Multi-candidate Political Competition and the Industrial Organization of Politics

Avner Seror*, Thierry Verdier†

August 15, 2018

Abstract

In this paper, we present a microfounded theory of multi-candidate political competition taking an "industrial organization" perspective of politics. The analytical framework is shown to be flexible enough to address several applications on the topics of special interest politics, coalition formation in the legislature in proportional elections, and redistribution under alternative electoral rules.

Keywords: Political Economy, Elections, Probabilistic Voting Models, Fréchet distributions, Industrial Organization, Redistribution, Public policy, Duverger, runoff, Plurality, Proportional.

JEL Classification: D71, D72, L11.

1 Introduction

Across the 235 countries listed in the World Factbook in 2016, the average number of active political parties is nine. While autocracies such as Cuba, Eritrea, Kuwait, Qatar, Saudi Arabia and the UAE have a single political party, some countries count more than twenty active political parties.

*Smith Institute for Political Economy and Philosophy, Chapman University, seror@chapman.edu

†Paris School of Economics, Ecole des Ponts ParisTech, PUC-Rio, and CEPR London, thierry.verdier@ens.fr. We are grateful to Gani Aldashev, Alberto Bisin, Laurent Bouton, Micael Castanheira, Alan Drazen, Rachel Kranton, Rebecca Morton and Ragnar Torvik for their discussions and insightful comments that greatly improved the paper. We thank the participants to the ECARES, PSI-PSE and IOEA 2018 seminars. Avner Seror and Thierry Verdier acknowledge the financial support from the ERC Grant TECTACOM n° 324004.

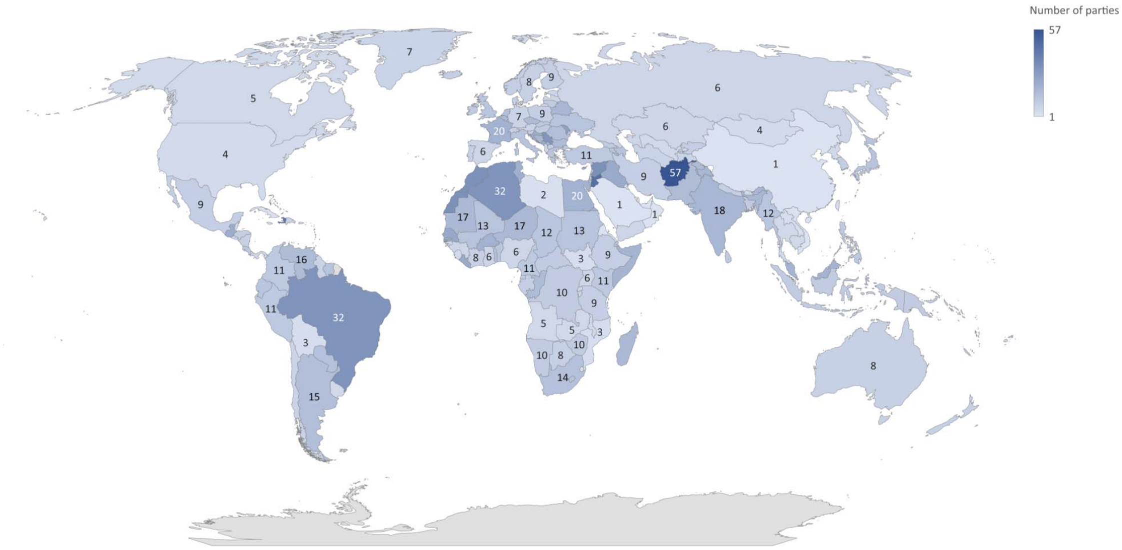

Figure 1: Number of active political parties per country (data from the World Factbook, 2016).

This is for example the case for Brazil (32 parties), Serbia (31 parties), Hong Kong (21 parties) and France (20 parties).

In spite of the obvious party fragmentation across countries, political competition has more than often been modeled by political economists as a two-party contest for power. This is for instance exemplified by the celebrated Hotelling model of deterministic voting or the seminal probabilistic voting model applied so successfully to different economic contexts (Persson and Tabellini (2002)). While some analytical frameworks of multi-candidate competition have been developed, in most cases these models are not easy to handle. On the one hand, they may face problems of existence of equilibria (deterministic spatial voting models), or may on the contrary involve many equilibria (citizen-candidates models), or the applications are exogenously limited to a few number of candidates.

The purpose of this paper is to provide a simple analytical framework on multi-candidate elections that is tractable enough to allow a systematic analysis of the endogenous structure of political competition in ways similar to the Industrial Organization models used to analyze endogenous market structures under economic competition. The versatility of our approach is demonstrated through several applications on classical topics in political economy: special interest politics under alternative voting rules and coalition formation in the legislature.

To analyze multi-candidate political competition, we propose a standard probabilistic voting model (e.g. Coughlin (1992) and Persson and Tabellini (2002)), assuming the noise in random voting decisions to be distributed according to some Fréchet (or extreme type II) distributions. This assumption is particularly convenient as the maximum of a finite sequence of random variables independently distributed according to Fréchet distributions is a Fréchet distribution.[1] This idea has already been exploited in the context of trade between multiple countries by Eaton and Kortum (2002). It is particularly relevant in the context of political competitions as well, since individuals cast their vote for their most preferred candidate out of a finite list of challengers. In various political competition contexts, this property of Fréchet distributions helps generate political objective functions (vote shares, probabilities of election) that can be expressed as relatively simple contest functions in terms of the policy platforms of the candidates under competition. These functions are then tractable enough to provide analytical characterizations of electoral equilibria and the market structure of political competition when there is an *endogenous* entry of candidates in election.

Within this framework, our first contribution is to provide simple conditions for the existence and determination of electoral equilibria in multi-candidate competitions. As well, following the seminal approach of McFadden (1974) of individual choice decisions, we outline (and show in the Appendix) that an axiomatic approach to probabilistic voting models may provide a micro-founded rationale for the use of Fréchet distributions in voting theories.

Our next contributions relate to several applications of the canonical model to discuss the endogenous fragmentation of political competition, and how it depends on fundamental parameters of the society (preferences, technologies, heterogeneities, institutional arrangements).

Specifically, the first application presents a study of the endogenous entry of political parties in the context of special interest politics with homogeneous entrants, heterogeneous voters and stochastic effective voting participation. We analytically characterize the equilibrium number of entrants in a symmetric equilibrium, and show how it is affected by the cost of party formation, the average voters' political responsiveness to political platforms, and the uncertainty of voting participation.

[1] In statistics language, the Fréchet distributions family satisfies the maximum stability postulate.

The second application extends the framework of the first application to the case of heterogeneous candidates. We study a case where parties can choose to invest resources in order to increase their popularity or their expected turn-out. This extension formalizes the effect of media campaigns or other popularity-enhancing investments on party fragmentation. We show that the presence of these political investments decreases the equilibrium degree of fragmentation of the polity and leads to more rent extraction from political parties.

The third application studies parliamentary systems with a proportional representation electoral system. Our model allows the study of the interactions between coalition formation, legislative bargaining, and party formation before the election. Importantly, our analysis accounts for both the endogenous entry of parties in proportional elections and the formation of coalitions in the legislature, and characterizes the resulting equilibrium degree of political fragmentation under a proportional system.

Finally, the fourth application discusses the effect of alternative electoral rules on political fragmentation in the context of special interest politics with heterogeneous sincere voters and homogenous political entrants. Specifically, we compare the endogenous entry of political parties in runoff and plurality elections. The Duverger's hypothesis or Duverger's second law (see Riker (1982, p. 754)) states that "the simple majority system with second ballot and proportional representation favors multipartyism". Duverger's Law instead predicts that plurality rule yields a two party system. Duverger (1964) gave two reasons for his predictions. The first is a"mechanical effect" associated to the fact that in a plurality electoral system, small parties are penalized by the fact that only the candidate obtaining the largest number of votes in each (single-seat) constituency gets elected. The second effect is a so-called "psychological effect", by which voters avoid wasting their votes on small parties and concentrate their votes on larger parties (Duverger (1964, p. 206-55)). There are, however, exceptions from this rule and Sartori (1986) in particular has shown that Duverger's statements can be regarded as 'laws' only if certain conditions are fulfilled. In our set-up, voting is sincere and stochastic, and in contrast to Duverger's hypothesis, we show that a plurality system should lead to *more* entrants than a runoff system. Specifically, the comparison between runoff and plurality elections hinges on how the two electoral systems affect the marginal benefits of parties for capturing rents. Each round of runoff elections creates a stronger incentive for parties to capture rents, since they imply a lower degree of competition. However, as it turns out, taken together, the two rounds create weaker incentives for parties to capture rents than under the plurality system.

As a consequence, there is less political fragmentation under a runoff system than under plurality elections. Related to the Duverger's hypothesis, this result suggests that the"psychological effect" or strategic voting dimension of the electorate may be an important factor for such a hypothesis to hold, while the "mechanical effect" of an institutional electoral rules translating votes into seats is not sufficient.

The paper is organized in the following way. The next section discusses how the paper relates to the current literature. Section 3 presents a benchmark model of multi-candidates competition with stochastic voting according to Fréchet distributions. We also discuss briefly an axiomatic approach to probabilistic voting models that may provide a rationale for the use of Fréchet distributions in voting theories. Section 4 presents the applications. In Section 4.1, we discuss the case of special interest politics with heterogeneous voters, and homogeneous political entrants. Section 4.2 analyzes an extension to heterogeneous entrants and political investments in popularity. In section 4.3, we consider the issue of endogenous entry, political competition and legislative bargaining in a proportional representation system. Section 4.4 discusses the issue of political competition in redistributive politics under runoff and plurality elections. Finally in section 5 we conclude and discuss future research avenues. The Appendix presents the details for an axiomatic rationale for the use of Fréchet distributions in stochastic voting models and provides as well all the proofs of the results of the main text.

2 Literature review

Our paper obviously builds upon the vast literature on electoral competition under probabilistic voting (Brams and O'Leary (1970), Hinich et al. (1972), Hinich (1977) and Coughlin and Nitzan (1981)).[2] As is well known, probabilistic voting theories are typically useful in dealing with the multidimensionality of political decisions and the issue of existence of electoral equilibria (Coughlin (1992) and Banks and Duggan (2005)).

Most of the existing studies on probabilistic voting focus on two-candidate elections. A notable exception is Lin et al. (1999) who provide an existence theorem for electoral equilibria for multi-

[2]See as well the excellent review of Coughlin (1992) on the early literature on probabilistic voting. More recent works on probabilistic voting include for example Dixit and Londregan (1996), Lindbeck and Weibull (1987) and Persson and Tabellini (2002).

candidate elections. They assume that voters' utility depends on the distance between their own ideal policy and the winning candidate's platform as well as on a random shock. They show that if the utility shocks have a high enough variance, then the expected objective functions of the candidates are concave, which implies the existence of an electoral equilibrium. Our axiomatic approach in the Appendix allows us as well to link the existence of a global equilibrium to the randomness of voting decisions. Consistently with Lin et al. (1999), we find that if voters are not too reactive to political platforms, i.e. if the randomness of voting behaviors is sufficiently large, then an electoral equilibrium exists and is unique. In contrast to Lin et al. (1999), our use of Fréchet distributions allows a tractable characterization of the electoral equilibrium and of the equilibrium endogenous degree of political fragmentation associated to political competition.

Closest to us is the model of multi-candidate election developed by Schofield (2007). The author uses Gumbel (extreme type I) distributions to build a model of multi-candidate elections. He notices that this type of distribution respects the "independence of irrelevant alternative property". Our axiomatic approach to probabilistic voting is consistent with this approach, but provides microfoundations for the use of extreme type II distributions in stochastic voting models. Schofield (2007) shows that candidates in election do converge to the mean platform when there is no large asymmetry in the electoral perception of the "quality" (or valence) or the parties, given that the variance in voters' ideal positions is not too large. In contrast to Schofield (2007), we are able to derive closed form results for electoral platforms, so we can directly link the valence of candidates in elections to the convergence of the electoral equilibrium. Specifically, we show in the second application that heterogeneity in the electoral perception of the "quality" of the candidates necessarily leads to diverging platforms in the case of plurality elections. Furthermore, the formalism implied by the extreme type II distributions allows us to study in a tractable way how candidates' heterogeneity affects the endogeneous entry of political parties.[3]

Our work also relates to spatial non stochastic voting literature that analyzes the existence and the convergence of electoral equilibria with endogenous entrants (see Cohen and Shepsle (1990)for a review on the early literature on the subject. See as well Shepsle (1991) and Osborne (1995) for

[3]Few works have estimated probabilistic voting models with extreme type I distributions, see for instance Schofield et al. (1998), Dow and Endersby (2004) and Schofield (2007). No study to our knowledge has relied on extreme type II distributions (see for example Eaton and Kortum (2002) for an estimation of a model with Fréchet noise in the context of international trade).

more recent overviews). In such models, policy options are represented by points on an Euclidian space and each voter's utility function is commonly assumed to be a decreasing function of the Euclidian distance between candidates' positions and voter's ideal point. In this literature, the policy motivations of candidates in elections are central in understanding the process of entry as for instance in the citizen-candidate models of Osborne and Slivinski (1996) and Besley and Coate (1997). An interesting extension of this line of research is Dickson and Scheve (2010) who consider a theory of electoral institutions with an endogenous number of candidates in a citizen-candidate framework adapted from Osborne and Slivinski (1996). They incorporate an identity-related political behavior in a model of electoral competition. In their theory, social identities provide a motivation for political behavior, including vote choice and decisions to seek office. Our theory differs from the standard spatial competition theories in two important ways. First, the existence of pure strategy equilibria is not always guaranteed in spatial competition models, while by contrast, in a stochastic voting theory, pure strategy equilibria typically exist, even in multidimensional and not Euclidian political spaces. As a matter of fact, the existence of a symmetric equilibrium is demonstrated in the various applications presented in the sequel. Second, we consider the entry of purely office motivated candidates, so entry decisions relate solely to the strength of electoral competition, not to policy preferences.

Our third application on political competition with endogenous entry and legislative bargaining in a proportional representation system connects to the large literature devoted to the formation of coalitions in the legislature (e.g. Schofield (1993), Baron (1993), Schofield (1997), Baron and Diermeier (2001) and Diermeier et al. (2002)). This literature typically studies the sustainability of coalitions in the legislature, as well as the convergence of electoral equilibria. Building on the framework of Baron and Diermeier (2001), our analysis more specifically focuses on the issue of endogenous political entry in a context that accounts for the formation of coalitions in the legislature in a proportional system. This way, we link the issue of endogenous political fragmentation to the institutional context of legislative bargaining, something that to our knowledge has not been formally investigated in the literature. Our main prediction in that respect is to show that a proportional system makes rent extraction more costly at the margin for existing parties relative to a plurality system. Consequently we should expect less parties to form under proportional rules.

