Political Selection under Alternative Electoral Rules*

Vincenzo Galasso‡ Tommaso Nannicini§

This version: June 2014

Abstract

This paper studies the different patterns of political selection in majoritarian versus proportional systems. Political parties choose the mix of high and low quality candidates. In doing so, parties face a trade-off between increasing the probability of winning the election and appointing low quality but loyal candidates. In majoritarian elections, the share of high quality politicians depends on the distribution of competitive versus safe (single-member) districts. This is not the case under proportional representation, where politicians' selection is determined by the number of swing voters in the entire electorate. We show that, when the share of competitive districts increases, the majoritarian system comes to dominate the proportional system in selecting high quality politicians. However, when the share of competitive districts becomes large enough, a non-linearity arises: the marginal (positive) effect of adding high quality politicians on the probability of winning the election is reduced, and proportional systems dominate even highly competitive majoritarian systems.

Keywords: electoral rules, political selection, probabilistic voting.
JEL codes: D72, D78, P16.

*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.

‡Bocconi University, Department of Public Policy, Via Rontgen 1, 20136 Milan (Italy); vincenzo.galasso@unibocconi.it.

§Bocconi University, Department of Economics, Via Rontgen 1, 20136 Milan (Italy); tommaso.nannicini@unibocconi.it.

1 Introduction

Electoral rules are known to affect politicians' behavior. Majoritarian and proportional systems shape the incentives of both voters and politicians, and therefore may 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 has however been devoted to the impact of electoral rules on political selection. Political scientists have studied how the political representation of women and ethnic minorities varies under different voting rules (Norris 2004). Yet, 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. The recruitment of good politicians has been shown to depend both on candidates' decision to run for office (Caselli and Morelli 2004) and on candidates' selection by political parties (Galasso and Nannicini 2011). But no study examines how electoral rules affect the valence of politicians, with the notable exception of Myerson (1993), who shows how higher entry barriers in majoritarian systems may lead to the election of low quality (dishonest) candidates.

This paper studies the different patterns of political selection in majoritarian versus proportional systems. Political parties select the candidates to be included in their electoral lists. Candidates can be either high or low quality. Parties face a tradeoff. On the one hand, high quality politicians are instrumental to win the election, because voters value their expertise. On the other hand, low quality politicians are loyal and hence valuable to the party. Experts are valued by voters under both majoritarian and proportional elections, because their relative share affects the national policy implemented by the winning party.

In majoritarian systems, experts are also valued for their influence on the district-level policy (e.g., constituency service). In order to increase the probability of winning in competitive (single-member) districts, parties have an incentive to allocate high quality politicians to these districts and to send loyalists to safe ones. Hence, the share of high quality politicians depends on the distribution of competitive versus safe districts. This is not the case under proportional representation, where political selection is simply affected by the number of swing voters in the entire electorate.

We show that, for a high concentration of safe districts, the proportional system is more effective in selecting good politicians. As the share of competitive districts increases, the majoritarian system becomes instead more effective. However, when this share is large enough, a non-linearity arises. Selecting a good politician to be allocated to a competitive district has still a positive effect on the probability of winning the general election, but the magnitude of this marginal effect is now lower. As a result, proportional systems end up dominating even highly competitive majoritarian 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 Section 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 less public goods than proportional systems (Persson and Tabellini 1999; Lizzeri and Persico 2001; Milesi-Ferretti et al. 2002). Electoral rules may also 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 (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 higher 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 welfare benefits, but decrease 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 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 in accountability between majoritarian and proportional systems. The impact of electoral rules on political selection when politicians may be of different types has received less attention. Political scientists have analyzed how the political representation of women and ethnic minorities varies under different voting rules (see Norris 2004). Iaryczower and Mattozzi (2014) have studied 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. No study, however, examines whether electoral rules affect the quality (or valence) of politicians, despite a large literature on valence issues (Stokes 1963) and a recent literature on the importance of political selection (Besley 2005).[1]

A prominent exception is represented by Myerson (1993), who builds a game-theoretic model showing that the proportional system 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 can still 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 rationally vote for the dishonest party whose ideology they share. This cannot happen under proportional representation. As the government 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, there are fewer dishonest (or low quality) politicians elected than in the majoritarian scenario. The crucial mechanism here is the magnitude of the electoral district, which affects the degree of entry barrier for high quality candidates.

In this paper, we tackle the same issue—namely, the impact of electoral rules on political selection—but in a different setup. We build on Galasso and Nannicini (2011), where we show how electoral competition within a majoritarian system can discipline political parties to select high valence candidates. Here, we go beyond the *allocation* of high and low quality politicians across majoritarian districts, and contrast the *selection* of politicians under alternative electoral systems.

Before moving to our model in the next section, 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 so informative for a number of reasons (for a discussion, see Acemoglu 2005). The Italian mixed electoral system in place between 1994 and 2006, however, allows for the comparison of politicians elected in different electoral tiers. 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 frequently changed over time.

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

During three legislative terms (1994-96, 1996-2001, 2001-06), 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. 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 to 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, where 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 income before being elected, and (4) their income after controlling for individual characteristics.[2] 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 features and job type. The use of administrative experience is linked to the idea that lower-level elections can be used by high quality politicians to build reputation and by voters to screen better candidates.

In Figure 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 horizontal line provides the benchmark characteristics of the average politician in the proportional 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 all of the measures, the quality of proportional politicians is higher than the quality of majoritarian politicians elected in safe districts, while the relationship is reverted in the case of (high quality) majoritarian politicians in competitive districts.

Indeed, focusing on the differences that are statistically significant at standard levels, 69% of proportional politicians had a college degree, against 74% of majoritarian politicians

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

elected in contestable districts and 67% of majoritarian politicians elected in safe districts, where contestability is captured by a lagged margin of victory lower than 10%. Preelection income was around 88 thousand euros for proportional politicians, against 99 for majoritarian politicians in contestable districts and 72 for majoritarian politicians in safe districts.

Figure 1 provides valuable information on the allocation decision into the different majoritarian districts, but does not allow to appreciate the overall selection decision by political parties. To focus on selection only, we can use as units of observations the larger macro-districts at the regional level (common to House and Senate members). This alternative strategy enables us to obtain an overall indicator of the contestability of the majoritarian environment, by calculating the share of contestable districts within each macro (i.e., regional) district. Contestable districts, again, are defined as those where the lagged margin of victory was below 10%. In Figure 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.[3] Although these non-linearities 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 level of contestability. Motivated by the above stylized facts, in the next section we propose a model of political recruitment under majoritarian versus proportional elections.

