Finland (Meriläinen and Tukiainen, 2018) and Sweden (Folke, Persson, and Rickne, 2016); gender quotas and female representation on party lists in Italy (Baltrunaite, Bello, Casarico, and Profeta, 2014), Spain (Esteve-Volart and Bagues, 2012) and Sweden (Besley, Folke, Persson, and Rickne, 2017); and, a party's trade-off between promoting loyalists and experts on the ballot paper (Galasso and Nannicini, 2015).

We contribute to a small but growing theoretical literature on candidate selection and internal party dynamics in list PR electoral systems. The most closely related is Matakos, Savolainen, Troumpounis, Tukiainen, and Xefteris (2018), who also distinguish between strategies that maximize a party's electoral appeal, and strategies that maximize its post-election success in legislative politics. In their setting, however, the trade-off turns not on competence but ideology: a list with broad ideological appeal appeals to a wider swathe of voters, but leads to a less cohesive post-election elected team. Crutzen, Flamand, and Sahuguet (2018) study how rank assignments can solve intra-party moral hazard problems and encourage candidate effort.

Organization. We develop our theoretical model of elections and legislative policymaking in

Section 2. We solve the model and obtain theoretical predictions in Section 3. Background information on our empirical setting is provided in Section 4, and Section 5 provides detail on our data. We develop measures of candidate quality and party competition in Section 6, and use these to test our theoretical predictions in Section 7. A discussion of alternative frameworks follows, and a conclusion. Proofs and additional results are contained in Appendix.

2. Model

Agents. We consider a two-date interaction between a unit mass of constituency voters, and two political parties, A and B. The parties compete in an election for control of a legislative assembly, which consists of five elected representatives. Each party consists of five candidates, who differ in their observed human capital q; without loss of generality we denote candidates by their human capital level: $q \in Q \equiv \{1, 2, 3, 4, 5\}$, where $q = 1$ denotes the lowest level, and $q = 5$ denotes the highest level.

The interaction proceeds in two stages: an *election* and a subsequent *legislative interaction*.

Election. Each party simultaneously chooses a *list* that determines the order in which its candidates appear on the ballot. Specifically, a list is $l_J \in L(Q)$—where $L(Q)$ is the set of strict orderings over Q—and its k^{th} element is denoted by $l_J(k)$.

Our model focuses on *closed list* contexts.[2] Thus, after the parties choose their lists, voters cast their ballots in an election for one of the two party lists. After votes are cast, each party $J \in \{A, B\}$ is allocated n_J seats, where $n_A + n_B = 5$. Seats are allocated proportionally across the parties: for each seat $k \in \{1, ..., 5\}$ there is a threshold vote share $\pi(k) \in (0, 1)$ that a party needs to clear to secure that seat. So, when π_J is between $\pi(k)$ and $\pi(k + 1)$, $n_J = k$, and $n_{-J} = 5 - k$. In keeping with real-world PR systems, we impose $\pi(3) = \frac{1}{2}$ and for all $k < 3$, $\pi(k) = 1 - \pi(6 - k)$.

Each party subsequently allocates its seats to the first n_J candidates that appear on its list—denoted by $S_J \equiv \{l_J(1), \ldots, l_J(n_J)\}$. Thus, the assembly is $S \equiv S_A \cup S_B$. The is headed by a chairperson—we call the *chief executive*, denoted by $m \in S$. For simplicity, we assume that the party obtaining a majority of the seats appoints the chief executive.

Voters care about the quality of governance. This depends positively on the total human

[2] Our data contains both closed and flexible list electoral settings. However, Buisseret and Prato (2018) fully analyze flexible list systems, and show that when list flexibility is non-zero—but not too large—outcomes are observationally equivalent to a fully closed list. This is reflected in our data: candidates virtually never receive sufficient preference votes to overturn their list ranking.

capital of the assembly—$\sum_{i \in S} i$—and the human capital of the chief executive (m).[3] These preferences are represented by the payoff function

$$H\left(\sum_{i \in S} i, m\right) : \{9, ..., 21\} \times Q \to \mathbb{R}.$$

We impose minimal structure by assuming only that governance quality increases in the aggregate human capital of the assembly and the human capital of the chief executive.

Assumption 1. H strictly increases in each of its arguments.

Legislative Interaction. After the election, the political parties divide a *surplus* whose total value we normalize to one. This surplus summarizes both material resources that can be channeled to specific individuals and groups (e.g., jobs) as well as policymaking authority. We assume that party $J \in \{A, B\}$'s share of the surplus depends on (i) the number of its allocated seats in the assembly (n_J), (ii) the human capital of the chief executive, and (iii) its share $\overline{q}_J(S) \equiv \frac{\sum_{i \in S_J} i}{\sum_{i \in S} i}$ of the aggregate human capital of the assembly. A party that obtains k seats in assembly S enjoys a payoff

$$G(k, m, \overline{q}_J) : \{1, \ldots, 5\} \times Q \times [0, 1] \to [0, 1].$$

Assumption 2 $G(k, m, \overline{q}_J)$ (i) strictly increases in k, (ii) weakly increases in $\overline{q}_J$, and (iii) weakly increases in m if $k \geq 3$ and weakly decreases in m if $k < 3$.

Assumption 2 states that a party's share of the surplus always increases with its share of the seats; it *increases* with the quality of the chief executive when it appoints the chief executive from its own ranks and does *not* increase when the opposing party appoints a higher-quality chief executive; and, it *does not* decrease when its share of the total skills and experience in the assembly increases relative to the opposing party's share.

Our formulation allows parties to attach different value to different seats. For example, in a highly majoritarian legislative context where all power derives from majority control, we have $G(3, \cdot, \cdot) - G(2, \cdot, \cdot) \approx 1$. Nonetheless, our formulation also allows for a more progressive distribution of political power as a function of the number of seats.

Voting Behavior. According to a process that we set out in detail, below, each voter i computes an equilibrium expected value of voting for each party V_J and votes for party B if and

[3] The model can be extended without any significant additional insights to multiple executive offices and to a probabilistic allocation of the chief executive across parties.

only if

$$V_B \geq V_A + \sigma_i + \xi \tag{1}$$

where σ_i is a voter-specific preference shock drawn from a zero-mean uniform with density ϕ and support $[-(2\phi)^{-1}, (2\phi)^{-1}]$, and ξ is an aggregate preference shock drawn from a distribution $F(x)$ with support $[-(2\psi)^{-1}, (2\psi)^{-1}]$ and density $f(\cdot) = \psi(1-\theta) + \theta\tilde{f}(\cdot)$, where $\tilde{f}(\cdot)$ is single-peaked, continuous, differentiable and symmetrically distributed around its peak, $\overline{\xi} > 0$. Thus, $\mathbb{E}[\xi] > 0$, i.e., the aggregate shock is expected to favor party A.[4] Larger $\psi > 0$ corresponds to voter preferences that are more responsive to variation in the quality of a party's candidates in the l elections; higher $\overline{\xi} > 0$ reflects a greater *structural advantage* enjoyed by party A that could derive from local or national conditions. For our empirical work, it is valuable to observe that—all else equal—party A expects to receive a larger share of seats than party B.

Reflecting real-world electoral contexts, we allow for the possibility that each party is sure to win at least one seat:

Assumption 1. *ϕ is sufficiently small:*

$$\underline{\pi} = \frac{1}{2} + \phi(-\overline{V} + \underline{V} - (2\psi)^{-1}) \in (\pi(1), \pi(2)),$$

where $\overline{V}$ and $\underline{V}$ denote the largest and smallest values from a party's list.

Assumption 1 implies that each party expects to receives at least one seat (an *uncontested* seat). We refer to the remaining three seats as *contested* seats.[5]

Summary of Timing. To summarize, the timing of the events is as follows:

1. Each party constructs an electoral ballot l_J
2. Each voter computes values V_A and V_B
3. The common (ξ) and idiosyncratic (σ_i) shocks are realized and votes are cast
4. Seats are allocated and the majority-winning party appoints the chief executive.

[4] That is, $\int_{-\frac{1}{2\psi}}^{\frac{1}{2\psi}} zg(z)dz > 0$.

[5] That Assumption 1 implies both parties have one uncontested seat follows from the assumption $\pi(k) = 1 - \pi(6-k)$.

Discussion. Our framework extends easily to an arbitrary district magnitude. As we detail below, however, we subsequently group together list ranks in our sample, maintaining three distinct categories of "contested ranks". Our hypotheses about the quality of a candidate in a party's *advantaged competitive* rank in our model can be interpreted as the average quality of candidates in the subset of contiguous ranks that we classify as advantaged competitive ranks in our data.

3. Theoretical Results

Chief Executive Appointment. Recall that whichever party wins a majority of seats appoints the chief executive. Since a party benefits from higher quality leadership when it enjoys majority status, it chooses its best elected politician to lead the assembly.

Lemma 1. *If $G(k, m, \overline{q})$ strictly increases in m, then the majority party appoints its highest-quality elected official to be chief executive.*

We may therefore denote the equilibrium identity of the chief executive is $\hat{m}(n_A, l_A, l_B)$, which depends on the lists l_A and l_B, as well as the number of seats n_A awarded to party A (recall that $n_B = 5 - n_A$):

$$\hat{m}(k, l_A, l_B) = \max\{l_A(1), \ldots, l_A(k)\}\mathbf{I}[k \geq 3] + \max\{l_B(1), \ldots, l_B(5-k)\}\mathbf{I}[k < 3]. \tag{2}$$

Instrumental Voting Behavior. Given the party's lists, l_A and l_B, voters decide whether to cast a vote for party A's list, or party B's list. This choice takes into account her conjectures about (i) the anticipated quality of the post-election assembly—including parties' appointment decisions—and (ii) her expectations about the impact that her vote will have on the outcome of the election.[6]

If there were only a single representative (i.e., a single-member system) in a two-candidate contest, the election outcome turns on a single event: one candidate winning a majority of the votes. Thus, in the hypothetical event that a voter *could* have been decisive for the outcome, she would always prefer to cast her ballot in favor of the candidate she most prefers. This motivates a restriction to *sincere* voting in single-member contexts.

In multi-member districts, by contrast, the consequences of a vote in favor of either ticket depend on how voters in the same district cast their ballots. In particular, conditional on her

[6] Since no voter can be decisive for the outcome, *any* voting behavior is consistent with optimality. The behavior that we develop below is a voting calculus that we believe is most faithful to the restriction to weak dominance in the single-member context.

vote being decisive, a voter that considers voting in favor of party B acknowledges that there are *three* relevant events . Each of these events corresponds to the election of a candidate positioned at rank $k \in \{2,3,4\}$ in each party's list (recall that we suppose that each party has a single *uncontested* seat). Moreover, the election of a candidate may also affect the identity of the chief executive, for example, via a change in majority control.

We develop a voting calculus in which an instrumental voter forms expectations about her relative likelihood of being decisive for each of these events. In particular, she anticipates that party B, being disadvantaged, will win relatively few seats. She therefore reasons that—conditional on being decisive—her vote is relatively more likely to affect the election prospects of a candidate that is ranked *higher* on party B's list. After all, if the party is relatively unlikely to win many votes, it is likely that only candidates located relatively high in the ballot order will have a chance of election. On the other hand, the voter anticipates that party A, being structurally advantaged, is likely to win a preponderant support in the electorate, and so its higher-ranked candidates are relatively likely to be elected. In that case, the voter anticipates that she is relatively more likely to affect the election prospects of a candidate that is ranked *lower* on party A's list.

Let $\tilde{H}(k, l_A, l_B)$ denote voters' payoff as a function of the number of seats awarded to A, the lists, and given $\hat{m}(k, l_A, l_B)$

$$\tilde{H}(k, l_A, l_B) = H\left(\sum_{i=1}^{k} l_A(i) + \sum_{i=1}^{5-k} l_B(i), \hat{m}(k, l_A, l_B)\right).$$

Our first result formalizes the intuition that voters focus relatively more on the characteristics of politicians for whom they expect their vote to count relatively more, and shows how voters endogenously assign their attention across each party's contested ranks.

Lemma 2. *For any party lists l_A and l_B, there exists a pair $V_A(l_A, l_B)$ and $V_B(l_A, l_B)$, and attention weights $\tau(k) \in (0,1)$ for $\kappa \in \{2,3,4\}$ satisfying $\sum_{k=2}^{4} \tau(k) = 1$, such that an instrumental voter's value from party B is:*

$$V_A(l_A, l_B) = \sum_{k=2}^{4} \tau(k)\ \tilde{H}(k, l_A, l_B),$$

while her value from party A is:

$$V_B(l_A, l_B) = \sum_{k=2}^{4} \tau(6-k)\ \tilde{H}(k-1, l_A, l_B).$$

Moreover, when θ is small enough, for each pair of lists (l_A, l_B), the difference $\Delta(l_A, l_B) = V_A(l_A, l_B) - V_B(l_A, l_B)$ is unique.

The key insight from Lemma 2 highlights how voters evaluate each party's list, not only as a function of the candidates and their order on the ballot, but also how they conjecture that other voters will behave. Each *attention weight*, $\tau(k)$, is equivalent to the probability that a voter is decisive for the k^{th}-ranked candidate of party A, *conditional on being decisive*.

Each voter allocates a fraction of his attention to each of the three *contested* ranks on each party's ballot: the weight $\tau(k)$ describes the fraction of his attention that is devoted to comparing the relative value from electing the k^{th}-ranked party A candidate, versus the $(6-k)^{th}$ party B candidate, i.e., raising party A's control of the assembly from $k = 1$ to k seats.

For example, $\tau(3)$ is the relative prospect that a voter expects to be decisive for whether party A wins the third seat. If she votes for A, she elects the *third*-ranked candidate on party A's list, deriving value $\tilde{H}(3; l_A, l_B)$; if she votes for B, she elects the *third*-ranked candidate on party B's list, deriving value $\tilde{H}(2; l_A, l_B)$. Note that these values take into account the payoff consequences from a change in the human capital of the assembly, as well as any change in the identity of the chief executive, say, due to a change in the identity of the majority party. We will subsequently highlight how our attention weights can be directly linked to our empirical work.