Our last application relates to the works of Myerson (1993), Lizzeri and Persico (2001) and

Lizzeri and Persico (2005).[4] Myerson (1993) studies electoral equilibria under different electoral rules when candidates simultaneously decide their political platform, so when pure strategy equilibria do not exist. Lizzeri and Persico (2001) and Lizzeri and Persico (2005) apply a similar framework to study the issue of public provision under alternative electoral regimes and the drawbacks of electoral competition, and find that public goods are provided less often in winner-take-all system relative to proportional systems. Relatedly, Bordignon et al. (2016) provide evidence that policy volatility is smaller under runoff elections, while Bordignon et al. (2017) formally demonstrate that result. Relatedly, we show that both proportional and runoff systems should lead to policies that are more favorable to the citizenry relative to plurality elections. This obtains because both the proportional and runoff systems make rent extraction more costly at the margin for existing parties relative to the plurality system.

The existing theoretical literature on Duverger's Hypothesis is mixed. Osborne and Slivinski (1996), Cox (1997) and Bordignon et al. (2017) derive results consistent with Duverger's predictions. Bouton (2013), Bouton and Gratton (2015) and Morelli (2004) establish various conditions sustaining the existence of equilibria that accord with and/or that contradict Duverger's predictions. The previous works examine the effect of strategic voting or strategic candidacy on party fragmentation under alternative electoral rules when there are few candidates with heterogeneous political preferences. Our complementary objective in this paper is to study how political competition affects party formation under alternative electoral rules in the context of sincere voting and office motivated parties.[5] On the empirical side, Bordignon et al. (2016) and Chamon et al. (2018) exploit regression discontinuity designs on Italian and Brazilian municipalities respectively, and provide evidence in line with Duverger's Hypothesis. Interestingly and using a regression discontinuity design in Brazilian mayoral races, Fujiwara (2011) provides evidence that strategic voting is the most likely driving force behind Duverger's Hypothesis. This is consistent with our result showing that when candidates are office motivated and political competition is accounted for, plurality systems should lead to more entrants than either runoff or proportional systems.

[4]While we study elections when the voters are perfectly informed, an interesting line of works studies multi-candidate elections under alternative rules in the context of imperfect information. See for instance Bouton and Castanheira (2012), and the experimental studies of Bouton et al. (2016) and Bouton et al. (2017).

[5]Callander (2005) shows that entry deterrence from incumbent parties can drastically reduce party fragmentation in runoff elections, thereby contradicting as well the Duverger's Hypothesis.

3 The canonical model

Suppose an individual in the population has a vector of measured attributes s that belongs to a convex and non-empty set S. The vectors $s \in S$ relate to the characteristics of the agents, say their education level, their ethnicity, their wealth, their region of birth and so forth. There are P candidates running for a single-district election indexed $i \in \mathcal{P} = \{1, \ldots, P\}$. We assume a plurality election rule, meaning that the party with the highest vote share wins the election. As only one candidate per party can run in the election, we will use interchangeably the terms party and candidate in the sequel.

We denote q_i the platform of candidate $i \in \mathcal{P}$ in the election, which we assume belongs to a closed and convex set Q_i. A set of feasible policies for the candidates is not necessarily an Euclidian space. Components of candidates' platforms can relate to tax collection, public good provision, redistribution, alternative institutional arrangements, allocation of natural resource revenues, of campaign resources and so forth. Furthermore, we allow the sets of feasible platforms to be candidate-specific in order to account for factors affecting policies such as differences in candidates' ability or more broadly idiosyncratic constraints that are not directly linked to the election under scrutiny but that weight on candidates' strategic decisions. Political parties for instance can impose such external constraints on their candidates in election.

We assume no commitment issue so the utility from electing candidate i for any individual with attributes s can be written in the form

$$U = V(s, i)\epsilon(s, i) \tag{1}$$

where V is non-stochastic. We posit that $V(s, i) \equiv V(s, q_i)$, meaning that the deterministic component of the utility of the agents with attributes s depends only on the quality of the platform of candidate i, and thus does not reflect some exogenous preference for candidate i. Alternatively, $\epsilon(., .)$ reflects the idiosyncracies of voting behaviors, which we allow to depend on candidate i rather than on the platform q_i as well as on the attributes $s \in S$. We assume that $\epsilon(s, i)$ is positive for any $s \in S$ and any platform $q_i, i \in \mathcal{P}$.

Any individual votes for the candidate that maximizes his utility. Thus, an agent with attributes $s \in S$ votes for candidate i if

$$V(s, q_i)\epsilon(s, i) > V(s, q_j)\epsilon(s, j) \text{ for any } j \in \mathcal{P} \setminus i. \tag{2}$$

The probability that candidate $i \in P$ is chosen by an agent with attribute s is denoted $i(s, q_i, q_{-i})$ and writes as:

$$i(s, q_i, q_{-i}) = \Pr[\epsilon(s, j) < V(s, q_i)/V(s, q_j)\epsilon(s, i) \text{ for any } j \in \mathcal{P} \setminus i], \tag{3}$$

with q_i the policy of candidate i and q_{-i} the vector of platforms of i's challengers. We denote $F_{(s,i)}$ the cumulative distribution of $\epsilon(s, i)$, which we assume to take support in the set of positive real numbers. The probability $i(s, q_i, q_{-i})$ can be rewritten in the form

$$i(s, q_i, q_{-i}) = \int_0^\infty \prod_{j \in \mathcal{P} \setminus i} F_{(s,j)}(V(s, q_i)/V(s, q_j)\epsilon) dF_{(s,i)}(\epsilon). \tag{4}$$

We assume that the idiosyncratic noises $\epsilon(i, s)$ is iid distributed according to Fréchet distributions,

$$F_{i,s}(\epsilon) = \exp(-t_i \epsilon^{-\theta_s}), \tag{5}$$

for $\epsilon > 0$, $t_i > 0$ and $\theta_s > 1$[6], for any $s \in S$ and $i \in \mathcal{P}$. The parameter $t_i > 0$ relates to the concept of valence in the political economy literature. Indeed, this parameter governs the location of the distribution. A higher t_i implies that candidate i on average has a high popularity among the citizenry, independently form the platform he promises. The parameter θ_s, which we assume independent from the set of candidates, reflects the amount of variation within the distribution. A higher value of θ_s means that the citizens of type s are highly reactive to the platforms announced by the candidates. We label θ_s the *political responsiveness* of the agents with attributes $s \in S$. From (4) and (5), we deduce that $i(s, q_i, q_{-i})$ rewrites as

$$i(s, q_i, q_{-i}) = \frac{t_i V(s, q_i)^{\theta^s}}{\sum_{j \in \mathcal{P}} t_j V(s, q_j)^{\theta^s}}. \tag{6}$$

Using the Law of large number, the vote share of candidate $i \in \mathcal{P}$ can be written as a function $vs_i : \prod_{j \in \mathcal{P}} Q_j \to [0, 1]$, with

$$vs_i(q_i, q_{-i}) = \sum_{s \in S} x_s i(s, q_i, q_{-i}), \tag{7}$$

given that $x_s \in [0, 1]$ is the fraction of agents with attributes $s \in S$ in the population and $\sum_{s \in S} x_s = 1$.

Definition 1. *An electoral equilibrium is such that any candidate promises a platform that maximizes his vote share and each candidate expects his challengers to do the same. The equilibrium*

[6]This restriction is necessary to ensure that the Fréchet distribution has a finite mean.

platform $w_i \in Q_i$ of any candidate $i \in \mathcal{P}$ is such that

$$w_i = \arg\max_{q_i \in Q_i} vs_i(q_i, w_{-i}) \tag{8}$$

for any $i \in \mathcal{P}$ with $w_{-i} \in Q_{-i} = \prod_{j \in \mathcal{P} \setminus i} Q_j$ the vector of optimal platforms of candidate i's challengers.

Theorem 1. *Suppose that there is an election in which (i) there is a finite set of attributes S and $\theta_s > 1$ for any $s \in S$ (ii) there is a finite set of candidates $\mathcal{P}$ and the set of feasible policies Q_i of any candidate i is compact and convex and (iii) the voters have $\mathcal{C}^1$ utility functions $V(s,.)$ such that $V(s,.)^{\theta_s}$ is concave over the set of feasible policies $Q = \prod_{i \in P} Q_i$. There exists an electoral equilibrium $\{w_i\}_{i \in \mathcal{P}} \in Q$. The electoral equilibrium is unique if there exists some $s \in S$ such that the function $V(s,q)^{\theta_s}$ is strictly concave on Q* [7]

Proof. The proof is available in Appendix B.1. □

Assuming that candidates maximize their margin of victory relative to their challengers (i.e. their plurality) will not change the results (see Coughlin (1992) or Coughlin and Nitzan (1981) for such a theory in two-candidate elections). The preceding theorem is a generalization of two results established in the case of two-candidate elections in random voting models with logit distributions by Coughlin (1992, p. 96-97) (theorem 4.2 and Corollary 4.2).

It has to be noted that this theorem applies outside the scope of redistribution strategies or Euclidian political spaces. Components of the candidates' platforms can relate to tax collection, public good provision, redistribution, alternative institutional arrangements, allocation of natural resource revenues, of campaign resources and so forth. Observe as well that the set of feasible platforms need not to be the same for all the candidates. The theorem finally requires the functions $V(s,.)^{\theta_s}$ to be quasi-concave (or strictly concave) for an equilibrium to exist (to be unique). This is stronger than the quasi-concavity (strict concavity) of the utility function $V(s,.)$ when $\theta_s > 1$ (ie. the Fréchet distributions are assumed to have a well defined finite mean). This will be satisfied when θ_s is not too much larger than 1 compared to the concavity of $V(s,.)$. This condition is a

[7] When the function $V(s,q)$ is twice differentiable, this will be the case when the following matrix

$$\left[\frac{\partial^2 V}{\partial q_v \partial q_h}\right] + \frac{(\theta^s - 1)}{V}\left[\frac{\partial V}{\partial q_v}\right] \cdot \left[\frac{\partial V}{\partial q_v}\right]^T$$

defines a semi-defini negative bilinear form.

general feature of a standard probabilistic voting model, namely the fact that the distribution of the stochastic element of the utility model has to be spread enough to ensure the existence of a (unique) equilibrium.

The proof of Theorem 1 easily follows from the convexity of the maximization problem of candidates, given that the functions $V(s,.)^{\theta_s}$ are concave and continuous (the standard existence theorem that is applied is derived from Fudenberg and Tirole (1991, p. 34)). Unicity follows from strict convexity, which is ensured whenever at least one function $V(s,.)^{\theta_s}$ is strictly concave.

As we will see in the sequel, imposing Fréchet distributions for the stochastic components of stochastic voting models, allows tractable applications of multi-candidates competition in several political economy contexts. Beyond the tractability property of these distributions, it is however also interesting to note that an axiomatic rationale can be provided for the use of Fréchet distributions in multi-candidate stochastic voting models. Appendix A presents a formal presentation of this aspect while we simply outline here the logic of the argument. Building on the insights ofMcFadden (1974), we consider a probabilistic voting model that satisfies three behavioral axioms (Independence of Irrelevant Alternatives, Positivity, Irrelevance of Alternative Set). The first two axioms naturally lead to the fact that the voting behavior of an individual can be represented as a Luce voting model (see Luce (1959), Becker et al. (1963)).[8] The third axiom essentially posits that the probability of voting any candidate j relative to some challenger k in pairwise elections depends on three dimensions that caracterizes individual voting behavior: the *popularity* or valence of the candidates, the *quality* of the electoral platforms as evaluated by the voters and the fact that there is a *constant rate of substitution* between popularity and quality. We then show that if we restrict ourselves to the class of distributions $F_{i,s}(\epsilon)$ of individual noises $\epsilon(i,s)$ that are invariant across candidates to multiplicative transformations[9], then a probabilistic voting model satisfies the three previous behavioral axioms if and only if it is a random utility model where the individual noise parameters $\epsilon(i,s)$ are i.i.d distributed according to Fréchet distributions $F_{i,s}(\epsilon) = \exp(-t_i\epsilon^{-\theta_s})$ for some positive

[8]A Luce voting model is such that there exist positive "utility indicator" functions $v(s,i)$ for any candidate $i \in \mathcal{C} \subseteq \mathcal{P}$ such that the probability of $i(s,\mathcal{C})$ being chosen out of the set of candidates $\mathcal{C}$ by the agents with attributes s can be expressed as

$$i(s,\mathcal{C}) = \frac{v(s,i)}{\sum_{j\in\mathcal{C}} v(s,j)}.$$

[9]Formally, we define $\mathcal{C}_d$ the class of distributions such that for any pair of candidates $i,j \in \mathcal{P}$ there exists a strictly positive constant $\alpha_{i,j}$ such that $F_{i,s}(\epsilon) = F_{j,s}(\alpha_{i,j}\epsilon)$ (see Appendix A for details)

parameters t_i and θ_s. This approach therefore provides some micro rationale for the use of random voting utility models with random parameters distributed according to some Fréchet (or extreme type II) distributions.

4 Applications

Our aim in this section is to demonstrate that our analytical framework provides a unifying and tractable approach for various topics of the political economy literature extended to multi-party competition. The two first subsections study special a general model of interest politics with endogenous entry. The third application considers the formation of coalitions in the legislature under proportional rules, while the fourth application compare the endogenous entry of parties under alternative electoral rules.

4.1 Special interest politics with heterogeneous voters and homogeneous entrants

We provide in this application a simple theory of special interest politics that accounts for the endogenous formation of political parties. We assume that the set of individual attributes $s \in S$ characterizes a set of interest groups. Let $S = \{1, \ldots, N\}$. The size of group s is denoted n_s, n denotes the size of the economy and $x_s = n_s/n$ is the fraction of agents of type s. We denote P^e the expected number of parties and posit that the political entrepreneurs creating the political parties have rational expectations. Since $P^e \equiv P$ in equilibrium, we will abuse the notations by denoting P the expected number of parties in the sequel. We will also use the notation $\mathcal{P} \equiv P$ when it is not confusing. Once formed, parties compete for the votes of citizens in a single district plurality election.

We assume that the utility of the agents in the interest group $s \in S$ only depends on their consumption that we denote c_s. Let y be the aggregate income, $y = \sum_{s \in S} n_s y_s$ We denote $\overline{\theta} = \sum_{s \in S} x_s \theta_s$ and $\overline{y} = y/n$ the average political responsiveness and average income respectively. Agents in group s are assumed to have iso-elastic utility function:

$$u(c_s) = c_s^{1-\epsilon}/(1-\epsilon) \tag{9}$$

with $0 < \epsilon < 1$ a parameter that captures the degree of diminishing returns to private consumption.