3 The model

Our model is populated by three types of players: voters, candidates, and parties. Two parties run for elections. The winner sets the national policy. Before the election, each party has to select the candidates, who are either party loyalists or experts. In the majoritarian system, parties have also to allocate their candidates into each district. 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. The share of loyalists and experts affects the parties' national policy. Voters can be of three types: core supporters of either party or independent, that is, not aligned to any party. Independent voters care about the national

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

policy, and in the majoritarian system about the valence of their district representative. We embed their voting decision in a standard probabilistic voting model (Lindbeck and Weibull 1987), so that, besides the national policy (and the quality of their local representative in the majoritarian system), they care about a popularity shock to the two parties, and have also an idiosyncratic ideological component towards the two parties.

Our model thus introduces two lines of conflicts: between the two parties—each one seeking to win the election and to implement its national policy—and among parties and independent voters on the national policy. This national policy identifies the cleavage between the welfare (pecuniary interest or ideology) of the winning party and of the independent voters. More specifically, it determines how public resources are split between the winning party and the general public. Parties commit before the election to their national policy by selecting loyalists and experts.

3.1 Parties and candidates

We consider two parties, D and R, which differ in their ideology, and thus in their core supporters. The two parties compete against one another in the political election, and the winner selects the national policy. The main role of the party (leaders) is to select the candidates to be included in the party list (and to allocate them into the different electoral districts in the case of a majoritarian system). This decision affects the national policy chosen by the winning party, which depends on the share of selected candidates.

Candidates can be of three types: party-D loyalists (D), party-R loyalists (R) or experts (E). Loyal candidates share their own party preferences, and do rent-seeking to secure public resources for their party. Regardless of their party of affiliation, experts instead act to devote resources to the general public, for instance through general interest policies. Moreover, in the majoritarian systems, experts have higher valence than loyalists in providing constituency service for their local district.

Each party chooses the share of experts and of party loyalists to include in the electoral list, respectively μ and $1-\mu$ (and how to allocate them to the single-member districts of the majoritarian system). The national policy consists of dividing the available public resources between the winning party and the public at large. We normalize the amount of available resources to one, and consider that this split is determined by the share of loyalists and experts selected by each party.

Hence, the utility to each party $j = D, R$ associated to 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:

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

In the former case, the party wins the election and enjoys the amount of public resources captured by its loyalists; while in the latter case, the party loses the elections and enjoys the resources made available to the general public.

3.2 Voters

We consider three groups of voters. Voters in group D and R are core supporters and hence always vote for party D and R. Independent voters (I) care instead about the national policy, and, in the majoritarian system, about the quality of the candidate in their electoral district.

In a proportional system, the independent voters utility from party i winning the election depends on the amount of resources dedicated to the general public:

$$V_I(\mu_i) = \mu_i \text{ for } i = D, R. \quad (3)$$

In a majoritarian system, political candidates play a double role for the voters. Besides affecting the national policy, they also provide constituency services, for instance by bringing the attention of the national government to local instances that affect their district. In a majoritarian system, the preferences of independent voters living in district k are thus summarized by the following utility function:

$$v_I\left(\mu_i, y_i^k\right) = (1-\rho)\,\mu_i + \rho\overline{V}\left(y_i^k\right) \quad (4)$$

with $i = \{D, R\}$, where y_i^k is the utility associated to the quality of party-i candidate in the electoral district k, and ρ measures the relative importance to the voters of local versus national policies. 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.

Besides the value attributed to the national policies, these 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 voters' decision is also affected by a common popularity shock δ to the parties that occurs before the election and that may modify the perception that all independent voters have about the image 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 $\delta \sim U\left[-\frac{1}{2\psi}, \frac{1}{2\psi}\right]$ with $\psi > 0$.

To summarize, an independent voter will support party D if the utility obtained from the national (and local in the majoritarian system) policy adopted by party D, 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 from party R. That is, an independent prefers party D if $V_I(\mu_D) - V_I(\mu_R) - s - \delta > 0$, in a proportional system, and if $v_I\left(\mu_D, y_D^k\right) - v_I\left(\mu_R, y_R^k\right) - s - \delta > 0$, in a majoritarian system.

3.3 Selection in a proportional system

The incentives for a party to select expert candidates depend largely on the behavior of the independent voters. In fact, while each party (leader) would prefer to distract all public resource for the party, in order to be able to implement a national policy a party needs first to win the election, and thus to convince the independent voters.

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 realizations. In particular, they select the share of loyal and expert candidates, which determines their national policy in case of electoral success. After the popularity shock has occurred, independent voters decide who to support between the two parties; while loyalist voters always support their own party. After the election, the winning party implements its national policy.

To understand the party decision, consider a party's probability of winning the election. Assume that loyalist voters are in equal size, 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 viceversa 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:

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

Hence, since independent voters value expert candidates, who however reduce the amount of resources made available to the party, a trade-off arises for the party between appropriating public resources, and winning the election. This emerges clearly from each party optimization problem. Consider party D, it will choose the share of experts, μ_D, in order to maximize the following expect utility:

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

where Π_D is defined at eq. 5, and $V_D\left(\mu_D\right)$ and $V_D\left(\mu_R\right)$ at eq. 1 and 2.

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

$$\mu_D^P = \mu_R^P = \frac{1}{2} - \frac{1}{4\psi}. \tag{7}$$

The share of experts in the proportional system thus depends positively on the density of the common shock. In other words, when large scandals are less likely to determine the elections, party leaders are more willing to invest in (costly) experts in order to increase their probability of winning the election.

3.4 Selection and allocation in a majoritarian system

Also in a majoritarian system, the incentives for the party to select their candidate, and to allocate them in the different electoral districts, depend largely on the behavior of the independent voters. Again, each party (leader) has preferences over the national policies; but to be able to implement the policy, a party needs first to win the election, and thus to convince the independent voters.

The party optimization problem is still to maximize the expected utility at eq. 6 (for party D), with the utilities defined at eq. 4. However, the parties probability of winning the election now depend both on their selection and allocation of experts.

This allocation of experts is best understood in a majoritarian system with uninominal 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 independent) into 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 $j = \{D, I, R\}$, in district k, and the share of independent voters is assumed to be constant across districts, $\lambda_k^I = \lambda^I \ \forall k$.

