Public Choice (2017) 171:257–281
DOI 10.1007/s11127-017-0436-x

Political selection under alternative electoral rules

Vincenzo Galasso[1] · **Tommaso Nannicini**[2]

Received: 26 August 2015 / Accepted: 17 February 2017 / Published online: 13 March 2017

Abstract We study the patterns of political selection in majoritarian versus proportional systems. Political parties face a trade off in choosing the mix of high- and low-quality candidates: high-quality candidates are valuable to voters, and thus help to win elections, but they crowd out loyal candidates, who are most preferred by the parties. Under proportional representation, politicians' selection depends on the share of swing voters in the entire electorate. In majoritarian elections, it depends also on the distribution of competitive versus safe (single-member) districts. We show that a majoritarian system with only a few competitive districts is less capable of selecting good politicians than a proportional system. As the share of competitive districts increases, the majoritarian system becomes more efficient than the proportional system. However, for a large enough share of competitive districts, a non-monotonic relation arises: the marginal (positive) effect of adding high-quality politicians on the probability of winning the election is reduced, and highly competitive majoritarian systems become less efficient than proportional ones in selecting good politicians.

Keywords Electoral rules · Political selection · Probabilistic voting

We thank Jim Snyder, Alois Stutzer, and seminar participants at the CSEF-IGIER Conference, the European Political Science Association Meetings in Dublin, the European Public Choice Society in Rennes, Bocconi University, and Erasmus University Rotterdam for their insightful suggestions. We also thank Andrea Di Miceli and Luca Riva for excellent research assistance. The remaining errors are ours and follow a random walk.

✉ Vincenzo Galasso
vincenzo.galasso@unibocconi.it

Tommaso Nannicini
tommaso.nannicini@unibocconi.it

[1] Department of Public Policy, Dondena, IGIER and CEPR, Bocconi University, Via Rontgen 1, 20136 Milan, Italy

[2] Department of Economics, IZA, IGIER and CEPR, Bocconi University, Via Rontgen 1, 20136 Milan, Italy

JEL codes D72 · D78 · P16

1 Introduction

Electoral rules are known to affect politicians' behavior. Majoritarian and proportional systems shape the incentives of both voters and politicians, and may thus lead to different policy outcomes (e.g., see Persson and Tabellini 2000; Voigt 2011). For instance, majoritarian systems have been shown to rely more on targeted redistribution and less on public goods than proportional systems, while rent-seeking tends to be higher in proportional systems (see Persson et al. 2003; Persson and Tabellini 2003; Gagliarducci et al. 2011).

Little emphasis, however, has been devoted to the impact of electoral rules on political selection. Norris (2004) studies how the political representation of women and ethnic minorities varies under different voting rules. In addition to achieving a more *equal* representation, electoral systems could also be designed to make political selection more *efficient*, namely, to increase the average quality of elected officials. Along these lines, the role of primary elections in selecting good candidates has been studied intensively (see Ashworth 2012 for a review). The recruitment of good politicians has been shown to depend both on the candidates' decisions to run for office (Caselli and Morelli 2004) and on the candidates' selection by political parties (Galasso and Nannicini 2011, 2015). In primary elections, both channels have been suggested to be at work (Hirano and Snyder 2014). To our knowledge, only a few studies examine directly how different electoral rules affect the competence (or valence) of politicians.[1] Myerson (1993) argues that higher entry barriers in majoritarian systems may lead to the election of low-quality (dishonest) candidates. Beath et al. (2014) propose a theoretical model in which district size may matter for the quality of the elected politicians, as voters have strategic reasons to prefer polarized politicians in district elections, but competent politicians in at-large elections. Mattozzi and Merlo (2015) study the recruitment of individuals by political parties and show that mediocre candidates are more likely to be chosen in proportional than in majoritarian systems.

Our paper highlights a channel of the political selection mechanism, which focuses on the conflict between voters and politicians, and which may vary in majoritarian and proportional electoral systems. Political parties select the candidates to be included on their electoral lists. Candidates differ in their valence: they can be either high- or low-quality. Valence is perfectly observable and valued by all voters under both electoral systems. Parties hence face a trade off. On the one hand, high-quality politicians are instrumental in winning elections, because voters value their expertise. On the other hand, low-quality politicians are loyal and therefore valuable to the party, to which they provide services (e.g., rent-seeking activities), but not to the voters.[2]

This trade off faced by the parties in selecting their candidates varies across electoral systems. In proportional systems, voters care about the average quality of the political candidates, and parties seek to win the electoral majority in the parliament by obtaining the support of the electorate's swing voters. In majoritarian systems, in addition to their positive impact on the average quality of political candidates, experts also are valued for their ability at the local (district) level. In order to increase the probability of winning in

[1] Throughout the paper, we follow the political science literature in using the term valence for competence of politicians.

[2] Low-quality politicians are more likely to be loyal to their party as they have a low value also in the labor market, and thus fewer outside options.

competitive (single-member) districts, and eventually to obtain an electoral majority in the parliament, parties thus have an incentive to allocate high-quality politicians to competitive districts and to send loyalists to safe ones. The overall share of high-quality politicians thus will be influenced by the distribution of competitive versus safe districts.

As a result, comparisons of the selection process between majoritarian and proportional systems depend on the share of competitive and safe districts in the majoritarian system. We show that, for a high concentration of safe districts, the majoritarian system is less capable of selecting good politicians than the proportional system. As the share of competitive districts increases, the majoritarian system becomes instead more effective. However, when this share is large enough, the selection process in the majoritarian system worsens, and becomes less efficient than the proportional system. The intuition for this non-monotonic effect is as follows. Placing a good politician in one of the many competitive districts in the majoritarian system still has a positive effect on the probability of winning the general election. However, the magnitude of this marginal effect is reduced, as too many districts are now competitive. Instead, the marginal cost to the party, in terms of foregone loyalists on the party list, remains constant. Hence, parties find it less convenient to select good politicians in a majoritarian system with too many competitive districts. As a result, majoritarian systems featuring large numbers of competitive districts become less effective in the political selection than proportional systems.

The paper is structured as follows. The next section reviews the related literature and provides some motivating evidence. Section 3 develops the theoretical model. We conclude with Sect. 4. All proofs are in the "Appendix".

2 Related literature and motivating evidence

A large theoretical and empirical literature has studied the effects of electoral rules. Majoritarian systems have been indicated to provide more targeted redistribution and fewer public goods than proportional systems (Persson and Tabellini 1999; Lizzeri and Persico 2001; Milesi-Ferretti et al. 2002). Electoral rules also may influence corruption and rent extraction by politicians. Theoretical predictions tend, however, to be ambiguous, with some models claiming that majoritarian elections increase the accountability of elected officials (Persson and Tabellini 1999, 2000), and others suggesting that proportional representation lowers entry barriers for honest competitors and therefore reduces rents from holding office (Myerson 1993).

The predictions of these models have been tested using cross-country aggregate data to find that proportional systems are associated with broader redistribution and more perceived corruption (see Persson and Tabellini 2003; Milesi-Ferretti et al. 2002; Persson et al. 2003). Funk and Gathmann (2013) use a difference-in-differences strategy with data on Swiss cantons to find that proportional systems shift spending toward education and social welfare benefits, but reduce spending on geographically targeted goods, such as roads. Gagliarducci et al. (2011) use a regression discontinuity design with data on the mixed-member Italian Parliament to find that politicians elected in majoritarian districts propose more targeted tax-and-spending bills and have lower absenteeism rates than politicians elected in proportional districts.

In most of the models mentioned above, politicians are homogeneous and the impact of electoral rules on policy outcomes is driven by the difference in incentives and accountability between majoritarian and proportional systems. The impact of electoral rules on political selection, wherein politicians are acknowledged to be of different types, has received less attention. Primaries have been suggested to promote selection of good candidates (Fearon 1999). Hirano and Snyder (2014) show that the magnitude of this positive selection effect of the primiry elections largely depends on the degree of competition of the final elections. Primaries select better candidates than open-seat elections in safe districts. How the political representation of women and ethnic minorities varies under different voting rules has been analyzed in Norris (2004). Iaryczower and Mattozzi (2013) study how alternative electoral rules affect the intensity of campaign competition, and thereby the number of candidates running for election and their degree of ideological differentiation. Despite a large literature on valence issues (Stokes 1963) and a recent literature on the importance of political selection (Besley 2005), however, only a few studies have examined whether electoral rules affect the quality (or valence) of politicians.[3]

In his seminal paper, Myerson (1993) builds a game-theoretic model showing that proportional systems may reduce entry barriers for honest politicians and, consequently, equilibrium rents (see also Myerson 1999). In his model, political parties differ along two dimensions: ideology and honesty. While voters may have different ideological preferences, they all favor honest parties. Honesty can thus be interpreted as a valence dimension. With plurality voting, a dishonest party still can clinch power, when the self-fulfilling prophecy of a close race between two dishonest politicians is realized. In this case, voters believe that their first-best choice has no chance of winning and thus vote rationally for the dishonest party whose ideology they share. This cannot happen under proportional representation. As governmental policy depends on whether a majority of the seats are allocated to leftist or rightist politicians, voters are free to vote for their first-best choice, thereby reducing corruption without affecting the balance between left and right in the parliament. Hence, fewer dishonest (or low-quality) politicians are elected than in the majoritarian scenario. The crucial mechanism here is the magnitude of the electoral district, which affects the height of entry barriers for high-quality candidates.

In Beath et al. (2014), district magnitude has a strategic effect on the voters' preferences over politician quality. Their theoretical model builds on the citizen candidate model and rationalizes a field experiment carried out in 250 villages in Afghanistan. In district elections, the village is split geographically into two districts. Each citizen votes for a candidate in his/her district and the final policy emerges from bargaining between the two elected officials. In at-large elections, each citizen has two votes to cast for any two candidates. Their model shows that, in district elections, citizens prefer to sacrifice competence in order to have more polarized politicians bargaining in their district's favor, whereas in at-large elections they vote for good politicians. The results from their field experiment support these theoretical predictions.

Mattozzi and Merlo (2015) introduce an equilibrium model of political recruitment under two electoral regimes. Candidates differ in their valence, and exert effort in party service in order to obtain their party's nomination. Parties face a trade off between electoral and organizational concerns. If only high valence candidates are nominated, a party has a greater chance of winning the election, but little effort (e.g., rent-seeking for the party) will be exerted by lower quality candidates, who become discouraged. To balance these two effects, parties prefer to recruit medium-quality politicians. Such lackluster choices are less likely to occur in majoritarian systems, wherein electoral competition is more vigorous, and thus the electoral returns to the party of selecting a good candidate are higher.

[3] On the mechanisms explaining political selection, also see Kotakorpi and Poutvaara (2011), Mattozzi and Merlo (2008, 2015), Caselli and Morelli (2004), and Gagliarducci and Nannicini (2013).