When ϵ is small, the marginal utility of consumption falls slowly as consumption rises. Therefore, even at high consumption levels, utility levels are significantly affected by variations in consumption. By contrast when ϵ is large, the utility levels of richer groups are less affected by variations in consumption levels.

Party $i \in P$ transfers $\tau_{s,i}$ units of income to any agent with attributes $s \in S$. We assume that these transfers can take negative values although they are necessarily bounded by the initial wealth of group s (i.e. $\tau_{s,i} \geq -y_s$). Thus, when party $i \in P$ is in office, the consumption of any agent in group s is equal to the sum of his revenues plus the group/party-specific transfers,

$$c_s(q_i) = y_s + \tau_{s,i} \tag{10}$$

We assume that the transfers do not sum to zero, as party $i \in P$ extracts a fraction $\chi^i \in [0,1]$ of the tax base as political rents,

$$\sum_{s \in S} n_s \tau_{s,i} = -\chi_i y. \tag{11}$$

In this setting, the policy vector of party i is $q_i = \{\chi_i, \{\tau_{s,i}\}_{s \in S}\}$ and belongs to a convex set Q_i since the feasible transfers are bounded ($\tau_{s,i} \in [-y_s, \sum_{p \neq s} n_p y_p / n_s]$, with the higher bound corresponding to a policy that transfers all the resources to group s with no political rents extracted from the citizenry). Observe additionally that aggregate consumption is such that

$$\sum_{s \in S} n_s c_s(q_i) = y + \sum_{s \in S} n_s \tau_{s,i} = y(1 - \chi_i). \tag{12}$$

Applying the canonical model of section 3, an agent with attributes $s \in S$ prefers candidate i if

$$u(c_s(q_i))\epsilon(s,i) > u(c_s(q_j))\epsilon(s,j) \text{ for any } j \in P \setminus i, \tag{13}$$

with $\epsilon(s,i)$ the i.i.d stochastic element following a Fréchet distribution $F_{i,s}(\epsilon) = \exp(-t_i \epsilon^{-\theta_s})$ for any $s \in S$ and any $i \in P$.

To allow for some aggregate uncertainty in terms of the outcome of the political competition, we assume furtthermore that political preferences do not directly translate into votes. More specifically, we posit that the probability that an agent with attributes s votes for candidate i writes:

$$P_i(s, q_i, q_{-i}) = \eta(s,i,P)\Pr[\epsilon(s,j) < u(c_s(q_i))/u(c_s(q_j))\epsilon(s,i) \text{ for any } j \in P \setminus i], \tag{14}$$

with $\eta(s,i,P) \in [0,1]$ the probability that an agent with attributes s that prefers candidate i in the set P effectively goes to the ballot. This additional parameter translates the idea that the probability that an agent goes to the ballot depends both on his characteristics (i.e. on s) and on the candidate he is expecting to support. Indeed, there is always some randomness associated to each candidate on how he can effectively mobilize people who support him, to go to the ballot box (this may reflect the discrepancy between opinion surveys and actual vote behavior and the fact effective political participation has a stochastic component, independent from policies and social outcomes).

As a matter of simplification, we assume that $\eta(s, i, P)$ is independent of $s \in S$, meaning that any agent that intends to vote for candidate $i \in P$ has the same likelihood of casting a ballot. We use the notation $\eta(s, i, P) \equiv \eta(i, P)$ hereafter. Using the formalization of the preceding section, we can rewrite the probability that an agent with attributes s votes for candidate i as

$$P_i(s, q_i, q_{-i}) = \eta(i, P) \frac{t_i u(c_s(q_i))^{\theta_s}}{t_i u(c_s(q_i))^{\theta_s} + \sum_{j \in P \setminus i} t_j u(c_s(q_j)),^{\theta_s}} \tag{15}$$

and using the law of large number, we can deduce that the vote share of candidate i takes the form

$$vs_i(q_i, q_{-i}) = vs_i^t(q_i, q_{-i}) \eta(i, P), \tag{16}$$

with

$$vs_i^t(q_i, q_{-i}) = \sum_{s \in S} \frac{n_s}{n} \frac{t_i u(c_s(q_i))^{\theta_s}}{t_i u(c_s(q_i))^{\theta_s} + \sum_{j \in P \setminus i} t_j u(c_s(q_j))^{\theta_s}}. \tag{17}$$

As per (16), the vote share of candidate i is the product of a "theoretical" vote share vs_i^t, which is the fraction of people that prefer candidate i in the population, with the probability $\eta(i, P)$ that those people effectively go to the ballot box. We posit that the probability $\eta(i, P)$ takes the following form:

$$\eta(i, P) = \frac{\mu_i}{\sum_{k \in P} \mu_k}, \tag{18}$$

with μ_i a warm glow for casting a ballot for candidate i that we also assume distributed according to a Fréchet distribution, $G(\mu) = \exp(-\mu^{-K})$, with $K > 1$. Given this, one may express the probability for a candidate i to get the largest fraction of effective votes as

$$G_i(q_i, q_{-i}) = \Pr[\mu_j < vs_i^t(q_i, q_{-i}) / vs_j^t(q_i, q_{-i}) \mu_i) \text{ for any } j \in \mathcal{P} \setminus i], \tag{19}$$

which can be rewritten as:

$$G_i(q_i, q_{-i}) = \int_0^\infty \left[G((vs_i(q_i, q_{-i})/vs_j(q_i, q_{-i})\eta))\right]^{P-1} dG(\mu) \tag{20}$$

$$= \frac{\left[vs_i^t(q_i, q_{-i})\right]^K}{\sum_{j\in\mathcal{P}} \left[vs_j^t(q_i, q_{-i})\right]^K}. \tag{21}$$

An equilibrium $(q_i^*)_{i\in P} = \{\chi_i^*, \{\tau_{s,i}^*\}_{s\in S}\}_{i\in P}$ is then a set of platform such that each party seeks to maximize its expected rents given what the other parties propose to the voters. So we have

$$q_i^* = \{\chi_i^*, \{\tau_{s,i}^*\}_{s\in S}\} = \underset{q_i=\{\chi_i, \{\tau_{s,i}\}_{s\in S}\}}{\arg\max}\; y\chi_i \cdot G_i(q_i, q_{-i}^*), \tag{22}$$

with q_{-i}^* the vector of optimal platforms of i's challengers.

We assume homogeneous candidates in the election, meaning that $t_i = t_j$ for any $i, j \in P$. Hence

$$vs_i^t(q_i, q_{-i}) = \sum_{s\in S} \frac{n_s}{n} \frac{u(c_s(q_i))^{\theta_s}}{u(c_s(q_i))^{\theta_s} + \sum_{j\in P\setminus i} u(c_s(q_j))^{\theta_s}} \tag{23}$$

and we characterize a symmetric equilibrium in policies q_i. The first-order condition with respect to the transfers to group s writes:

$$-n_s G_i(q_i, q_{-i}^*) + y\chi_i \left[\frac{\partial G_i}{\partial vs_i}\frac{\partial vs_i}{\partial \tau_{s,i}} + \sum_{j\neq i} \frac{\partial G_i}{\partial vs_j}\frac{\partial vs_j}{\partial \tau_{s,i}}\right] = 0 \tag{24}$$

Condition (24) says that when party i marginally increases the transfers to group s, on the one hand, it reduces the expected level of extraction by an amount $n_s G_i(q_i, q_{-i}^*)$. On the other hand, party i grabs an increased vote share from his challengers (the second term in bracket), which comes from an increase in his own vote share vs_i and a decrease in the vote shares of all of his the challengers for a given interest group s. We show the following result in the Appendix.

Proposition 1. *A symmetric equilibrium necessarily exists when parties are homogeneous and is such that for any party i, and any $k, p \in S$*

$$\frac{\theta_s}{c_s(q_i)} = \frac{\theta_p}{c_p(q_i)} = \frac{1}{\frac{y}{n}\chi_i K(1-\epsilon)(1-G_i)}. \tag{25}$$

and the consumption of the agents in group s as a function of the rents extracted is

$$c_s(q_i) = \frac{\theta_s}{\overline{\theta}}\overline{y}(1-\chi_i). \tag{26}$$

In equilibrium, each group $s \in S$ gets a fraction $\theta_s/\overline{\theta}$ of the transferable revenues. This result derives directly from the fact that - notwithstanding the effect of the responsiveness parameters $\{\theta_s\}_{s\in S}$ - there always is a higher marginal benefit at targeting the poorest groups, because they have lower consumption levels.[10] The optimal redistribution scheme can then be understood as proceeding in two steps. First, it consists in neutralizing the effect of the income distribution on the vote share by means of transfers. Then, it redistributes resources according to the pattern of responsiveness parameters across income groups. Ex-post and ex-ante income inequalities are consequently independent since the latter reflects the political behaviors of the existing interest groups. This result obtains because parties have no vested interests with respect to the various interest groups, and turn-out is independent from the income distribution. The transfers are not necessarily directed from the high income to the low income groups. Indeed, if the rich are significantly more politically responsive than the poor, then we should exactly expect the opposite.

Given that the electoral equilibrium is convergent, by substituting (26) in the first-order condition, we find that the rents χ_i are such that

$$\chi_i = \frac{1}{1+\overline{\theta}K(1-\epsilon)(P-1)/P}. \tag{27}$$

The political rents χ_i decrease with the number of parties competing in the election. Indeed, the marginal loss of the parties from extracting more rents out of the citizenry is higher when each party has more competitors.

In order to derive a closed form result for the number of parties competing in this election, we assume a fixed cost of party formation c. Parties decide to compete for an election as long as the expected utility from doing so is higher than their cost of formation c. Consequently, the equilibrium number of parties P^* is the highest integer below the solution of

$$\frac{1}{P}\chi_i y = c, \tag{28}$$

with χ_i given in (27). Indeed, the left hand side of (28) is the utility derived by a given party i from entering the race when P parties are expected to compete (including i). As represented by the blue curve in figure 2, this utility decreases with P, as the rents χ_i and the vote share $1/P$ are both decreasing with P. The blue curve crosses only once the horizontal line c, hence the unicity of the equilibrium number of parties P^*.

[10] This effect has been noticed already by Dixit and Londregan (1996).

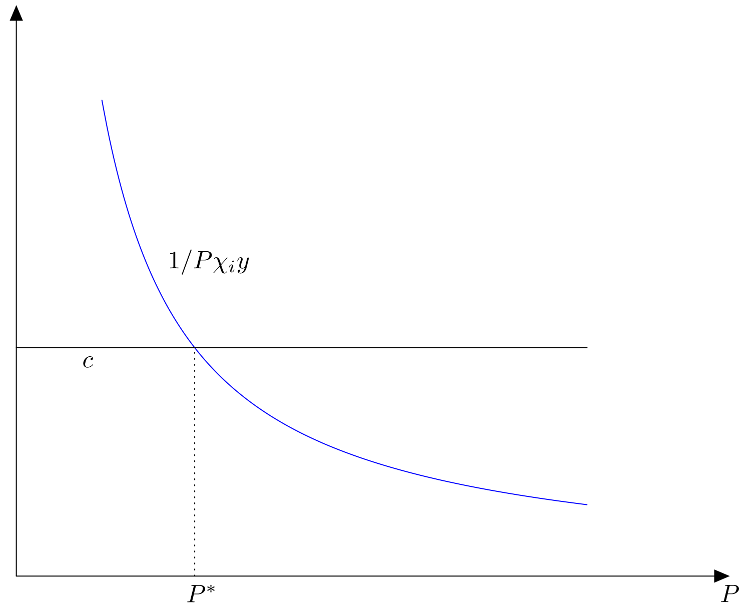

Figure 2: Characterization of the equilibrium number of entrants P^*.

Proposition 2. *s The number of parties competing in the election is the highest integer below P^* with*

$$P^* = \frac{1}{1 + K\overline{\theta}(1-\epsilon)}[K\overline{\theta}(1-\epsilon) + \frac{y}{c}] \tag{29}$$

if $y > c$, and $P^* = 0$ otherwise.

- *P^* decreases with the cost of party formation c, with the average responsiveness of the citizenry $\overline{\theta}$ and with the shape of the aggregate uncertainty K of political participation. Alternatively, P^* increases with the income of the citizenry y, and with the degree of diminishing returns to private consumption ϵ.*

- *An increase in the size of any group n_s that has a lower than average responsiveness $\theta_s < \overline{\theta}$ increases the number of parties while the effect of an increase in the size of a group with a higher than average responsiveness is ambiguous.*

Proof. The proof is available in Appendix B.4. □

Note first that if $y < c$, there is no party formation because the entry cost c is higher than the maximum level of rents that can be extracted by any party from the citizenry. We assume that the condition $y < c$ is respected in the sequel.

Relative to the comparative statics, a higher value of the cost of party formation straightforwardly affects negatively the fraction of parties P^*. Additionally, a higher $\overline{\theta}$ increases the marginal cost at capturing rents, since it implies that voters are more responsive on average to political platforms in their voting behavior. By the same token, an increase in the responsiveness of any group θ_s affects negatively the number of parties in equilibrium. Similarly, whenever the degree of diminishing returns to private consumption ϵ increases, then the agents care less about higher consumption levels and the marginal political cost of capturing rents decreases for the existing parties. This, in turn, affects positively the number of parties willing to enter the political arena in equilibrium. Finally, if the average income increases, then so does the amount of rents that can be extracted and party fragmentation increases.

Regarding the last point of the proposition, observe that an increase in the size of a group n_s affects both the tax base and the average responsiveness. Indeed, on the one hand, when n_s increases, so does the tax base and this tends to increase P^*. On the other hand, an increase in n_s also affects the average responsiveness $\overline{\theta}$, which will increase when $\theta_s > \overline{\theta}$ and decrease otherwise. Thus, when the size of a group that has a lower than average responsiveness increases, then the number of parties in equilibrium necessarily increases because (i) the taxable income increases and (ii) the average responsiveness decreases, so the marginal cost of capturing rents goes down. Alternatively, when $\theta_s > \overline{\theta}$, the taxable income still increases, while the average responsiveness now tends to decrease. The overall effect of n_s on P^* is then ambiguous.

We close this first application by substituting P^* in (27). We find the equilibrium political rents η^* and the consumption levels c_s^*, $s \in S$ when the endogenous structure of the political market in accounted for:

$$\eta^* = y\chi_i = \frac{1}{1 + (1-\epsilon)K\overline{\theta}}((1-\epsilon)K\overline{\theta}c + y), \tag{30}$$

and

$$c_s^* = \frac{K\theta_s(1-\epsilon)}{n(1 + K\overline{\theta}(1-\epsilon))}(y - c) \tag{31}$$

for any $s \in S$.