In some circumstances, the attention weights $\tau(2)$, $\tau(3)$ and $\tau(4)$ can be ordered for any possible pair of lists. In particular:

Lemma 3. *If party A is sufficiently advantaged, i.e., $\overline{\xi} > 0$ is large enough, then $\tau(4) > \tau(3) > \tau(2)$ for any pair l_A and l_B. Moreover, $\tau(2) < \min\{\tau(3), \tau(4)\}$.*

When assessing party A's list, voters place the smallest amount of attention on party A's second-ranked candidate, relative to its remaining candidates in the contested ranks. To see why, notice that a voter assesses that her relative prospect of being decisive for the A's highest-ranked candidate amongst the contested ranks is smaller than her relative prospect of being decisive for one of A's lower-ranked candidates.

Since $\tau(2)$ is the attention that a voter places on party A's *second* rank, it is also the attention that she places on party B's *fourth* rank. To see why, recall that if a voter is decisive for the election of party A's second-ranked candidate, a vote for B results in the election of B's fourth-ranked candidate. Thus, Lemma 3 states that voters place the smallest amount of attention on party B's fourth-ranked candidate, relative to its remaining candidates in the contested ranks. The reason is that party B is structurally disadvantaged and not likely to do well in the election; a voter assesses that her relative prospect of being decisive for B's lowest-ranked candidate amongst the contested ranks is smaller than her relative prospect of being decisive for one of B's higher-ranked candidates.

Notice that in the special case of $\theta = 0$, voters anticipate that they are equally likely to be decisive for each of the three contested ranks, i.e., the probability weights are $\tau(k) = \frac{1}{3}$ for each $k \in \{2, 3, 4\}$.

Party Nomination Strategies. Our two main results characterize equilibrium party lists. The results are obtained under alternative perspectives on (i) how voters care about different aspects of the aggregate quality of the assembly, i.e., the function H, and (ii) aspects of a party's elected members that are most effective for the pursuit of legislative goals, e.g., majority control or maximizing the talent of its elected representatives. These aspects are determined by the function G.

Marginal Rank Hypothesis. We start with a benchmark result that we call the Marginal Rank Hypothesis. It is derived under a stark view that what parties consider most important to pursuing legislative goals—that they *exclusively* care about—is maximizing their share of seats in the assembly.

Definition 1. Parties are *purely seat-motivated* if, for any seats $k \in \{1, ..., 5\}$ and for all pairs $(q, \overline{q}), (q', \overline{q}') \in Q \times [0, 1]$: $G(k, q, \overline{q}) - G(k, q', \overline{q}') = 0$.

We treat this objective as a benchmark: it states that the only way in which parties value electoral outcomes is through the number of seats that they are allotted in the council. However, the definition still leaves a great deal of flexibility for precisely how parties value additional seats. For example, the definition holds in contexts where parties care exclusively about achieving majority status, i.e., if $G(3, \cdot, \cdot) - G(2, \cdot, \cdot) = 1$. In classical models of electoral competition, for example, this represents a context in which parties compete under a majoritarian electoral rule and care solely about winning the election. The definition also applies if parties care solely about maximizing their vote share.

In order to explicitly connect our predictions to our subsequent empirical results, we refer to the second rank on a each party's list in our theoretical framework as its *advantaged* competitive rank.

Proposition 1. *(Marginal Rank Hypothesis) When parties are purely seat-motivated, for θ not too* *large, the competitive ranks have the highest-quality candidates; that is, for each party* $i \in \{A, B\}$*:*

$$\max\{l_A(1), l_A(5)\} < \min\{l_i(2), l_i(3), l_i(4)\}.$$

Moreover, within the contested ranks:

1. *A's best candidates are not located in its* advantaged *ranks:* $l_A(2) < \max\{l_A(3), l_A(4)\}$.
2. *B's best candidates are located in its* advantaged *ranks:* $l_B(4) > \min\{l_B(2), l_B(3)\}$.

Finally, when A's advantage, i.e., $\overline{\xi}$, is large enough:

$$l_A(2) < l_A(3) < l_A(4) \text{ and } l_B(2) > l_B(3) > l_B(4).$$

The proposition identifies properties of the party lists that maximize each party's appeal to voters, highlighting two distinct properties.

First, each party reserves its best candidates for contested ranks, to which voters assign positive attention weights. By putting the best candidate in ranks for which voters assign a positive prospect of being decisive, the party raises voters' relative value from supporting its ticket.

Second, within the contested ranks, party A maximizes its appeal by putting its very best candidate in either the third or fourth rank. The reason is that each voter places less attention on the party's advantaged rank, correctly anticipating that her vote is relatively more likely to affect party A's chances of winning a third or fourth seat. By putting its strongest candidate in one of these seats, the party makes the election of the candidate for which voters anticipate that they are relatively more likely to be decisive more valuable. Thus, a purely seat-oriented party A prefers not to place its best candidate in its advantaged competitive rank. For the same reason, party B does not place its best candidate in its fourth rank.

The final part of the proposition provides even stronger results, for a setting in which party A is very likely to win a large share of seats. Anticipating a near-certain prospect that the party will win its advantaged rank, and even very likely to win at least three seats, voters expect that their choice of ballots is relatively likely to determine whether A also wins a fourth seat, or instead B wins a second seat. Proposition 1 implies that the parties put their strongest candidates in these ranks in order to maximize their lists' appeal to voters.

Top-Down Rank Order Hypothesis. An alternative view of parties and legislative institutions, however, places much greater emphasis on the role of individual politicians—in particular, legislative *leaders*—rather than seats, alone, as the route to legislative accomplishment. We now consider this perspective, and ask how it affects the organization of party lists.

Recall that a party's payoff is $G(k, m, \overline{q}_J)$, and which depends on k, the number of seats it wins, m, the quality of the chief executive, and $\overline{q}_J$, its share of the legislative human capital. Recall our assumption that G weakly increases in k and in $\overline{q}_J$. We derived the marginal rank hypothesis under an especially stark form of this assumption: purely seat-motivated parties care only about their share of seats (k).

We consider an alternative view of how parties achieve their legislative goals—not merely by seats alone, but also by virtue of having skilled and able legislators, by assuming that G strictly increases in *both* $\overline{q}$ *and* k. This implies that—all else equal—a party is strictly more likely to accomplish its legislative goals when its share of the total human capital in the assembly increases. Consider the list l^* satisfying

$$l^*_J(1) = 5,\ l^*_J(2) = 4,\ l^*_J(3) = 3,\ l^*_J(4) = 2,\ l^*_J(5) = 1.$$

This list maximizes a party's average share of the assembly's human capital, for *every* realized allocation of seats between the parties.

We formalize the perspective that legislative leaders matter via two conditions: our first condition focuses on how *voters* view the determinants of policy outcomes.

Condition 1. *There exists κ small enough that $H_1(\cdot, 5) < \kappa$, where H_1 denotes the derivative of H with respect to its first argument.*

Suppose that the assembly leadership—the chief executive—is the highest quality candidate (i.e., with quality 5) from either of the two parties. Condition 1 states that—given the identity of the assembly's leadership—the incremental value of more able rank-and-file politicians is not too large. Alternatively stated: voters recognize that high quality leaders are the primary drivers of good policy outcomes. This condition is very likely to hold in parliamentary contexts, where agenda-setting powers typically reside with the executive, from which the vast majority of legislative initiatives originate.

While Condition 1 in and of itself generates a testable restriction on the data, we consider a related condition that describes how—from the perspective of parties—the skills of a legislative team interact with a party's agenda-setting authority to facilitate the achievement of legislative goals. We do not require the next condition for our main result, but it generates an even stronger empirical prediction that we take to the data. The condition states that variation in the quality of a party's legislative team matters relatively more when the party holds a majority of seats.

Condition 2. *For any $k \geq 3 > k'$, and any list $l' \neq l^*$:*

$$\tilde{G}(k, l^*, l^*) - \tilde{G}(k, l', l^*) > \tilde{G}(k', l^*, l^*) - \tilde{G}(k', l', l^*).$$

This reflects legislative contexts in which majority parties are most empowered to achieve their policy goals, at which point the skills and quality of their legislative team becomes relatively more important. Conversely, a party that holds a minority of seats is unable to play nearly as a much of a role in legislative outcomes; variation in the quality of the minority team may be important, but it is second-order to contexts in which that team held control of key committees and the legislative timetable. We emphasize that this condition is not needed for our second main result, but it provides a valuable 'comparative static' that generates additional empirical implications, and thus more exacting tests on our data.

The expected payoff to party J when its list is l_J and the opposing party $-J$ offers a list l_{-J} is:

$$\Pi_J(l_j, l_{-j}) = \sum_{k=1}^{4} \Pr(\pi(k) < \pi_J < \pi(k+1) | l_J, l_{-J}) \tilde{G}(k; l_J, l_{-J}). \tag{3}$$

Our second result establishes that when Condition 1 holds, the list l^* is a best response to *every possible list* that could be selected from the opposing party. If, in addition, Condition 2 holds, the incentive to promote higher quality candidates to higher ranks is always stronger for party A, the advantaged party. In other words: the party with the greatest relative value from promoting better candidates is that party that is most likely to win majority control of the assembly.

Proposition 2. *(Top-Down Rank Order Hypothesis) If Condition 1 holds, the list l^* is the unique best response to any opposing party's list. If Condition 2 also holds, then for any list $l \neq l^*$:*

$$\Pi_A(l^*, l^*) - \Pi_A(l', l^*) > \Pi_B(l^*, l^*) - \Pi_B(l', l^*).$$

Conditions 1 and 2 have a common theme: they emphasize that policy outcomes and goals are not accomplished by seats, alone, but also by the skill and experience of parties that hold majority control, i.e., legislative leaders.

Propositions 1 and 2 generate empirically testable predictions, which we proceed to test using Swedish register data on the competence characteristics of the complete universe of Swedish politicians from 1982-2014.

4. Empirical Context

Sweden has three levels of political representation: the parliament, 21 counties and 290 municipalities. All three levels are concurrently elected, and voter turnout is usually between 80 and 90 percent. Local elections are politically and economically important: municipalities have significant political autonomy and control large shares of expenditures and employment sectors in the Swedish welfare state. Notably, their public expenditures constitute between 15 and 20 percent of GDP, and they employ roughly 20 percent of the country's labor force. Election to a municipal council also serves as the principal route into national politics (Lundqvist, 2013; Dal Bó, Finan, Folke, Persson, and Rickne, 2017).

The municipal council varies in size between 31 and 101 members with an average of 46. Representation is not subject to an explicit electoral threshold. A municipality is effectively a parliamentary democracy, in which elections are solely to the political assembly, and in which all executive appointments are made after the election. Nearly all executive positions are appointed by the governing majority, i.e. the political coalition or single political party that receives more than 50% of the council seats.

The chief executive is the municipal council board chair—for simplicity we henceforth refer to the office as the *mayor*. The mayor is always appointed by the largest party in the ruling coalition. It is the most powerful position in municipal politics and with the exception to large municipalities, the only political position that offers a full-time salary, as opposed to piece-rate compensation for meetings (Montin, 2015).

The Swedish party system is very stable, and most political parties have local branches in all municipalities. These local parties enjoy significant autonomy from their national party organizations when determining who should be on their ballot paper and at which rank. The only rules are that candidates must be above 18 years old and be residents of the municipality in which they stand for election. Some parties also employ gender quotas that are dictated by the central party organizations. Purely local parties, i.e. without representation in the national parliament, account for less than four percent of the party-election observations in our data.

The precise procedure for selecting and ranking candidates varies across parties. Parties on the ideological left mainly use internal nominations.[7] Nominations for the ballot are made by the party's clubs and are then aggregated by an election committee. Clubs are organized around neighborhoods in the municipality and around factions such as the *Women's League* or the *Youth League*. The election committee consists of senior politicians—either for-

[7] Authors' interviews.

mer or current—who can make adjustments to the final list before it receives final approval at a member meeting.

Parties on the ideological center-right are more likely to use procedure of internal member primaries to determine the rank-order of their ballot.[8] Members vote for their preferred candidates, and votes are accumulated to form the rank-order on the ballot. Notably, internal primaries are similar to internal nominations by having a large element of top-down coordination. Center-right parties are also organized in party branches and geographic clubs, and members loosely belong to factions that coordinate votes around their political interests.

5. Data

Our data come from Swedens administrative register, which covers every candidate on every electoral ballot in every election from 1982 to 2014. Variables in the dataset come from various government agencies such as the tax agency, the electoral agency, and the military defense ministry. We observe every candidates rank on the ballot, and whether he or she was elected. These variables are perfectly measured and have no missing data. Although data is linked via the personal ID code of each individual in the dataset, this code is scrambled by Statistics Sweden and researchers work with the anonymized data.

In what follows, we describe how we create our measurement of ballot rank contestability, which then forms the basis for our sample restrictions. We also describe our four measurements of politician valence.

6. Measurement

To test our theoretical predictions, we need two key types of measurements: measurements of the contestability of different list ranks *and* measurements of politician valence. We first develop a novel methodology to compute rank contestability, and then discuss our construction of valence measures.

Computing contestability of ballot ranks. In closed and flexible party-list PR systems, a candidate's prospect of election increases with her list rank. We propose a data-driven and flexible approach to calculate the probability that a candidate located at each ballot rank is elected. Our methodology reflects the idea that parties—and, possibly, voters—can use information about the electoral history of the party to predict its performance in the next election.

[8] Authors' interviews.

For example, a party that has consistently won ten municipal seats is likely to win approximately ten seats, in the future. That party's fifth-ranked candidate occupies a relatively safe seat, while its twentieth-ranked candidate is almost sure to lose; neither seat is relatively likely to be contested. By contrast, the tenth and eleventh seats are more likely to hang in the balance. Using a simple regression framework, we predict the probability that those contested ranks, and each of the other ranks, are converted into a seat.