In our paper, we tackle the same issue—namely, the impact of electoral rules on political selection—but in a different setup, which builds on Galasso and Nannicini (2011, 2015). Galasso and Nannicini (2011) study the selection of politicians in a majoritarian system, and use Italian data from 1994 to 2006, to show that party loyalists are sent to safe districts, while high-quality politicians compete in contestable districts. Galasso and Nannicini (2015) examine the allocation of candidates onto party lists in a closed-list proportional system, and exploit data from the 2013 election to show that party loyalists are overrepresented in the safe positions at the top of the list. Our paper differs from these contributions in two ways. First, we extend Galasso and Nannicini's theoretical model to give politicians in majoritarian systems the double role of providing constituency service (as in Galasso and Nannicini 2011) and of participating in designing national policy. Analogously, we depart from Galasso and Nannicini (2015) to study the political selection process in proportional systems, in which politicians affect only national policy outcomes. Second, we introduce a comparison of the effectiveness of the political selection process across electoral systems, based on the share of competitive districts in the majoritarian system.

Before moving to our theoretical model, we discuss some motivating evidence. Empirical findings on the different patterns of political selection under majoritarian versus proportional elections are scarce at best. Cross-country comparisons are not very informative for a number of reasons (for a discussion, see Acemoglu 2005). The Italian mixed electoral system in place between 1994 and 2006, allows us to compare politicians elected in different electoral tiers, but in the same country.[4] We now provide some stylized facts on the selection of politicians in the majoritarian versus the proportional tier of the Italian system.

The rules for the election of the Italian Parliament have changed frequently over time. During three legislative terms (1994–1996, 1996–2001, 2001–2006), members of parliament were elected with a two-tier system, 75% majoritarian and 25% proportional. In the House of Representatives, composed of 630 members, voters received two ballots on election day: one to cast a vote for a candidate in their single-member district, and another to cast a vote for a party list in their larger proportional district. Overall, 75% of House members were elected with plurality voting in 475 single-member districts, while 25% were elected using proportional representation with closed party lists in 26 multiple-member districts (2–12 seats per district). In the Senate, composed of 315 members, voters received one ballot to cast their vote for a candidate in a single-member district, and the best losers in the 232 majoritarian districts were assigned to the remaining 83 seats according to the proportional rule. Hence, for our analysis we drop senators elected in the proportional tier. In the House, instead, the two tiers of the mixed system represented separate playing fields, wherein politicians made different electoral promises and were then called to answer for them.

We compare the characteristics of politicians elected in the majoritarian tier with those of politicians elected in the proportional tier. We focus on four measures of ex ante quality of the members of parliament: (1) whether they have a college degree or not, (2) whether they have local government experience or not, (3) their market incomes before being

[4] The drawback of this within-country analysis is that, while the theoretical model introduced in Sect. 3 examines and compares two separate electoral systems, this empirical evidence refers instead to a mixed system election, in which voters actually face elements of both systems.

elected, and (4) their incomes after controlling for individual characteristics.[5] The rationale for each measure is simple. College degree captures the acquisition of formal human capital and skills. Preelection income is a measure of market success and ability, especially after conditioning on demographic characteristics and job types, so that we can measure the ideosyncratic market ability of an individual when compared with her peers—those of the same sex, same age group, same type of job. The use of administrative experience is linked to the idea that lower-level elections can be used by high-quality politicians to build reputations and by voters to screen better candidates.

In Fig. 1, we report the running-mean smoothing of the above individual characteristics as a function of the contestability of (single-member) districts in the majoritarian tier. The degree of political contestability of a single electoral district is equal to one minus the margin of victory in the previous political election. For the proportional tier, in line with the theoretical model introduced in Sect. 3, we report the average characteristics of the politicians for the entire country, as shown by the horizontal line. For all measures, the average quality of the politicians in the proportional tier is higher than that of politicians elected in safe majoritarian districts, while the relationship is reversed in the case of (high-quality) politicians in majoritarian competitive districts.

Indeed, focusing on the differences that are statistically significant at standard levels, 69% of the politicians in the proportional tier had a college degree, against 74% of those elected in majoritarian contestable districts; 67% of politicians elected in majoritarian safe districts, where contestability is captured by a lagged margin of victory less than 10%. Preelection income was around 88,000 euros for proportional politicians, against 99,000 for majoritarian politicians in contestable districts and 72,000 for majoritarian politicians in safe districts.

Figure 1 provides valuable information on the allocation decision into the different majoritarian districts, but does not allow one to appreciate the overall selection decision by political parties. To focus on selection only, we turn to macro-districts at the regional level (common to House and Senate members) as units of observations. Within each macro (i.e., regional) district, we can calculate the share of contestable districts, defined as those where the lagged margin of victory was below 10%,[6] and obtain an overall indicator of the contestability of the majoritarian environment. In Fig. 2, we report the same running-mean smoothing exercise at this aggregate level. For three out of our four quality measures, the proportional system dominates the majoritarian system when the share of contestable districts is either small or large; the opposite happens for intermediate levels.[7] Although these non-monotonic relation clearly emerge from the figure, the small sample size prevents us to precisely test them.

Overall, this evidence suggests that, in order to compare majoritarian and proportional system, we need to take into account the pre-existing political environment, such as the distribution of majoritarian districts by their degree of contestability. Motivated by the above stylized facts, in the next section we propose a model of political recruitment under majoritarian versus proportional elections.

[5] Specifically, we regress preelection income on sex, age, education, and job dummies, and use the OLS residuals as our fourth quality measure.

[6] No such measure is calculated for the proportional tier, which is defined at national level, as in the model at Sect. 3.

[7] Only administrative experience always is higher for majoritarian politicians, owing to the fact that the small geographical magnitude of majoritarian districts favors local candidates in both safe and contestable districts.

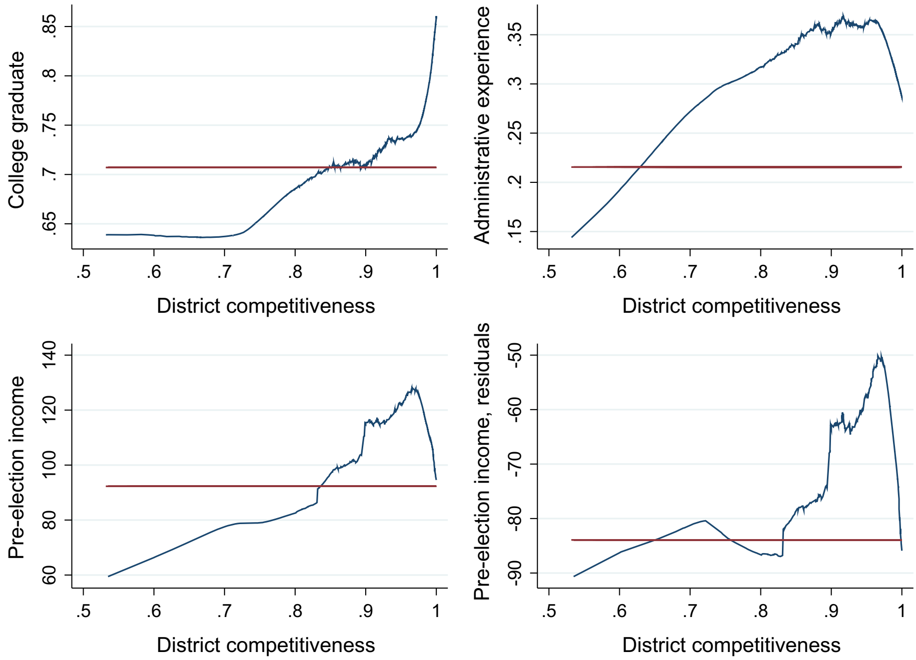

Fig. 1 Quality of politicians based on district competitiveness. Italian mixed-member Parliament; terms XII, XIII, and XIV; ministers excluded. Running-mean smoothing of the characteristics of *majoritarian* members of Parliament as a function of the competitiveness of the (single-member) district where they have been elected. District competitiveness is measured as one minus the lagged margin of victory of the past incumbent. The *horizontal line* represents the average characteristics of *proportional* members of Parliament

3 The model

Our model is populated by three types of players: voters, candidates and political parties. Two parties contest elections. Before the election, each party has to select its candidates. In the majoritarian system, parties also must allocate their candidates into each district. Candidates differ in quality or valence. Voters prefer high- to low-valence politicians. We characterize the low-quality politicians as party loyalists, and the high-quality types as experts.[8] Candidates are selected from a large pool, so that parties are assumed not to be supply-constrained, for instance, in being able to recruit experts. Parties seek to win the elections, and to have their party loyalists elected. Voters can be of three types: core supporters of either party or independent, that is, not aligned to any party. In proportional systems, independent voters care about the average quality of the politicians on the party list. In the majoritarian system, they care about the average quality as well as about the valence of the representative of their own district. We embed voting decisions into a standard probabilistic voting model (Lindbeck and Weibull 1987), so that, besides the quality of the politicians, independent voters care about a popularity shock to the two parties, and also have idiosyncratic ideological attitudes towards the two parties.

[8] This distinction is meant to capture the idea that high-valence politicians—the experts—have better outside options than the low quality ones, and therefore can exert effort to be more independent in their policy decisions, and less loyal to their party positions.

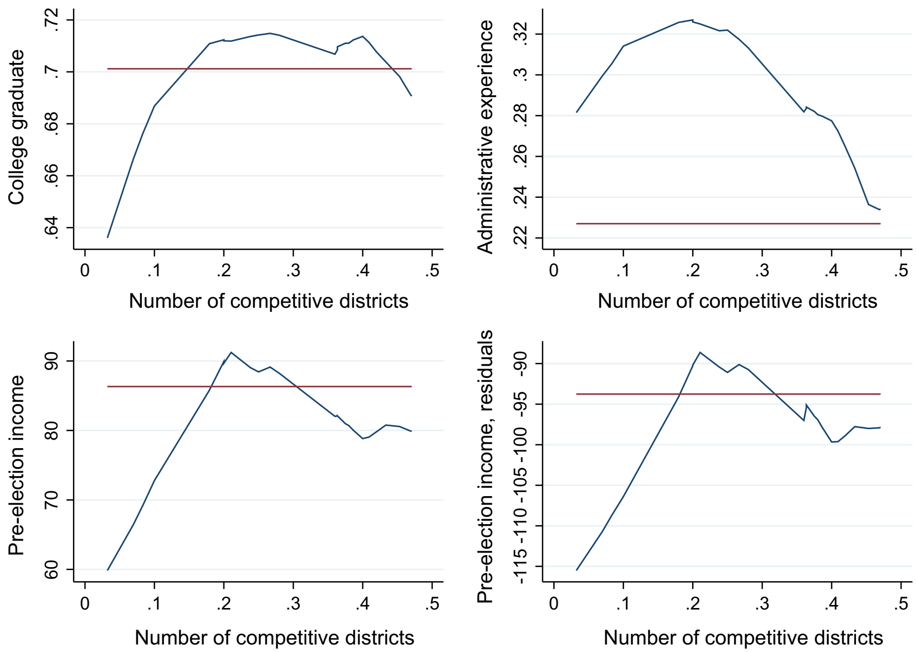

Fig. 2 Quality of politicians based on the share of competitive districts. Italian mixed-member Parliament; terms XII, XIII, and XIV; ministers excluded. Running-mean smoothing of the characteristics of *majoritarian* members of Parliament as a function of the share of competitive (single-member) districts in the region of election. Competitive districts are defined as those where the lagged margin of victory of the past incumbent was below 10%. The *horizontal line* represents the average characteristics of *proportional* members of Parliament

Our model thus introduces two lines of conflict: between the two parties on winning the elections, and among parties and independent voters on politicians' qualities. Parties value loyalists; independent voters prefer experts.