First, when the voters in a given interest group are more reactive, they attract more transfers, reduce the transfers made to the other interest groups, and discipline the existing parties who capture less political rents. This result however is not as obvious as in a model with an exogenous number of entrants. Indeed, when a group becomes more reactive, there are two effects to consider. First, the parties are incentivized to direct more resources toward that group, which is a standard effect. Second, less parties are expected to form in equilibrium. A weaker political competition, in turn, decreases the incentive of the parties to transfer resources to the citizenry. Relative to a model with an exogenous political structure, the current model shows that assuming an exogenous political structure over-estimates the effect of voters' reactivity on public choices.

Similarly, if it is true that a higher expected turn-out (i.e. a higher value of K) increases the transfers made to the citizenry, the effect is attenuated in a theory with an endogenous political structure, since a higher turn-out also decreases the fragmentation of the polity.

Finally, this model allows to study the effect of party formation costs on public policies. It is clear that the parameter c has a positive effect on rent capture. Indeed, when the cost of party formation is higher, then parties are willing to enter the political arena only if it allows them to capture higher rents. In equilibrium, this leads to weaker political competition and this lowers transfers to the citizenry. Observe that the parties have an interest in deterring further entrance in the political arena by raising c, while the citizenry would prefer low entry costs so as to strengthen the political competition.

As a simple illustration of the preceding result, Scherlis (2014) argues that a legitimacy crisis and popular discontent led to important reforms aiming at reducing party formation costs across Latin American countries in the 1990s. The new Colombian Constitution enacted in 1991 virtually abolished entry barriers to democratic competition by recognizing social movements and other groups of citizens as equivalents to parties (Scherlis (2014)). As of 1994, 50,000 signatures or 50,000 votes in the preceding elections were required to obtain legal recognition. Furthermore, anyone could register a candidacy, even without legal party recognition, by paying a sum to be refunded if a threshold of 50,000 votes is reached. Finally, multiple lists from the same party were authorized as well. As a result, the number of lists competing for the Senate and the House of Representatives steadily increased from 1990 to 2002. Interestingly, once policymakers and parties regained legitimacy, new laws aiming at reinstating high party formation costs where voted. Indeed, President Uribe, supported by conservatives, liberals, and a part of the leftist Democratic Pole reformed the party system so as to reduce the fragmentation of the polity. The threshold for legal recognition was increased to 2% of the votes for example. Similar reversal of the reforms on the costs of party formation are observed in periods of economic growth in Argentina, Mexico and Peru (Scherlis (2014)).

Our framework may also be related to the recent political reform voted in Brazil by the Congress in October 2017. Brazil has a multiparty political system, mixing both presidential/parliamentary dimensions. While executives and senators are elected according to a presidential system,[11] lower house deputies are elected by proportional representation with an open list system[12]. A well known feature of Brazil is the extreme party fragmentation of its political system (Figueiredo and Limongi (2000)). For instance, after the 2014 elections, the number of parties represented in Congress grew from 23 to 28 (Vaz de Melo (2015)). This highly fragmentation situation has obviously attracted the attention of several political science studies (see Junior et al. (2015) for a recent review). Part of the explanations provided by this literature highlights specific institutional features, such as open-list proportional representation, incumbents' guaranteed re-election rights, possibility of having more candidates than contested seats, cross party mobility with no cost, and minimal threshold for electoral representation,[13] that create system weaknesses with low political entry and moving costs, leading to a myriad of catch-all parties with consequent weak legislative disciplinary, political instabilities and policy inefficiencies between the executive and the legislature (Ames (2001), Mainwaring (1999), Mainwaring and Shugart (1997)). Other authors have emphasized countervailing stabilizing forces, such as the role of coalition voting discipline (Amorim Neto (2002)), the importance of cabinet formation or executive appointment powers (Martinez-Gallardo (2005), Amorim Neto (2006)), and the use of rent sharing and 'pork' barrels mechanisms to glue legislative coalitions (Raile et al. (2011)). Following nationwide corruption scandals[14], and a subsequent legitimacy crisis of the main political parties, the current government of President Michel Temer recently sanctioned a series of electoral reforms approved by the federal congress, with the aim at reducing the number of political parties operating in Brazil, putting cap campaign financing, and limiting the power of coalitions.

[11] More precisely, presidents, governors, mayors win by majority runoff. Senators are elected by a plurality of the vote.

[12] Seats are awarded in proportion to the votes that each coalition wins, but the candidates who win seats are those who win the most votes within each coalition.

[13] such as open-list proportional representation, incumbents with guaranteed re-election rights, possibility of having more candidates than contested seats, cross party mobility with no cost, minimal threshold for attaining one seat.

[14] Among which is the notorious Petrobras *Lava Jato* (ie. *Car Wash*) scandal.

In particular, elements of the reform include the creation of a public fund to finance campaigns, in compensation for a Supreme Court ban on corporate donations . The reform also introduces a cap on candidates' spendings, as well as the abolition of party coalitions by 2020.

Interestingly, to reduce political fragmentation and filter out some of the smaller parties, a minimum performance is introduced in order to be given TV and radio time as well as access to the public fund. More precisely, for the general election in 2018, parties will need to obtain at least 1.5% of valid votes across at least nine states, with at least 1% of the valid votes in each one of them to gain access to the shared campaign fund plus their quota of free radio and TV propaganda during electoral campaigns. This threshold will then increase incrementally over the years to 3% of the vote by 2030 (Melo (2017)).

While big parties might first suffer from the loss of corporate donations, the introduction of this new electoral threshold is likely to increase the cost of entry in politics and to favor in the long run the biggest parties in the legislature. It is argued that such reform from an efficiency point of view, is expected to lessen the need for pork and patronage associated with legislative coalition building with multiple parties. At the same time, this reform is also perceived as a reaction by the main parties to restore a competitive political advantage against the legitimacy crisis associated to corruption scandals and the subsequent judicial ban of their main source of financing through corporations. Interestingly in the process, the new electoral funding rules may also raise differentially the cost of entry across parties. In particular, it may benefit incumbents and parties that have already a loyal basis (like the Evangelical parties) at the expense of candidates who need cash to promote themselves.[15]

4.2 Special interest politics with heterogeneous entrants

In this section, we consider two extensions of the preceding framework. First, we examine a case where parties can decide to increase their popularity by investing resources during the electoral campaign. Second, we consider a relatively analogous situation where parties differ in their ability to mobilize their electorate. More specifically, we study an extension of the preceding framework where parties can increase their expected turn-out through higher investments during the electoral

[15] See *Elections on a shoestring: Brazil's congress starts to reform itself*, The Economist (2017). It is also argued that the current political reform misses other important dimensions such as divide statewide congressional districts, and correct for over-representation of small states in the lower house (The Economist (2017)).

campaign. Our objective is to investigate how heterogeneity in terms of popularity or turn-out affects public choices and the fragmentation of the polity. To concentrate on the degree of heterogeneity on the supply side of the political market, we assume that voters are homogeneous (ie. $\mathbf{S} = 1$). Because of this, the platform of any party i reduces to a level of rent extraction χ_i.

We consider that the parties can choose before the election to invest a high or a low amount of resources in media campaigns, meetings and other investments. These resources allow them to gain popularity or to mobilize their electorate. Formally, we can consider two cases:

a) In the first case, a party pays a formation cost c_h (resp. c_l), and benefits from a popularity t_h (resp. t_l) in the election, with $c_l < c_h$ and $t_l < t_h$. We denote P_l and P_h the number of parties with a low and a high reputation respectively, and χ_l (respectively χ_h) the level of rent extraction of a low (respectively a high) reputation candidate. Following the same procedure as in the previous section, the equilibrium probability of elections are easily written as:

$$\begin{aligned} G_h(t_l, t_h, P_l, P_h) &= \frac{t_h^K}{P_h t_h^K + P_l t_l^K Z(\chi_h, \chi_l)} \quad \text{and} \\ G_l(t_l, t_h, P_l, P_h) &= \frac{t_l^K}{P_l t_l^K + P_h t_h^K / Z(\chi_h, \chi_l)}, \end{aligned} \tag{32}$$

with

$$Z(\chi_h, \chi_l) = (\frac{1-\chi_l}{1-\chi_h})^{(1-\epsilon)\theta K}. \tag{33}$$

Parties maximize their expected rents, and the optimal levels of extraction χ_l and solve the following system of first-order conditions:

$$1 - \frac{1}{n}\frac{\chi_h}{(1-\chi_h)} K(1-\epsilon)\theta \left(1 - \frac{t_h^K}{P_h t_h^K + P_l t_l^K Z(\chi_h, \chi_l)}\right) = 0 \tag{34}$$

for the P_h high reputation parties, and

$$1 - \frac{1}{n}\frac{\chi_l}{(1-\chi_l)} K(1-\epsilon)\theta \left(1 - \frac{t_l^K}{P_l t_l^K + P_h t_h^K / Z(\chi_h, \chi_l)}\right) = 0 \tag{35}$$

for their P_l low reputation challengers. As well, the free entry conditions for the high and low reputation candidates write respectively as:

$$\begin{aligned} G_h(\chi_h, \chi_l, P_l, P_h)\chi_l y &= c \\ G_l(\chi_h, \chi_l, P_l, P_h)\chi_l y &= c \end{aligned} \tag{36}$$

We then deduce two loci $\chi_h(\chi_l, t_l, t_h, P_l, P_h)$ and $\chi_l(\chi_h, t_l, t_h, P_l, P_h)$ from the first-order conditions (34) and (35). From (36), we are finally able to characterize the free entry equilibrium degree of political fragmentation.

b) Alternatively, one may also consider the case where a party may increase his expected turn-out. Specifically, we may assume that when a party pays a formation cost c_h (resp. c_l), then the warm glow for casting a ballot for that party is distributed according to a Fréchet distribution $G(\mu) = \exp(-\alpha_h \mu^K)$ (resp. $G(\mu) = \exp(-\alpha_l \mu^K)$), with $\alpha_h > \alpha_l$.[16] In this case, the equilibrium probabilities of elections for the high and low types are respectively:

$$\begin{aligned} G_h(t_l, t_h, P_l, P_h) &= \frac{\alpha_h t_h^K}{P_h \alpha_h t_h^K + P_l \alpha_l t_l^K Z(\chi_h, \chi_l)} \quad \text{and} \\ G_l(t_l, t_h, P_l, P_h) &= \frac{\alpha_l t_l^K}{P_l \alpha_l t_l^K + P_h \alpha_h t_h^K / Z(\chi_h, \chi_l)}, \end{aligned}$$

with

$$Z(\chi_h, \chi_l) = (\frac{1-\chi_l}{1-\chi_h})^{(1-\epsilon)\theta K}. \tag{37}$$

As can be seen, these are precisely the same formulas as in the previous case with $\alpha_i t_i^K$ substituted for t_i^K. The derivation of the electoral equilibrium with endogenous party entry follows therefore the same steps as to those of case a).

We then establish the following result (described more formally in the Appendix).

Proposition 3. •

- *When the parties can increase their popularity by means of campaign investments, the equilibrium is uniquely determined and the equilibrium fragmentation of the polity is expected to decrease. Popular and unpopular parties coexist in equilibrium, but popular parties trump out unpopular parties and capture more political rents.*
- *Similarly, when the parties can increase their expected turn-out by means of campaign investments, the equilibrium is uniquely determined and the equilibrium fragmentation of the polity is expected to decrease. High and low turn-out parties coexist, while high turn-out parties trump out low turn-out parties and capture more political rents in equilibrium.*

Proof. The proof is available in Appendix B.5. □

Relative to the first point of the proposition, the parties that invest in increasing their popularity pay a higher cost for entering the race, while they become more popular than their challenger. This

[16] The parameter α_i positively affects the average value of a random variable μ distributed according to a Fréchet distribution with c.d.f. $G(\mu) = \exp(-\alpha_i \mu^K)$.

has two main consequences on the structure of the political competition. First, the parties that choose to run a “low-cost” campaign expect to face a tougher competition and they decrease the rents that they extract from the citizenry. Second, and symmetrically, the popular parties face a weaker competition and extract more rents from the citizenry. As the low-cost parties have a lower cost for entering the election, it can still be profitable for them to enter, despite the fact that they capture lower rents. Furthermore, the heterogeneity implies that relative to a case where the parties can not run high-cost political campaigns, the party fragmentation is lower. Observe however that it does not necessarily mean that the citizenry is negatively affected, since the decrease in the political competition that is induced by the presence of popular parties is compensated by lower rent capture from the unpopular ones.

The intuition for the second point of the proposition is very similar. The parties that pay an additional cost to mobilize their electorate are able to capture more rents, and they face weaker political competition. Because of this, when the parties have different abilities to mobilize their electorate, we should expect a lower degree of fragmentation of the polity.

4.3 Endogenous entry and coalition formation in the legislature in proportional systems

In this section we examine the formation of coalitions in the legislature under proportional electoral rules. Arguably, the possibility to form coalitions in the legislature should affect parties' incentive to run as well as their political platforms. As a simple illustration, we should expect the incentive of small parties to run in proportional elections to be higher despite their low expected seat shares in the legislature when they anticipate that they might be able to participate to governing coalitions. This application will show that this intuition is not entirely true. Besides, to our knowledge, the “first stage” of the coalition formation game in the legislature that accounts for the formation of parties has not been apprehended yet in the literature. This application fills this gap.

We assume that the parties get a fraction of the seats in the legislature that equalizes their vote share in a single election. We do not introduce thresholds for simplicity. One party is chosen in the legislature to form a government. This party is labeled the *formateur* hereafter. We assume that the probability that a given party is chosen to form a governing coalition is equal to its seat share in the legislature. This assumption has been made by Baron and Diermeier (2001) for instance, and finds empirical support in the analysis of Diermeier and Merlo (2004). If the formateur does not manage to form a government, then each legislator gets some exogenous rents r and the game ends. We posit that those rents are independent from the seat share of the parties in the legislature. Once chosen, the formateur makes a take-it-or-leave-it offer to some other parties in the legislature so as to form a minimum winning coalition. A legislator will therefore accept to be part of the governing coalition if the formateur makes an offer that is above and arbitrarily close to r.