The maximum electoral contestability, $\lambda_k = 0$, is obtained in a district k, where the share of R and D core supporters is the same, $\lambda_k^R = \lambda_k^D$. Increases in absolute value of λ_k indicate lower contestability of the district. In particular, districts such that $\lambda_k < -1/2$ or $\lambda_k > 1/2$ are safe, since respectively party D or R win for sure. Hence, only intermediate districts with $\lambda_k \in [-1/2, 1/2]$ are contestable. We consider a continuum of districts, distributed according to a uniform function $\lambda_k \sim U\left[-\frac{1-\lambda^I}{2\lambda^I}, \frac{1-\lambda^I}{2\lambda^I}\right]$, 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 ideology $s < \widetilde{s}$, where $\widetilde{s}$ is the ideology of the (independent) swing voter: $\widetilde{s} = v_I\left(\mu_D, y_D^k\right) - v_I\left(\mu_R, y_R^k\right) - \delta$. Since the share of votes from the independent is $\lambda^I(\widetilde{s} - 1/2)$, party D obtains more than 50% of the votes, and hence wins district k, if $\widetilde{s} > \lambda_k$, which occurs with probability:

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

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 to affect their winning probability in district k. They can modify the relative share of experts and loyalists, μ_i, in order to influence the national policy, and they can choose which candidate to allocate to district k. Hence, the selection decision affects the national policy, while the allocation affects the local policies.

It is convenient to separate these two decisions by considering that parties first select their share of experts, and then how to allocate them to the electoral districts.

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 probability of winning the election. In fact, the difference in utility provided to the independent voters in district k by the two parties can be written as

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

where the former depends on the national policy through the selection of candidates, while the latter is determined by the candidate allocation by the two parties. 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 $W = \rho\left[V_I(E) - V_I(L)\right]$. Hence, parties will compete on good politicians (the experts) to win the contestable districts.

What is the parties strategic behavior in this simultaneous allocation game? Let us begin with only loyal candidates being allocated by both parties to the contestable districts, so that $\overline{V}\left(y_R^k\right) - \overline{V}\left(y_D^k\right) = V_C(L) - V_C(L) = 0$ for all districts $\lambda_k \in [-1/2, 1/2]$. The party 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 a district k for a shock $\delta < d_k = (1-\rho)\left(\mu_D - \mu_R\right) - \lambda_k$. 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 thus the elections, if the shock is strictly in its favor: $\delta < d_0 = 0$. If instead party D has selected more experts, $\mu_D - \mu_R = z > 0$, party D wins the elections, if the shock is $\delta < (1-\rho)z$, again winning all the districts with $\lambda_k < 0$, as shown in Figure 3 (and viceversa for party R). This suggests that the pivotal districts to win the election 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 probability of winning these districts, and thus the 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 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 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 the district λ_0 equal to the probability of winning district $\lambda_W = W$, when both parties allocate a loyalist.

The districts λ_w and λ_W define the range of contestable districts that party D and R will consider for their experts allocation. In fact, if party R allocates only loyalists in $[\lambda_w, \lambda_W]$, party D best response would be to place its experts in $[\lambda_w, \lambda_\Xi]$ in order to win the election if $\delta \leq d_w = \lambda_w + (1-\rho)\, z$. It is important to emphasize that party D could not increase its probability of winning the election 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 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 next proposition characterizes the winning probabilities corresponding to the equilibrium allocation for given party selections (μ_D, μ_R).

Proposition 1. *In a majoritarian system, the winning probabilities* (Π_i, Π_j) *corresponding to the equilibrium allocation 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* $\mu_j = \eta/2$, $\Pi_i = 1/2 + \psi [(1-\rho)(\mu_i - \mu_j) + W/2]$ *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. See Appendix. □

This proposition generalizes the result in Galasso and Nannicini (2011), which characterizes the equilibrium allocation of a fixed share of candidates into districts, to an environment in which parties choosing different share of experts leads to differences in their national policies. The above proposition reports the winning probabilities associated with the equilibrium allocations.

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 $i = D, R$, difference in the winning probabilities may only emerge from the national policies. 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 the election. In all other cases, the probability of winning the elections will depend on the national policies, as well as on the different allocation strategies that may emerge. In particular, the party enjoying an advantage in the share of experts will typically 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 probability are reported in the above proposition.

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. More experts move the national policy away from the party most preferred policy, thereby reducing the party (leaders) utility in the case of victory at the election, and thus of implementation of the policy. However, experts are valuable in attracting the votes of the independents. Thus, a larger share of experts increases the probability of election, as characterized at proposition 1.

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

Proposition 2. *In a majoritarian system, there exist two values of the share of independent voters,* $0 < \lambda_1^I < \lambda_2^I < 1$*, such that the share of experts chosen by both parties is*

$$\mu_D^M = \mu_R^M = \begin{cases} \frac{1}{2} - \frac{1}{4\psi(1-\rho)} \text{ for } \lambda^I \leq \lambda_1^I \\ \frac{1}{2} - \frac{1}{4\psi\left(1-\rho+\frac{1-\lambda I}{\lambda^I}\right)} \text{ for } \lambda^I \geq \lambda_2^I \end{cases}$$

Proof. See Appendix. □

In the former case, $\lambda^I \leq \lambda_1^I$, the share of independent voters is small – and hence only few districts are highly contestable, i.e., $\eta/2$ is also small. Both parties will hence be willing to select enough experts to span their crucial competitive districts, respectively $[\lambda_w, \lambda_\varepsilon]$ for party D and $[\lambda_\varepsilon, \lambda_W]$ for party R (see figure A1 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 would thus find it costly – in terms of deviation from their most preferred policy – to fill all their crucial competitive districts with experts. 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 $[\lambda_w, \lambda_\varepsilon]$ for party D and $[\lambda_\varepsilon, \lambda_W]$ for party R – and the allocation strategy will follow case III in proposition 1 (see also figure A3 in the appendix). 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. This is because the marginal impact of selecting and allocating an expert to a competitive district on the probability of winning the election decreases as the share of competitive districts increases, while the cost—in terms of deviating from the party most preferred national policy—remains constant.[4]

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 the election, and thereby of setting the national policy, but at the cost of pushing the national policy away from the party (leader) most preferred point. Yet, this trade-off differs across electoral systems. In a proportional system, it is entirely based on the impact of the experts on the determination of the national policy. Since they induce more distribution of resources to the general public, they appeal to independent voters, and hence increase their party winning probability, but at the same time they move the national policy away from the party bliss point. The incentives to select experts candidates in a majoritarian system are different. Besides the impact on the national policy, 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 at Figure 4.