3.1 Parties and candidates

We consider two parties, D and R, which differ in their ideologies, and thus in their core supporters. The two parties compete against one another in democratic elections. The main role of the party (leaders) is to select the candidates to be included on the party list. Those choices determine the average qualities of the party lists. In the majoritarian system, the party also allocates candidates into the different electoral districts.

Candidates can be of three types: party-D loyalists (D), party-R loyalists (R) or experts (E). Loyal candidates provide rent-seeking activities only for their own party, and are of no value to independent voters. Regardless of their party affiliation, experts instead act in favor of the general interest. Moreover, in the majoritarian systems, experts also have higher valence than loyalists in providing constituency services to their local districts.

Each party chooses the share of experts and of party loyalists to include on the electoral list, respectively μ and $1-\mu$, and how to allocate them to the single-member districts of the majoritarian system. The utility to each party $j = D, R$ associated with party i winning the election can be written as a function of party $i = D, R$ share of experts. In particular, for $j = D, R$ and $i = D, R$, we have:

$$V_j(\mu_i) = 1 - \mu_i \quad \text{for} \quad i = j \tag{1}$$

$$V_j(\mu_i) = \mu_i - 1 \quad \text{for} \quad i \neq j \tag{2}$$

with $\mu_i \in [0, 1)$. In the former case, the party wins the election and enjoys a utility that depends on the share of loyalists on the party list; while in the latter case, the party loses the elections and its utility depends positively on the share of the other (winning) party's experts.[9]

3.2 Voters

We consider three groups of voters. Voters in groups D and R are core supporters and, hence, always vote for party D or R. Independent voters (I) care instead about the average quality of the politicians, and, in the majoritarian system, also about the quality of the candidates running in their electoral district.

In a majoritarian system, political candidates play a double role for the voters, at the national and at the local level. The preferences of independent voters living in district k for party i thus are summarized by the following utility function:

$$v_I\left(\mu_i, y_i^k\right) = (1 - \rho)\mu_i + \rho\overline{V}\left(y_i^k\right) \tag{3}$$

with $i = \{D, R\}$, where $\overline{V}\left(y_i^k\right)$ is the utility associated with the quality of party-i's candidate in electoral district k, and ρ measures the relative importance to the voters of local versus national issues. Notice that $y_i^k = \{L, E\}$, respectively, for party loyalists and experts, with $\overline{V}(E) > \overline{V}(L)$, so that expert candidates provide higher utility at the local level.

In a proportional system, no single district representative is elected, and thus the value of the local politician plays no role (i.e., $\rho = 0$). The independent voter's utility from party i winning the election depends only on its share of expert candidates:

$$V_I(\mu_i) = \mu_i \text{ for } i = D, R. \tag{4}$$

Besides the value attributed to the valance of the politicians, independent voters may feel ideologically closer to one party or another. The ideological characteristic of each independent voter is indexed by s, with $s > 0$ if the voter is closer to party R, and vice versa. The distribution of ideology among independent voters is assumed to be uniform; in particular, $s \sim U[-1/2, 1/2]$. The independent voter's decision also is affected by a common popularity shock δ to the parties that occurs before the election and may modify the perception that all independent voters have about the images of the two parties. In particular, if $\delta > 0$, party R gains popularity from this pre-electoral image shock, and vice versa for $\delta < 0$. Again, it is customary in this class of probabilistic voting models to assume that δ is uniformly distributed, so that $\delta \sim U\left[-\frac{1}{2\psi}, \frac{1}{2\psi}\right]$. And it is convenient to assume that $\psi > 1/4$.

To summarize, an independent voter will support party D if the utility obtained from the quality of party D's politicians, which depends on μ_D, is larger than the sum of the ideological idiosyncratic component, s , of the common shock, δ, and of the utility obtained

[9] By assuming the share of experts to be less than one, we ensure that winning always provides a higher utility than losing the elections.

from party R. That is, an independent voter prefers party D if $v_I(\mu_D, y_D^k) - v_I(\mu_R, y_R^k) - s - \delta > 0$, in a majoritarian system, and if $V_I(\mu_D) - V_I(\mu_R) - s - \delta > 0$, in a proportional system.

3.3 Selection in a proportional system

The incentives for a party to select expert candidates depend on the behavior of the independent voters. In fact, while each party (leader) prefers to select loyal candidates, experts are more valuable in convincing independent voters, and thus in winning the electoral majorities in the parliament.

As in a standard probabilistic voting model, before the election, parties independently and simultaneously make their moves, knowing the distribution of the popularity shock that takes place before the election, but not its realization. In particular, they select the share of loyal and expert candidates. After the popularity shock has occurred, independent voters decide which of the two parties to support, while loyalist voters always support their own party.

To understand party decisions, consider a party's probability of winning the election. Assume that the number of loyalist voters for each party is the same, so that winning the election depends entirely on the independent voters. Call $\widetilde{s}$ the ideology of the swing voter, that is, of the independent voter who is indifferent between party D or R. Hence, $\widetilde{s} = V_I(\mu_D) - V_I(\mu_R) - \delta$. All independent voters with ideology $s < \widetilde{s}$ will support party D, and vice versa for party R. To win the election, the sum of votes from the party D loyalists and of the votes that party D obtains from the independent voters has to exceed 50%. It is easy to see that the probability of party D winning the election (Π_D) can be expressed as a function of the popularity shock, δ. Since the popularity shock is uniformly distributed with density ψ, we have:

$$\Pi_D = \Pr\{\delta < V_I(\mu_D) - V_I(\mu_R)\} = \frac{1}{2} + \psi(\mu_D - \mu_R). \tag{5}$$

Hence, since independent voters value expert candidates, a trade off arises for the party between getting loyalists elected, and winning a majority of parliamentary seats. That trade off emerges clearly from each party's optimization problem. Consider party D. It will choose the share of experts, μ_D, in order to maximize the following expected utility:

$$\Pi_D V_D(\mu_D) + (1 - \Pi_D) V_D(\mu_R) \tag{6}$$

where Π_D is defined at Eq. 5, and $V_D(\mu_D)$ and $V_D(\mu_R)$ at Eqs. 1 and .

It is easy to see that, for both parties, this optimization problem yields the following solution:

$$\mu_D^P = \mu_R^P = 1 - \frac{1}{4\psi}. \tag{7}$$

The share of experts in the proportional system thus depends positively on the density of the common shock. In other words, when the random component (i.e., the scandal, δ) is less likely to determine the outcome of the national election, party leaders are more willing to invest in (costly) experts in order to increase their probability of winning an electoral majority in the parliament.

3.4 Selection and allocation in a majoritarian system

In a majoritarian system, the incentives for the party to select their candidates, and to allocate them to the different electoral districts, likewise depend on the behavior of the independent voters. Again, each party (leader) has a preference for getting loyalists elected. But to win the election, it needs to convince the independent voters.

The party optimization problem still is to maximize the expected utility at Eq. 6 (for party D), with the utilities defined at Eq. 3. However, the party's probability of obtaining a winning electoral majorities in the parliament now depends both on their selection and allocation of experts.

This allocation of experts is best understood in a majoritarian system with single-member electoral districts. The degree of competitiveness of each electoral district will depend on the distribution of the three groups of voters (R-supporters, D-supporters, and independents) across districts. We defined the degree of ex-ante contestability of a district k as

$$\lambda_k = \frac{1}{2}\frac{\lambda_k^R - \lambda_k^D}{\lambda^I} \tag{8}$$

where λ_k^j is the share of type-j voters, with $j \in \{D, I, R\}$, in district k; the share of independent voters is assumed to be constant across districts, $\lambda_k^I = \lambda^I \ \forall k$.

Maximum electoral contestability, $\lambda_k = 0$, is obtained in district k when the share of R and D core supporters is the same, $\lambda_k^R = \lambda_k^D$. Increases in the absolute value of λ_k indicate lower district contestability. In particular, districts such that $\lambda_k < -1/2$ or $\lambda_k > 1/2$ are safe, since, respectively, party D or R wins for sure. Hence, only intermediate districts with $\lambda_k \in [-1/2, 1/2]$ are contestable. We consider a continuum of districts, distributed uniformly according to $\lambda_k \sim U\left[-\frac{1-\lambda^I}{2\lambda^I}, \frac{1-\lambda^I}{2\lambda^I}\right]$, with a cumulative distribution $G(\lambda_k)$.

What is the probability that a party—say party D—wins a contestable district k? Party D will obtain the votes of its core voters (λ_k^D), and of the independent voters with ideologies $s < \tilde{s}$, where $\tilde{s}$ is the ideology of the (independent) swing voter: $\tilde{s} = v_I(\mu_D, y_D^k) - v_I(\mu_R, y_R^k) - \delta$. Since the share of votes from the independents is $\lambda^I(\tilde{s} + 1/2)$, party D obtains more than 50% of the votes and, hence, wins district k, if $\tilde{s} > \lambda_k$, which occurs with probability:

$$\Pi_D^k = \Pr\{\delta < v_I(\mu_D, y_D^k) - v_I(\mu_R, y_R^k) - \lambda_k = d_k\} = \frac{1}{2} + \psi d_k \tag{9}$$

where d_k can be interpreted as a measure of the ex-post contestability (i.e., after parties' decisions) of district k. When the two parties have the same selection and allocation of candidates, we have $d_k = -\lambda_k$. However, parties will act to modify d_k, and thus to increase their chances of winning district k. Parties have two instruments for affecting their winning probabilities in district k. They can modify the relative shares of experts and loyalists, μ_i, and they can choose which candidate to allocate to each district k. Hence, the selection decision affects national policy issues, while the allocation affects the local ones.[10]

It is convenient to analyze these two decisions separately by considering, first, how parties allocate given shares of experts to electoral districts, and then how many experts are selected.

[10] The theoretical framework in Galasso and Nannicini (2011) considers only the latter political instrument.