We focus on the case of special interest politics with heterogeneous voters and homogeneous candidates (see the first application). The set of individual attributes $s \in S$ characterizes a set of interest groups. The utility of the agents in the interest group $s \in S$ only depends on their consumption that we denote c_s. Party $i \in P$ transfers $\tau_{s,i}$ units of income to any agent with attributes $s \in S$ and extracts a fraction $\chi_i \in [0,1]$ of the tax base as rents. The strategy of a party consists in setting a policy vector $q_i = \{\chi_i, \{\tau_{s,i}\}_{s\in S}\}$ and in building a minimum winning coalition C_i in case it is chosen to be the formateur. C_i is a subset of the set of parties in the legislature that we still denote P. The objective of party i can be written in the form

$$\max_{C_i,q_i} W(C_i,q_i) = vs_i(q_i,C_i,q^*_{-i},C^*_{-i})[\chi_i y - rP \sum_{j\neq i, j\in C_i} vs_j(q^*_j,C^*_j,q_{-j},C_{-j})]+ \\ \sum_{j\neq i} vs_j(q^*_j,C^*_j,q_{-j},C_{-j})\Pr(i\in C^*_j) rPvs_i(q_i,C_i,q^*_{-i},C^*_{-i}), \quad (38)$$

under the constraints

$$\begin{cases} -\sum_{s\in S} x_s\tau_{s,i} \geq \chi_i\overline{y} + rP\sum_{j\neq i,j\in C_i} vs_j(q^*_j,C^*_j,q_{-j},C_{-j}) \\ \sum_{j\neq i,j\in C_i} vs_j(q^*_j,C^*_j,q_{-j},C_{-j}) + vs_i(q_i,C_i,q^*_{-i},C^*_{-i}) \geq 1/2 \\ \max_{C_i,q_i} W(C_i,q_i) \geq \max_{C_i,q_i} rPvs_i(q_i,C_i,q^*_{-i},C^*_{-i})\{1+\sum_{j\neq i} vs_j(q^*_j,C^*_j,q_{-j},C_{-j})\Pr(i\in C^*_j)\}, \end{cases} \quad (39)$$

where q^*_j and C^*_j are respectively the optimal policy and the minimum winning coalition of party j. The vector q_{-j} is such that $q_{-j} = \{q^*_1,\ldots,q_i,\ldots,q^*_{j-1},q^*_{j+1},\ldots,q^*_P\}$. Similarly, C_{-j} denotes the vector of optimal coalitions of j's challengers given that i does not have an optimal coalition.

The first constraint is a standard budget constraint. The formateur extracts resources from the citizenry, takes some rents and transfers r units of revenues to a set of legislators that belongs to the parties in the coalition C_i. The second constraint says that a minimum winning coalition C_i is such that the seat share of C_i is at least equal to one half. Indeed, assume that the policies are enacted by majority voting in the legislature. It is strictly suboptimal for the legislators that are not part of the governing coalition to vote for the platform of the formateur because it incentivizes him to buy their support. Furthermore, observe that the formateur will not make offers below r in equilibrium, precisely because any party that belongs to the coalition is median in that if it leaves, then the coalition loses the majority in the legislature. Finally, there is still an incentive compatibility constraint, since building a minimum winning coalition should provide higher rents. Indeed, it could be that the formateur prefers not to build a coalition, in which case he gets an exogenous level of rents r. This incentive compatibility constraint is described in the third line of (39). Observe nevertheless that if r is sufficiently low relative to the perks of passing a tax policy in the legislature, then this outcome is unlikely. We will posit hereafter that the third constraint is always respected, so that formateurs prefer to form coalitions.

Relative to the objective (38), if party i is chosen to be the formateur with probability vs_i - and given that it builds a minimum winning coalition, it will be able to set the tax policy and to fix the level of rents to χ_i. Furthermore, an amount r will be transferred to the $rP\sum_{j\neq i,j\in C_i} vs_j$ legislators. Relative to the second line of (38), with probability vs_j, party j is chosen to be the formateur, in which case the Pvs_i legislators of party i receive a transfer r if they are chosen to be part of j's governing coalition. We denote $\Pr(i \in C_j^*)$ the probability that i is chosen to be part of C_j^*, for $j \neq i, j \in P$

For simplicity, we assume that the parties have the same popularity, i.e. $t_i = t_j$ for any $i, j \in P$. Under those conditions, parties are homogeneous. We moreover focus on the determination of the symmetric electoral equilibrium, if it exists, since the global concavity of any party's objective cannot be obtained in a simple way.

In a symmetric electoral equilibrium, parties have the same vote share and thus any party $j \in P$ has an equal chance of participating to a coalition initiated by any party $i \in P$. As before, the number of parties will be the largest integer below a real number P that is endogenously determined. We denote $E(P)$ the largest integer below P and assume that there are at least two parties represented in the legislature. In those conditions, it is direct to have that

$$\Pr(i \in C_j^*) = \binom{E(P)-2}{E(E(P)/2)-2} \Big/ \binom{E(P)-1}{E(E(P)/2)-1}. \tag{40}$$

The denominator gives the total number of subgroups of size $E(E(P)/2)-1$ in a set of size $E(P)-1$. Thus, the denominator of (40) gives the number of possible coalition of size $E(E(P)/2)$ that includes party j. It represents the number of minimum winning coalitions that include a given party j. By analogy, the numerator is the set of winning coalitions that include both j and i. Indeed, if i belongs to the coalition formed by party j, then party j still has to choose $E(E(P)/2)-2$ other parties for the coalition among the set of remaining $E(P)-2$ parties. $\Pr(i \in C_j^*)$ simplifies to

$$\Pr(i \in C_j^*) = \frac{E(P)/2-1}{E(P)-1}. \tag{41}$$

Combining (38) and (39), and using the symmetry assumption, we can rewrite the objective of party i as

$$\max_{C_i,q_i} W(C_i,q_i) = vs_i(q_i,C_i,q_{-i}^*,C_{-i}^*)[\chi_i y - rP(1/2 - vs_i(q_i,C_i,q_{-i}^*,C_{-i}^*)]+$$
$$(1-vs_i(q_i,C_i,q_{-i}^*,C_{-i}^*))\frac{E(P)/2-1}{E(P)-1} rPvs_i(q_i,C_i,q_{-i}^*,C_{-i}^*). \tag{42}$$

We assume that P is an even integer for simplicity in the sequel. In order to gain intuitions on the mechanisms at play, we follow the steps of the first application and write the first-order condition with respect to the transfers $\tau_{s,i}$:

$$-\frac{1}{P}n_s + (1-\epsilon)\frac{n_s}{n}\frac{\theta_s}{c_s}\{\chi_i y + \frac{1}{2}r\frac{P}{P-1}\}\frac{P-1}{P^2} = 0. \tag{43}$$

Increasing the transfers to the interest group s still decreases the level of extraction (first term in the LHS of (43)). Relative to the first application however, the marginal benefits from increasing the transfers to group s changes. Indeed, when party i increases marginally its vote share, (i) it increases the likelihood of being the formateur (ii) reduces marginally the cost of building a winning coalition and (iii) increases the transfers that can be gained from participating to a governing coalition without being the formateur. This is why the second term in the LHS of (43) is higher than in the first application, since when there is no coalition formation, only effect (i) is at play. As in the first application, it follows immediately from the first-order condition that

$$\theta_s/c_s = \theta_k/c_k \tag{44}$$

for any pair of interest group $s,k \in S$. Using the budget constraint then, we deduce that

$$c_s = \frac{\theta_s}{n\bar{\theta}}(y(1-\chi_i) - r(\frac{P}{2}-1)). \tag{45}$$

Not surprisingly, since parties have to invest resources in order to build a winning coalition, the consumption level c_s is decreasing in the number of parties in the legislature and in the exogenous rents r that have to be redistributed to the parties in governing coalitions. Observe then from (43) that we should expect the marginal benefit of transferring resources to any interest group s to be larger than in the case of the first application because consumption levels are anticipated to be lower. Following the steps of the first application, we can deduce from this point the level of rent extraction χ_i and the number of parties P_c^* in equilibrium. The computations are detailed in Appendix B.6.

Proposition 4. *In equilibrium, the number of parties is the largest integer below P_c^*, with*

$$P_c^* = \frac{1}{1 + (1-\epsilon)\overline{\theta} + r/2c}[(1-\epsilon)\overline{\theta} + \frac{y + r(1 - \overline{\theta}/2)}{c}]. \tag{46}$$

The comparative statics of the first application are robust to the formation of coalitions. Furthermore, the number of parties decreases with the rents of legislators r. The formation of coalitions in the legislature reduces party fragmentation and there are strictly more parties under plurality than proportional electoral rules.

Proof. The proof is available in Appendix B.6. □

Our analysis of stochastic and sincere voting contradicts Duverger's prediction that proportional systems should favor higher party fragmentation. At the center of our analysis lies the interaction between public policy, coalition formation and party formation. In particular, introducing coalitions creates stronger incentives for parties to increase their vote share (and then to decrease rent extraction). Furthermore, the cost of coalition formation is partially born by the citizenry as it lowers the feasible transfers (and in turn strengthen even more parties' incentive to reduce their extraction level). This result suggests that, out of the two reasons given by Duverger for his prediction, the second so-called "psychological effect" by which voters strategically avoid wasting their votes by abandoning small parties and concentrating their votes to larger parties (Duverger, 1964: 206-55) may actually be an important component for the result to hold. The other "mechanical effect" related to how electoral rules translate votes into seats may not be enough.

To conclude with an illustration, take the case where $r/c = 3$, $y/c = 5$ and $\overline{\theta} = \epsilon = 0.5$. In that case, we should expect only 2 parties to form under the proportional system, while in a plurality system, there should be as much as 6 parties from Proposition 2.

4.4 Redistributive policies under runoff and plurality elections

The Duverger's hypothesis (second law), as formulated by Riker (1982), states that runoff elections should produce more candidates than plurality elections. Furthermore, the Duverger's first law states that simple majority single ballot plurality favors a two-party system whereas simple majority with a second ballot or proportional representation favors multipartysm (Riker (1982)). As said, Duverger (1964) gave two reasons for his hypotheses to hold. The first is a "mechanical effect" associated to the fact that in a plurality electoral system, small parties are penalized by the fact that only the candidate obtaining the largest number of votes in each (single-seat) constituency gets elected. The second effect is a so-called "psychological effect", by which voters avoid wasting their votes on small parties and concentrate their votes on larger parties (Duverger, 1964: 206-55). In this application, we intend to study whether our theory of stochastic and sincere voting is consistent with Duverger's hypothesis. By preventing voters from being strategic, we investigate whether the "mechanical effect" related to institutional factors translating votes into seats, is central to Duverger's argument.

We compare the endogenous entry of political parties in runoff and plurality elections. The case of plurality elections has been developed in the canonical model of Section 3 and applied to redistributive politics with homogeneous candidates in Section 4.1. The runoff system by contrast proceeds in two rounds. In the first round, the two candidates with the highest vote shares are selected and are allowed to run for the second round. This is the system used in the French presidential election for example. We first develop a general framework for the study of runoff elections and then apply it to a simple case of redistributive politics with homogeneous candidates and heterogeneous voters. As a matter of simplification, we assume that in runoff elections, parties promise platforms before the first round and cannot back pedal on their promises between the two rounds and citizens have to show up to the ballot twice.

Let $ij(q_i, q_{-i})$ be the probability that the citizenry ranks i and j first given the list of possible candidates P. Following the same steps as in the first application in section 3, we denote $vs_i(q_i, q_{-i})$ the fraction of voters that cast a ballot for candidate i. This is given by:

$$vs_i(q_i, q_{-i}) = vs_i^t(q_i, q_{-i})\eta(i, P),$$

with $vs_i^t(q_i, q_{-i})$ the share of individuals in society who prefer party i in the set of P candidates, and $\eta(i, P)$ is the probability that an agent who prefers candidate i in the set P effectively goes to the ballot to cast his vote. The probability $ij(q_i, q_{-i})$ can then be expressed as:

$$ij(q_i, q_{-i}) = \Pr[\min(vs_i^t(q_i, q_{-i})\eta(i, P), vs_j^t(q_j, q_{-j})\eta(j, P)) > \\ vs_k^t(q_k, q_{-k})\eta(k, P) \text{ for any } k \in P \setminus \{i, j\}].$$

Using the same expressions for the probabilities $\eta(i, P)$, $i \in P$ as in section 3, we show the following result in the Appendix.

Lemma 1. *The probability that i and j are ranked in the two first positions in the first round of a runoff election can be expressed as*

$$ij(q_i, q_{-i}) = \frac{\left[vs_i^t\right]^K}{\sum_{k \neq j} \left[vs_k^t\right]^K} + \frac{\left[vs_j^t\right]^K}{\sum_{k \neq i} \left[vs_k^t\right]^K} - \frac{\left[vs_i^t\right]^K + \left[vs_j^t\right]^K}{\sum_k \left[vs_k^t\right]^K}. \tag{47}$$

Proof. The proof is available in Appendix B.7. □

Lemma 2 follows from a standard result in probability theory on the distribution of the minimum of a list of random variables, and from the property that the product of two Fréchet cdf is also a Fréchet cdf.

Similarly, in the secound round between two candidates i and j, the probability of candidate i winning the election is

$$\frac{\left[\tilde{vs}_i^t(q_i, q_j)\right]^K}{\left[\tilde{vs}_i^t(q_i, q_j)\right]^K + \left[\tilde{vs}_j^t(q_j, q_i)\right]^K}, \tag{48}$$

with $\tilde{vs}_i^t(q_i, q_j)$ the share of individuals in society who prefer candidate i in a pairwise election against j.