Proposition 3. *There exists a threshold value of independent voters, $\lambda_3^I = 1/(1+\rho)$, such that*

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

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

(II) if $\lambda_2^I < \lambda_3^I$, for $\lambda^I \in \left(\lambda_2^I, \lambda_3^I\right)$, 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 low political competition. Most districts are indeed safe, and the party leader need not to 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, is above a certain threshold, $\lambda^I \in \left(\lambda_2^I, \lambda_3^I\right)$, the degree of political competition in the majoritarian system is high, and this becomes the better electoral system to select experts.[5] 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" high. Parties have no incentive to continue to select expert politicians since an additional expert has little impact on the probability of winning the election, but has a cost in terms of national policy for the party. 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 politicians. As recognized in the literature, proportional systems provide broad, nation-wide incentives, while majoritarian systems also entail a local, district-level component. Several studies have shown that this leads to the adoption of different public policies under alternative electoral rules. We suggest that a similar difference may emerge for political selection. In majoritarian systems, the relevance of the local dimension induces 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 please the 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 low or high shares of competitive districts, the proportional system provides better incentives to select good politicians; for

[5]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 $\frac{W}{\rho} > 1 - \frac{1}{4\psi}$.

intermediate levels, the majoritarian systems is instead more effective. These results are in line with our suggestive empirical evidence on Italian mixed-member elections.

References

Acemoglu, Daron (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.

Besley, Timothy (2005). "Political Selection," *Journal of Economic Perspectives* 19, 43–60.

Caselli, Francesco, and Massimo Morelli (2004). "Bad Politicians," *Journal of Public Economics* 88, 759–782.

Funk, Patricia, and Christina Gathmann (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, Stefano, Tommaso Nannicini, and Paolo Naticchioni (2011). "Electoral Rules and Politicians' Incentives: A Micro Test," *American Economic Journal: Economic Policy* 3, 144–174.

Gagliarducci, Stefano, and Tommaso Nannicini (2013). "Do Better Paid Politicians Perform Better? Disentangling Incentives from Selection," *Journal of the European Economic Association* 11, 369–398.

Galasso, Vincenzo, and Tommaso Nannicini (2011). "Competing on Good Politicians," *American Political Science Review* 105, 79–99.

Iaryczower, Matias, and Andrea Mattozzi (2014). "On the Nature of Competition in Alternative Electoral Systems," *Journal of Politics*, forthcoming.

Kotakorpi, Kaisa, and Panu Poutvaara (2011). "Pay for Politicians and Candidate Selection: An Empirical Analysis," *Journal of Public Economics* 95, 877–885.

Lindbeck, A., and J. Weibull (1987). "Balanced-Budget Redistribution as the Outcome of Political Competition," *Public Choice* 52, 273–97.

Lizzeri, Andrea, and Nicola Persico (2001). "The Provision of Public Goods under Alternative Electoral Incentives," *American Economic Review* 91(1), 225–239.

Mattozzi, Andrea, and Antonio Merlo (2008). "Political Careers or Career Politicians?" *Journal of Public Economics* 92, 597–608.

Milesi-Ferretti, Gian-Maria, Roberto Perotti, and Massimo Rostagno (2002). "Electoral Systems and Public Spending," *Quarterly Journal of Economics* 117(2), 609–657.

Myerson, Roger B. (1993). "Effectiveness of the Electoral Systems for Reducing Government Corruption: A Game-Theoretic Analysis," *Games and Economic Behaviour* 5, 118–132.

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

Norris, Pippa (2004). *Electoral Engineering: Voting Rules and Political Behavior*, Cambridge University Press.

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

Persson, Torsten, and Guido Tabellini (2000). *Political Economics*, Cambridge, MA: MIT Press.

Persson, Torsten, and Guido Tabellini (2003). *The Economic Effects of Constitutions*, Cambridge, MA: MIT Press.

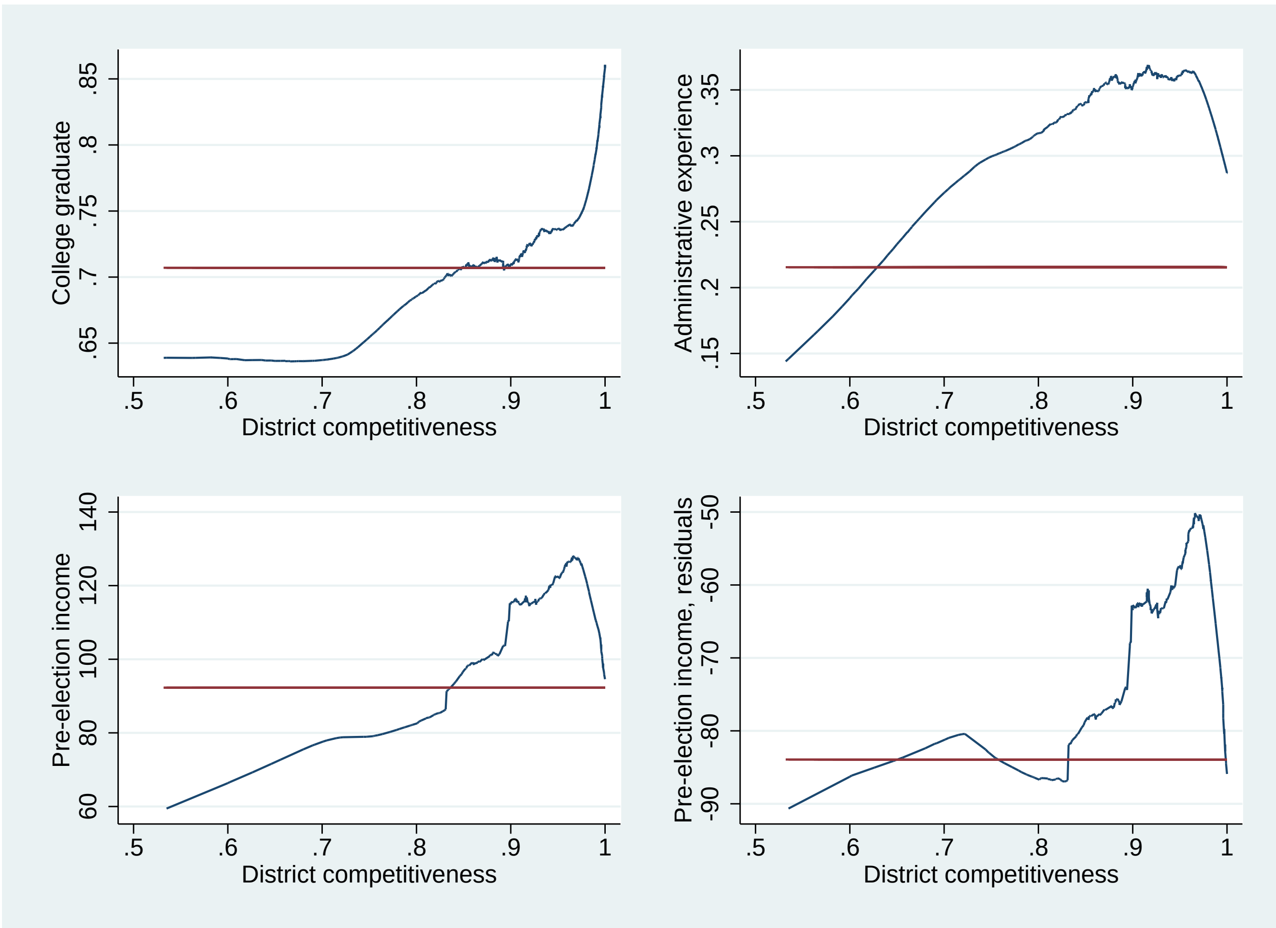
Persson, Torsten, Guido Tabellini, and Francesco Trebbi (2003). "Electoral Rules and Corruption," *Journal of the European Economic Association* 1(4), 958–989.