3.4.1 Allocation of experts

For given shares of experts for the two parties (μ_D, μ_R), the two national policies are determined, and parties can concentrate on allocating their experts into districts in order to increase their probabilities of winning an electoral majority in the parliament. In fact, the difference in utility provided to the independent voters in district k by the two parties can be written as

$$v_I(\mu_D, y_D^k) - v_I(\mu_R, y_R^k) = (1-\rho)(\mu_D - \mu_R) + \rho(\overline{V}(y_D^k) - \overline{V}(y_R^k))$$

where the former term on the right-hand side depends on the average quality of the politicians obtained through the selection of candidates, while the latter term is determined by the parties' allocations of candidates to district k. Experts are more valuable than loyalists to independent voters. Having an expert rather than a party loyalist in the electoral district increases independent voters' utility by $W = \rho[\overline{V}(E) - \overline{V}(L)]$. Hence, parties will compete on allocating good politicians (the experts) to win the contestable districts.

What are the parties' strategic behaviors in this simultaneous allocation game?[11] Let us begin with only loyal candidates being allocated by both parties to the contestable districts, so that $\overline{V}(y_D^k) - \overline{V}(y_R^k) = \overline{V}(L) - \overline{V}(L) = 0$ for all districts $\lambda_k \in [-1/2, 1/2]$. One party's probability of winning any of these districts will depend on the national policies (μ_D, μ_R), and on the district characteristics, λ_k. For instance, party D will win district k after a shock $\delta < d_k = (1-\rho)(\mu_D - \mu_R) - \lambda_k$. Hence, given the distribution of districts (λ_k), if both parties have selected the same share of experts, $\mu_D = \mu_R$, party D wins more than 50% of the districts (those with $\lambda_k < 0$), and thereby obtains the parliamentary majority, only if the shock is strictly in its favor, i.e., $\delta < d_0 = 0$. If instead party D has selected more experts, $\mu_D - \mu_R = z > 0$, party D wins the elections, even if the shock is mildly against it, i.e., for $\delta < (1-\rho)z$, again winning all districts with $\lambda_k < 0$, as shown in Fig. 3 (and vice versa for party R). This result suggests that the pivotal districts for winning an electoral majority in the parliament are in a small interval around $\lambda_k = 0$. We will refer to a small district interval around $\lambda_k = 0 = \lambda_0$ as $[\lambda_\varepsilon, \lambda_\Xi]$ with $\lambda_0 - \lambda_\varepsilon = \lambda_\Xi - \lambda_0 = \varepsilon$ small enough.

Consider party D sending experts to the district interval $[\lambda_0, \lambda_\Xi]$. This increases party D's probability of winning those districts, and thus the national elections. In particular, a party D expert in district λ_0, matched by a party R loyalist, allows party D to win this district even for a less favorable realization of the shock, namely, for $\delta < W + (1-\rho)z$. This event occurs with the same probability that party D has of winning district $\lambda_w = -W$, which is ex-ante biased in its favor, when both parties send a loyalist in λ_w. Hence, by aligning experts in districts $[\lambda_0, \lambda_\Xi]$, matched by party R loyalists, for $\delta = (1-\rho)z$, party D would win the national election, rather than just tying it. The same reasoning applies to party R. By sending an expert to the most contestable district, λ_0, matched by a party D loyalist, party R has a probability of winning district λ_0 equal to the probability of winning district $\lambda_W = W$, when both parties allocate loyalists.

Districts λ_w and λ_W define the range of contestable districts to which party D and R will consider allocating their experts. In fact, if party R allocates only loyalists in $[\lambda_w, \lambda_W]$, party D's best response would be to place its experts in $[\lambda_w, \lambda_\Xi]$ in order to win the election if

[11] This simultaneous allocation decision by the two parties resemble the Colonel Blotto game, in which two colonels fight a war over a number of battlefields, and have to decide how to allocate their troops to the different battlefields. Our allocation game has two peculiar features. First, it entails a binary allocation choice, since either a high- or low-valence candidate is allocated to a district. Second, our battlefields—i.e., the electoral districts—vary ex-ante in their political contestability, as measured by λ_k.

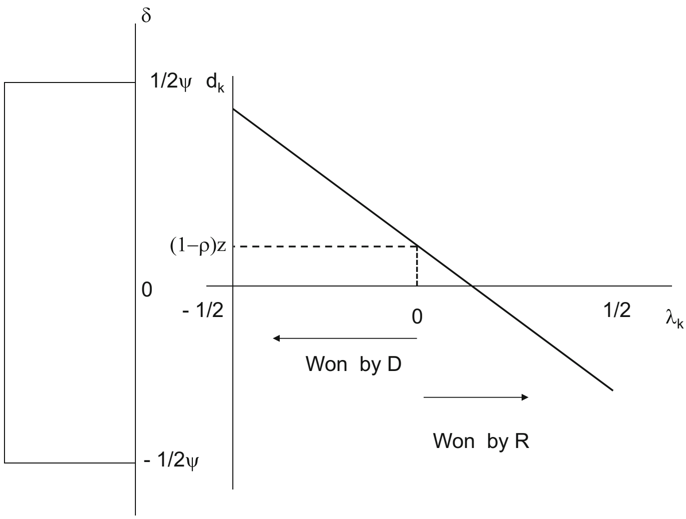

Fig. 3 Allocation of experts in the majoritarian system

$\delta \leq d_w = \lambda_w + (1-\rho)z$. It is important to emphasize that party D could not increase its probability of winning an electoral majority in the parliament by placing additional experts in any district. We identify with $\eta/2$ the mass of districts between λ_w and λ_0, i.e., $\eta/2 = G(\lambda_0) - G(\lambda_w)$. Hence, party D would need $\eta/2$ experts to span the districts $[\lambda_w, \lambda_0]$. Symmetrically, for party R, we have $\eta/2 = G(\lambda_W) - G(\lambda_0)$.

Since the distribution of districts is assumed to be uniform, we have $\eta = G(\lambda_W) - G(\lambda_w) = \frac{2\lambda^I}{1-\lambda^I} W$. The share of experts needed to cover all of the contestable districts between λ_w and λ_W thus depends positively on the mass of independent voters, λ^I, and on the intrinsic value of an expert to the independent voters, W.

The characterization of the probabilities of winning an electoral majority in the parliament which correspond to the equilibrium allocation for given party selections (μ_D, μ_R) is provided at Proposition 3 in the Appendix. This proposition generalizes the result in Galasso and Nannicini (2011) to an environment in which parties choose the shares of experts in order to affect the voters' utility, according to the first term in Eq. 3. The intuition of this proposition is the following. When both parties select a sufficiently large share of experts to cover the most competitive districts that are biased in their favour, $\mu_i > \eta/2$ for $i = D, R$, a difference in the probabilities of winning an electoral majority in the parliament may emerge only from a different average quality in the party lists. In particular, if a party—say party D—has more experts than the other, it will provide additional utility, equal to $(1-\rho)(\mu_D - \mu_R)$, to all independent voters, and this will increase its probability of winning an electoral majority in the parliament. In all other cases, the probability of winning the an electoral majority will depend on the average quality of candidates on the party list, as well as on the different allocation strategies that may emerge. In particular, the party enjoying an advantage in the share of experts typically will adopt an "offensive" strategy, by allocating its experts to the contestable districts, which ex ante favor its opponent; and the party with fewer experts will respond with an equally offensive strategy. The resulting winning probabilities are reported in the Proposition 3 in the Appendix.

3.4.2 Selection of experts

Before deciding where to allocate their candidates, parties have to choose how many experts to select. This selection process entails a clear trade off. Having more experts reduces the party's (leaders') utility in the case of electoral victory. However, experts are valuable in attracting the votes of the independents. Thus, a larger share of experts increases the probability of winning an electoral majority in the parliament, as described at Proposition 3 in the Appendix.

The next proposition characterizes the equilibrium selection of experts by the two parties.

Proposition 1 *In a majoritarian system, there exist two values of the proportion of independent voters,* $0 < \lambda_1^I < \lambda_2^I < 1$, *such that the share of experts chosen by both parties is*

$$\mu_D^M = \mu_R^M = \begin{cases} 1 - \dfrac{1}{4\psi(1-\rho)} & \text{for} \quad \lambda^{\mathrm{I}} \leq \lambda_1^{\mathrm{I}} \\ 1 - \dfrac{1}{4\psi\left(1 - \rho + \dfrac{1-\lambda I}{\lambda^I}\right)} & \text{for} \quad \lambda^{\mathrm{I}} \geq \lambda_2^{\mathrm{I}} \end{cases}$$

Proof See Appendix. □

In the former case, $\lambda^I \leq \lambda_1^I$, the proportion of independent voters is small—and hence only few districts are highly contestable, i.e., $\eta/2$ also is small. Both parties therefore will be willing to select enough experts to span their crucial competitive districts, respectively $[\lambda_w, \lambda_\Xi]$ for party D and $[\lambda_\varepsilon, \lambda_W]$ for party R (see Fig. 5 in the appendix). In the latter case, $\lambda^I \geq \lambda_2^I$, the existence of a large proportion of independent voters makes many districts highly contestable ($\eta/2$ is large). Party leaders thus would find it costly to fill all of their crucial competitive districts with experts, since only a few loyalists would find place on the party list. Although in equilibrium the share of experts will be greater than in the former case, parties will not select enough experts to cover the district interval $[\lambda_w, \lambda_\Xi]$ for party D and $[\lambda_\varepsilon, \lambda_W]$ for party R—and the allocation strategy will follow case III in Proposition 3 in the appendix (see also Fig. 7). Moreover, in this case, an increase in the share of independent voters, λ^I, and thus of the highly competitive districts, $\eta/2$, reduces the share of experts selected in equilibrium by both parties. That is because the marginal impact of selecting and allocating an expert to an additional competitive district on the probability of winning an electoral majority in the parliament is declining in the share of competitive districts, while the cost to the parties—in terms of fewer loyalists being elected—remains constant.[12]

3.5 Selection in proportional versus majoritarian systems

The selection of loyalist and expert candidates by the parties gives rise to a clear trade off: experts enhance the party's probability of winning an electoral majority, but at the cost of

[12] Suppose that party D is deciding whether to select and allocate one more expert, given an initial situation in which $\mu_D = \mu_R < \eta/2$. From case IV of Proposition 3, the marginal increase in Party D's probability of winning the election is equal to $1 - \rho + \frac{1-\lambda^I}{\lambda^I}$, which clearly is decreasing in λ^I.

reducing the share of elected loyalists. Yet, this trade off differs across electoral systems. In a proportional system, it depends entirely on the share of experts. They appeal to independent voters and, hence, increase their party's probability of winning an electoral majority in the parliament, but the share of elected loyalists is reduced. The incentives to select expert candidates in a majoritarian system are different. Besides the relevance of the share of experts at the national level, their allocation to the different electoral districts also affects the parties' winning probabilities.

We are now in a position to compare the selection of political candidates, as measured by the share of experts, in these two alternative electoral systems. The next proposition summarizes our results, which are also displayed in Fig. 4.