We begin by constructing a dataset for the number of ranks won by each local party in each election. We then create a dummy variable, D^r, for each ballot rank $R \in \{0, ..., 50\}$. This dummy takes the value 1 if the party p won this rank in the election t, and zero otherwise. For each of these dummies, we run a regression where that dummy is regressed on all fifty dummies in the most recent election, $t-1$. For the case of the fifth rank, for example, the equation becomes:

$$D^5_{pt} = \sum_{r=1}^{50} \delta^r D^r_{p,t-1} + X'_{p,t-1}\gamma + \varepsilon_{p,t}. \tag{4}$$

The fifty regressions can then be used to give a predicted value for each rank, and for each party, and at each election. The predicted probabilities of winning five seats is thus:

$$\hat{D}^5_{pt} = \sum_{r=1}^{50} \hat{\delta}^r D^r_{p,t-1} + X'_{p,t-1}\hat{\gamma}. \tag{5}$$

Repeating this process for each of the fifty regressions gives us a new dataset of predicted probabilities for each list rank in each local party, and in each election. Note that we can choose which variables are included in equations Equation 4 and Equation 5—for example, we can extend the dummies for the party's electoral history to further back in time. In the Swedish context, anecdotal evidence suggests that parties set the baseline for the expected number of seats at the number of seats in the previous election.[9]

Our setup incorporates additional predictors (X') that a party might use when calculating its expected number of seats. Our estimation accounts for the large variation in municipality size, and its consequence for the variability of seat shares: in a large municipality, a smaller percentage point vote shift is required for an additional seat than in a small municipality. We therefore interact council size with each rank-dummy and include these in X'. Additional and potential variables could include indicators for party size, for example approximated by seat shares. Equation 4 is estimated by OLS for clarity of interpretation, but recognizing that some predicted probabilities may fall outside of the unit range.

[9] Authors' interviews.

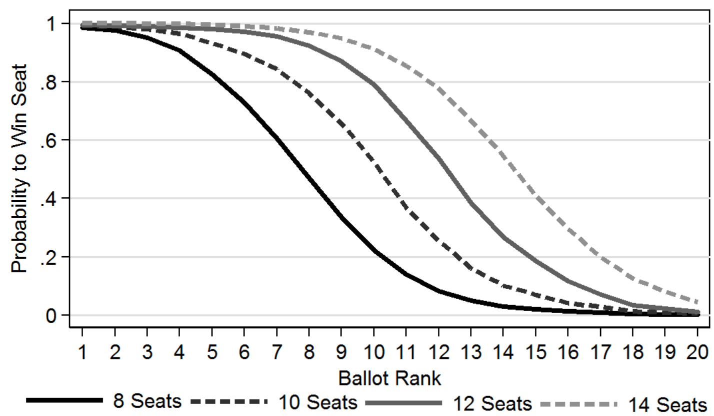

Figure 1 – Average predicted win probabilities for ballot ranks. The figure shows average predicted win probabilities (Y-axis) for each ballot rank (X-axis) and for political parties with either 8, 10, 12 or 14 seats. Probabilities are calculated by estimating equations (4) and (5) at the level of the party, election, and municipality, and using data for all parties, in all municipalities and all years between 1990 and 2014.

To illustrate our approach, Figure 1 plots the averages of the predicted probabilities that a party wins each list rank, given that it won $k \in \{8, 10, 12, 14\}$ seats in the previous election: each line corresponds to each of the four hypothetical seat shares (i.e., each value of k) in the previous election. As expected, the smaller parties' prospects fall away at relatively higher ballot ranks. For each pre-election party size (k), the predicted probabilities show the expected decline over ballot ranks.

We use the predicted win probabilities to categorize ballot ranks into six categories of electoral safety. The intention is, of course, to then compare the average valence of candidates across ranks. The point of departure for the categorization is the contested ranks, which we define as ranks with a win margin in the interval of 2% to 98%, illustrated in Figure 2. We further subdivide the contested ranks into "*Advantaged Contested*" ranks, "*Highly Contested*" ranks, and "*Disadvantaged Contested*" ranks. Our theoretical predictions about the quality of a candidate in each of the three contested ranks in our theoretical model can be interpreted as predictions about the *average* quality of candidates in each of the three categories of contested ranks, in the data.

We define the top name on the list as a separate category, the "*Capolista*", and the interim ranks between the the capolista and the contested ranks to be "*Safe*". Finally, we denote by "*Certain Loss*" the candidates with a predicted probability of election below two percent. Figure 2 clarifies the six categories on a hypothetical ballot, in which the party won thirteen seats in the previous legislative cycle—the median number of previously elected politicians across all the party lists in our data.

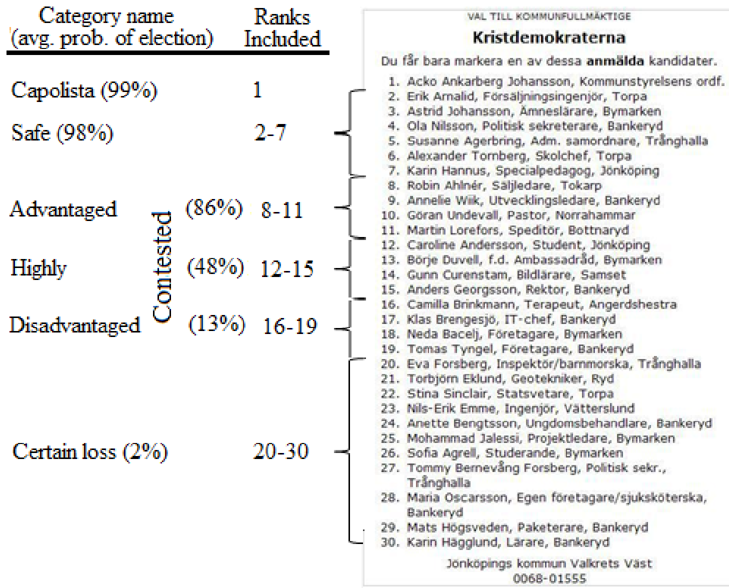

Figure 2 – **Ballot rank categorization of an example ballot**. The figure shows the six categories of ballot ranks for an example ballot of a party with 13 incumbent politicians. Category names are listed in the leftmost column together with the average predicted win probability in each category. Win probabilities are computed from OLS estimation of equations (4) and (5).

In our data there are 372,072 candidate-election observations and 15,819 municipal party-election observations. After restricting the data to our estimation sample, we maintain 217,734 candidate-election observations and 6,561 municipal party-election observations. The municipal parties excluded are typically small: their median size is between 1 and 2 seats, with 14 percent of the excluded local parties winning no seats, and 90 percent winning 5 seats or less. The parties in our estimation sample are much larger, with a median number of 8 seats, and 75 percent having four seats or more. The estimation sample includes 64,085 of the 92,530 councilors elected during the time period that we study.

Measurements of politician valence. Since there is no unified approach to quantifying the skills and experience—"valence"—that is at the heart of our theoretical framework, we use a broad approach that includes the most common and non-controversial measurements. The first two come from data collected in Sweden's military enlistment—in particular, a mandatory draft that existed for men born between 1955-1979. The enlistment process included a test of cognitive abilities, as well as an interview to evaluate recruits' capacity for leadership, conducted by a trained psychologist. Both skills, cognitive and leadership, are normalized by the enlistment agency on a 1-9 stanine scale in the cohort, and are linked to our dataset via the personal ID code. For a detailed description of the enlistment variables and their strong correlations with private-sector careers, see Lindqvist and Vestman (2011) and Dal Bó et al. (2017).

A third, income-based valence measure was first suggested by Keane and Merlo (2010) and more thoroughly developed by Besley et al. (2017). It captures deviations in labor income among politicians with highly similar education, age and occupation. The idea is that an ideal measure of political competence should capture key governing abilities, but not be correlated with socioeconomic status. Politicians' earnings are therefore benchmarked against the totality of the Swedish working-age population in the same education-occupation-cohort and municipality cell. If politicians are paid a wage in their political roles, we use their earnings in the positions they held prior to becoming a full-time politician.[10]

We use education as a fourth and final valence variable. Recent related scholarly work has relied heavily on measures of candidate education, on the basis that it instills civic values and builds skills that improve political performance—see, for example, De Paola and Scoppa (2011); Besley and Reynal-Querol (2011); Galasso and Nannicini (2011); Ferraz and Finan (2009); Schwindt-Bayer (2011); Franceschet and Piscopo (2012) and Baltrunaite et al. (2014). We acknowledge, nonetheless, that education is a problematic measure of valence because of its strong correlation with family background and social background.[11]

All four valence variables are standardized in pooled data for the 1991-2014 elections, and all local parties that meet our sample restrictions. The unit of measurement is in standard deviations from the variable mean. In an Appendix, we report raw means and standard deviations of the variables and preview their averages across our six categories of list ranks.

7. Empirical Results

Regression specification and baseline results. We first run a simple regression to quantify the average valence of politicians across the six categories of ballot ranks. Each of the four valence variables is regressed on dummy variables for the rank categories, excluding the below

[10] This measure is based on the idea that the private and political spheres are complements rather than substitutes. Besley et al. (2017) show that it is correlated with various measures of politician career success, voter support measured in preference votes, and that the proportion of municipal councilors who are defined as above-average in competence is correlated with better government outcomes.

[11] Some other candidate measurements of politician valence were not included in the analysis. We did not use occupation background, for example in law, business or medicine, since it risk an even stronger correlation with family background and social class than education (Carnes 2013). Political experience could capture learned skills, for example bill drafting, bargaining and building coalitions, that is valuable in a political organization. We also exclude political experience. Although experience could capture learned skills, such as bill drafting, bargaining and building coalitions, it has also been argued to measure insider status (Galasso and Nannicini, 2015). Finally, we exclude local ties, being born and/or raised inside or outside the electoral district (e.g. Tavits, 2010, Shugart, Valdini, and Suominen, 2005), since this measure would lack variation in our data on all-local politics.

marginal ranks which serve as the reference.

$$Y_i = \alpha_{p,t} + \beta_1 Capo_i + \beta_2 Safe_i + \beta_3 Adv\ cont_i + \beta_4 Highly\ cont_i + \beta_5 Certain\ loss_i + \varepsilon_i. \tag{6}$$

Equation 6 also contains one dummy variable for each local party and election period (a total of 3,950 dummies). Our estimates hence capture average differences across rank categories within local parties, whilst isolating the comparison from variation across parties, municipalities, or even over time within the same local party.

The estimates on the dummy variables β_1 through β_5 capture the difference in average valence between politicians in each category and the reference category of *Disadvantaged Contested* ranks below the highly contested ranks. The estimates from the four regressions are plotted in Figure 3. Across *all* party lists, i.e., in both closed and flexible list contexts, averaging across all parties and municipalities, average politician valence consistently increases toward the top of the ballot.

Our theoretical results generate predictions about the relationship between a candidate's quality and his or her rank on the party's ballot, within the contested ranks. We proceed to evaluate these predictions.

Evaluating the Marginal Rank Hypothesis. Proposition 1 makes two distinct predictions. First, all parties should locate their best candidates in the contested ranks. Second, if a party is expected to win a large share of seats, it should not place its best candidates in its *advantaged* contested ranks; and, a party that expects to win a small share of seats should not place its best candidates in its *disadvantaged* ranks.

Figure 3 highlights that across all party lists, the average politician valence in each of the non-contested ranks at the top of the ballot strictly exceeds the average politician valence in each of the competitive ranks. This contradicts the first prediction of the marginal rank hypothesis. The figure also highlights that the average politician valence in the advantaged competitive ranks strictly exceeds the corresponding average politician valence in each of the highly competitive and disadvantaged competitive ranks. This does not contradict Proposition 1, to the extent that the full sample contains parties that are expected to win a small share of seats. However, the proposition states that if we were to focus exclusively on parties that are expected to win a large share of seats, we should observe the opposite pattern: average valence within the contested ranks should maximized in either the highly competitive ranks or the disadvantaged competitive ranks.

To examine this prediction, directly, Figure 4 divides party lists into two categories: those

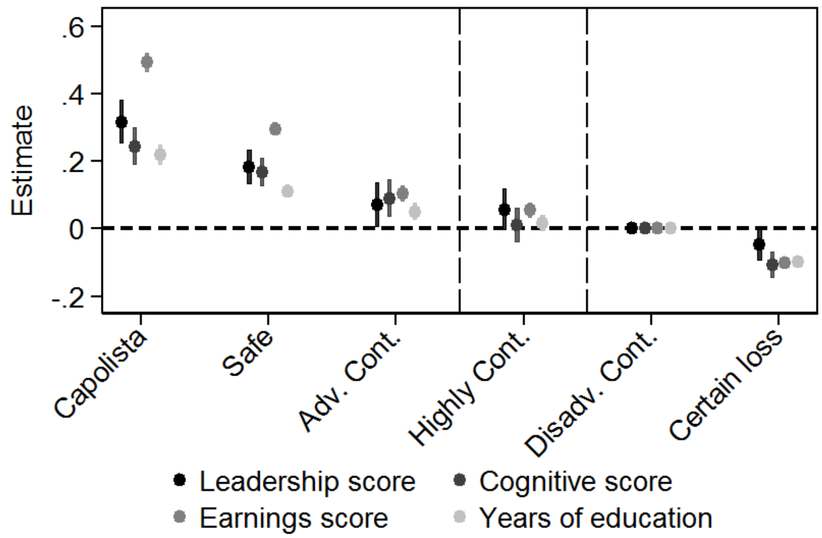

Figure 3 – **Average valence levels across ballot rank categories**. The figure shows estimated coefficients from Equation (6), where the average valence difference between each category of ballot ranks is estimated relative to the reference category of the disadvantaged competitive ranks. The outcome variables are the four valence measurements, transformed into Z-scores. Vertical lines show 95% confidence intervals. All regressions include fixed effects for every local party in every election period (3,950 dummies). The data is all candidates in local parties with at least one candidate in each of the six ballot rank categories. The number of observations in the regression analysis is: Leadership Score: $N = 35,662$; Cognitive Score: $N = 48,088$, Earnings Score: $N = 209,995$, Years of Education: $N = 185,382$.

which, in their respective districts, enjoy a seat share above and below 17% at the time of the election—the median seat share in our sample. Across each of the rank categories and valence measures, the pattern is the same: the average valence amongst the advantaged contested ranks strictly exceeds the average valence in *each* of the remaining contested ranks. In fact, the correlation between rank and valence is *even stronger* in relatively larger seat share parties than it is in parties with a relatively smaller seat share—decisively rejecting the predictions of the marginal rank hypothesis.