Stokes, Donald (1963). "Spatial Models of Party Competition," *American Political Science Review* 57(2), 368–377.

Voigt, Stefan (2011). "Positive Constitutional Economics II—A Survey of Recent Developments," *Public Choice* 146, 205–256.

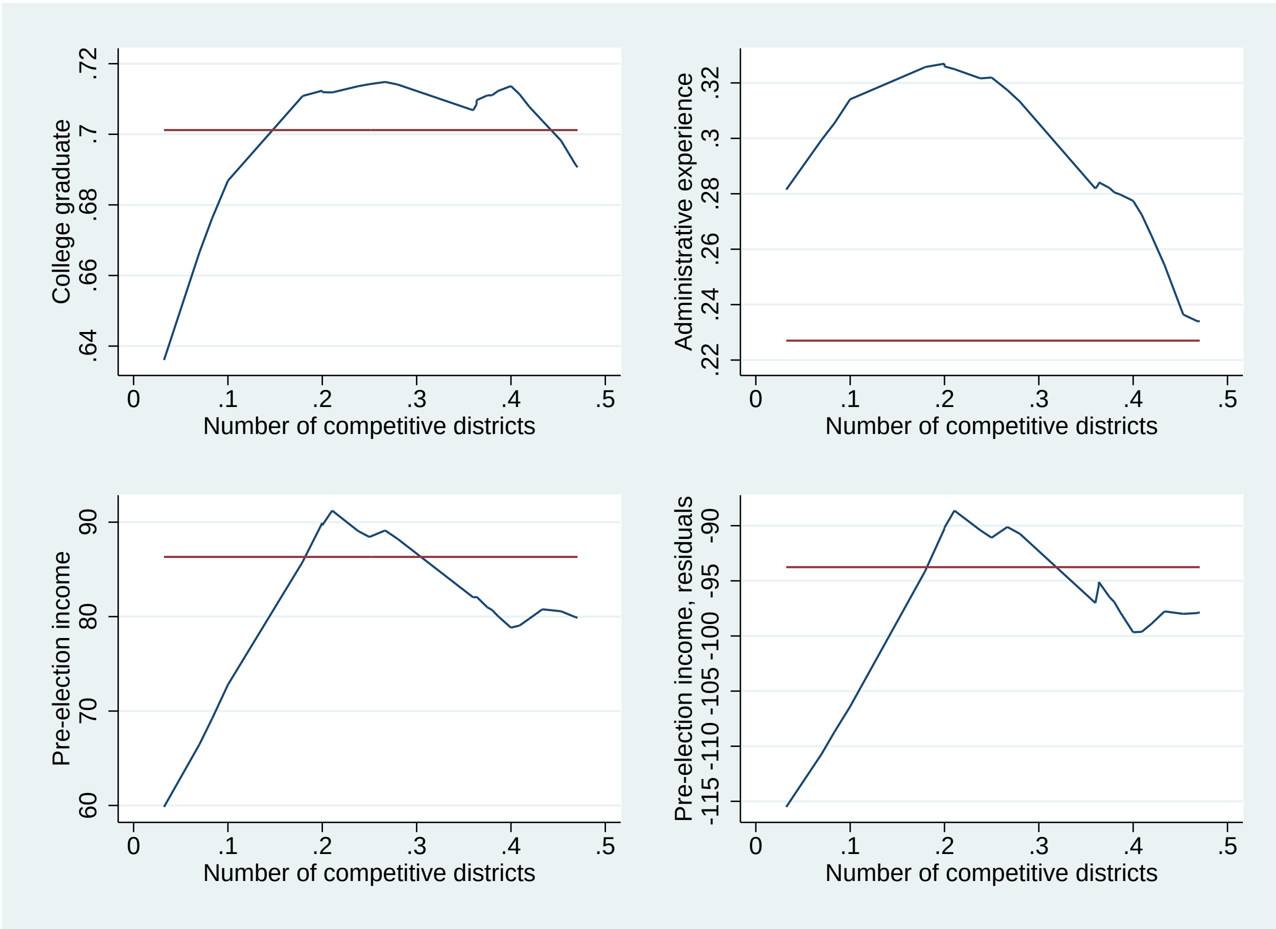
Figures

Figure 1: Quality of politicians based on district competitiveness



Notes. 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.

Figure 2: Quality of politicians based on the share of competitive districts



Notes. 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 percent. The horizontal line represents the average characteristics of *proportional* members of Parliament.

Figure 3: Allocation of experts in the majoritarian system

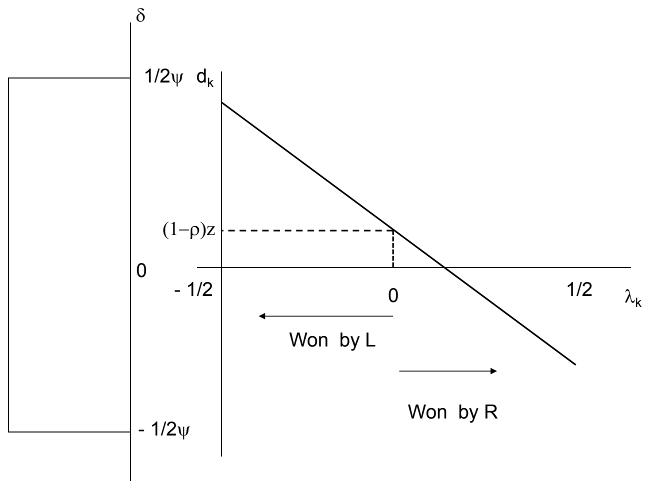
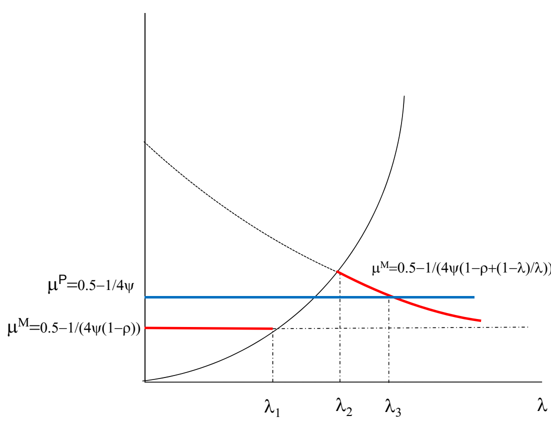