Proposition 2 *There exists a threshold value of independent voters,* $\lambda_3^I = 1/(1+\rho)$, *such that*

(I) *if* $\lambda^I \leq \lambda_1^I$ *or if* $\lambda^I > \lambda_3^I$, *more experts are selected under a proportional than under a majoritarian system*: $\mu_i^P > \mu_i^M$ *with* $i = D, R$; *and*
(II) *if* $\lambda_2^I < \lambda_3^I$, *for* $\lambda^I \in (\lambda_2^I, \lambda_3^I)$, *more experts are selected under a majoritarian than under a proportional system,* $\mu_i^P < \mu_i^M$ *with* $i = D, R$.

Proof See Appendix. □

For a small share of independent voters, λ^I, and thus of contestable districts, $\eta/2$, the majoritarian system yields weak political competition. Most districts are indeed safe, and the party leaders need not pay the cost of allocating experts there. A proportional system is thus a better alternative for selecting good politicians. If the proportion of independent voters and, hence, of contestable districts, exceeds a certain threshold, $\lambda^I \in (\lambda_2^I, \lambda_3^I)$, the degree of political competition in the majoritarian system is strong, and this becomes the better electoral system for selecting experts.[13] However, as the share of contestable districts continues to increase and reaches a certain threshold, $\lambda^I > \lambda_3^I$, the level of political competition in the majoritarian system becomes "too" vigorous. Parties have no incentive to continue to select expert politicians since an additional expert has little impact on the probability of winning an electoral majority in the parliament, but has a cost in terms of reducing the number of elected party loyalists. In this region, proportional systems perform better in selecting politicians than majoritarian systems, despite the latter having many highly competitive districts.

4 Conclusion

This paper models how electoral rules may influence the selection of candidates for political office. As recognized in the literature, proportional systems provide broad, nationwide incentives, while majoritarian systems also entail a local, district-level component. Several studies have shown that this difference leads to the adoption of different public policies under alternative electoral rules. We suggest that a similar difference may emerge in political selection. In majoritarian systems, the relevance of the local dimension induces

[13] Notice that for this region to exist, the value to the independent voters of having an expert assigned to their district—rather than a loyalist—has to be large. In fact, we have $\lambda_2^I < \lambda^I$, if $\frac{W}{\rho} > 1 - \frac{1}{4\psi}$.

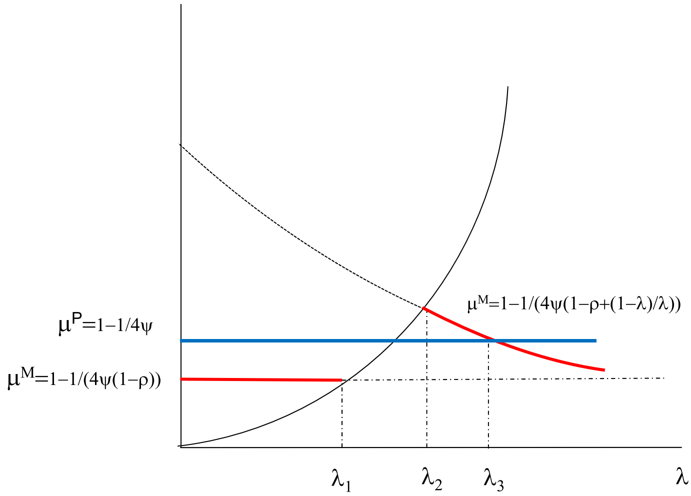

Fig. 4 Selection in proportional versus majoritarian systems

political parties to allocate high-quality candidates to competitive districts. This allocation mechanism affects the party's selection decision. In proportional systems, the local component plays no role, and thus parties simply choose the overall share of high- versus low-quality politicians in order to attract swing voters at the national level. As a result, the comparison between the two systems hinges on the share of competitive (majoritarian) districts. For either small or large shares of competitive districts, the proportional system provides strong incentives for parties to select good politicians; for intermediate levels, the majoritarian systems is instead more effective. These results are in line with our suggestive empirical evidence on Italian mixed-member elections.

Appendix

Proposition 3 *In a majoritarian system, the winning probabilities* (Π_i, Π_j) *corresponding to the equilibrium allocations for given party selections* (μ_i, μ_j) *with* $i = D, R$ *and* $j = R, D$ *are*

(I) *For* $\mu_i > \eta/2$ *and* $\mu_j > \eta/2$, $\Pi_i = 1/2 + \psi(1-\rho)(\mu_i - \mu_j)$ *and* $\Pi_j = 1 - \Pi_i$; *while for* $\mu_i > \eta/2$ *and* $\mu_j = \eta/2$, $\Pi_i = 1/2 + \psi\left[(1-\rho)(\mu_i - \mu_j) + W/2\right]$ *and* $\Pi_j = 1 - \Pi_i$;

(II) *For* $\mu_i > \eta/2 > \mu_j$, $\Pi_i = 1/2 + \psi\left(1 - \rho + \frac{1-\lambda^I}{2\lambda^I}\right)(\mu_i - \mu_j)$ *and* $\Pi_j = 1 - \Pi_i$;

(III) *For* $\mu_i = \mu_j \leq \eta/2$, $\Pi_i = \Pi_j = 1/2$

(IV) *For* $\mu_j < \mu_i \leq \eta/2$ *and* $\mu_i < \frac{1}{2}(\mu_j + \eta/2)$, $\Pi_i = 1/2 + \psi\left(1 - \rho + \frac{1-\lambda^I}{\lambda^I}\right)(\mu_i - \mu_j)$ *and* $\Pi_j = 1 - \Pi_i$;

(V) *For* $\mu_j < \mu_i \leq \eta/2$ *and* $\mu_i > \frac{1}{2}(\mu_j + \eta/2)$, $\Pi_i = 1/2 + \psi\left[(1-\rho)(\mu_i - \mu_j) + W/2 - \frac{1-\lambda^I}{2\lambda^I}\mu_j\right]$ *and* $\Pi_j = 1 - \Pi_i$;

Proof of Proposition 3 Define Λ^D, party-D's allocation of experts, as the union of the district intervals $\Lambda_i^D = \left[\lambda_I^i, \lambda_{II}^i\right]$, where party D allocates its experts, $\Lambda^D = \cup_i \Lambda_i^D$, and analogously Λ^R for party R. Define $z = \mu_D - \mu_R \in (-1, 1)$, as the difference in the share of experts between party D and R. Finally, define $H\left(\Lambda_i^D\right) = G\left(\lambda_{II}^i\right) - G\left(\lambda_I^i\right)$ as the mass of districts in the interval Λ_i^D.

Given their shares of experts, μ_D and μ_R, parties' objectives in allocating their experts is to maximize the probability of winning the election, i.e., of winning more than 50% of the districts. Consider party D. Its probability of winning a district k is $\delta < d_k = (1-\rho)(\mu_D - \mu_R) + \rho\left(V_I\left(y_D^k\right) - V_I\left(y_R^k\right)\right) - \lambda_k$. Thus, given μ_D and μ_R, party D allocates experts to districts in order to modify $V_I\left(y_D^k\right)$ in the marginal districts. These are the district(s) such that, given the shock, winning the district(s) increases the probability of winning the election.

Case (I) Both parties have enough experts to span the interval between λ_w and λ_0, i.e., $\mu_D > \eta/2$ and $\mu_R > \eta/2$. Consider an allocation Λ^D by party D that includes Λ_i^D s.t. $[\lambda_w, \lambda_\Xi] \subset \Lambda_i^D$. An allocation Λ^R by party R that includes Λ_i^R s.t. $[\lambda_\varepsilon, \lambda_W] \subset \Lambda_i^R$ is a best response to Λ^D. In fact, given Λ^D, by sending its experts to the interval $[\lambda_\varepsilon, \lambda_W]$, party R restores its probability of winning the election to $\frac{1}{2} + \psi(1-\rho)(\mu_R - \mu_D)$, so that only the (given) different shares of experts (μ_R, μ_D) matters. In particular, party R wins the election for $\delta > (1-\rho)(\mu_D - \mu_R)$ and party D for $\delta < (1-\rho)(\mu_D - \mu_R)$. Allocating additional experts may modify the share of seats won by party R, but not its probability of winning the election. The same reasoning can be used to show that Λ^D with Λ_i^D s.t. $[\lambda_w, \lambda_\Xi] \subset \Lambda_i^D$ is a best response to Λ^R with Λ_i^R s.t. $[\lambda_\varepsilon, \lambda_W] \subset \Lambda_i^R$. Hence, a pair of allocations Λ^D and Λ^R that includes (i) Λ_i^D s.t. $[\lambda_w, \lambda_\Xi] \subset \Lambda_i^D$ and $H\left(\Lambda^D\right) = \sum_i H\left(\Lambda_i^D\right) = \mu_D$, and (ii) Λ_i^R s.t. $[\lambda_\varepsilon, \lambda_W] \subset \Lambda_i^R$ and $H\left(\Lambda^R\right) = \sum_i H\left(\Lambda_i^R\right) = \mu_R$ is a Nash equilibrium of the allocation game. This allocation is displayed in Fig. 5.

To prove that any equilibrium allocation Λ^D must include Λ_i^D s.t. $[\lambda_w, \lambda_\Xi] \subset \Lambda_i^D$, consider first an allocation $\widehat{\Lambda}^D$ with $\widehat{\Lambda}_i^D = \left[\lambda_I^i, \lambda_{II}^i\right]$ s.t. $0 > \lambda_I^i > \lambda_w$ and $\lambda_{II}^i > \lambda_\Xi$, so that no other experts are in $[\lambda_w, \lambda_I]$. Party-R's best response is to allocate its experts in $[\lambda_w, \lambda_I] \cup [\lambda_0, \lambda_{II}]$. Following that strategy, party R wins the election with a probability greater than $\Pr\{\delta > (1-\rho)(\mu_D - \mu_R)\}$, since for $\delta = (1-\rho)(\mu_D - \mu_R)$ party R wins all districts with $\lambda > 0$ (and hence 50%), but also the districts in $[\lambda_w, \lambda_I]$. Hence, $\widehat{\Lambda}^D$ cannot be part of an equilibrium since simply matching the previous best response by party R would give party D a probability $\frac{1}{2} + \psi(1-\rho)(\mu_D - \mu_R)$ of winning the election. It remains to be shown that an equilibrium allocation Λ^D has to include the interval $[\lambda_\varepsilon, \lambda_\Xi]$. Consider $\widehat{\Lambda}^D = \widehat{\Lambda}_i^D \cup \widehat{\Lambda}_j^D$ with $\widehat{\Lambda}_i^D = \left[\lambda_I^i, \lambda_{II}^i\right] \in [\lambda_w, \lambda_\varepsilon]$ and $\widehat{\Lambda}_j^D = \left[\lambda_I^j, \lambda_{II}^j\right] \in [\lambda_\Xi, \lambda_W]$. Party-$R$'s best response would be Λ^R, such that $\Lambda_i^R = [\lambda_\varepsilon, \lambda_W]$, which yields party R a winning probability greater than $\frac{1}{2} + \psi(1-\rho)(\mu_R - \mu_D)$. Hence, $\widehat{\Lambda}^D$ cannot be part of an equilibrium. The same reasoning applies also when $\mu_R = \mu_D = \eta/2$, in which case parties will allocate experts respectively to $[\lambda_w, \lambda_0]$ for party D and to $[\lambda_0, \lambda_W]$ for party R. Notice also that if $\mu_R = \eta/2$ and $\mu_D > \eta/2$ (or vice versa), the party having more experts will win the election with probability $\Pi_D = \frac{1}{2} + \psi\left[(1-\rho)(\mu_D - \mu_R) + \frac{W}{2}\right]$. That is because party D wins the election for shocks such that $\delta < (1-\rho)(\mu_D - \mu_R)$, but it also ties the elections for $\delta \in [(1-\rho)(\mu_D - \mu_R), (1-\rho)(\mu_D - \mu_R) + W]$.