Alternative Test. Our tests compare *average* candidate quality across rank categories—the natural empirical referent for our theoretical predictions. Nonetheless, these averages may obscure variation in candidate quality *within* each of the rank categories. For example, while the average quality of candidates in the advantaged ranks strictly exceeds the corresponding average in each of the remaining contested ranks, there may be instances in which individual candidates that are located in the advantaged ranks are nonetheless lower quality than individual candidates in the remaining contested ranks.

We consider an alternative test in which we report for each rank k on a given list the share of candidates in ranks $k + 1$ through n whose valence is less than or equal to the valence of

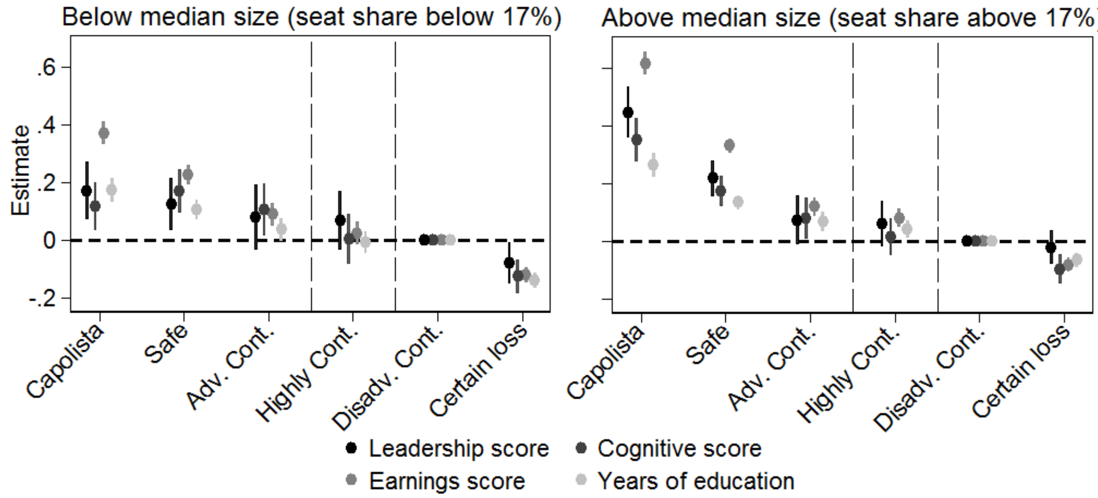

Figure 4 – **Average valence levels across ballot rank categories in *small* and *large* political parties**. The figure replicates the analysis in Figure 3 in two sub-samples of data, split by median party size.

candidate at rank k. Under the marginal rank hypothesis, this share should be *smallest* for candidates in non-contested ranks at the top of the ballot for *all* parties. And within their contested ranks, parties that anticipate a large share of seats should see the smallest share for candidates in the advantaged contested ranks. In the Appendix, we show that both predictions are contradicted under this alternative test

Evaluating the *Top-Down Rank Order* Hypothesis. Our second theoretical result also yields two distinct predictions. First, if Condition 1 is satisfied, i.e., voters care mainly about the quality of a council's leaders, rather than its rank-and-file, then *all* party lists should place higher quality candidates in higher ranks. Second, if Condition 2 is satisfied, i.e., that the quality of a party's legislative team is most important when the party controls the council, then the relative value from choosing this list versus any alternative list is greatest for parties that enjoy a large prospect of holding majority status.

Figure 3 highlights that, indeed, across all party lists and quality measures, average candidate quality increases with the rank category, consistent with the first prediction. To examine the second prediction, we ask: in which contexts is Condition 2 most likely to hold? We argue that it is most likely to hold for parties that have historically appointed the mayor in *low competition* districts, followed by opposition parties in *high competition* districts, followed by opposition parties in *low competition* districts. To see why, notice that low competition districts confer the greatest prospect of majority status to parties that have historically held power, and the weakest prospect to parties that have languished in opposition. And, competitive districts offer the leading parties a lower prospect of majority control, but conversely offer opposition

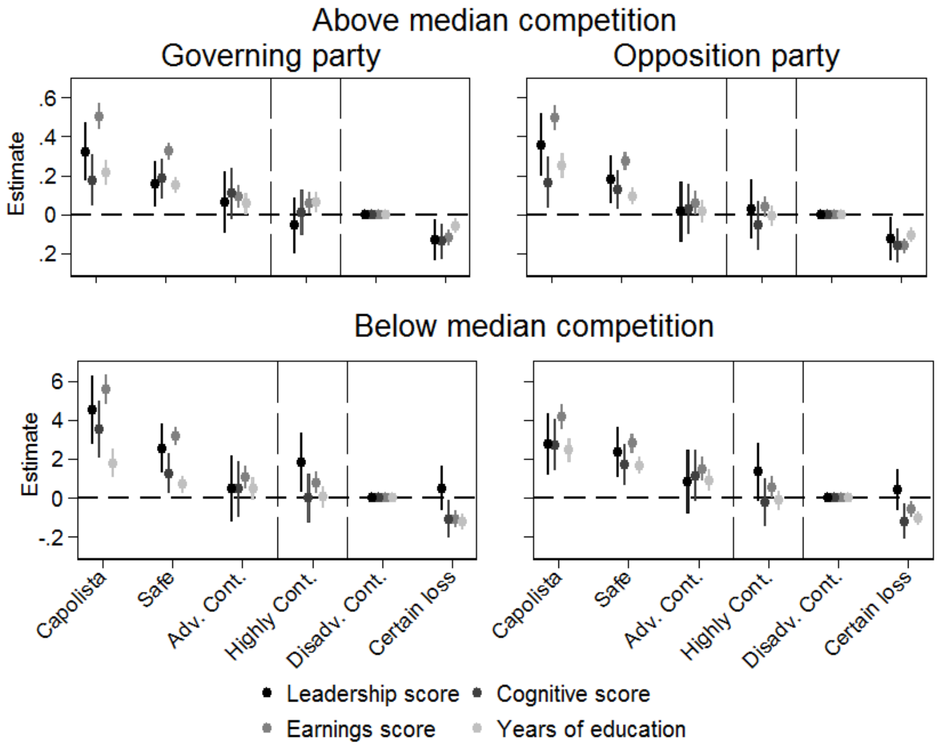

Figure 5 – **Average valence levels across ballot rank categories in high and low political competition, and in governing and opposition parties.** The figure replicates the analysis in Figure 3 in four sub-samples of data. The top two plots show the estimates in municipalities with above-median political competition, and the bottom two plots for below-median political competition. The two plots on the left include political parties in the governing majority, and the two plots to the right include opposition parties.

parties a stronger prospect of power than non-competitive contexts.

To test this prediction, we split our sample into two categories: *above median competition* and *below median competition* in the bottom panels. Here, we define the extent of competition in a municipal as the average vote share difference between the two party blocs over the previous three elections: a small vote share difference reflects higher competition between the blocs, since the distribution of votes is more even across them. We also focus on two classes of political parties: those that appointed the mayor in the previous electoral cycle, and the largest opposition party from the previous electoral cycle, i.e., the largest party that was not part of the governing coalition. Figure 5 shows estimates from Equation 6 using this subset.

Our framework predicts that we should observe the strongest decreasing quality-rank relationship amongst majority parties in relatively low competition settings, i.e., the bottom left-hand panel in Figure 5. As we move from that baseline to both mayoral and opposition parties in relatively high competition settings, we observe a similarly strong relationship. However—consistent with our framework—there is a significant decrease in the rank-quality relationship when we isolate the largest opposition party in relatively low-competition settings. This is consistent with the idea that in low competition contexts, parties that have appointed the mayor in the past have a high degree of confidence that they will continue to wield executive authority, while opposition parties in these contexts anticipate that they are very unlikely to wield executive authority.

Discussion. Our theoretical results incorporate assumptions about both parties' goals and voters' behavior. Thus, our rejection of the marginal rank hypothesis is a rejection of a joint hypothesis about parties' objectives (that they solely care about seats) *and* voter behavior—i.e., that voters formulate their induced preferences over party lists according to Lemma 2. To what extent is this view of voter behavior reasonable? Perhaps voters have limited attention and knowledge about political candidates on the ballot, and simply focus deterministically only on candidates that are placed right at the top of the ballot. In our framework, this is equivalent to replacing the equilibrium attention weights $\tau(k)$ for each rank k in Lemma 2 with exogenous weights satisfying $\tau(1) > \tau(2) \geq ...\tau(5)$, with, $\tau(k) = 0$ for some $2 \leq k \leq 4$.

A Swedish electoral reform allows us to examine this alternative hypothesis about voting behavior. Preference voting (i.e., flexible list PR) was introduced in all local elections in 1998, replacing the previous closed system, and thereby allowing voters to cast a single preference vote for any politician on a party's list. Our data includes the tally of preference votes for every politician in every rank in each of the five local elections that were held between 1998 and 2014, i.e., after the reform.

In the top panel of Figure 7 we plot the distribution of local parties' preference votes across list ranks. In the bottom left panel we plot estimates of Equation 6, with the proportion of preference votes as the outcome variable; finally, the bottom right panel highlights pairwise correlations between preference votes and valence among politicians in each of our six ballot rank categories. The regression that captures the pairwise correlations includes fixed effects for the interaction between local party, election, and municipality. This restricts our variation to voter behavior within a given list in a given election, meaning also that we cannot get an estimate for top-ranked politicians (as there is only a single such candidate in a list). Both the fixed effects and data limitation to a subset of male cohorts generate substantial noise in our military enlistment variables, and we omit these.

In the top and bottom left panels, we observe that the vast majority of preference votes are awarded to the top portion of party lists. The top-ranked politician alone receives about a third of the total votes in the average party list. The top three politicians, together, account for approximately one half of all votes. On the one hand, this is consistent with the idea that

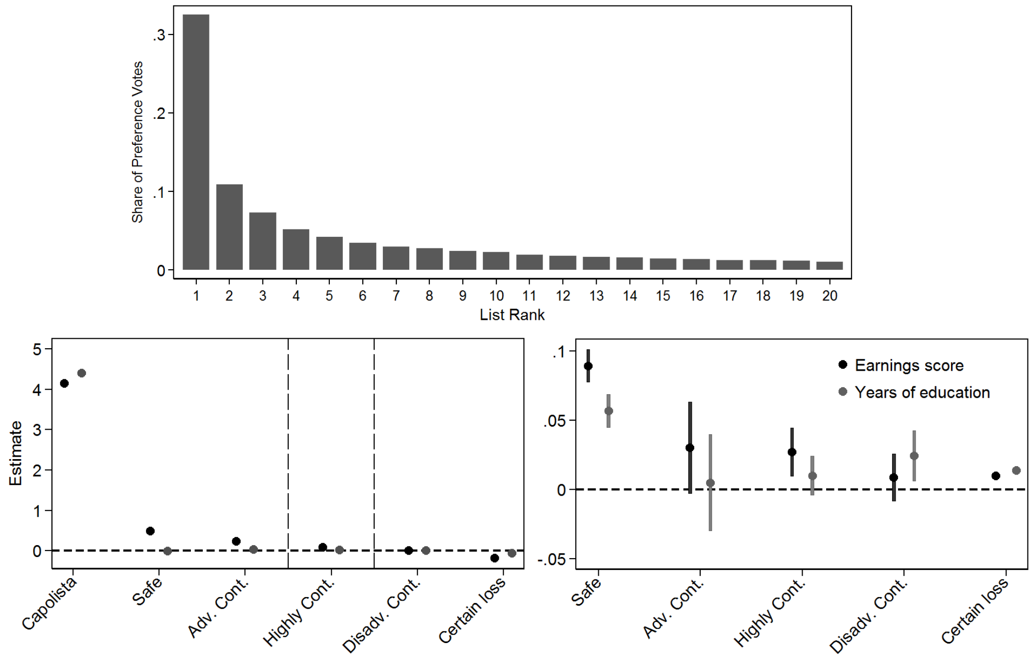

Figure 6 – Preference votes by list rank (top), preference votes relative to disadvantaged contested seats (bottom left); and pairwise correlations between valence and preference votes in each ballot rank category (bottom right). All figures use data for parties with at least one politician in each ballot rank category and data for all elections between 1998 and 2014. The upper figure reports the average share of preferences vote for each ballot rank, 1 to 20. The lower left-hand side figure reports estimated coefficients with 95% confidence intervals from equation (3) where the disadvantaged contested ranks are the reference category, and the share of preference votes is the outcome variable. The regression includes fixed effects for every local party in every election period, and is estimated both with and without fixed effects for ballot rank. The lower right-hand side figure shows estimates from an OLS regression for the relationship between valence and preference votes; and using the two valence measures available for the full population (earnings score and education length). A separate regression is estimated for each category of ballot ranks and each valence measure, all including fixed effects for every local party in every election period.

the highest ranks are focal for voters.

On the other hand, the bottom right panel highlights that the correlation between candidate quality and preference does not vanish as one moves out of the safe seats, i.e., the highest ranks. Preference votes are still responsive to quality within the contested ranks. So, while voters generally cast their ballots for higher-ranked candidates *and* their preference votes are also more sensitive to candidate quality within higher ranks, their votes are also responsive to variation in candidate quality outside of these ranks. This casts doubt on the empirical validity of a simple heuristic in which voters solely focus on the very highest ranks.

8. Conclusion

We analyze party list assignments both in theory and in real-world contexts. We develop a theoretical framework to examine party list assignments in multi-member contexts, under alternative specifications of how voters value the skills and abilities of their politicians, and also how party goals are facilitated by the conjunction of legislative seats and the skills of individual politicians. We test two predictions generated by the framework, leveraging a dataset containing the universe of Swedish politicians on every ballot in 290 municipal elections over seven election periods.