Figure 4: Selection in proportional versus majoritarian systems



Appendix

Proof of Proposition 1

Define Λ^D, party-D allocation of experts, as the union of the district intervals $\Lambda_i^D = [\lambda_I^i, \lambda_{II}^i]$ 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 share of experts, μ_D and μ_R, parties' objective 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^L by party D which includes Λ_i^D s.t. $[\lambda_w, \lambda_\Xi] \subset \Lambda_i^D$. It is easy to see that an allocation Λ^R by party R which 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 different share of experts matters, due to its impact on the national policy. 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 shows 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 include (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 figure A1.

To prove that any equilibrium allocation Λ^D has to include Λ_i^D s.t. $[\lambda_w, \lambda_\Xi] \subset \Lambda_i^D$, consider first an allocation $\widehat{\Lambda}^D$ with $\widehat{\Lambda}_i^D = [\lambda_I^i, \lambda_{II}^i]$ s.t. $0 > \lambda_I^i > \lambda_w$ and $\lambda_{II}^i > \lambda_\Xi$, and no other experts are in $[\lambda_w, \lambda_I]$. Party-R best response is to allocate its experts in $[\lambda_w, \lambda_I] \cup [\lambda_0, \lambda_{II}]$. Following this 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. Finally, it is trivial to show that an equilibrium allocation Λ^D has to include the interval $[\lambda_\varepsilon, \lambda_\Xi]$. Consider $\widehat{\Lambda}^D$ with $\widehat{\Lambda}_i^D = [\lambda_I^i, \lambda_{II}^i] \in [\lambda_w, \lambda_\varepsilon]$ and $\widehat{\Lambda}_j^D = \left[\lambda_I^j, \lambda_{II}^j\right] \in [\lambda_\Xi, \lambda_W]$. Party-$R$ 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. It is easy to see that this holds also for $\mu_R = \mu_D = \eta/2$, 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 viceversa), the party having more experts will win the election with probability $\Pi_D = \frac{1}{2} + \psi\left[(1-\rho)(\mu_D - \mu_R) + \frac{W}{2}\right]$. This is because party D wins the election for population shocks such that $\delta < (1-\rho)(\mu_D - \mu_R)$, but it also ties the elections for $\delta \in [(1-\rho)(\mu_D - \mu_R), (1-\rho)(\mu_D - \mu_R) + W]$.

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 Figure A2. What is party-R best response? Party-R does not have enough experts to match party-D experts and re-establish its probability of winning the election to $\Pr\{\delta > (1-\rho)(\mu_D - \mu_R)\} = \frac{1}{2} + \psi(1-\rho)(\mu_R - \mu_D)$, , but it can reduce party-D probability of winning the election. To see how, 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 targets 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 $[\lambda_j, \lambda_m]$ with $\lambda_j = \max\{\lambda_a, \lambda_w\}$, since it never pays off to send experts outside the interval $[\lambda_w, \lambda_W]$, and λ_m s.t. $G(\lambda_m) - G(\lambda_j) = \mu_R$. Moreover, it is trivial to see that for party-R sending experts to $[\lambda_j, \lambda_m]$, party-D best response is to span the interval $[\lambda_j, \lambda_W]$. 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 Figure A2. Party D wins the election when more than 50% of the districts are in its favor; 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 Figure A2 shows that all 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 $\Pi_D = \frac{1}{2} + \psi\left(1 - \rho + \frac{1-\lambda^I}{2\lambda^I}\right)(\mu_D - \mu_R)$.

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 higher than $\frac{1}{2} + \psi(1-\rho)(\mu_R - \mu_D)$. But with this allocation by party-R, party-D 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 $[\lambda_0, \lambda_B]$. What is party-R best response? To re-establish its probability of winning the election to 1/2, party R can send its experts to $[\lambda_b, \lambda_0]$. As displayed in Figure A3, 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 $\delta \in [-\lambda_B, -\lambda_b]$. Finally, notice that party R cannot increase its probability of winning the election above 1/2 by allocation experts in other districts. Hence, party-D allocation in $[\lambda_0, \lambda_B]$ and party-R allocation in $[\lambda_b, \lambda_0]$ is an equilibrium, and each party has 50% probability of winning the election.

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 election, which can instead be achieved by allocating experts in this interval. Consider party-D allocation $\Lambda^D = [\lambda_b, \lambda_0]$. Party-$R$ 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., $z = \mu_D - \mu_R < (\frac{\eta}{2} - \mu_R)/2$), allocates its experts to $[\lambda_0, \lambda_B]$. What is party-R best response? Having fewer experts, party R is unable to match party-D action with a symmetric allocation (i.e., with $[\lambda_b, \lambda_0]$ as in Figure A3) and to restore the probability to win the election to $\frac{1}{2} + \psi(1-\rho)(\mu_R - \mu_D)$. But it can reduce party-D probability of winning the election. To see how, 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 targets 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 $[\lambda_w, \lambda_g]$, such that $G(\lambda_g) - G(\lambda_w) = \mu_R$ (or alternatively to the right of λ_0), as shown in figure A4. Moreover, notice that for this allocation by party R, party D best response is to allocate its experts to $[\lambda_0, \lambda_B]$ (or alternatively to $[\lambda_w, \lambda_g]$ and the remaining part to the right of λ_0). Hence, this allocation constitutes an equilibrium.