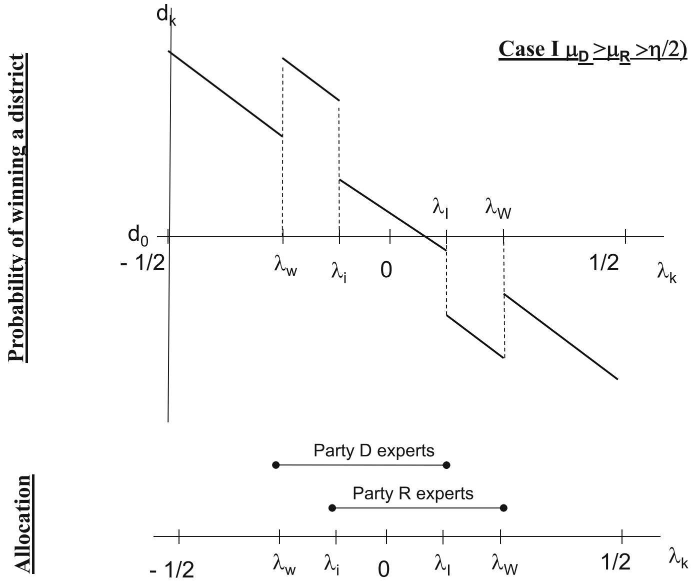

Fig. 5 Equilibrium allocation, case I

Case (II) One party (say party D) has enough experts to span the crucial interval, but the other does not, i.e., $\mu_D > \eta/2 > \mu_R$. Suppose that party D allocates its experts to $[\lambda_a, \lambda_W]$, as displayed in Fig. 6. Party-R's best response is to reduce as much as possible party-D's probability of winning the election, given that party-R does not have enough experts to match party-D's experts and re-establish its probability of winning the election to $\Pr\{\delta > (1-\rho)(\mu_D - \mu_R)\} = \frac{1}{2} + \psi(1-\rho)(\mu_R - \mu_D)$. Consider the largest (positive) realization of the shock, δ, that still allows party-D to win the election, given that party-D has allocated experts as described above, and party-R has not allocated any. Party-R will have to use its experts to target those districts that are marginally in favor of party-D, for this level of the shock. This can be done by sending experts to $\left[\lambda_j, \lambda_m\right]$ with $\lambda_j = \max\{\lambda_a, \lambda_w\}$, since it never pays to send experts outside the interval $[\lambda_w, \lambda_W]$, and λ_m s.t. $G(\lambda_m) - G\left(\lambda_j\right) = \mu_R$, so that party-$R$ has exhausted its experts. For party-R sending experts to $\left[\lambda_j, \lambda_m\right]$, party-$D$'s best response is to span the interval $\left[\lambda_j, \lambda_W\right]$. Hence this allocation constitutes an equilibrium.

To see why under this allocation party D wins the election with probability $\Pi_D = \frac{1}{2} + \psi\left(1 - \rho + \frac{1-\lambda^I}{2\lambda^I}\right)z$, consider Fig. 6. Party D wins the election when more than 50% of the districts support it; these districts are $\left[-\frac{1-\lambda^I}{2\lambda^I}, -\lambda_m + x\right] \cup [\lambda_m, \lambda_m + x]$, such that $\frac{\lambda^I}{1-\lambda^I}\left[-\lambda_m + x + \frac{1-\lambda^I}{2\lambda^I}\right] + \frac{\lambda^I}{1-\lambda^I}[\lambda_m + x - \lambda_m] = 1/2$. Hence, $x = \lambda_m/2$, where $\lambda_m = \frac{1-\lambda^I}{\lambda^I}z$, since $G(\lambda_W) - G(\lambda_a) = \mu_D = \mu_R + z$ and $G(\lambda_m) - G(\lambda_a) = \mu_R$. A simple inspection of Fig. 6 shows that all of these districts are won by party-D if $\delta < -x + \lambda_m + (1-\rho)(\mu_D - \mu_R) = \left(1 - \rho + \frac{1-\lambda^I}{2\lambda^I}\right)(\mu_D - \mu_R)$, that occurs with probability $\Pi_D = \frac{1}{2} + \psi\left(1 - \rho + \frac{1-\lambda^I}{2\lambda^I}\right)(\mu_D - \mu_R)$.

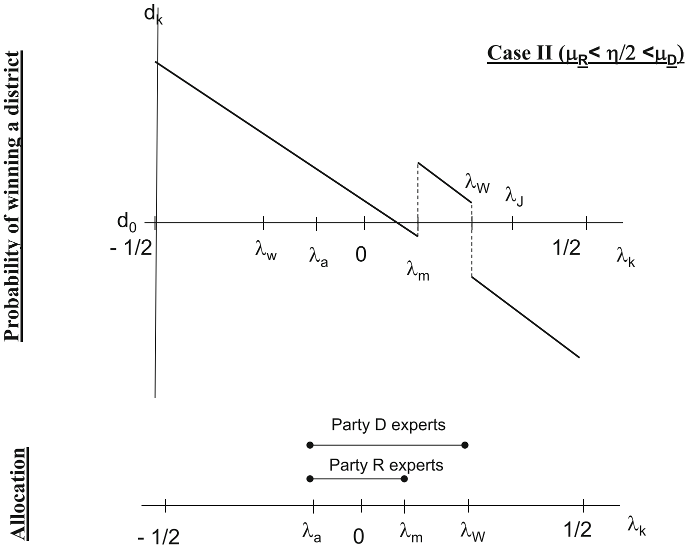

Fig. 6 Equilibrium allocation, case II

Finally, to see that no other equilibrium allocation is possible, consider party-*D* allocating experts to $\Lambda^D = [\lambda_w, \lambda_s]$, such that $G(\lambda_s) - G(\lambda_w) = \mu_D$. Party-*R* would have an incentive to allocate experts to $\Lambda^R = [\lambda_\varepsilon, \lambda_s]$, thereby winning the elections with a probability exceeding $\frac{1}{2} + \psi(1-\rho)(\mu_R - \mu_D)$. But with this allocation by party-*R*, party-*D*'s best response would be to allocate its experts to $[\lambda_a, \lambda_W]$.

Case (III) Parties have equal shares of experts, but are unable to span the crucial districts, $\mu < \eta/2$. Suppose that party *D* allocates its experts to $[\lambda_0, \lambda_B]$. Party-*R*'s best response is to send its experts to $[\lambda_b, \lambda_0]$, which re-establishes its probability of winning the election to 1 / 2. As displayed in Fig. 7, party *D* wins the election for $\delta < \max[-\lambda_B, -\lambda_b - W]$, party *R* for $\delta > \min[-\lambda_b, -\lambda_B + W]$, and the election is tied for $\delta \in [-\lambda_B, -\lambda_b]$. Party *R* cannot increase its probability of winning the election above 1 / 2 by allocating experts to other districts. Hence, party-*D*'s allocation in $[\lambda_0, \lambda_B]$ and party-*R*'s allocation in $[\lambda_b, \lambda_0]$ is an equilibrium, and each party has a 50% probability of winning the election. Notice that party-*R* could allocate experts in $[\lambda_0, \lambda_B]$ and still re-establish its winning probability to 50%. However, with party *R* experts in $[\lambda_0, \lambda_B]$, party *D*'s best response would be to send experts to $[\lambda_B, \lambda_W]$ to increase its winning probability above 50%. Hence, both parties allocating experts in $[\lambda_0, \lambda_B]$ is not an equilibrium.

To prove that no other equilibrium allocation exists, first notice that allocating experts outside the interval $[\lambda_w, \lambda_W]$ is never part of an equilibrium, since it does not modify the probability of winning the election, which can instead be achieved by allocating experts in this interval. Consider party-*D* allocation $\Lambda^D = [\lambda_b, \lambda_0]$. Party-*R*'s best response would be to allocate experts to $[\lambda_w, \lambda_b]$, which would yield party *R* a winning probability above 1 / 2, since for $\delta = 0$ party *R* would win in districts with $\lambda > 0$ and in $[\lambda_w, \lambda_b]$. The same reasoning applies to any $\widehat{\Lambda}^D = [\lambda_I, \lambda_{II}]$ s.t. $\lambda_I \in [\lambda_w, \lambda_0)$, $\lambda_{II} \in [\lambda_w, \lambda_b)$ and $G(\lambda_{II}) - G(\lambda_I) = \mu$. And to $\widehat{\Lambda}^D = [\lambda_I, \lambda_W]$ and $G(\lambda_W) - G(\lambda_I) = \mu$.

Case (IV) Parties are unable to span the crucial districts, and have marginally different shares of experts. Suppose that party *D*, which has few more experts than party *R* (i.e.,

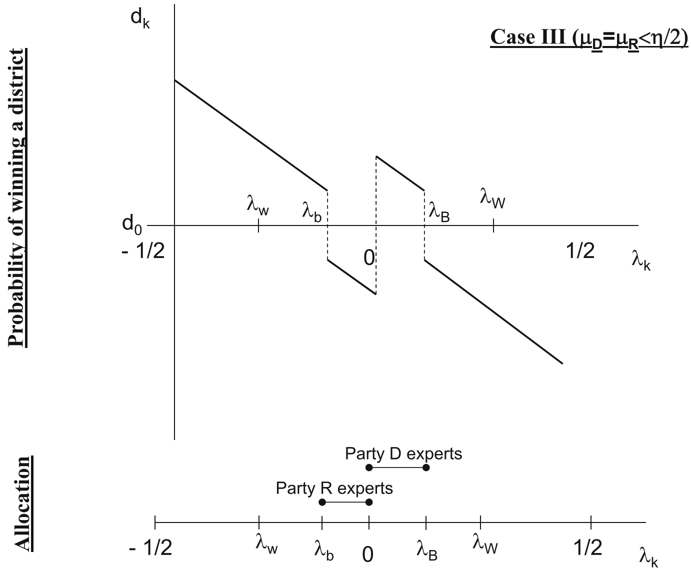

Fig. 7 Equilibrium allocation, case III

$z = \mu_D - \mu_R < (\frac{\eta}{2} - \mu_R)/2)$, allocates its experts to $[\lambda_0, \lambda_B]$. Party-*R*'s best response is to send experts to $\left[\lambda_w, \lambda_g\right]$, such that $G\left(\lambda_g\right) - G(\lambda_w) = \mu_R$ (or alternatively to the right of λ_0), as shown in Fig. 8. To see why, consider the largest (positive) realization of the shock, δ, that still allows party-*D* to win the election, given that party-*D* has allocated experts as described above, and party-*R* has not allocated any. Party-*R* will have to target with its experts those districts that are marginally in favor of party-*D*, for this level of the shock. This can be done by sending experts to $\left[\lambda_w, \lambda_g\right]$, such that $G\left(\lambda_g\right) - G(\lambda_w) = \mu_R$ (or alternatively to the right of λ_0). For this allocation by party *R*, party *D*'s best response is to allocate its experts to $[\lambda_0, \lambda_B]$ (or alternatively to $\left[\lambda_w, \lambda_g\right]$ and the remaining part to the right of λ_0). Hence, this allocation constitutes an equilibrium.