We derive our first prediction in an environment where parties care only about their share of seats, and the second in an environment where voters and parties both care predominantly about the skills of legislative leaders. We decisively reject the first environment as an accurate representation of the context in which parties choose their electoral strategies, but find strong support for the second.

Our framework is intended to facilitate the development of theoretical hypotheses that are applicable to many different list PR contexts—both within and across countries. Nonetheless, we recognize that this requires us to abstract from subtle and important variation in the way these electoral rules operate in real-world contexts, which is explored in existing and ongoing scholarly efforts that focus on particular country-contexts, such as Cheibub and Sin (2018). We hope that our attempt to develop a broad theoretical framework will play a complementary role in shedding further theoretical and empirical light on candidate selection in this important class of electoral rules.

References

BALTRUNAITE, A., P. BELLO, A. CASARICO, AND P. PROFETA (2014): “Gender quotas and the quality of politicians,” *Journal of Public Economics*, 118, 62–74.

BESLEY, T., O. FOLKE, T. PERSSON, AND J. RICKNE (2017): “Gender quotas and the crisis of the mediocre man: Theory and evidence from Sweden,” *American Economic Review*, 107, 2204–2242.

BESLEY, T. AND M. REYNAL-QUEROL (2011): “Do democracies select more educated leaders?” *American Political Science Review*, 105, 552–566.

BUISSERET, P. AND C. PRATO (2018): “Competing Principals? Legislative Representation in List PR Electoral Systems,” Unpublished.

CAILLAUD, B. AND J. TIROLE (2002): "Parties as Political Intermediaries," *Quarterly Journal of Economics*, 1453–1489.

CHEIBUB, J. AND G. SIN (2018): "Intraparty Competition and Open-List PR Elections: What Can We Learn from an Extremely Permissive System?" *working paper*.

CRUTZEN, B., S. FLAMAND, AND N. SAHUGUET (2018): "On the Optimality of Closed Lists under Proportional Representation," Working paper.

CRUTZEN, B. S., M. CASTANHEIRA, AND N. SAHUGUET (2010): "Party organization and electoral competition," *Journal of Law, Economics, and Organization*, 26, 212–242.

CRUZ, C., P. KEEFER, AND C. SCARTASCINI (2016): "The database of political institutions 2015 (DPI2015)," .

DAL BÓ, E., F. FINAN, O. FOLKE, T. PERSSON, AND J. RICKNE (2017): "Who Becomes a Politician?" *The Quarterly Journal of Economics*, 132, 1877–1914.

DE PAOLA, M. AND V. SCOPPA (2011): "Political competition and politician quality: evidence from Italian municipalities," *Public Choice*, 148, 547–559.

DEWAN, T. AND R. HORTALA-VALLVE (2013): "Teams of rivals: Learning about a cabinet and its shadow," Tech. rep., working paper.

DUVERGER, M. (1959): *Political parties: Their organization and activity in the modern state*, Methuen.

EGUIA, J. (2011): "Endogenous Parties in an Assembly," *American Journal of Political Science*, 55, 16–26.

ESTEVE-VOLART, B. AND M. BAGUES (2012): "Are women pawns in the political game? Evidence from elections to the Spanish Senate," *Journal of Public Economics*, 96, 387–399.

FERRAZ, C. AND F. FINAN (2009): "Motivating politicians: The impacts of monetary incentives on quality and performance," National Bureau of Economic Research.

FOLKE, O., T. PERSSON, AND J. RICKNE (2016): "The primary effect: Preference votes and political promotions," *American Political Science Review*, 110, 559–578.

FRANCESCHET, S. AND J. M. PISCOPO (2012): "Gender and political backgrounds in Argentina," in *The Impact of Gender Quotas*, Oxford University Press New York, 43–56.

GALASSO, V. AND T. NANNICINI (2011): "Competing on good politicians," *American Political Science Review*, 105, 79–99.

——— (2015): "So closed: Political selection in proportional systems," *European Journal of Political Economy*, 40, 260–273.

HIRANO, S. AND J. SNYDER (2014): "Primary elections and the quality of elected officials," *Quarterly Journal of Political Science*, 9, 473–500.

KARLSSON, D. AND M. GILLJAM (2014): *Svenska politiker: om de folkvalda i riksdag, landsting och kommun*, Santérus.

KEANE, M. P. AND A. MERLO (2010): "Money, political ambition, and the career decisions of politicians," *American Economic Journal: Microeconomics*, 2, 186–215.

KRASA, S. AND M. POLBORN (2018): "Political Competition in Legislative Elections," *American Political Science Review*, forthcoming.

LINDQVIST, E. AND R. VESTMAN (2011): "The labor market returns to cognitive and noncognitive ability: Evidence from the Swedish enlistment," *American Economic Journal: Applied Economics*, 3, 101–128.

LUNDQVIST, H. (2013): "Is it worth it? On the returns to holding political office," .

MATAKOS, K., R. SAVOLAINEN, O. TROUMPOUNIS, J. TUKIAINEN, AND D. XEFTERIS (2018): "Electoral Institutions and Intraparty Cohesion," Working paper.

MERILÄINEN, J. AND J. TUKIAINEN (2018): "Rank effects in political promotions," *Public Choice*, 1–23.

MONTIN, S. (2015): "Municipalities, Regions, and County Councils," *The Oxford Handbook of Swedish Politics*, 367.

SCHWINDT-BAYER, L. A. (2011): "Women who win: Social backgrounds, paths to power, and political ambition in Latin American legislatures," *Politics & Gender*, 7, 1–33.

SHUGART, M. S., M. E. VALDINI, AND K. SUOMINEN (2005): "Looking for locals: voter information demands and personal vote-earning attributes of legislators under proportional representation," *American Journal of Political Science*, 49, 437–449.

TAVITS, M. (2010): "Effect of local ties on electoral success and parliamentary behaviour: The case of Estonia," *Party Politics*, 16, 215–235.

9. Appendix: Descriptive Statistics

	Leadership Score	Cognitive Score	Earnings Score	Years of Education
Top Ranked	6,22 (1,73)	6,20 (1,67)	0,74 (0,95)	13,6 (3,0)
Safe	5,93 (1,74)	5,90 (1,67)	0,64 (0,88)	12,9 (3,0)
Advantaged Contested	5,76 (1,78)	5,85 (1,80)	0,40 (0,89)	13,0 (2,9)
Highly Contested	5,76 (1,79)	5,69 (1,82)	0,37 (0,91)	12,9 (2,9)
Disadvantaged Contested	5,72 (1,76)	5,69 (1,82)	0,30 (0,92)	13,0 (2,9)
Certain Loss	5,62 (1,75)	5,55 (1,85)	0,23 (0,94)	12,7 (3,0)
All Categoriess	5.72 (1,76)	5.67 (1,76)	0,34 (0,94)	12,8 (3,0)

Figure 7 – **Descriptive statistics for valence measures across ballot rank categories**. The table shows means and standard deviations for four valence measures within six categories of ballot ranks, and for all categories combined.

10. Appendix: Proofs of Propositions

For ease of exposition, we include the statements of each result.

Lemma 2 *For any party lists l_A and l_B, there exists a pair $V_A(l_A, l_B)$ and $V_B(l_A, l_B)$ such that an instrumental voter's value from party B is:*

$$V_A(l_A, l_B) = \sum_{k=2}^{4} \tau(k)\ \tilde{H}(k, l_A, l_B),$$

while her value from party A is:

$$V_B(l_A, l_B) = \sum_{k=2}^{4} \tau(5-k)\ \tilde{H}(k-1, l_A, l_B).$$

Moreover, when θ is small enough, for each pair of lists (l_A, l_B), the difference $\Delta(l_A, l_B) = V_A(l_A, l_B) - V_B(l_A, l_B)$ is uniquely determined.

Proof of Lemma 2. For a given voter, the probability of being pivotal for the assignment of A's k^{th} seat ($k \in \{2,3,4\}$), conditional on being pivotal, and given other voters' computed $\Delta \equiv V_A - V_B$ is given by

$$\frac{\Pr\left(\pi_A = \pi(k)\right)}{\sum_{r=2}^{4} \Pr\left(\pi_A = \pi(r)\right)} = \frac{f\left(\phi^{-1}\pi(k) - \phi^{-1}\frac{1}{2}) - \Delta\right)}{\sum_{r=2}^{4} f\left(\phi^{-1}\pi(r) - \phi^{-1}\frac{1}{2} - \Delta\right)} = \frac{f\left(\tilde{\pi}(k) - \Delta\right)}{\sum_{r=2}^{4} f\left(\tilde{\pi}(r) - \Delta\right)} \equiv \tau_k(\Delta)$$

where $\tilde{\pi}(k) = \phi^{-1}\left(\pi(k) - \frac{1}{2}\right)$ is such that, by symmetry $\sum_{r=2}^{4} \tilde{\pi}(k) = 0$. We can then rewrite the value of voting for A for a voter given other voters' computed $V_A - V_B$ as follows:

$$V_A(\Delta) = \sum_{k=2}^{4} \tau_k(\Delta)\tilde{H}(k; l_A, l_B) \tag{7}$$

$$V_B(\Delta) = \sum_{k=2}^{4} \tau_k(\Delta)\tilde{H}(k-1; l_A, l_B) \tag{8}$$

First, we show that there exists $\kappa > 0$ such that when $|\tilde{f}'(\cdot)| \leq \kappa$, the mapping $\mathcal{V} : [\underline{V} - \overline{V}, \overline{V} - \underline{V}] \to [\underline{V} - \overline{V}, \overline{V} - \underline{V}]$ defined as

$$\mathcal{V}(\Delta) = \sum_{k=2}^{4} \tau_k(\Delta)\left[\tilde{H}(k; l_A, l_B) - \tilde{H}(k-1; l_A, l_B)\right]$$

has a unique fixed point. To see this, first notice that $\tilde{H}(k; l_A, l_B) - \tilde{H}(k-1; l_A, l_B) < \overline{V} - \underline{V}$ (since the two values differ only by one elected councilor) and $\mathcal{V}$ is a convex combination of three such values, which implies that $\mathcal{V}(\underline{V} - \overline{V}) > \underline{V} - \overline{V}$ and $\mathcal{V}(\overline{V} - \underline{V}) < \overline{V} - \underline{V}$. Second, notice that, by construction, $\sum_{k=2}^{4} \tau_k(\Delta) = 1$, which implies that $\sum_{k=2}^{4} \frac{\mathrm{d}}{\mathrm{d}\Delta}\tau_k(\Delta) = 0$. Hence, we can write (suppressing the dependence of $\tilde{H}(k; l_A, l_B)$ on (l_A, l_B) for clarity

$$\begin{aligned}
\frac{\mathrm{d}}{\mathrm{d}\Delta}\mathcal{V}(\Delta) &= [2\tilde{H}(2) - \tilde{H}(1) - \tilde{H}(3)]\frac{\mathrm{d}}{\mathrm{d}\Delta}\tau_2(\Delta) + [\tilde{H}(4) + \tilde{H}(2) - 2\tilde{H}(3)]\frac{\mathrm{d}}{\mathrm{d}\Delta}\tau_4(\Delta) \\
&< (\overline{V} - \underline{V})\left(\left|\frac{\mathrm{d}}{\mathrm{d}\Delta}\tau_2(\Delta)\right| + \left|\frac{\mathrm{d}}{\mathrm{d}\Delta}\tau_4(\Delta)\right|\right) \\
&< \frac{2\kappa\theta(\overline{V} - \underline{V})}{\sum_{r=2} f\left(\tilde{\pi}(r) - \Delta\right)}
\end{aligned}$$

since

$$\left|\frac{\mathrm{d}}{\mathrm{d}\Delta}\tau_k(\Delta)\right| = \left|\frac{f\left(\tilde{\pi}(k)-\Delta\right)\sum_{r\neq k}^{4} f'\left(\tilde{\pi}(r)-\Delta\right) - f'\left(\tilde{\pi}(k)-\Delta\right)\sum_{r\neq k} f\left(\tilde{\pi}(r)-\Delta\right)}{\left(\sum_{r=2} f\left(\tilde{\pi}(r)-\Delta\right)\right)^2}\right|$$
$$< \frac{\kappa\theta\sum_{r=2}^{4} f\left(\tilde{\pi}(r)-\Delta\right)}{\left(\sum_{r=2} f\left(\tilde{\pi}(r)-\Delta\right)\right)^2} = \frac{\kappa\theta}{\sum_{r=2} f\left(\tilde{\pi}(r)-\Delta\right)}$$

and $f'(x) = \theta\tilde{f}'(x)$ and $\kappa \equiv \max \tilde{f}'(x)$, $x \in \left[-\frac{1}{2\psi}, \frac{1}{2\psi}\right]$. As a consequence, whenever θ is small enough, $\frac{\mathrm{d}}{\mathrm{d}\Delta}\mathcal{V}(\Delta) < 1$ and $\mathcal{V}(\Delta)$ admits a unique fixed point. □

Lemma 3. If party A is sufficiently advantaged, i.e., $\bar{\xi}$ is large enough, then $\tau(4) > \tau(3) > \tau(2)$ for all $(l_A, l_B) \in L(Q)^2$.

Proof. Let $\bar{\xi} \geq \overline{V} - \underline{V} + \tilde{\pi}(4)$. In this case, we have that $\forall\ \Delta \in (\underline{V} - \overline{V}, \overline{V} - \underline{V})$,

$$\tilde{\pi}(2) - \Delta < \tilde{\pi}(3) - \Delta < \tilde{\pi}(4) - \Delta < \bar{\xi}$$

which implies

$$f(\tilde{\pi}(2) - \Delta) < f(\tilde{\pi}(3) - \Delta) < f(\tilde{\pi}(4) - \Delta)$$

This completes the proof. □

Proposition 1. (Marginal Rank Hypothesis) When parties are purely seat-motivated, for θ not too large, both parties place their best candidates in their competitive ranks.