To see why under this allocation party-D wins the election with probability $\Pi_D = \frac{1}{2} + \psi\left(1 - \rho + \frac{1-\lambda^I}{\lambda^I}\right)(\mu_D - \mu_R)$, consider again Figure A4. Party-D wins the election when more than 50% of the districts are in its favor; these districts are $\left[-\frac{1-\lambda^I}{2\lambda^I}, \lambda_w\right] \cup [\lambda_g, \lambda_w + x] \cup [\lambda_0, \lambda_B]$, such that $\frac{\lambda^I}{1-\lambda^I}\left[\lambda_w + \frac{1-\lambda^I}{2\lambda^I}\right] + \frac{\lambda^I}{1-\lambda^I}[\lambda_w + x - \lambda_g] + \frac{\lambda^I}{1-\lambda^I}[\lambda_B - \lambda_0] = 1/2$. Hence, $x = W - \lambda_B$, where $\lambda_B = \frac{1-\lambda^I}{\lambda^I}\mu_D$. A simple inspection of Figure A4 shows that all these districts are won by party-D if $\delta < d_x = -(x + \lambda_g) + (1-\rho)(\mu_D - \mu_R) = \left(1 - \rho + \frac{1-\lambda^I}{\lambda^I}\right)(\mu_D - \mu_R)$, that occurs with probability $\Pi_D = \frac{1}{2} + \psi\left(1 - \rho + \frac{1-\lambda^I}{\lambda^I}\right)(\mu_D - \mu_R)$.

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 best respond 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 an equilibrium either, since, regardless of party R 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 largely different shares of experts. Suppose that party D, which has more experts than party R (i.e., $z = \mu_D - \mu_R > (\frac{\eta}{2} - \mu_R)/2$), allocates its experts to $[\lambda_0, \lambda_B]$. What is party-R best response? In this case, having much fewer experts, party R can only try to reduce party-D probability of winning the election. To see how, 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 targets with its experts those districts that are marginally in favor of party-D, for this level of the shock. This is easily done by sending its few experts to $[\lambda_0, \lambda_P]$, such that $G(\lambda_P) - G(\lambda_0) = \mu_R$.

Moreover, notice that 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.

To see why 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 Figure A5. Party-D wins the election when more than 50% of the districts are in its favor; 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 Figure A5 shows that all 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 best respond 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 an equilibrium either, since, regardless of party R response, party D could always do at least as well by sending them to $[\lambda_0, \lambda_B]$. QED

Proof of Proposition 2

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 simultaneously performed 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 the expected utility at eq. 8, where $V_D(\mu_D)=1-\mu_D$, $V_D(\mu_R)=\mu_R$, 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 $\mu_D=\frac{1}{2}-\frac{1}{4\psi(1-\rho)}$. 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 $\mu_D=\frac{1}{2}-\frac{1}{4\psi\left(1-\rho+\frac{1-\lambda^I}{2\lambda^I}\right)}$.

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 $\mu_D=\frac{1}{2}-\frac{1}{4\psi\left(1-\rho+\frac{1-\lambda^I}{2\lambda^I}\right)}$. 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 $\mu_D=\frac{1}{2}-\frac{1}{4\psi\left(1-\rho+\frac{1-\lambda^I}{\lambda^I}\right)}$. 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 $\mu_D = \frac{1}{2} - \frac{1+\psi\left(W+\frac{1-\lambda^I}{\lambda^I}\mu_R\right)}{4\psi(1-\rho)}$.

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 best response is $\mu_D = \frac{1}{2} - \frac{1}{4\psi(1-\rho)} = \mu^*$. Notice that $\mu^* > \eta/2$, if $\lambda^I \leq \lambda^I_1 = \frac{0.5-[4\psi(1-\rho)]^{-1}}{0.5-[4\psi(1-\rho)]^{-1}+W}$, 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 $\lambda^I \leq \lambda^I_1$. Therefore, for $\lambda^I \leq \lambda^I_1$, $\mu_D = \mu_R = \mu^* = \frac{1}{2} - \frac{1}{4\psi(1-\rho)}$ is an equilibrium.

(ii) For $\mu_R < \eta/2$, party D best response – considering that $\mu_D < \frac{1}{2}(\mu_R + \eta/2)$ – is $\mu_D = \frac{1}{2} - \frac{1}{4\psi\left(1-\rho+\frac{1-\lambda^I}{\lambda^I}\right)} = \mu^{**}$. Notice that $\mu^{**} < \eta/2$, for $\lambda^I \geq \lambda^I_2$, where λ^I_2 is such that $\frac{1}{2} - \frac{1}{4\psi\left(1-\rho+\frac{1-\lambda^I}{\lambda^I}\right)} < \frac{\lambda^I}{1-\lambda^I}W$. A graphical representation of this inequality is provided at figure A6, which shows how the term on the left hand side is decreasing in λ^I (and converging to $\frac{1}{2} - \frac{1}{4\psi(1-\rho)}$ 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 thus the inequality is satisfied for $\lambda^I \geq \lambda^I_2$. For $\lambda^I \geq \lambda^I_2$, if $\mu_D = \mu^{**} < \eta/2$, party R best response would also 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 $\lambda^I \geq \lambda^I_2$.

Finally, notice that no other equilibrium (with $\mu_R > \eta/2$ and $\mu_D < \eta/2$, or viceversa) may emerge. In fact, for $\mu_R > \eta/2$, party D could choose $\mu_D = \frac{1}{2} - \frac{1}{4\psi\left(1-\rho+\frac{1-\lambda^I}{2\lambda^I}\right)}$, which is less than $\eta/2$ for $\lambda^I \geq \lambda^I_4$, where λ^I_4 is such that $\frac{1}{2} - \frac{1}{4\psi\left(1-\rho+\frac{1-\lambda^I}{2\lambda^I}\right)} < \frac{\lambda^I}{1-\lambda^I}W$. However, for $\mu_D = \frac{1}{2} - \frac{1}{4\psi\left(1-\rho+\frac{1-\lambda^I}{2\lambda^I}\right)} < \eta/2$, party R best response (with $\mu_R > \eta/2$) would be $\mu_R = \frac{1}{2} - \frac{1}{4\psi\left(1-\rho+\frac{1-\lambda^I}{2\lambda^I}\right)}$, which is greater than $\eta/2$ for $\lambda^I < \lambda^I_4$. Hence, a selection with $\mu_R > \eta/2$ and $\mu_D < \eta/2$ cannot be an equilibrium. QED