Under this allocation, party-*D* wins the election with probability $\Pi_D = \frac{1}{2} + \psi\left(1 - \rho + \frac{1-\lambda^l}{\lambda^l}\right)(\mu_D - \mu_R)$. Consider again Fig. 8; party-*D* wins the election when more than 50% of the districts vote in its favor; these districts are $\left[-\frac{1-\lambda^l}{2\lambda^l}, \lambda_w\right] \cup \left[\lambda_g, \lambda_w + x\right] \cup [\lambda_0, \lambda_B]$, such that $\frac{\lambda^l}{1-\lambda^l}\left[\lambda_w + \frac{1-\lambda^l}{2\lambda^l}\right] + \frac{\lambda^l}{1-\lambda^l}\left[\lambda_w + x - \lambda_g\right] + \frac{\lambda^l}{1-\lambda^l}[\lambda_B - \lambda_0] = 1/2$. Hence, $x = W - \lambda_B$, where $\lambda_B = \frac{1-\lambda^l}{\lambda^l}\mu_D$. A simple inspection of Fig. 8 shows that all of these districts are won by party-*D* if $\delta < d_x = -(x + \lambda_g) + (1-\rho)(\mu_D - \mu_R) = \left(1 - \rho + \frac{1-\lambda^l}{\lambda^l}\right)(\mu_D - \mu_R)$, that occurs with probability $\Pi_D = \frac{1}{2} + \psi\left(1 - \rho + \frac{1-\lambda^l}{\lambda^l}\right)(\mu_D - \mu_R)$.

To prove that no other equilibrium allocation exists, notice that party *D* has no incentive to allocate experts anywhere in the interval $[\lambda_w, \lambda_0]$, since party *R* would respond best by sending experts to the subset of the interval $[\lambda_w, \lambda_0]$, where party *D* has instead sent loyalists, and would thus win the election with a higher probability than $\Pi_R = \frac{1}{2} + \psi(1 - \rho)(\mu_R - \mu_D)$. Party *D* sending experts to the interval $[\lambda_z, \lambda_W]$ is not part of

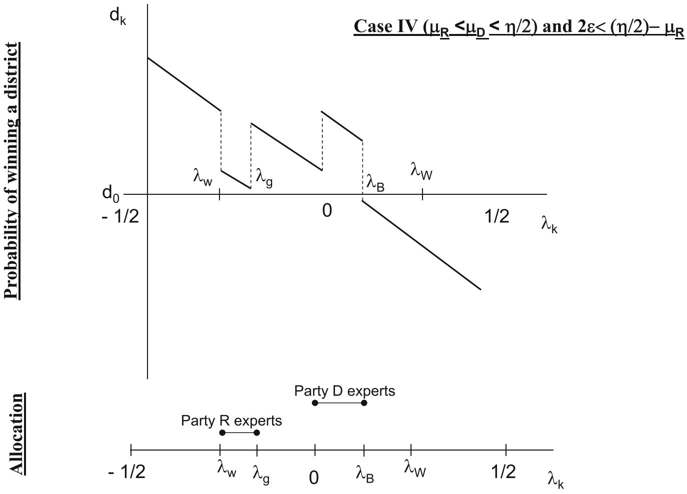

Fig. 8 Equilibrium allocation, case IV

an equilibrium either, since, regardless of party *R*'s response, party *D* could always do at least as well by sending them to $[\lambda_0, \lambda_B]$.

Case (V) Parties are unable to span the crucial districts, and have considerably different shares of experts. Suppose that party *D*, which has many more experts than party *R* (i.e., $z = \mu_D - \mu_R > (\frac{\eta}{2} - \mu_R)/2$), allocates them to $[\lambda_0, \lambda_B]$. Party-*R*'s best response is to send them to $[\lambda_0, \lambda_P]$, such that $G(\lambda_P) - G(\lambda_0) = \mu_R$. To see why, consider the largest (positive) realization of the shock, δ, that still allows party-*D* to win the election, given that party-*D* has allocated experts as described above, and party-*R* has not allocated any. Party-*R* will target with its experts those districts that marginally favor party-*D*, for this level of the shock. This is easily done by sending its few experts to $[\lambda_0, \lambda_P]$, such that $G(\lambda_P) - G(\lambda_0) = \mu_R$. For this allocation by party *R*, party *D* is indifferent between allocating its experts to $[\lambda_0, \lambda_B]$ (or alternatively to the right of λ_P). Hence, this allocation constitutes an equilibrium.

Under this allocation, party *D* wins the election with probability $\Pi_D = \frac{1}{2} + \psi\left[(1-\rho)(\mu_D - \mu_R) + \frac{W}{2} - \mu_R \frac{1-\lambda^I}{2\lambda^I}\right]$. Consider Fig. 9, party-*D* wins the election when more than 50% of the districts favor it; these districts are $\left[-\frac{1-\lambda^I}{2\lambda^I}, \lambda_w\right] \cup [\lambda_w, \lambda_w + x] \cup [\lambda_P, \lambda_w + x]$, such that $\frac{\lambda^I}{1-\lambda^I}\left[\lambda_w + \frac{1-\lambda^I}{2\lambda^I}\right] + \frac{\lambda^I}{1-\lambda^I}[\lambda_w + x - \lambda_w] + \frac{\lambda^I}{1-\lambda^I}[\lambda_w + x - \lambda_P] = 1/2$. Hence, $x = (W + \lambda_P)/2$, where $\lambda_P = \frac{1-\lambda^I}{\lambda^I}\mu_R$. A simple inspection of Fig. 9 shows that all of these districts are won by party-*D* if $\delta < -(x + \lambda_w) + (1-\rho)(\mu_D - \mu_R) = \frac{W}{2} - \mu_R \frac{1-\lambda^I}{2\lambda^I} + (1-\rho)(\mu_D - \mu_R)$; that occurs with probability $\Pi_D = \frac{1}{2} + \psi\left[(1-\rho)(\mu_D - \mu_R) + \frac{W}{2} - \mu_R \frac{1-\lambda^I}{2\lambda^I}\right]$.

To prove that no other equilibrium allocation exists, notice that party *D* has no incentive to allocate experts anywhere in the interval $[\lambda_w, \lambda_0]$, since party *R* would respond best by sending experts to the subset of the interval $[\lambda_w, \lambda_0]$, where party *D* has instead sent loyalists, and would thus win the election with a probability exceeding $\Pi_R = \frac{1}{2} + \psi(1-\rho)(\mu_R - \mu_D)$. Party *D* sending experts to the interval $[\lambda_z, \lambda_W]$ is not part of

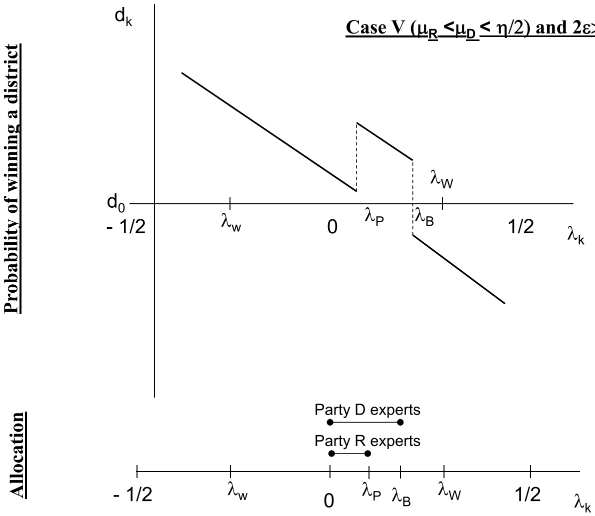

Fig. 9 Equilibrium allocation, case V

an equilibrium either, since, regardless of party R's response, party D could always do at least as well by sending them to $[\lambda_0, \lambda_B]$. $\square$

Proof of Proposition 1 Each party will choose the share of experts—to be allocated according to the results in Proposition 1—in order to maximize its expected utility, given the selection and allocation undertaken simultaneously by the other party. Since the selection problem—just as the allocation problem described at Proposition 1—is symmetric, we can concentrate on the decision of one party—say party D.

Party D selects μ_D experts, given μ_R, in order to maximize expected utility at Eq. 6, where $V_D(\mu_D) = 1 - \mu_D$, $V_D(\mu_R) = -(1 - \mu_R)$, and Π_D depends on μ_D and μ_R as described at Proposition 1.

Consider that party R selects $\mu_R > \eta/2$. For $\mu_D > \eta/2$, then $\Pi_D = 1/2 + \psi(1 - \rho)(\mu_D - \mu_R)$ (case I in Proposition 1), and the optimization problem yields $\mu_D = 1 - \frac{1}{4\psi(1-\rho)}$. For $\mu_D < \eta/2$, then $\Pi_D = 1/2 + \psi\left(1 - \rho + \frac{1-\lambda^I}{2\lambda^I}\right)(\mu_D - \mu_R)$ (case II in Proposition 1), and we have $\mu_D = 1 - \frac{1}{4\psi\left(1-\rho+\frac{1-\lambda^I}{2\lambda^I}\right)}$.