1. Within these ranks, A's list does not put its strongest candidate at the top and B's list does not place its highest quality candidate at the bottom:

$$l_A(2) < \max\{l_A(3), l_A(4)\}\,; \quad l_B(4) > \min\{l_B(2), l_B(3)\}.$$

2. When A's advantage, i.e., $\bar{\xi}$, is large enough, then for any pair of contested ranks $k' > k$,

$$l_A(k') > l_A(k)\,; \quad l_B(k') < l_B(k).$$

From Lemma 2, the mapping $\mathcal{V}$ admits a unique fixed point and thus there exists a unique $\Delta(l)$ for each pure strategy profile (l_A, l_B). Let λ_J denote a mixed strategy for party J (a probability distribution over the elements of $L(Q)$) and $\mathcal{L}$ denote the set of individually rational mixed strategy profiles. An equilibrium is a profile $(\lambda_A^*, \lambda_B^*)$ that solves

$$\min_{\lambda_B} \max_{\lambda_A} \sum_{y \in L(Q)} \sum_{z \in L(Q)} \Delta(y,z)\lambda_A(y)\lambda_B(z)$$

Since the set $L(Q)$ is finite, by the Minimax Theorem, an equilibrium exists, and it produces a unique Δ^*. This immediately implies that there is a unique vector of attentions weights $\{\tau^*(2), \tau^*(3), \tau^*(4)\}$.

Part 1. We can rewrite the mapping $\mathcal{V}$ implicitly defining Δ as follows:

$$\mathcal{V}(\Delta) = \frac{\tilde{H}(4; l_A, l_B) - \tilde{H}(1; l_A, l_B)}{3} + \sum_{k=2}^{4} \varepsilon_k(\Delta)\big[\tilde{H}(k; l_A, l_B) - \tilde{H}(k-1; l_A, l_B)\big]$$

where $\varepsilon_k(\Delta) = \tau_k(\Delta) - \frac{1}{3}$. Since $\varepsilon_k(\Delta) \underset{\theta \to 0}{\longrightarrow} 0$, there exists θ small enough for which

$$\tilde{H}(4; l_A, l_B) - \tilde{H}(1; l_A, l_B) > (<) 0 \Rightarrow \Delta(l_A, l_B) > (<) 0$$

Notice that, for any pair (l_A, l_B):

$$\tilde{H}(4; l_A, l_B) - \tilde{H}(1; l_A, l_B) = H\left(\sum_{i=1}^{4} l_A(i) + l_B(1), \max_{i \le 4} l_A(i)\right) - H\left(l_A(1) + \sum_{i=1}^{4} l_B(i), \max_{i \le 4} l_B(i)\right). \tag{9}$$

We next claim that each party must put its worst candidate, i.e., with quality 1, in an uncontested rank, when θ is small enough. We prove the argument for party A. Suppose not, i.e., that a list l_A is chosen with positive probability that satisfies $1 \in \{l_A(2), l_A(3), l_A(4)\}$; let $i \in \{2,3,4\}$ denote the rank such that $l_A(i) = 1$. Consider an alternative list l'_A that sets $l'_A(5) = 1$ and $l'_A(i) = l_A(5)$ and otherwise replicates l_A. For any realized l_B, we have that $\tilde{H}(1; l_A, l_B) = \tilde{H}(1; l'_A, l_B)$, and also $\max_{j \le 4} l'_A(j) = 5$, and thus:

$$\begin{aligned} &\tilde{H}(4; l'_A, l_B) - \tilde{H}(1; l'_A, l_B) - [\tilde{H}(4; l_A, l_B) - \tilde{H}(1; l_A, l_B)] \\ =&H\left(\sum_{j=1}^{4} l_A(j) + \underbrace{l_A(5) - l_A(i)}_{>0} + l_B(1), 5\right) - H\left(\sum_{i=1}^{4} l_A(i) + l_B(1), \max_{i \le 4} l_A(i)\right) > 0, \end{aligned} \tag{10}$$

where $\max_{i \le 4} l_A(i) \le 5$. Since (10) holds for *any* list l_B, we conclude that l_A is strictly dominated, for θ small.

We next claim that each party must put its second-worst candidate, i.e., with quality 2, in an uncontested rank, when θ is small enough. We prove the argument for party A. Suppose not, i.e., that a list l_A is chosen with positive probability satisfying $2 \in \{l_A(2), l_A(3), l_A(4)\}$; let $i \in \{2,3,4\}$ denote the rank such that $l_A(i) = 2$. There are two cases, depending on whether $l_A(1) = 1$, or instead $l_A(5) = 1$.

Case 1: $l_A(1) = 1$. Consider an alternative list l'_A that sets $l'_A(5) = 2$ and $l'_A(i) = l_A(5)$ and otherwise replicates l_A. For any realized l_B, we have that $\tilde{H}(1; l_A, l_B) = \tilde{H}(1; l'_A, l_B)$, and also $\max_{j \leq 4} l'_A(j) = 5$, and thus the difference $\tilde{H}(4; l'_A, l_B) - \tilde{H}(1; l'_A, l_B) - [\tilde{H}(4; l_A, l_B) - \tilde{H}(1; l_A, l_B)]$ is again given by expression (10) for *any* list l_B. We conclude that l_A is strictly dominated, for θ small.

Case 2: $l_A(1) \neq 1$. Then, our previous claim implies that $l_A(5) = 1$, and thus $l_A(1) > 2$. Consider an alternative list l'_A that sets $l'_A(1) = 2$ and $l'_A(i) = l_A(1)$. For any realized l_B, we have that $\tilde{H}(4; l'_A, l_B) - \tilde{H}(4; l_A, l_B) = 0$, and thus

$$\begin{aligned}
&\tilde{H}(4; l'_A, l_B) - \tilde{H}(1; l'_A, l_B) - [\tilde{H}(4; l_A, l_B) - \tilde{H}(1; l_A, l_B)] \\
&= \tilde{H}(1; l_A, l_B) - \tilde{H}(1; l'_A, l_B) \\
&= \tilde{H}\left(l_A(1) + \sum_{i=1}^{4} l_B(i), \max_{j \leq 4} l_B(j)\right) - \tilde{H}\left(2 + \sum_{i=1}^{4} l_B(i), \max_{j \leq 4} l_B(j)\right) > 0,
\end{aligned} \tag{11}$$

since $l_A(1) > 2$. Since this inequality holds for *any* list l_B, we therefore conclude that l_A is strictly dominated, for θ small. This proves that, for $\theta > 0$, a list l_J is not strictly dominated only if $l_J(1) \in \{1,2\}$ and $l_J(5) \in \{1,2\}$.

We conclude that, for any pair (l_A, l_B) that arises with positive probability, if θ is small enough:

$$\begin{aligned}
&\tilde{H}(4; l_A, l_B) - \tilde{H}(1; l_A, l_B) \\
&\quad = H(l_A(1) + l_B(1) + \underbrace{3+4+5}_{=\sum_{r=2}^{4} l_A(r)}; 5) - H(l_A(1) + l_B(1) + \underbrace{3+4+5}_{=\sum_{r=2}^{4} l_B(r)}; 5) = 0,
\end{aligned} \tag{12}$$

i.e., Δ^* is close enough to zero, such that for every $\delta > 0$ we can find a θ small enough so that for *any* list (l_A, l_B) that is played with positive probability:

$$\tau_k(\Delta(l_A, l_B)) - \tau_k(0) < \delta.$$

This, in turn, implies that for θ small enough, $\tau_2(\Delta(l_A, l_B)) < \tau_3(\Delta(l_A, l_B))$ and $\tau_2(\Delta(l_A, l_B)) <$

$\tau_4(\Delta(l_A, l_B))$.

We now show that a list l_A satisfying $l_A(2) = 5$ is a strictly dominated action for party A. Suppose that in equilibrium $l_A(2) = 5$ with positive probability and consider an alternative list l'_A that sets $l'_A(2) = l_A(3)$ and $l'_A(3) = l_A(2)$ and otherwise satisfies $l_A(k) = l'_A(k)$. It is immediate to see that for any list l_B and rank $k \in \{3, 4\}$, $\tilde{H}(k; l_A, l_B) = \tilde{H}(k; l'_A, l_B)$. First, we observe that:

$$\begin{aligned}&\tilde{H}(2; l'_A, l_B) - \tilde{H}(2; l_A, l_B)\\ =&H\left(l_A(1) + l_A(2) + \underbrace{l_A(3) - l_A(2)}_{<0} + \sum_{j=1}^{3} l_B(j), 5\right) - H\left(l_A(1) + l_A(2) + \sum_{j=1}^{3} l_B(j), 5\right) < 0.\end{aligned}$$

Second, we show that $\Delta(l_A, l_B)$ is strictly decreasing in $\tilde{H}(2; l'_A, l_B)$, which implies that $\Delta(l'_A, l_B) > \Delta(l_A, l_B)$, completing the argument. Notice that, using the fact that in any equilibrium $\tilde{H}(4; l_A, l_B) = \tilde{H}(1; l_A, l_B)$, we can rewrite $\mathcal{V}(\Delta)$ as

$$\begin{aligned}\mathcal{V}(\Delta) =&\tilde{H}(4; l_A, l_B)[\tau_4(\Delta) - \tau_2(\Delta)]\\ &+\tilde{H}(3; l_A, l_B)[\tau_3(\Delta) - \tau_4(\Delta)]\\ &+\tilde{H}(2; l_A, l_B)[\tau_2(\Delta) - \tau_3(\Delta)]. \end{aligned} \tag{13}$$

Hence, by the implicit function theorem,

$$\frac{d\Delta(l_A, l_B)}{d\tilde{H}(2; l_A, l_B)} = \frac{\tau_2(\Delta) - \tau_3(\Delta)}{1 - \frac{\partial \mathcal{V}}{\partial \Delta}} < 0$$

which follows form $\tau_2(\Delta) - \tau_3(\Delta) < 0$, which we established earlier, and the the fact that, by Lemma 2, $\frac{\partial \mathcal{V}}{\partial \Delta} < 1$. The argument for why $l_B(4) < 5$ follows a similar logic.

Part 2. For each θ satisfying the assumptions of parts 1-3, we may find $\bar{\xi}$ large enough so that for every pair of individually rational lists (l_A, l_B), $\tilde{\pi}(4) + \Delta(l_A, l_B) < \bar{\xi}$. This implies that for all $\Delta(l_A, l_B)$,

$$\tau_2(\Delta) < \tau_3(\Delta) < \tau_4(\Delta).$$

Inspection of Equation 13 implies that—all else equal—Δ *increases* in $\tilde{H}(4; l_A, l_B)$; likewise, it *decreases* in $\tilde{H}(3; l_A, l_B)$ and in $\tilde{H}(2; l_A, l_B)$.

First, we show that $l_A(2) = 3$. Suppose not, that is $l_A(2) > 3$, and let $i > 2$ denote the rank of 3, so $l_A(i) = 3$. There are three cases: (i) $i = 3$, (ii) $i = 4$ and $l_A(3) = 4$, (iii) $i = 4$ and $l_A(3) = 5$.

Case 1: $i = 3$. Consider the list l'_A that sets $l'_A(2) = 3$ and $l'_A(i) = l_A(2)$ and otherwise replicates l_A. Then, $\tilde{H}(3; l_A, l_B) = \tilde{H}(3; l'_A, l_B)$ and $\tilde{H}(3; l_A, l_B) = \tilde{H}(3; l'_A, l_B)$; however:

$$\begin{aligned}
&\tilde{H}(2; l'_A, l_B) - \tilde{H}(2; l_A, l_B) \\
&\qquad = H\left(l_A(1) + l_A(2) + \underbrace{3 - l_A(2)}_{<0} + \sum_{j=1}^{3} l_B(j), \max_{j \in \{1,2,3\}} l_B(j)\right) \\
&\qquad - H\left(l_A(1) + l_A(2) + \sum_{j=1}^{3} l_B(j), \max_{j \in \{1,2,3\}} l_B(j)\right) < 0.
\end{aligned}$$

This implies that l'_A is strictly preferred to l_A for any l_B, i.e., that l_A is strictly dominated.

Case 2: $i = 4$ *and* $l_A(3) = 4$. This also implies that $l_A(2) = 5$. Consider the list l'_A that sets $l'_A(2) = 4$ and $l'_A(3) = 5$ and otherwise replicates l_A. Then, $\tilde{H}(3; l_A, l_B) = \tilde{H}(3; l'_A, l_B)$ and $\tilde{H}(4; l_A, l_B) = \tilde{H}(4; l'_A, l_B)$; however:

$$\begin{aligned}
&\tilde{H}(2; l'_A, l_B) - \tilde{H}(2; l_A, l_B) \\
&= H\left(l_A(1) + 5 + \underbrace{4 - 5}_{<0} + \sum_{j=1}^{3} l_B(j), \max_{j \in \{1,2,3\}} l_B(j)\right) - H\left(l_A(1) + 5 + \sum_{j=1}^{3} l_B(j), \max_{j \in \{1,2,3\}} l_B(j)\right) < 0.
\end{aligned}$$

This implies that l'_A is strictly preferred to l_A for any l_B, i.e., that l_A is strictly dominated.

Case 3: $i = 4$ *and* $l_A(3) = 5$. This implies that $l_A(2) = 4$. Consider the list l'_A that sets $l'_A(2) = 3$ and $l'_A(4) = 4$ and otherwise replicates l_A. In terms of induced $\tilde{H}(k; \cdot, l_B)$, the two lists differ in terms of both $\tilde{H}(2; l_A, \cdot)$ and $\tilde{H}(3; l_A, \cdot)$. It is immediate to see that

$$\begin{aligned}
&\tilde{H}(2; l'_A, l_B) - \tilde{H}(2; l_A, l_B) \\
&= H\left(l_A(1) + 4 + \underbrace{3 - 4}_{<0} + \sum_{j=1}^{3} l_B(j), \max_{j \in \{1,2,3\}} l_B(j)\right) - H\left(l_A(1) + 4 + \sum_{j=1}^{3} l_B(j), \max_{j \in \{1,2,3\}} l_B(j)\right) < 0
\end{aligned}$$

and

$$\begin{aligned}
&\tilde{H}(3; l'_A, l_B) - \tilde{H}(3; l_A, l_B) \\
&\quad = H\left(l_A(1) + 4 + 5 + \underbrace{3 - 4}_{<0} + \sum_{j=1}^{2} l_B(j), 5\right) - H\left(l_A(1) + 4 + 5 + \sum_{j=1}^{2} l_B(j), 5\right) < 0
\end{aligned}$$

This implies that l'_A is strictly preferred to l_A for any l_B, i.e., that l_A is strictly dominated.