From this, the probability $P_i(q_i, q_{-i})$ that i wins the runoff election can finally be expressed as

$$P_i(q_i, q_{-i}) = \sum_{j \neq i} ij(q_i, q_{-i})\left[\frac{\left[\tilde{vs}_i^t(q_i, q_j)\right]^K}{\left[\tilde{vs}_i^t(q_i, q_j)\right]^K + \left[\tilde{vs}_j^t(q_j, q_i)\right]^K}\right]. \tag{49}$$

The bracketed term in (49) gives the probability that i wins a pairwise election against candidate j.[17]

We may now apply the model to the case of special interest politics with heterogeneous voters and homogeneous entrants of section 4.1 and focus on the symmetric electoral equilibrium. Assume

[17]One can observe that in a symmetric equilibrium, the probability that i is elected simplifies to $1/P$.

that the utility function of the agents in group s is given by (9). The platform of any party i consists in a vector of transfers $\{\tau_{s,i}\}_{s\in S}$ and a level of extraction χ_i. In this setting $vs_i^t(q_i, q_{-i})$, the share of individuals who prefer candidate i in the set of P candidates, writes as

$$vs_i^t(q_i, q_{-i}) = \sum_{s\in S} \frac{n_s}{n} \frac{t_i u(c_s(q_i))^{\theta_s}}{t_i u(c_s(q_i))^{\theta_s} + \sum_{j\in P\backslash i} t_j u(c_s(q_j))^{\theta_s}}.$$

while $\tilde{vs}_i^t(q_i, q_j)$, the share of individuals in society who prefer candidate i in a pairwise election against candidate j writes as

$$\tilde{vs}_i^t(q_i, q_j) = \sum_{s\in S} \frac{n_s}{n} \frac{t_i u(c_s(q_i))^{\theta_s}}{t_i u(c_s(q_i))^{\theta_s} + t_j u(c_s(q_j))^{\theta_s}}$$

Remember that we were able to characterize a closed form result for the equilibrium number of parties in plurality elections in the first application as :

$$P_p^* = \frac{1}{1+\overline{\theta}K(1-\epsilon)}[\overline{\theta}K(1-\epsilon) + \frac{y}{c}]. \tag{50}$$

Consider now the case of runoff elections. In an electoral equilibrium, i's platform solves

$$q_i^* = \arg\max_{q^i} P_i(q_i, q_{-i}^*)\chi_i y, \tag{51}$$

with

$$\chi y \leq -\sum_{s\in S} n_s \tau_{s,i}. \tag{52}$$

In a symmetric equilibrium, the optimal level of rents extracted by political parties in runoff elections is determined by the following condition:

$$-\frac{1}{P} + (1-\epsilon)\overline{\theta}K\frac{\chi_m}{1-\chi_m}[\frac{1}{2}\frac{(P-2)(2P-1)}{P^2(P-1)} + \frac{1}{2}\frac{1}{P}] = 0. \tag{53}$$

As a matter of comparison, it is useful to recall that in the case of plurality elections, the condition that determines the level of extraction writes

$$-\frac{1}{P} + (1-\epsilon)\overline{\theta}K\frac{\chi_p}{1-\chi_p}[\frac{P-1}{P^2}] = 0. \tag{54}$$

When increasing marginally the transfers to the interest groups, party i decreases its rents by a corresponding amount in the two systems. The marginal benefits from doing so are however different. In plurality elections, from (53), by increasing marginally the transfers, party i grabs a fraction $1/P^2$ of the vote share of each of its $P-1$ challengers, abstracting from the effects of ϵ, $\overline{\theta}$ and of the level of extraction χ_p.

The first and the second terms in the bracket in (53) reflect respectively the effect of increasing the transfers on party i's vote share in the first round and in the second round. First, observe that the probability of being in the pair of candidates that passes the first round is proportional to $P-2$, since a pair faces $P-2$ challengers. It is also proportional to one half, the probability of being elected in the second round in a symmetric equilibrium. The term $(2P-1)/(P^2(P-1))$ reflects the marginal probability of being in the pair of candidates selected for the second round.

In order to interpret the second term in the bracket of (53), it is useful to note that it writes as

$$\frac{1}{2}\frac{1}{P} = \frac{1}{4} \cdot (P-1)(\frac{2}{P-1} - \frac{2}{P})$$

where $1/4$ is the marginal vote share in a pairwise election and $(P-1)(2/(P-1) - 2/P)$ is the probability of being in a pair of candidates selected for the second round of the election in a symmetric equilibrium.

Taken separately, the first and the second rounds of a runoff system induce lower marginal benefits from transferring resources to the citizenry relative to a plurality system. Indeed, it can easily be shown that when $P \geq 2$, then

$$\begin{cases} \frac{1}{2}\frac{(P-2)(2P-1)}{P^2(P-1)} \leq \frac{P-1}{P^2} \text{ and} \\ \frac{1}{2}\frac{1}{P} \leq \frac{P-1}{P^2}, \end{cases} \tag{55}$$

meaning that each round in a runoff election incentivizes *less* the parties to transfer resources to the citizenry. Indeed, when comparing the first round of a runoff election and the plurality election, it is clear that in the former case, parties face less competition (and similarly for the second round as long as $P \geq 2$). Interestingly however, taken together, the first and the second rounds of a runoff election creates higher marginal benefits from transferring resources to the citizenry, since

$$\frac{1}{2}\frac{(P-2)(2P-1)}{P^2(P-1)} + \frac{1}{2}\frac{1}{P} \geq \frac{P-1}{P^2} \tag{56}$$

when $P \geq 2$. In other words, the repetition of electoral competition in two consecutive rounds dominates the lower electoral competition that parties face in each round taken separately relative to plurality elections. In sum, this also implies that parties should capture *less* rents under the runoff system than under the plurality system. The direct consequence of this is that *less* parties should form in the runoff system, assuming that the cost of party formation is independent of the electoral rule. The following result summarizes the previous discussion.

Proposition 5.

- *The equilibrium number of parties in a runoff electoral system P_m^* is uniquely determined by the equation*
$$\frac{1}{P}\chi_m(P)y = c,$$
with $\chi_m(P)$ a decreasing function for $P \geq 1$ such that
$$\chi_m(P) = \frac{1}{1 + \overline{\theta}K(1-\epsilon)/2\{1 + \frac{(P-2)(2P-1)}{P(P-1)}\}}. \tag{57}$$
- *With stochastic and sincere voting, the Duverger's hypothesis does not hold, since $P_m^* > P_p^*$.*

Proof. The proof is available in Appendix B.8. □

In a model with sincere and stochastic voting, the difference between runoff and plurality elections boils down to the differential effect of the two electoral systems on the marginal benefits of parties from capturing rents (see (53) and (54)). Taken separately, the two rounds of a runoff election create a stronger incentive for parties to capture rents, because competition is lowered relative to plurality elections. However taken together, the two rounds create weaker incentives for parties to capture rents. This explains why we should expect less parties to form under runoff elections. Given that we abstracted from strategic voting, this result does not necessarily undermine the existence of the Duverger's hypothesis. Rather it underlines again the importance of the so-called "psychological effect" for the Duverger's hypothesis to have a chance to hold.

As a numerical example, we illustrate the determination of the number of parties in the two electoral systems in figure 3 in the case where $y = 1.5$, $c = 0.1$ and $(1-\epsilon)\overline{\theta} = 0.95$. In this example, we should expect respectively 8 and 6 parties in plurality and runoff systems.

5 Conclusion

We have provided in this paper a new stochastic voting model for multi-candidate elections. We show that the systematic use of Fréchet (or extreme type II) distributions for the stochastic components of individual voting decisions and participation, significantly ease the issue of computing

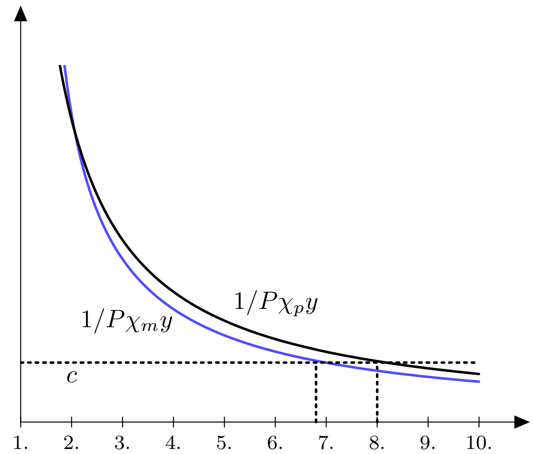

Figure 3: Determination of the number of parties in plurality elections (black curve) and runoff elections (blue curve).

candidates' objective functions in plurality, runoff and proportional elections when the number of candidates is arbitrary[18].

Our central objective in this paper was to show that our theory is flexible and provides a unifying framework to study various topics of the political economy literature. Four applications of the canonical model were developed on the topics of special interest politics, coalition formation in the legislature, and Duverger's hypothesis. Those applications yielded various interesting predictions in terms of the equilibrium degree of political fragmentation that can be observed in a society, and how it can be affected by specific socio-economic fundamentals (technology, endowments, preferences, popularity structure, sources of heterogeneity) and institutional electoral rules that frame the pattern of political competition. Among the key contributions of the paper, we have established that both proportional and runoff systems should lead to a higher party fragmentation than a plurality system. Furthermore, our analysis suggests that when the parties freely decide to run costly electoral campaigns so as to increase their popularity, or to mobilize their electorate, the fragmentation of the polity is expected to decrease.

Important issues have been left aside, which could certainly be approached within the stochas-

[18]Moreover, relying on a simple axiomatic approach to probabilistic voting inspired of McFadden (1974) , we were able to provide some rationale for the use of Fréchet distributions in these models.

tic model presented in this paper. For instance, our examples remain in the confines of political competition in democratic systems with sincere voting. We have not therefore apprehended the issue of the formation of opposition movements in other sytems such as autocracies or theocracies. Arguably, a theory on the formation of an endogenously fragmented opposition could help researchers understand strategies of entrenched elites willing to avoid a democratic transition.[19] Furthermore, it could be particularly interesting to extend this theory in order to account for more complex technologies of party formation and study the process of selection of candidates within parties (e.g. Caillaud and Tirole (2002)). As well, the issue of the policy motivations, ideologies and political identities (e.g. Snyder and Ting (2002)) of parties and candidates has been left aside in this paper, since we have focused on the effect of political competition on the fragmentation of the polity. Future works may account for both policy and office motivations. Such extensions could help disentangle the effect of preferences from that of pure competition on the motives of political entry under various electoral rules. We hope that the framework developed here can be used as a stepping stone to these extensions and others in future research.

Appendix

A Axiomatic approach to probabilistic voting

This section provides an axiomatic rationale for the use of Fréchet distributions in multi-candidate stochastic voting models. As we focus on agents with attributes $s \in S$, we will denote $i(s, \mathcal{P}) \equiv i(\mathcal{P})$ the probability that an agent of type s votes for candidate i when the set of candidates is $\mathcal{P}$. Similarly, for any subset of candidates $\mathcal{C} \subseteq \mathcal{P}$, we denote $C(s, \mathcal{P}) \equiv C(\mathcal{P})$ the probability that the candidate chosen by the agents with attributes s belongs to the subset $\mathcal{C}$ of candidates.

Axiom 1. *(Independence of Irrelevant Alternatives Axiom). For all possible alternative set of candidate $C \subseteq P$ and vectors of measured attributes $s \in S$,*

$$i(\mathcal{C})/j(\mathcal{C}) = i(\mathcal{P})/j(\mathcal{P}), \tag{A.1}$$

Axiom 1, adapted from McFadden (1974), says that the odds of i being chosen relative to candidate j out of $\mathcal{P}$ candidates $i(s, \mathcal{P})/j(s, \mathcal{P})$ are equal to the odds of choosing i relative to j out

[19]This issue has been studied by Acemoglu et al. (2004) and Padró i Miquel (2007) for example.

of any subset of candidates $\mathcal{C}$, $i(s,\mathcal{C})/j(s,\mathcal{C})$.

Axiom 2. *(Positivity). For all possible alternative set of candidates $\mathcal{C} \subseteq \mathcal{P}$ and vectors of measured attributes $s \in S$, $i(s,\mathcal{C}) > 0$.*

For any set of candidates $\mathcal{C}$, an agent with attributes $s \in S$ has a strictly positive probability of voting for *any* candidate $i \in \mathcal{C}$. The main consequence of Axiom 2 is that deterministic voting models are not consistent with an axiomatic approach.

We now define a special class of probabilistic voting models originally introduced by Luce (1959), which will be particularly useful in the subsequent analysis. Using the framework introduced by Becker et al. (1963), a probabilistic voting model will be called a *Luce or strict voting utility model* if there exist positive "utility indicator" functions $v(s,i)$ for any candidate $i \in \mathcal{C} \subseteq \mathcal{P}$ such that the probability of i being chosen out of $\mathcal{C}$ by the agents with attributes s can be expressed as

$$i(s,\mathcal{C}) = \frac{v(s,i)}{\sum_{j\in\mathcal{C}} v(s,j)}. \tag{A.2}$$

Theorem 2. *A probabilistic model satisfies axioms 1 and 2 if and only if it is a Luce voting model.*

Instead of giving a full proof of the preceding theorem - which can be found in Becker et al. (1963) (their Theorem II) - we follow the ingenuous method of McFadden (1974) in order to characterize $j(\mathcal{P})$. Assume that there are only two candidates $\{i,j\}$ competing in the election. From (A.1) with $\mathcal{C} = \{i,j\}$,

$$i(\mathcal{P}) = \frac{i(\{i,j\})}{j(\{i,j\})} j(\mathcal{P}). \tag{A.3}$$

Thus,

$$\frac{i(\{i,j\})}{j(\{i,j\})} = \frac{i(\mathcal{P})/k(\mathcal{P})}{j(\mathcal{P})/k(\mathcal{P})} \tag{A.4}$$

for some third candidate $k \in \mathcal{P} \setminus \{i,j\}$, implying that

$$\frac{i(\{i,j\})}{j(\{i,j\})} = \frac{i(\{i,k\})/k(\{i,k\})}{j(\{j,k\})/k(\{j,k\})}. \tag{A.5}$$

Moreover,

$$\sum_{i\in\mathcal{P}} i(\mathcal{P}) = 1 = \sum_{i\in\mathcal{P}} \frac{i(\{i,j\})}{j(\{i,j\})} j(\mathcal{P}). \tag{A.6}$$

Thus,

$$j(P) = \frac{1}{\sum_{i\in P} i(\{i,j\})/j(\{i,j\})}. \tag{A.7}$$

From (A.4), we deduce that

$$j(\mathcal{P}) = \frac{j(\{j,k\})/k(\{j,k\})}{\sum_{i \in P} i(\{i,k\})/k(\{i,k\})}. \tag{A.8}$$

We denote $j(s,\{j,k\})/k(s,\{j,k\}) = v(j,k,s)$ so that (A.4) rewrites

$$j(\mathcal{P}) = \frac{v(j,k,s)}{\sum_{i \in \mathcal{P}} v(i,k,s)}, \tag{A.9}$$

meaning that the probability of an agent with attribute s from voting j out of $\mathcal{P}$ candidates is equal to how well j fares against some candidate k in a pairwise election relative to how well all the candidates fare against the same candidate k in pairwise elections. Our last axiom gives a specification for $v(j,k,s)$, which is the main departure from McFadden (1974).

Axiom 3. *(Irrelevance of Alternative Set). The function $v(j,k,s)$ determining the selection probability in pairwise elections has a product separable form and there exist utility indicators $u(q_j,s)$ and $u(q_k,s)$ and some constants $t_j > 0$, $t_k > 0$ and $\theta_s > 0$ for any pair of candidates $j,k \in \mathcal{P}$ and any attributes $s \in S$ such that*

$$v(j,k,s) = \frac{t_j u(q_j,s)^{\theta_s}}{t_k u(q_k,s)^{\theta_s}}. \tag{A.10}$$

We posit in Axiom 3 that in pairwise elections, there are three dimensions in individual voting decisions. The first dimension of voting decisions is linked to the *popularity* or valence of the candidates. The parameters t_i and t_k model the popularity of the two candidates and are assumed independent from individual attributes $s \in S$ and from the promised platforms q_j and q_k.