Consider that party R selects $\mu_R < \eta/2$. For $\mu_D > \eta/2$, then $\Pi_D = 1/2 + \psi\left(1 - \rho + \frac{1-\lambda^I}{2\lambda^I}\right)(\mu_D - \mu_R)$ (case II in Proposition 1), and the optimization problem yields $\mu_D = 1 - \frac{1}{4\psi\left(1-\rho+\frac{1-\lambda^I}{2\lambda^I}\right)}$. For $\mu_D < \eta/2$ and $\mu_D < \frac{1}{2}\left(\frac{\eta}{2} + \mu_R\right)$, then $\Pi_D = 1/2 + \psi\left(1 - \rho + \frac{1-\lambda^I}{\lambda^I}\right)(\mu_D - \mu_R)$ (case IV in Proposition 1), and we have $\mu_D = 1 - \frac{1}{4\psi\left(1-\rho+\frac{1-\lambda^I}{\lambda^I}\right)}$. For $\mu_D < \eta/2$ and $\mu_D > \frac{1}{2}\left(\frac{\eta}{2} + \mu_R\right)$, then $\Pi_D = 1/2 + \psi\left[(1 - \rho)(\mu_D - \mu_R) + \frac{W}{2} + \frac{1-\lambda^I}{2\lambda^I}\mu_R\right]$ (case V in Proposition 1), and we have $\mu_D = 1 - \frac{1+\psi\left(W+\frac{1-\lambda^I}{\lambda^I}\mu_R\right)}{4\psi(1-\rho)}$.

Recall that the selection game is symmetric, so that party R has the same reaction function as party D.

(i) Hence, for $\mu_R > \eta/2$, party D's best response is $\mu_D = 1 - \frac{1}{4\psi(1-\rho)} = \mu^*$. Notice that $\mu^* > \eta/2$, if $\lambda^I \leq \lambda_1^I = \frac{0.5-[4\psi(1-\rho)]^{-1}}{0.5-[4\psi(1-\rho)]^{-1}+W}$, since $\frac{\eta}{2} = \frac{\lambda^I}{1-\lambda^I}W$. And analogously for party R, $\mu_R = \mu^* > \eta/2$, if $\mu_D = \mu^* > \eta/2$, and $\lambda^I \leq \lambda_1^I$. Therefore, for $\lambda^I \leq \lambda_1^I$, $\mu_D = \mu_R = \mu^* = 1 - \frac{1}{4\psi(1-\rho)}$ is an equilibrium.

(ii) For $\mu_R < \eta/2$, party D's best response – considering that $\mu_D < \frac{1}{2}(\mu_R + \eta/2)$ – is $\mu_D = 1 - \frac{1}{4\psi\left(1-\rho+\frac{1-\lambda^I}{\lambda^I}\right)} = \mu^{**}$. Notice that $\mu^{**} < \eta/2$, for $\lambda^I \geq \lambda_2^I$, where λ_2^I is such that $1 - \frac{1}{4\psi\left(1-\rho+\frac{1-\lambda^I}{\lambda^I}\right)} < \frac{\lambda^I}{1-\lambda^I}W$. A graphical representation of this inequality is provided in Fig. 10, which shows how the term on the left-hand side is decreasing in λ^I (and converging to $1 - \frac{1}{4\psi(1-\rho)}$ for $\lambda^I = 1$), while the term on the right-hand side is increasing in λ^I (from zero for $\lambda^I = 0$ to infinity for $\lambda^I = 1$), and the inequality thus is satisfied for $\lambda^I \geq \lambda_2^I$. For $\lambda^I \geq \lambda_2^I$, if $\mu_D = \mu^{**} < \eta/2$, party R's best response also would be $\mu_R = \mu^{**} < \eta/2$, and thus $\mu_D < \frac{1}{2}(\mu_R + \eta/2)$ is satisfied. Hence, $\mu_R = \mu_D = \mu^{**}$ is an equilibrium for $\lambda^I \geq \lambda_2^I$.

Finally, notice that no other equilibrium (with $\mu_R > \eta/2$ and $\mu_D < \eta/2$, or vice versa) may emerge. In fact, for $\mu_R > \eta/2$, party D could choose $\mu_D = 1 - \frac{1}{4\psi\left(1-\rho+\frac{1-\lambda^I}{2\lambda^I}\right)}$, which is less than $\eta/2$ for $\lambda^I \geq \lambda_4^I$, where λ_4^I is such that $1 - \frac{1}{4\psi\left(1-\rho+\frac{1-\lambda^I}{2\lambda^I}\right)} < \frac{\lambda^I}{1-\lambda^I}W$. However, for $\mu_D = 1 - \frac{1}{4\psi\left(1-\rho+\frac{1-\lambda^I}{2\lambda^I}\right)} < \eta/2$, party Rś best response (with $\mu_R > \eta/2$) would be $\mu_R = 1 - \frac{1}{4\psi\left(1-\rho+\frac{1-\lambda^I}{2\lambda^I}\right)}$, which is greater than $\eta/2$ for $\lambda^I < \lambda_4^I$. Hence, a selection with $\mu_R > \eta/2$ and $\mu_D < \eta/2$ cannot be an equilibrium. □

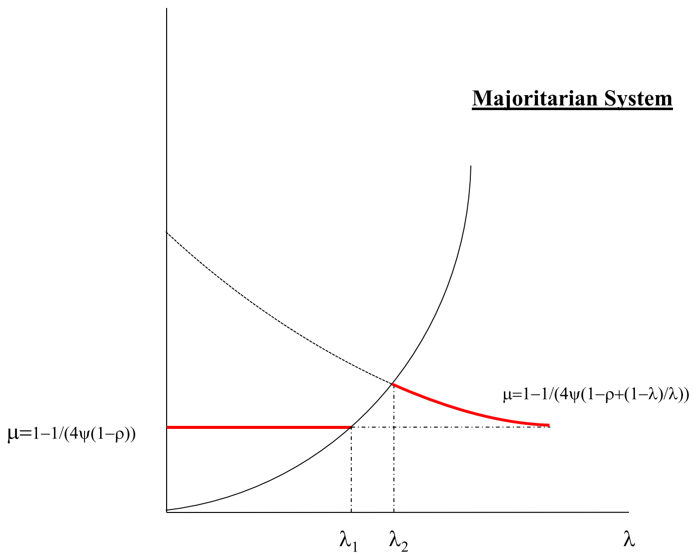

Fig. 10 Political selection in majoritarian systems

Proof of Proposition 2 For $\lambda^I < \lambda_1^I$, it is straightforward to see that $\mu_i^P = 1 - \frac{1}{4\psi} > \mu_i^M = 1 - \frac{1}{4\psi(1-\rho)}$ (for $i = D, R$). The threshold $\lambda_3^I = 1/(1+\rho)$ is such that $\mu_i^P = \mu^{**} = 1 - \frac{1}{4\psi\left(1-\rho+\frac{1-\lambda^I}{\lambda^I}\right)}$. Hence, for $\lambda^I > \lambda_3^I$, $\mu_i^P > \mu^{**}$ and vice versa. Notice that, by Proposition 2, $\mu_i^M = \mu^{**}$ if $\lambda^I > \lambda_2^I$. Hence, if $\lambda_2^I < \lambda_3^I$, we have that $\mu_i^P = 1 - \frac{1}{4\psi} < \mu_i^M = \mu^{**}$ for $\lambda^I \in \left(\lambda_2^I, \lambda_3^I\right)$, and $\mu_i^P = 1 - \frac{1}{4\psi} > \mu_i^M = \mu^{**}$ for $\lambda^I > \lambda_3^I$. If instead $\lambda_2^I > \lambda_3^I$, then $\mu_i^P = 1 - \frac{1}{4\psi} > \mu_i^M$ always. $\square$

References

Acemoglu, D. (2005). Constitutions, politics and economics: A review essay on Persson and Tabellini's the economic effects of constitutions. *Journal of Economic Literature*, *43*, 1025–1048.

Ashworth, S. (2012). Electoral accountability: Recent theoretical and empirical work. *Annual Review of Political Science*, *15*, 183–201.

Beath, A., Christia F., Egorov G., & Enikolopov R. (2014). *Electoral rules and the quality of politicians: Theory and evidence from a field experiment in Afghanistan*. NBER Working Paper 20082.

Besley, T. (2005). Political selection. *Journal of Economic Perspectives*, *19*, 43–60.

Caselli, F., & Morelli, M. (2004). Bad politicians. *Journal of Public Economics*, *88*, 759–782.

Fearon, J. D. (1999). *Electoral accountability and the control of politicians: Selecting good types versus sanctioning poor performance*. Cambridge: Cambridge University Press.

Funk, P., & Gathmann, C. (2013). How do electoral systems affect fiscal policy? Evidence from state and local governments, 1890 to 2005. *Journal of the European Economic Association*, *11*, 1178–1203.

Gagliarducci, S., & Nannicini, T. (2013). Do better paid politicians perform better? Disentangling incentives from selection. *Journal of the European Economic Association*, *11*, 369–398.

Gagliarducci, S., Nannicini, T., & Naticchioni, P. (2011). Electoral rules and politicians' incentives: A micro test. *American Economic Journal Economic Policy*, *3*, 144–174.

Galasso, V., & Nannicini, T. (2011). Competing on good politicians. *American Political Science Review*, *105*, 79–99.

Galasso, V., & Nannicini, T. (2015). So closed: Political selection in proportional systems. *European Journal of Political Economy*, *40*, 260–273.

Hirano, S., & Snyder, J. M, Jr. (2014). Primary elections and the quality of elected officials. *Quarterly Journal of Political Science*, *9*, 473–500.

Iaryczower, M., & Mattozzi, A. (2013). On the nature of competition in alternative electoral systems. *Journal of Politics*, *75*, 743–756.

Kotakorpi, K., & Poutvaara, P. (2011). Pay for politicians and candidate selection: An empirical analysis. *Journal of Public Economics*, *95*, 877–885.

Lindbeck, A., & Weibull, J. (1987). Balanced-budget redistribution as the outcome of political competition. *Public Choice*, *52*, 273–297.

Lizzeri, A., & Persico, N. (2001). The provision of public goods under alternative electoral incentives. *American Economic Review*, *91*(1), 225–239.

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

Mattozzi, A., & Merlo, A. (2015). Mediocracy. *Journal of Public Economics*, *130*, 32–44.

Milesi-Ferretti, G., Perotti, P., & Rostagno, M. (2002). Electoral systems and public spending. *Quarterly Journal of Economics*, *117*(2), 609–657.

Myerson, R. B. (1993). Effectiveness of the electoral systems for reducing government corruption: A game-theoretic analysis. *Games and Economic Behaviour*, *5*, 118–132.

Myerson, R. B. (1999). Theoretical comparisons of electoral systems. *European Economic Review*, *43*, 671–697.

Norris, P. (2004). *Electoral engineering: Voting rules and political behavior*. Cambridge: Cambridge University Press.

Persson, T., & Tabellini, G. (1999). The size and scope of government: Comparative politics with rational politicians. *European Economic Review*, *43*, 699–735.

Persson, T., & Tabellini, G. (2000). *Political economics*. Cambridge, MA: MIT press.

Persson, T., & Tabellini, G. (2003). *The economic effects of constitutions*. Cambridge, MA: MIT Press.

Persson, T., Tabellini, G., & Trebbi, F. (2003). Electoral rules and corruption. *Journal of the European Economic Association*, *1*(4), 958–989.

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

Voigt, S. (2011). Positive constitutional economics II—A survey of recent developments. *Public Choice*, *146*, 205–256.