Having established that $l_A(2) = 3$, we just need to show that $l_A(3) = 4$. Suppose not, that is $l_A(3) = 5$, and consider the list l'_A that sets $l'_A(3) = 4$ and $l'_A(4) = 5$ and otherwise replicates l_A. Then, $\tilde{H}(2; l_A, l_B) = \tilde{H}(2; l'_A, l_B)$ and $\tilde{H}(4; l_A, l_B) = \tilde{H}(4; l'_A, l_B)$; however:

$$\begin{aligned}
&\tilde{H}(3; l'_A, l_B) - \tilde{H}(3; l_A, l_B) \\
&= H\left(l_A(1) + 3 + 5 + \underbrace{4 - 5}_{<0} + \sum_{j=1}^{2} l_B(j), 5 + \underbrace{4 - 5}_{<0}\right) - H\left(l_A(1) + 3 + 5 + \sum_{j=1}^{2} l_B(j), 5\right) < 0
\end{aligned}$$

This implies that l'_A is strictly preferred to l_A for any l_B, i.e., that l_A is strictly dominated. The argument that we must have $l_B(2) = 5$, $l_B(3) = 4$, $l_B(4) = 3$ is similar.

For our next result, define l^*:

$$l^*_J(1) = 5,\ l^*_J(2) = 4,\ l^*_J(3) = 3,\ l^*_J(4) = 2,\ l^*_J(5) = 1.$$

Proposition 2 *If Condition 1 is satisfied, and θ is not too large*:

1. *for any list l_{-J}, l^* is party J's unique best response;*
2. *if Condition 2 is also satisfied, then for any $l' \neq l^*$:*

$$\Pi_A(l^*, l^*) - \Pi_A(l', l^*) > \Pi_B(l^*, l^*) - \Pi_B(l', l^*). \tag{14}$$

Proof of Proposition 2. *Part (i)* Let

$$p_J(k|l_J, l_{-J}) \equiv \Pr(\pi(k) < \pi_J < \pi(k+1)).$$

Consider the relative value to party J from employing list l^*, rather than a list $l' \in L^*$, when party $-J$ chooses list l_{-J}:

$$\begin{aligned}
\Pi_J(l^*, l_{-J}) - \Pi_J(l', l_{-J}) &= \sum_{k=1}^{4} p_J(k|l^*, l_{-J})\tilde{G}(k; l^*, l_{-J}) - \sum_{k=1}^{4} p_J(k|l', l_{-J})\tilde{G}(k; l', l_{-J}) \\
&= \sum_{k=1}^{4} [p_J(k|l^*, l_{-J})(\tilde{G}(k; l^*, l_{-J}) - \tilde{G}(k; l', l_{-J})) + R^\theta_J(k; l^*, l', l_{-J})\tilde{G}(k; l', l_{-J})],
\end{aligned} \tag{15}$$

where

$$R^\theta_J(k; l^*, l', l_{-J}) = -p_J(k|l', l_{-J}) + p_J(k|l^*, l_{-J}). \tag{16}$$

Suppose that $\Delta(l^*, l_{-J}) \geq \Delta(l', l_{-J})$. Then $R^\theta_J(k; l^*, l', l_{-J}) \geq 0$, and since for each $k \in \{1, \ldots, 4\}$, $\tilde{G}(k; l^*, l_{-J}) - \tilde{G}(k; l', l_{-J}) \geq 0$, l^* is optimal.

Suppose, instead, $\Delta(l^*, l_{-J}) < \Delta(l', l_{-J})$. Notice that $p_J(k|l_J, l_{-J}) = p_{-J}(5-k|l_J, l_{-J})$ and recall that $\tilde{\pi(3)} = 0$ and

$$\tilde{\pi(2)} = -\tilde{\pi(4)} \equiv \tilde{\pi} < 0.$$

Let $\Delta^0(l_A, l_B) = \lim_{\theta\to 0} \Delta(l_A, l_B)$. We have

$$\lim_{\theta\to 0} p_A(1|l_A, l_B) = \frac{1}{2} + \psi\tilde{\pi} - \psi\Delta^0(l_A, l_B) \tag{17}$$

$$\lim_{\theta\to 0} p_A(2|l_A, l_B) = -\psi\tilde{\pi} \tag{18}$$

$$\lim_{\theta\to 0} p_A(3|l_A, l_B) = -\psi\tilde{\pi} \tag{19}$$

$$\lim_{\theta\to 0} p_A(4|l_A, l_B) = \frac{1}{2} + \psi\tilde{\pi} + \psi\Delta^0(l_A, l_B). \tag{20}$$

Hence

$$\begin{aligned}
\lim_{\theta\to 0} \Big\{\Pi_J(l^*, l_{-J}) - \Pi_J(l', l_{-J})\Big\} &= \Big[\frac{1}{2} + \psi\tilde{\pi}\Big]\Big[\tilde{G}(4; l^*, l_{-J}) - \tilde{G}(4; l', l_{-J}) + \tilde{G}(1; l^*, l_{-J}) - \tilde{G}(1; l', l_{-J})\Big] \\
&\quad -\psi\tilde{\pi}\Big[\tilde{G}(3; l^*, l_{-J}) - \tilde{G}(3; l', l_{-J}) + \tilde{G}(2; l^*, l_{-J}) - \tilde{G}(2; l', l_{-J})\Big] \\
&\quad + \psi\Delta^0(l^*, l_{-J})\Big[\tilde{G}(4; l^*, l_{-J}) - \tilde{G}(1; l^*, l_{-J})\Big] - \psi\Delta^0(l', l_{-J})\Big[\tilde{G}(4; l', l_{-J}) - \tilde{G}(1; l', l_{-J}).\Big]
\end{aligned}$$

Suppose that $l'(1) = l^*(1) = 5$. As θ approaches zero, only the aggregate quality of the candidates in the contested ranks matter for Δ:

$$\Delta^0(l_J, l_{-J}) = \frac{1}{3}\sum_{k=2}^{4}\Big\{\tilde{H}(k, l_J, l_{-J}) - \tilde{H}(k-1, l_J, l_{-J})\Big\}$$

As a consequence,

$$\begin{aligned}
\Delta^0(l', l_{-J}) - \Delta^0(l^*, l_{-J}) &\propto \tilde{H}(4, l', l_{-J}) - \tilde{H}(1, l', l_{-J}) - \tilde{H}(4, l^*, l_{-J}) + \tilde{H}(1, l^*, l_{-J}) \\
&= \tilde{H}(4, l', l_{-J}) - \tilde{H}(4, l^*, l_{-J}) \leq 0,
\end{aligned}$$

which is a contradiction. Hence, we must have $l'(1) \neq 5$, which in turn implies that $\tilde{G}(1; l^*, l_{-J}) - \tilde{G}(1; l', l_{-J}) > 0$.

Moreover, we can restrict l' to (i) $l'(5) = 1$ and (ii) l' strictly decreases in candidate quality in the contested ranks. The reason is that for any list l'' that does not satisfy either restrictions, we can find a list l''' that is preferred by party J to l'', i.e., yields weakly higher Δ^0 and weakly higher $\tilde{G}$, with one strict inequality:

$$\lim_{\theta\to 0}\Big\{\Pi_J(l''', l_{-J}) - \Pi_J(l'', l_{-J})\Big\} > 0.$$

Since $l'(5) = 1$, l^* and l' yield the same set of elected councilors when party A receives four seats:

$$\tilde{H}(4, l^*, l_{-J}) = \tilde{H}(4, l', l_{-J})$$

Combining the latter fact with the definition of Δ^0 yields

$$\begin{aligned}
\Delta^0(l', l_{-J}) - \Delta^0(l^*, l_{-J}) &= \frac{1}{3}\Big[\tilde{H}(1, l^*, l_{-J}) - \tilde{H}(1, l', l_{-J})\Big] \\
&= \frac{1}{3}\left[H\left(5 + \sum_{i=1}^{4} l_{-J}(i), \max_{i<5} l_{-J}(i)\right) - H\left(l'(1) + \sum_{i=1}^{4} l_{-J}(i), \max_{i<5} l_{-J}(i)\right)\right] \\
&\leq \frac{1}{3}\left[H\left(5 + 14, 5\right) - H\left(2 + 14, 5\right)\right] = \frac{1}{3}\int_{16}^{19} H_1(z, 5)\, dz \leq \frac{1}{3}\kappa\int_{16}^{19} dz = \kappa.
\end{aligned}$$

As a consequence—and using again $\tilde{G}(4; l^*, l_{-J}) = \tilde{G}(4; l', l_{-J})$—we can write

$$\begin{aligned}
\lim_{\theta\to 0}\Big\{\Pi_J(l^*, l_{-J}) - \Pi_J(l', l_{-J})\Big\} &\geq \Big[\frac{1}{2} + \psi\tilde{\pi}\Big]\Big[\tilde{G}(1; l^*, l_{-J}) - \tilde{G}(1; l', l_{-J})\Big] \\
&\quad + \psi\Delta^0(l^*, l_{-J})\Big[\tilde{G}(4; l^*, l_{-J}) - \tilde{G}(1; l^*, l_{-J})\Big] \\
&\quad - \psi\Big[\Delta^0(l^*, l_{-J}) + \kappa\Big]\Big[\tilde{G}(4; l^*, l_{-J}) - \tilde{G}(1; l', l_{-J})\Big] \\
&= \Big[\frac{1}{2} + \psi\tilde{\pi} - \psi\Delta^0(l^*, l_{-J})\Big]\Big[\tilde{G}(1; l^*, l_{-J}) - \tilde{G}(1; l', l_{-J})\Big] \\
&\quad - \psi\kappa\Big[\tilde{G}(4; l^*, l_{-J}) - \tilde{G}(1; l', l_{-J})\Big].
\end{aligned}$$

Since the distribution of mayoral assignments under (l^*, l_{-J}) and (l', l_{-J}) does not change, Condition 1 implies that κ and $\Delta^0(l^*, l_{-J})$ are small enough relative to the term $\tilde{G}(1; l^*, l_{-J}) - \tilde{G}(1; l', l_{-J}) > 0$. This implies that for each list l_{-J} there exists a finite number $\kappa(l_{-J})$ such that

$$\frac{\kappa(l_{-J})}{\frac{1}{2} + \psi\tilde{\pi} - \psi\Delta^0(l^*, l_{-J})} = \frac{\tilde{G}(1; l^*, l_{-J}) - \tilde{G}(1; l', l_{-J})}{\tilde{G}(4; l^*, l_{-J}) - \tilde{G}(1; l', l_{-J})}$$

we can then define $\kappa \equiv \min_{l_{-J}\in L(Q)} \kappa(l_{-J})$.

Part (ii) Consider a potential deviation l' that for all θ yields a higher winning probability to J. By a reasoning similar to the one in the previous part of this proof, it must be that $l'(1) \neq 5$, $l'(5) = 1$ and that l' is strictly decreasing in the contested ranks.

To prove the second claim, recall that $p_B(k|l,l') = p_A(5-k|l,l')$ and consider the difference:

$$\begin{aligned}
&\Pi_A(l^*,l^*) - \Pi_A(l',l^*) - (\Pi_B(l^*,l^*) - \Pi_B(l',l^*)) \\
&= \sum_{k=1}^{4}[p_A(k|l^*,l^*)(\tilde{G}(k;l^*,l_{-J}) - \tilde{G}(k;l',l_{-J})) + R_A^\theta(k;l^*,l',l^*)\tilde{G}(k;l',l^*)] \\
&- \sum_{k=1}^{4}[p_B(k|l^*,l^*)(\tilde{G}(k;l^*,l_{-J}) - \tilde{G}(k;l',l_{-J})) + R_B^\theta(k;l^*,l',l^*)\tilde{G}(k;l',l^*)] \\
&= \sum_{k=1}^{4}[p_A(k|l^*,l^*) - p_B(k|l^*,l^*)](\tilde{G}(k;l^*,l_{-J}) - \tilde{G}(k;l',l_{-J})) \\
&+ \sum_{k=1}^{4}[R_A^\theta(k;l^*,l',l^*) - R_B^\theta(k;l^*,l',l^*)]\tilde{G}(k;l',l^*) \\
&= \sum_{k=1}^{4}[p_A(k|l^*,l^*) - p_A(5-k|l^*,l^*)](\tilde{G}(k;l^*,l_{-J}) - \tilde{G}(k;l',l_{-J})) \\
&+ \sum_{k=1}^{4}[R_A^\theta(k;l^*,l',l^*) - R_B^\theta(k;l^*,l',l^*)]\tilde{G}(k;l',l^*). \qquad (21)
\end{aligned}$$

We therefore obtain:

$$\begin{aligned}
&\Pi_A(l^*,l^*) - \Pi_A(l',l^*) - (\Pi_B(l^*,l^*) - \Pi_B(l',l^*)) \\
&= p_A(4|l^*,l^*) - p_A(1|l^*,l^*)[\tilde{G}(4|l^*,l^*) - \tilde{G}(4|l',l^*) - (\tilde{G}(1|l^*,l^*) - \tilde{G}(1|l',l^*))] \\
&+ p_A(3|l^*,l^*) - p_A(2|l^*,l^*)[\tilde{G}(3|l^*,l^*) - \tilde{G}(3|l',l^*) - (\tilde{G}(2|l^*,l^*) - \tilde{G}(2|l',l^*))] \\
&+ \sum_{k=1}^{4}[R_A^\theta(k;l^*,l',l^*) - R_B^\theta(k;l^*,l',l^*)]\tilde{G}(k;l',l^*). \qquad (22)
\end{aligned}$$

We wish to show $\lim_{\theta\to 0}[\Pi_A(l^*,l^*) - \Pi_A(l',l^*) - (\Pi_B(l^*,l^*) - \Pi_B(l',l^*))] > 0$. Under Condition 2, we have

$$\tilde{G}(4|l^*,l^*) - \tilde{G}(4|l',l^*) - (\tilde{G}(1|l^*,l^*) - \tilde{G}(1|l',l^*)) > 0$$
$$\tilde{G}(3|l^*,l^*) - \tilde{G}(3|l',l^*) - (\tilde{G}(2|l^*,l^*) - \tilde{G}(2|l',l^*)) > 0.$$

Let $\Delta^* \equiv \Delta(l^*,l^*)$, $\Delta'_A \equiv \Delta(l',l^*)$, and $\Delta'_B \equiv \Delta(l^*,l')$. We have, since $\Delta^* \xrightarrow[\theta\to 0]{} 0$, that

$$p_A(4|l^*,l^*) - p_A(1|l^*,l^*) = 1 - F(-\tilde{\pi} - \Delta^*) - F(\tilde{\pi} - \Delta^*) \xrightarrow[\theta\to 0]{} 0$$
$$p_A(3|l^*,l^*) - p_A(2|l^*,l^*) = F(-\tilde{\pi} - \Delta^*) + F(\tilde{\pi} - \Delta^*) - 2F(-\Delta^*) \xrightarrow[\theta\to 0]{} 0$$

We argue that both of these terms are strictly positive, for all $\theta > 0$. To see this, note first that for θ sufficiently small, $-\Delta^* < \overline{\xi}$. Then, observe that $F(\cdot)$ is strictly convex on the interval $[-(2\psi)^{-1}, \overline{\xi}]$, and strictly concave on the interval $[\overline{\xi}, (2\psi)^{-1}]$, and recall that $\tilde{f}(\cdot)$ is symmetric around its mode $\overline{\xi}$. There are two possible cases.