The second aspect of voting decisions that is accounted for in (A.10) is linked to the *quality* of the promised platforms, as evaluated by the agents with attributes $s \in S$. This dimension of voting decisions is dealt with through utility indicator functions $u(.,s)$ that are defined over the set of feasible platforms $Q = \bigcup_{i \in \mathcal{P}} Q_i$ for any vector of attributes $s \in S$.

The last parameter θ_s is linked to the rate of substitution between popularity and quality. If $u(q_j,s)/u(q_k,s)$ - the relative quality of the platform of candidate j - increase by a percentage point, then the relative popularity of candidate j must decrease by θ_s percentage point so that the odds of electing j stay constant. Thus, when θ_s is high, even popular candidates need to provide quality policies because popularity does not affect much voting decisions.

Axiom 3 sets a specification for the probability of voting any candidate j relative to some challenger k in pairwise elections. To illustrate, McFadden (1974) assumes an additively separable form in his Axiom of Irrelevance of Alternative Set. The author then establishes that the only distribution of the noise parameters affecting individual decisions that are compatible with his three behavioral axioms are Weibull distributions. With the product functions of (A.10), we will show that the noise in voting decisions must be distributed according to Fréchet distributions, which are as well labeled *inversed* Weibull functions.

A probabilistic voting model satisfies the three preceding Axioms if and only if for any vector of attributes $s \in S$ and for any candidate $j \in \mathcal{P}$ there exists a utility indicator $u(., s)$ defined over the set of feasible policies Q, a popularity parameter $t_j > 0$ and a political responsiveness parameter θ_s such that

$$j(\mathcal{P}) = \frac{t_j u(q_j, s)^{\theta_s}}{\sum_{i \in \mathcal{P}} t_i u(q_i, s)^{\theta_s}}. \tag{A.11}$$

This result is obtained by applying theorem 2 and substituting (A.10) in (A.8).

We next derive the family of distributions that satisfy the behavioral assumptions given in the three preceding Axioms. Summarizing our previous findings, from (4), we know that a random utility voting model with a specification given in (1) is such that the probability for an individual s to vote for candidate i out of $\mathcal{P}$ is

$$i(s, \mathcal{P}) = \int_0^\infty \prod_{j \in \mathcal{P} \setminus i} F_{(s,j)}(V(s,i)/V(s,j)\epsilon) dF_{(s,i)}(\epsilon)$$

where $F_{(s,i)}(.)$ is the distribution of the random parameter $\epsilon(s, i)$.

From the preceding formal development inspired of McFadden (1974), we know that a probabilistic voting model respecting Axioms 1, 2 and 3 must be such that for any vector of attributes $s \in S$ and for any candidate $i \in \mathcal{P}$, the probability of an individual s to vote for candidate i out of $\mathcal{P}$ takes the form

$$j(s, P) = \frac{t_j u(q_j, s)^{\theta_s}}{\sum_{i \in P} t_i u(q_i, s)^{\theta_s}}.$$

with $u(q_k, s)$, $k \in \mathcal{P}$ some utility indicators, $t_k > 0$ a popularity indicator and θ_s the political responsiveness of the agents with attributes $s \in S$.

We define the following class of distributions.

Definition 2. *Let $\mathcal{C}_d$ the class of distributions such that for any pair of candidates $i, j \in \mathcal{P}$ there exists a strictly positive constant $\alpha_{i,j}$ such that $F_{i,s}(\epsilon) = F_{j,s}(\alpha_{i,j}\epsilon)$.*

This class of distributions includes more specifically the case where the noise parameters are all distributed according to the same distribution (when $\alpha_{i,j} = 1$ for any pair $i, j \in \mathcal{P}$) and allows for some differences between the distributions which are a matter of translation.

Theorem 3. *In the class of distribution $\mathcal{C}_d$, a probabilistic voting model satisfies axioms 1, 2 and 3 if and only if it is a random utility model where the noise parameters $\epsilon(i, s)$ are distributed according to Fréchet distributions $F_{i,s}(\epsilon) = exp(-t_i \epsilon^{-\theta_s})$ for some positive parameters t_i and θ_s.*

Proof. The proof is available in Appendix B.2. □

This concludes the micro-foundations of random voting utility models with random parameters distributed according to some Fréchet (or extreme type II) distributions. As in the approach to economic decisions of McFadden (1974), we showed that voting probabilities can be interpreted as deriving from representative utilities, which are affected by the popularity of the candidates and by the quality of the political platforms they offer.

This theory makes it simple to ascertain the effect of an increased number of candidates in election on voting decisions and therefore provides an approach to the industrial organization of politics. It is worth mentioning two caveats. From the Luce model implied by the two first axioms, it inherently builds in the model a particular effect of competition on voting behavior. Indeed, an increase in the set of competing parties necessarily leads to proportional decreases in the vote shares of the old candidates, and a corresponding increase of the vote share of new candidates. This is a first limitation of this theory.

The second main limitation of our approach lies in the specification of selection probabilities in pairwise election given in Axiom 3. Indeed, although we are able to specify probabilistic voting behaviors in pairwise elections, our approach assumes that voting decisions respond to three main dimensions that have found support in the political economy literature. The first one is the valence of candidates in election (the parameters that we denoted t_i, $i \in \mathcal{P}$). The second is the quality of the political platforms that are proposed by the candidates in election. We modeled this by assuming that the agents of type $s \in S$ derive some utility $u(q, s)$ from electing a candidate that implements some platform q. Finally, we have assumed that the last dimension of probabilistic voting decisions is linked to the elasticity of substitution between political platforms' quality and the valence of candidates in election through the parameters θ_s, $s \in S$. More complex approaches may account for other dimensions of probabilistic voting decisions and may correspondingly find different distributions for the randomness in voting decisions. Our approach is however sufficiently simple and flexible to be adapted to different topics of the political economy literature as we demonstrate in the main text.

B Mathematical proofs:

B.1 Proof of Theorem 1:

Proof. Existence: Consider the function

$$h_{si}(y) = \frac{t_i y}{t_i y + K_i} \tag{B.1}$$

defined on a convex space E_i that contains $V(s, Q_i)^{\theta_s}$ (such a space exists since Q_i is convex and $V(s, .)$ is continuous) is straightforwardly concave.

As the vote share of any candidate i can be expressed as

$$vs_i(q_i, q_{-i}) = \sum_{s \in S} h_{si}(V(s, q_i)^{\theta_s}), \tag{B.2}$$

$vs_i(., q^{-i})$ is concave on $Q_i \subseteq Q$ whenever $q_i \to V(s, q_i)^{\theta_s}$ is concave for any s. Finally, since $vs_i(., .)$ is continuous on $\prod_{j \in P} Q_j$, we can apply a standard result of equilibrium existence (e.g. Fudenberg and Tirole (1991, p. 34)) that we state below.

Theorem. *(Adapted from Fudenberg and Tirole (1991, p. 34) to fit the concept of electoral equilibrium given in Definition 1). Consider a strategic-form game whose strategic spaces Q_i are non-empty compact convex subsets of an Euclidian space. If the vote share functions vs_i are continuous in $\prod_{j \in P} Q_j$ and quasi-concave in Q_i there exists a pure-strategy voting equilibrium.*

Unicity: Whenever there exists $s \in S$ such that $V(s, .)^{\theta_s}$ is strictly concave over the set of feasible policies Q, then it follows that the vote share of any candidate i in the subset of voters with attributes $s \in S$ is strictly concave on q_i. Thus, the vote share of any candidate i is strictly concave over Q_i, which implies unicity. □

B.2 Proof of Theorem 2:

This theorem is again inspired of the seminal approach of Daniel McFadden, and more specifically the first two Lemma of McFadden (1974) although it extends a bit his result to account for some heterogeneity in the distributions of the noise parameters.

Assume first that the agents use a random utility voting model with a noise distributed according to Fréchet distributions $F_{i,s}(\epsilon) = \exp(-t_i \epsilon^{-\theta_s})$ for some strictly positive parameters t_i and θ_s. Thus, we find that

$$i(s, \mathcal{P}) = \frac{t_i V(s, q_i)^{\theta^s}}{\sum_{j \in P} t_j V(s, q_j)^{\theta^s}}, \tag{B.3}$$

so the random utility model is a Luce voting model from Theorem 2. Furthermore, applying Axiom 3, there exists "utility indicators" $u(q_j, s) = V(q_j, s)$,"popularity indicators" $t_j > 0$ and political responsiveness parameters θ_s for any candidate j for any vector of attributes $s \in S$ such that

$$i(s, \mathcal{P}) = \frac{t_i u(s, i)^{\theta_s}}{\sum_{j \in P} t_j u(s, j)^{\theta_s}}. \tag{B.4}$$

By Lemma A then, a random utility model with noise parameters distributed according to the Fréchet distributions $F_{i,s}(.)$ is a probabilistic voting model satisfying the three Axiom of the previous subsection.

Proving the other implication of the equivalence is a little more demanding. Assume that a probabilistic voting model satisfies axioms 1, 2 and 3. Take some attributes $s \in S$ and a candidate $i \in \mathcal{P}$. We know from Lemma A that there exists utility indicators $u(q_j, s)$, popularity indicators $t_j > 0$ for any $j \in \mathcal{P}$ and a political responsiveness parameter θ_s such that

$$i(\mathcal{P}) = \frac{t_i u(q_i, s)^{\theta_s}}{\sum_{j \in P} t_j u(q_j, s)^{\theta_s}}.$$

Consider the choice between either candidate i with a popularity $t_i > 0$ with a representative utility $u(i, s)$ and n candidates with a popularity $t_j > 0$ yielding $u(j, s)$. Thus,

$$i(P) = \frac{t_i u(q_i, s)^{\theta_s}}{t_i u(q_i, s)^{\theta_s} + n t_j u(q_j, s)^{\theta_s}}.$$

Assume now that there exists two distributions $G_i(\epsilon)$ and $G_j(\epsilon)$ in $\mathcal{C}_d$ such that

$$i(\{i, j, \dots, j\}) = \frac{t_i u(q_i, s)^{\theta_s}}{t_i u(q_i, s)^{\theta_s} + n t_j u(q_j, s)^{\theta_s}} = \int_0^\infty G_j(u(q_i, s)/u(q_j, s)\epsilon)^n dG_i(\epsilon). \tag{B.5}$$

On the other hand, consider a binary choice between candidate i and an alternative candidate k with $t_k > 0$ and $u(q_k, s)$.

$$i(\{i, k\}) = \frac{t_i u(q_i, s)^{\theta_s}}{t_i u(q_i, s)^{\theta_s} + t_k u(q_k, s)^{\theta_s}} \tag{B.6}$$

Assume that there exists some distribution $G_k(\epsilon) \in \mathcal{C}_d$ such that

$$i(\{i,k\}) = \frac{t_i u(q_i,s)^{\theta_s}}{t_i u(q_i,s)^{\theta_s} + t_k u(q_k,s)^{\theta_s}} = \int_0^\infty G_k(u(q_i,s)/u(q_k,s)\epsilon)dG_i(\epsilon). \tag{B.7}$$

Assume finally that $nt_j u(q_j,s)^{\theta_s} = t_k u(q_k,s)^{\theta_s}$ so that

$$i(\{i,j,\ldots,j\}) = i(\{i,k\}). \tag{B.8}$$

Thus, for any set of attributes $s \in S$, any platform $q_i \in Q_i$, $q_j \in Q_j$ and $q_k \in Q_k$,

$$\int_0^\infty G_k(u(q_i,s)/u(q_k,s)\epsilon)dG_i(\epsilon) - \int_0^\infty G_j(u(q_i,s)/u(q_j,s)\epsilon)^n dG_i(\epsilon) = 0. \tag{B.9}$$

As the distributions belong to the class $\mathcal{C}_d$, we can rewrite the preceding expression as

$$\int_0^\infty [G_i(\alpha_{ki}u(q_i,s)/u(q_k,s)\epsilon) - G_i(\alpha_{ji}u(q_i,s)/u(q_j,s)\epsilon)^n]dG_i(\epsilon) = 0. \tag{B.10}$$

The integrand in (B.10) must then be equal to zero for a non-zero value of ϵ. This means that for any set of attributes $s \in S$ and any platform $q_i \in Q_i$, $q_j \in Q_j$ and $q_k \in Q_k$, there exist some ϵ such that

$$G_i(\alpha_{ki}u(q_i,s)/u(q_k,s)\epsilon) = G_i(\alpha_{ji}u(q_i,s)/u(q_j,s)\epsilon)^n, \tag{B.11}$$

As $nt_j u(q_j,s)^{\theta_s} = t_k u(q_k,s)^{\theta_s}$, we obtain

$$G_i(\alpha_{ki}(\frac{t_k}{nt_j})^{1/\theta_s}\frac{u(q_i,s)\epsilon}{u(q_k,s)}) = G_i(\frac{\alpha_{ji}u(q_i,s)\epsilon}{u(q_k,s)})^n, \tag{B.12}$$

which must hold for any vector of attributes $s \in S$ and any platform $q_i \in Q_i$, $q_j \in Q_j$. We now take q_i such that $\alpha_{ji}\epsilon u(q_i,s)/u(q_j,s) = 1$. Then the preceding equation implies

$$G_i(\frac{\alpha_{ki}}{\alpha_{ji}}(\frac{t_k}{nt_j})^{1/\theta_s}) = G_i(1)^n. \tag{B.13}$$