Proof of Proposition 3

For $\lambda^I < \lambda^I_1$, it is straightforward to see that $\mu^P_i = \frac{1}{2} - \frac{1}{4\psi} > \mu^M_i = \frac{1}{2} - \frac{1}{4\psi(1-\rho)}$ (for $i = D, R$). The threshold $\lambda^I_3 = 1/(1+\rho)$ is such that $\mu^P_i = \mu^{**} = \frac{1}{2} - \frac{1}{4\psi\left(1-\rho+\frac{1-\lambda^I}{\lambda^I}\right)}$. Hence, for $\lambda^I > \lambda^I_3$, $\mu^P_i > \mu^{**}$ and viceversa. Notice that, by Proposition 2, $\mu^M_i = \mu^{**}$ if $\lambda^I > \lambda^I_2$. Hence, if $\lambda^I_2 < \lambda^I_3$, we have that $\mu^P_i = \frac{1}{2} - \frac{1}{4\psi} < \mu^M_i = \mu^{**}$ for $\lambda^I \in \left(\lambda^I_2, \lambda^I_3\right)$, and $\mu^P_i = \frac{1}{2} - \frac{1}{4\psi} > \mu^M_i = \mu^{**}$ for $\lambda^I > \lambda^I_3$. If instead $\lambda^I_2 > \lambda^I_3$, then $\mu^P_i = \frac{1}{2} - \frac{1}{4\psi} > \mu^M_i$ always. QED

Figure A1

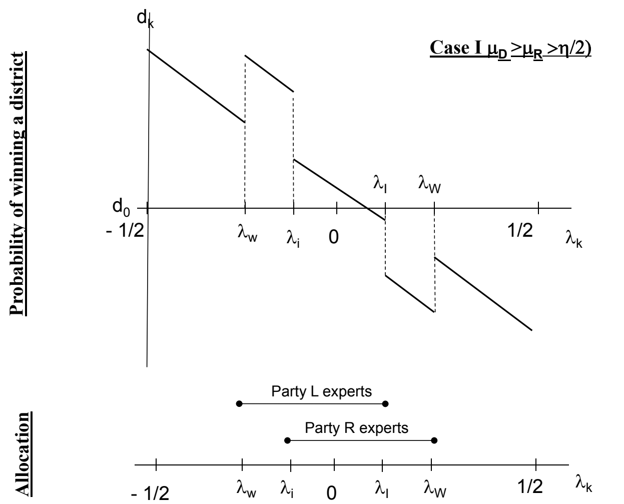


Figure A2

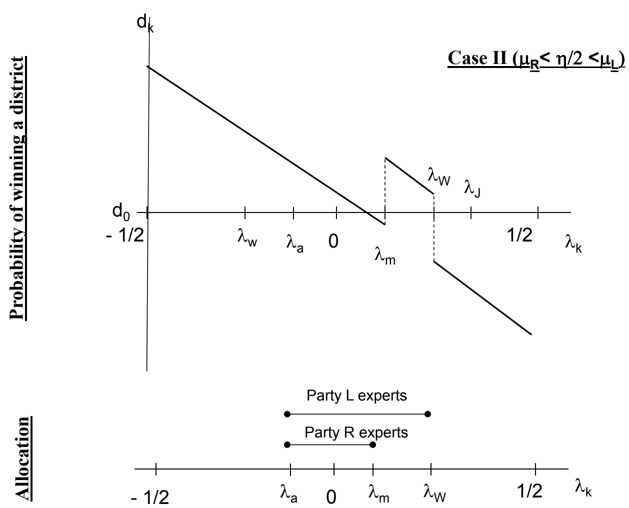


Figure A3

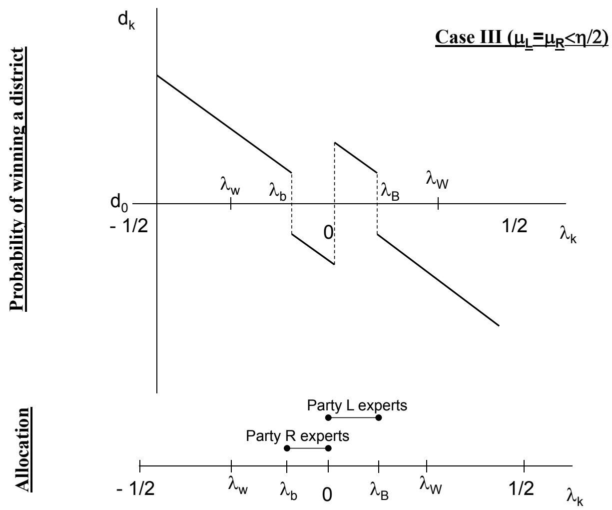


Figure A4

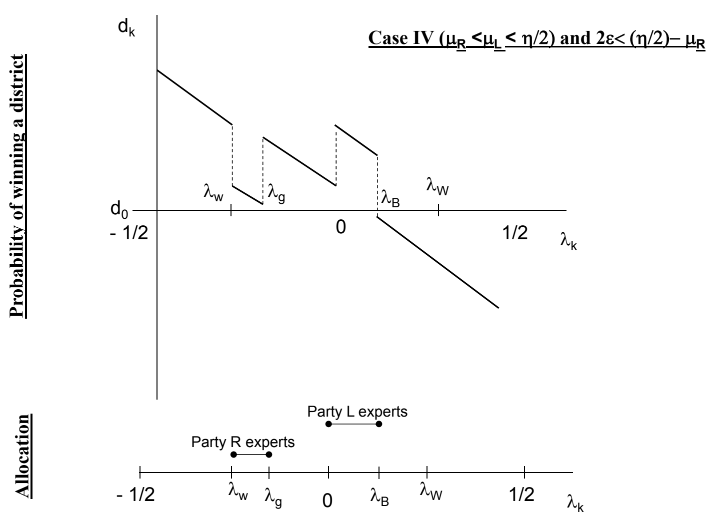


Figure A5

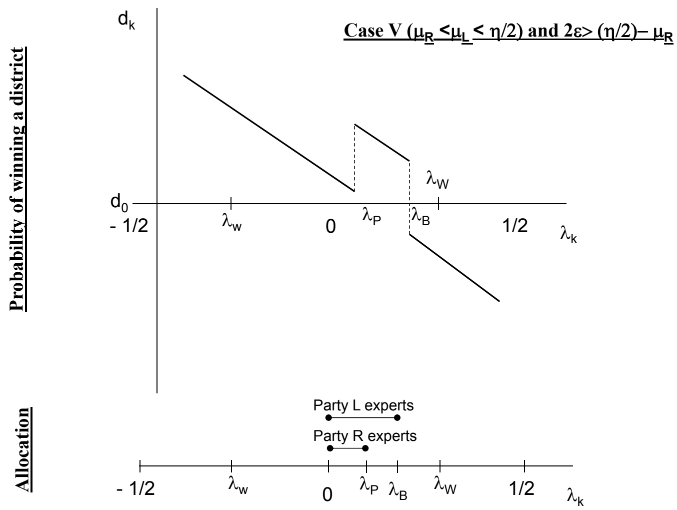


Figure A6