Case 1: Suppose, first, $-\tilde{\pi} - \Delta^* \leq \overline{\xi}$. Then,

$$\begin{aligned} 1 - F(-\tilde{\pi} - \Delta^*) - F(\tilde{\pi} - \Delta^*) >& F(\overline{\xi} + \overline{\xi} + \tilde{\pi} + \Delta^*) - F(\tilde{\pi} - \Delta^*) \\ >& F(\overline{\xi}) - F(\tilde{\pi} - \Delta^*) > \frac{1}{2} - F(\tilde{\pi} - \Delta^*) > 0 \end{aligned}$$

so that $p_A(4|l^*, l^*) - p_A(1|l^*, l^*)$ is strictly positive. And, the convexity of $F(x)$ in $x < \overline{\xi}$ further implies

$$\frac{1}{2}F(-\tilde{\pi} - \Delta^*) + \frac{1}{2}F(\tilde{\pi} - \Delta^*) > F\left(\frac{1}{2}\Big(- \tilde{\pi} - \Delta^*\Big) + \frac{1}{2}\Big(\tilde{\pi} - \Delta^*\Big)\right) = F(-\Delta^*),$$

so that $p_A(3|l^*, l^*) - p_A(2|l^*, l^*)$ is strictly positive.

Case 2: Suppose, instead, that $-\tilde{\pi} - \Delta^* > \overline{\xi}$. Then, since $\overline{\xi} > 0$, we have that $-\tilde{\pi} - \Delta^* - \overline{\xi} < \overline{\xi} - (\tilde{\pi} - \Delta^*)$, and symmetry of $\tilde{f}(\cdot)$ yields

$$1 - F(-\tilde{\pi} - \Delta^*) - F(\tilde{\pi} - \Delta^*) > F(\tilde{\pi} - \Delta^*) - F(\tilde{\pi} - \Delta^*) = 0, \tag{23}$$

so that $p_A(4|l^*, l^*) - p_A(1|l^*, l^*)$ is strictly positive. Finally, consider the expression:

$$\Lambda(x) = F(x - \Delta^*) - F(-x - \Delta^*) - 2F(-\Delta^*). \tag{24}$$

We observe that $\Lambda'(x) = f(x - \Delta^*) - f'(-x - \Delta^*)$ which is strictly positive for $x > 0$. Since $\Lambda(0) = 0$, we conclude that $\Lambda(-\tilde{\pi}) > 0$ since $\tilde{\pi} = \tilde{\pi}(2) < 0$.

Next, we observe:

$$R_A^\theta(k; l^*, l', l^*) - R_B^\theta(k; l^*, l', l^*) = p_A(k|l^*, l^*) - p_A(5 - k|l^*, l^*) - p_A(k|l', l^*) + p_A(5 - k|l^*, l')$$

$$= \begin{cases} (1-\theta)\psi(\Delta'_A - \Delta^*) + \theta[\tilde{F}(\tilde{\pi} - \Delta^*) - \tilde{F}(\tilde{\pi} - \Delta'_A)] & \\ -(1-\theta)\psi(\Delta^* - \Delta'_B) - \theta[\tilde{F}(-\tilde{\pi} - \Delta'_B) - \tilde{F}(-\tilde{\pi} - \Delta^*)] & k=1 \\ \theta[\tilde{F}(-\Delta^*) - \tilde{F}(\tilde{\pi} - \Delta^*) - \tilde{F}(-\Delta'_A) + \tilde{F}(\tilde{\pi} - \Delta'_A)] & \\ -\theta[\tilde{F}(-\tilde{\pi} - \Delta^*) - \tilde{F}(-\Delta^*) - \tilde{F}(-\tilde{\pi} - \Delta'_B) + \tilde{F}(-\Delta'_B)] & k=2 \\ -\theta[\tilde{F}(-\Delta^*) - \tilde{F}(\tilde{\pi} - \Delta^*) - \tilde{F}(-\Delta'_B) + \tilde{F}(\tilde{\pi} - \Delta'_B)] & \\ +\theta[\tilde{F}(-\tilde{\pi} - \Delta^*) - \tilde{F}(-\Delta^*) - \tilde{F}(-\tilde{\pi} - \Delta'_A) + \tilde{F}(-\Delta'_A)] & k=3 \\ -(1-\theta)\psi(\Delta'_B - \Delta^*) - \theta[\tilde{F}(\tilde{\pi} - \Delta^*) - \tilde{F}(\tilde{\pi} - \Delta'_B)] & \\ +(1-\theta)\psi(\Delta^* - \Delta'_A) + \theta[\tilde{F}(-\tilde{\pi} - \Delta'_A) - \tilde{F}(-\tilde{\pi} - \Delta^*)]. & k=4 \end{cases}$$

Since $\lim_{\theta\to 0} \Delta'_A = \lim_{\theta\to 0} \Delta'_B = \lim_{\theta\to 0} \Delta^*$, we obtain $R^\theta_A(k; l^*, l', l^*) - R^\theta_B(k; l^*, l', l^*) \xrightarrow[\theta\to 0]{} 0$ for each $k \in \{1, 2, 3, 4\}$. In light of this, to complete the proof we need to show that for $k \in \{3, 4\}$, $p_A(k|l^*, l^*) - p_A(5-k|l^*, l^*)$ converges to zero more slowly than

$$R^\theta_A(k; l^*, l', l^*) - R^\theta_B(k; l^*, l', l^*)$$

and

$$R^\theta_A(5-k; l^*, l', l^*) - R^\theta_B(5-k; l^*, l', l^*).$$

Now, we argue that (we make the dependence of Δ on θ explicit)

$$\Delta'_A(\theta) + \Delta'_B(\theta) - 2\Delta^*(\theta) = O(\Delta^*(\theta)) \quad \text{as} \quad \theta \to 0.$$

To see this, notice that using L'Hopital's rule and the Implicit Function Theorem, we obtain:

$$\lim_{\theta\to 0} \frac{\Delta'_A(\theta)}{\Delta^*(\theta)} = \lim_{\theta\to 0} \frac{d\Delta'_A(\theta)/d\theta}{d\Delta^*(\theta)/d\theta} = \frac{\sum_{k=2}^4 [2\tilde{f}(\tilde{\pi}(k)) - \sum_{r\neq k} \tilde{f}(\tilde{\pi}(r))][\tilde{H}(k; l', l^*) - \tilde{H}(k-1; l', l^*)]}{\sum_{k=2}^4 [2\tilde{f}(\tilde{\pi}(k)) - \sum_{r\neq k} \tilde{f}(\tilde{\pi}(r))][\tilde{H}(k; l^*, l^*) - \tilde{H}(k-1; l^*, l^*)]} < \infty$$

which follows from

$$\frac{d\Delta(\theta)}{d\theta} = -\frac{\sum_{k=2}^4 \frac{d\tau_k(\Delta,\theta)}{d\theta} [\tilde{H}(k; l_A, l_B) - \tilde{H}(k-1; l_A, l_B)]}{\sum_{k=2}^4 \frac{d\tau_k(\Delta,\theta)}{d\Delta} [\tilde{H}(k; l_A, l_B) - \tilde{H}(k-1; l_A, l_B) - 1}, $$

$\lim_{\theta\to 0} \frac{d\tau_k(\Delta,\theta)}{d\Delta} = 0$, and

$$\frac{d\tau_k(\Delta,\theta)}{d\theta} = \frac{3\psi[\tilde{f}(\tilde{\pi}(k)) - \psi] - \left(\sum_{r=2}^4 \{\tilde{f}(\tilde{\pi}(r))\} - 3\psi\right)\psi}{9\psi^2} = \frac{2\tilde{f}(\tilde{\pi}(k)) - \sum_{r\neq k} \tilde{f}(\tilde{\pi}(r))}{9\psi}.$$

The argument for $\Delta'_B(\theta)$) is analogous. In light of this, and since $\lim_{\theta\to 0}\Delta'_A = \lim_{\theta\to 0}\Delta'_B = \lim_{\theta\to 0}\Delta^*$, for each we can write

$$\begin{aligned} R^{\theta}_A(k;l^*,l',l^*) - R^{\theta}_B(k;l^*,l',l^*) &= O\big(\Delta^*(\theta) + \theta\tilde{D}^k_F(\theta)\big) \quad \text{as} \quad \theta\to 0 \quad k\in\{1,4\} \\ R^{\theta}_A(k;l^*,l',l^*) - R^{\theta}_B(k;l^*,l',l^*) &= O\big(\theta\tilde{D}^k_F(\theta)\big) \quad \text{as} \quad \theta\to 0 \quad k\in\{2,3\} \end{aligned}$$

where

$$\tilde{D}^k_F(\theta) = \begin{cases} \tilde{F}(\tilde{\pi}-\Delta^*) - \tilde{F}(\tilde{\pi}-\Delta'_A) - \tilde{F}(-\tilde{\pi}-\Delta'_B) + \tilde{F}(-\tilde{\pi}-\Delta^*)] & k=1 \\ \tilde{F}(-\Delta^*) - \tilde{F}(\tilde{\pi}-\Delta^*) - \tilde{F}(-\Delta'_A) + \tilde{F}(\tilde{\pi}-\Delta'_A) & \\ -\tilde{F}(-\tilde{\pi}-\Delta^*) + \tilde{F}(-\Delta^*) + \tilde{F}(-\tilde{\pi}-\Delta'_B) - \tilde{F}(-\Delta'_B)] & k=2 \\ -\tilde{F}(-\Delta^*) + \tilde{F}(\tilde{\pi}-\Delta^*) + \tilde{F}(-\Delta'_B) - \tilde{F}(\tilde{\pi}-\Delta'_B) & \\ +\tilde{F}(-\tilde{\pi}-\Delta^*) - \tilde{F}(-\Delta^*) - \tilde{F}(-\tilde{\pi}-\Delta'_A) + \tilde{F}(-\Delta'_A)] & k=3 \\ -\tilde{F}(\tilde{\pi}-\Delta^*) + \tilde{F}(\tilde{\pi}-\Delta'_B) + \tilde{F}(-\tilde{\pi}-\Delta'_A) - \tilde{F}(-\tilde{\pi}-\Delta^*) & k=4 \end{cases}$$

goes to zero as $\theta\to 0$. This implies that there is no positive constant C for which

$$\begin{aligned} R^{\theta}_A(k;l^*,l',l^*) - R^{\theta}_B(k;l^*,l',l^*) &= O\big(\Delta^*(\theta) + \theta C\big) \quad \text{as} \quad \theta\to 0 \quad k\in\{1,4\} \\ R^{\theta}_A(k;l^*,l',l^*) - R^{\theta}_B(k;l^*,l',l^*) &= O\big(\theta C\big) \quad \text{as} \quad \theta\to 0 \quad k\in\{2,3\} \end{aligned}$$

On the other hand, since for all θ

$$\begin{aligned} [1-\tilde{F}(\tilde{\pi}-\Delta^*) - \tilde{F}(-\tilde{\pi}-\Delta^*)] &\xrightarrow[\theta\to 0]{} 1-\tilde{F}(\tilde{\pi}) - \tilde{F}(-\tilde{\pi}) > 0 \\ [\tilde{F}(\tilde{\pi}-\Delta^*) + \tilde{F}(-\tilde{\pi}-\Delta^*) - 2\tilde{F}(-\Delta^*)] &\xrightarrow[\theta\to 0]{} \tilde{F}(\tilde{\pi}) + \tilde{F}(-\tilde{\pi}) - 2\tilde{F}(0) > 0 \end{aligned}$$

there exists a constant $C>0$ such that

$$\begin{aligned} p_A(k|l^*,l^*) - p_A(5-k|l^*,l^*) &= O(\Delta^*(\theta)+\theta C) \quad \text{as} \quad \theta\to 0 \quad k\in\{1,4\} \\ p_A(k|l^*,l^*) - p_A(5-k|l^*,l^*) &= O(\theta C) \quad \text{as} \quad \theta\to 0 \quad k\in\{2,3\}. \end{aligned}$$

This completes the proof. □