As this must be true for any $i,j,k \in P$ and any $s \in S$, it must be that the positive constants α_{ki} and α_{ji} are such that

$$\frac{\alpha_{ki}}{\alpha_{ji}}(\frac{t_k}{t_j})^{1/\theta_s} = \alpha_0 \tag{B.14}$$

for some constant α_0 independent from i, j, k and s. Thus, (B.13) rewrites

$$G_i(\alpha_0 n^{-1/\theta_s}) = G_i(1)^n. \tag{B.15}$$

Since $G_i(1) < 1$, there exists some constant $r_i > 0$ such that

$$G_i(1) = exp(-r_i). \tag{B.16}$$

Thus,

$$G_i(\alpha_0 n^{-1/\theta_s}) = e^{-nr_i}. \tag{B.17}$$

Let $\gamma = \alpha_0 n^{-1/\theta_s}$. Injecting γ in the last equation implies

$$G_i(\gamma) = exp(-r_i(\gamma/\alpha_0)^{-\theta_s}). \tag{B.18}$$

From that point, it is straightforward that

$$\alpha_{i,j} = (r_j/r_i)^{1/\theta_s} \tag{B.19}$$

Thus, injecting (B.19) in (B.14), we find that

$$\alpha_0 = 1, \tag{B.20}$$

which proves that $G_i(.)$ follows a Fréchet distribution,

$$G_i(\gamma) = exp(-r_i(\gamma)^{-\theta_s}). \tag{B.21}$$

B.3 Proof of Proposition 1

Proof. The first-order condition writes

$$-n_s G_i(q_i, q^*_{-i}) + y\chi_i \left[\frac{\partial G_i}{\partial vs_i} \frac{\partial vs_i}{\partial \tau_{s,i}} + \sum_{j \neq i} \frac{\partial G_i}{\partial vs_j} \frac{\partial vs_j}{\partial \tau_{s,i}} \right] = 0 \tag{B.22}$$

with

$$\frac{\partial G_i}{\partial vs_i} = \frac{K\left[vs_i\right]^K \left[\sum_{j \neq i} \left[vs_j\right]^K\right]}{\left[\sum_{j \in \mathcal{P}} \left[vs_j\right]^K\right]^2} \frac{1}{vs_i} = \frac{K}{vs_i} G_i\left(1 - G_i\right) > 0 \tag{B.23}$$

$$\frac{\partial G_i}{\partial vs_j} = -\frac{K\left[vs_i\right]^K \left[vs_j\right]^K}{\left[\sum_{j \in \mathcal{P}} \left[vs_j\right]^K\right]^2} \frac{1}{vs_j} = -\frac{K}{vs_j} G_i G_j < 0 \tag{B.24}$$

and

$$\frac{\partial vs_i}{\partial \tau_{s,i}} = (1-\epsilon)\frac{n_s}{n}\frac{\theta_s}{c^i_s} vs_{i,s}(1 - vs_{i,s}) \tag{B.25}$$

$$\frac{\partial vs_j}{\partial \tau_{s,i}} = -(1-\epsilon)\frac{n_s}{n}\frac{\theta_s}{c^i_s}\frac{u(c_s(q_j))^{\theta_s} u(c_s(q_i))^{\theta_s}}{\left[\sum_{l \in P} u(c_s(q_l))^{\theta_s}\right]^2} = -(1-\epsilon)\frac{n_s}{n}\frac{\theta_s}{c^i_s} vs_{i,s} vs_{j,s} \tag{B.26}$$

This finally writes as:

$$-n_s G_i + y\chi_i \left[\frac{\partial G_i}{\partial vs_i}\frac{\partial vs_i}{\partial \tau_{s,i}} + \sum_{j\neq i} \frac{\partial G_i}{\partial vs_j}\frac{\partial vs_j}{\partial \tau_{s,i}} \right] = 0 \tag{B.27}$$

$$-n_s G_i + y\chi_i \left[\frac{K}{vs_i} G_i (1-G_i) \frac{\partial vs_i}{\partial \tau_{s,i}} - K \sum_{j\neq i} G_i G_j \frac{K}{vs_j}\frac{\partial vs_j}{\partial \tau_{s,i}} \right] = 0 \tag{B.28}$$

Manipulation of the first order condition provides for $G_i > 0$ the conditions for each interest group s

$$-n_s + y\chi_i K(1-\epsilon)\frac{n_s}{n}\frac{\theta_s}{c_s^i}\left[(1-G_i)\frac{vs_{i,s}(1-vs_{i,s})}{vs_i} + \sum_{j\neq i} G_j \frac{vs_{j,s} vs_{i,s}}{vs_j} \right] = 0$$

$$\frac{y}{n}\chi_i K(1-\epsilon)\frac{\theta_s}{c_s^i}(1-G_i) = 1$$

Thus in a symmetric equilibrium for any party i, and any $k, p \in S$

$$\frac{\theta_s}{c_s^i} = \frac{\theta_p}{c_p^i} = \frac{1}{\frac{y}{n}\chi_i K(1-\epsilon)(1-G_i)}. \tag{B.29}$$

Using the budget constraint (11), we deduce the consumption of the agents in group s as a function of the rents extracted,

$$c_s^i = \frac{\theta_s}{\overline{\theta}}\overline{y}(1-\chi_i). \tag{B.30}$$

□

B.4 Proof of Proposition 2

By substituting (27) in (28), it is easy to establish that the optimal number of parties is given by

$$P^* = \frac{1}{1+K\overline{\theta}(1-\epsilon)}[K\overline{\theta}(1-\epsilon) + \frac{y}{c}]. \tag{B.31}$$

From this point, the effects of c, K, ϵ $\overline{\theta}$, y on P^* are straightforward.

Relative to the second point of the proposition, observe y and $\overline{\theta}$ both depend on n_s, for any $s \in S$. Furthermore, $d\overline{\theta}/dn_s = (\theta_s - \overline{\theta})/n$, so $d\overline{\theta}/dn_s > 0$ when $\theta_s > \overline{\theta}$ and $d\overline{\theta}/dn_s \leq 0$ otherwise. From this point, since P^* decreases with $\overline{\theta}$, it is direct that when $\theta_s < \overline{\theta}$, then P^* increases with n_s because (i) the average reactivity of the citizenry becomes lower and (ii) the tax base y increases. By contrast, when $\theta_s \geq \overline{\theta}$, then the variations of P^* with n_s are ambiguous since on the one hand the average reactivity of the citizenry becomes larger, which creates a downward pressure on P^* while on the other hand the aggregate income of the citizenry y still becomes larger, which creates an upward pressure on P^*.

B.5 Proof of Proposition 3:

B.5.1 Heterogeneous entrants: a simple introductory case

Before detailing the more complex case of heterogeneous entrants with an endogenous heterogeneity, it is useful to study a simple case where a single party with high popularity t_h faces $P-1$ challengers with popularity $t_l < t_h$, with P endogenous. The entry cost for the parties with popularity t_l is denoted c. Since we assume homogeneous voters, the platform of any party i reduces to a level of extraction χ_i. Assuming that the cost of party formation is c, the free entry condition still writes as $G_l(\chi_h, \chi_l)\chi_l y = c$. We establish the following preliminary result:

Proposition 6.

- *There exists a unique electoral equilibrium, where the parties with a low popularity t_l extracts a level of rents $\chi_l^*(t_l, t_h, P)$, with $\chi_l^*(t_l, t_h, P) < \chi_h^*(t_l, t_h, P)$.*
- *$\chi_l^*(t_l, t_h, P)$ increases with t_l and decreases with t_h, while $\chi_h^*(t_l, t_h, P)$ decreases with t_l and increases with t_h. Consequently, $\chi_l^*(t_l, t_h, P) < \chi^s(P) < \chi_h^*(t_l, t_h, P)$ for a given value of P, with $\chi^s(P)$ the level of extraction in the equilibrium where the candidates are homogeneous.*
- *Since $\chi_l^*(t_l, t_h, P)$ decreases with P, the number of entrants is uniquely determined and less parties should form in equilibrium relative to a case where all the parties have a low popularity.*

Proof. The consumption of any citizen when i wins the office simplifies to

$$c(\chi_i) = y(1 - \chi_i). \tag{B.32}$$

Following the steps of the preceding section, we can easily show that the probability of party $i \in \{h, l\}$ being elected now writes as

$$G_h(\chi_h, \chi_l) = \frac{[vs_h(\chi_h, \chi_l)]^K}{[vs_h(\chi_h, \chi_l)]^K + (P-1)[vs_l(\chi_l, \chi_h)]^K} \tag{B.33}$$

$$G_l(\chi_h, \chi_l) = \frac{[vs_l(\chi_h, \chi_l)]^K}{[vs_h(\chi_h, \chi_l)]^K + (P-1)[vs_l(\chi_l, \chi_h)]^K}, \tag{B.34}$$

with

$$vs_h(\chi_h, \chi_l) = \frac{t_h u(c(\chi_h))^\theta}{t_h u(c(\chi_h))^\theta + (P-1) t_l u(c(\chi_l))^\theta} \tag{B.35}$$

$$vs_l(\chi_l, \chi_h) = \frac{t_l u(c(\chi_l))^\theta}{t_h u(c(\chi_h))^\theta + (P-1) t_l u(c(\chi_l))^\theta} \tag{B.36}$$

$$\frac{vs_l(\chi_l, \chi_h)}{vs_h(\chi_h, \chi_l)} = \frac{t_l u(c(\chi_l))^\theta}{t_h u(c(\chi_h))^\theta}. \tag{B.37}$$

Party i still seeks to maximize its expected rents, so

$$\chi_i^* = \arg\max_{\chi_i} \chi_i y G_i(\chi_i, \chi_{-i}^*). \tag{B.38}$$

Following the steps of the previous section, it can be shown that the first-order conditions associated with the preceding optimizations are

$$1 - \frac{\chi_h}{(1-\chi_h)} K(1-\epsilon)\theta \left(1 - \frac{t_h^K}{t_h^K + (P-1)\, t_l^K Z(\chi_h, \chi_l)^K}\right) = 0 \tag{B.39}$$

for the high popularity party, and

$$1 - \frac{\chi_l}{(1-\chi_l)} K(1-\epsilon)\theta \left(1 - \frac{t_l^K}{(P-1)\, t_l^K + t_h^K / Z(\chi_h, \chi_l)}\right) = 0 \tag{B.40}$$

for his $P-1$ challengers, with

$$Z(\chi_h, \chi_l) = \left(\frac{1-\chi_l}{1-\chi_h}\right)^{(1-\epsilon)\theta K}. \tag{B.41}$$

From this point, (B.39) gives a locus $\chi_h(\chi_l, t_l, t_h, P)$, while (B.40) provides a locus $\chi_l(\chi_h, t_l, t_h, P)$. Notice that the unicity of the electoral equilibrium is not straightforward, since there is a complementarity between rent extraction of the two types of parties in equilibrium. Indeed, if the high popularity party captures more rents, then it becomes less costly for low popularity parties to do the same and reciprocally. However, by combining the two first-order conditions, we can establish a third and simple relationship between χ_l and χ_h that proves the unicity of the electoral equilibrium.

The intuition of this proposition is represented in Figure 4. The idea is that by combining (B.39) and (B.40), we can establish a relatively simple relation between χ_h and χ_l in equilibrium that is represented by the blue curve $\chi_l(\chi_h)$ in figure 4. Since this relation is independent from t_h and t_l, the comparative statics are easily established. Of particular interest, the complementarity between χ_l^* and χ_h^* along the equilibrium path is always dominated. For instance, an increase in t_h will lead to higher rents captured by the popular party and lower rents captured by its challengers. This implies that as long as $t_h \neq t_l$, we should expect an electoral equilibrium where the platforms of the two

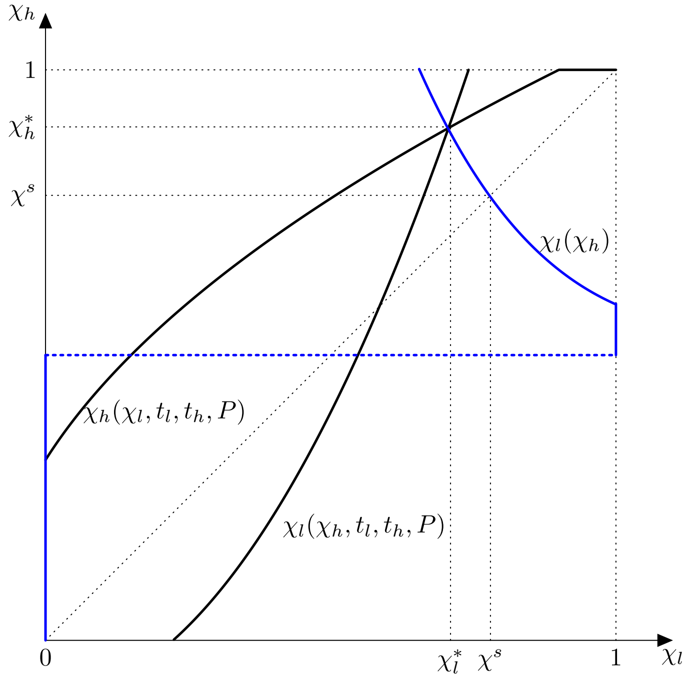

Figure 4: Determination of the Electoral Equilibrium

types of candidates are necessarily different. This result holds even if the difference in the valence of the two types of candidates is small, by contrast with the mean voter theorem of Schofield (2007).

By differentiating (B.39) and (B.40) with respect to χ_l and χ_h respectively, we can show that there is a complementarity between the rents extracted by the low and the high reputation parties. Thus, considering the first-order conditions separately is not sufficient to prove the unicity of the Nash equilibrium here. We need to establish a third relationship between χ_l and χ_h by combining the two FOCs. By substituting $Z(\chi_h, \chi_l)$ from (B.39) in (B.40), we find that

$$\chi_l = \frac{\chi_h(P-1)}{\chi_h((1+K\theta(1-\epsilon))(P-1)+1)-1}. \tag{B.42}$$

Notice that χ_l decreases with χ_h as long as the denominator is positive, which establishes the unicity of the intersection of the two loci.

We have represented the effect of an increase in t_h on the position of the equilibrium. Given that $\chi_l(\chi_h)$ is independent from the popularity parameters t_l and t_h in this third relationship, it is direct that $\chi_l^*(t_l, t_h, P)$ decreases with t_h, while $\chi_h^*(t_l, t_h, P)$ increases with t_h. The reasoning for the effect of an increase in t_l is similar.

As $\chi_h^*(t_l, t_h, P) = \chi^S(P) = \chi_l^*(t_l, t_h, P)$ when $t_h = t_l$, from the previous comparative statics, it is clear that $\chi_h^*(t_l, t_h, P) < \chi^S(P) < \chi_l^*(t_l, t_h, P)$ when $t_l < t_h$.

Finally, when the number of low popularity parties increase, then parties decrease the level of rent they extract (notice that the complementarity between χ_l and χ_h does not create an ambiguity here). This implies that $\chi_l^*(t_l, t_h, P)$ is a decreasing function of P. Furthermore, (B.39) can be rewritten as

$$1 - \frac{\chi_l}{1-\chi_l} K\theta(1-\epsilon)(1 - vs_l) = 0, \tag{B.43}$$

with vs_l the vote share of a low popularity party, so

$$vs_l(t_l, t_h, P) = 1 - \frac{1 - \chi_l^*(t_l, t_h, P)}{K\theta\chi_l^*(t_l, t_h, P)} \tag{B.44}$$

and the expected rents extracted by a low reputation party in equilibrium are equal to

$$W(t_l, t_h, P) = y\chi_l^*(t_l, t_h, P) vs_l(t_l, t_h, P). \tag{B.45}$$

Since both $\chi_l^*(t_l, t_h, P)$ and $vs_l(t_l, t_h, P)$ are positive and decreasing functions of P, the equation $W(t_l, t_h, P) = c$ admits a unique solution $P^*(t_l, t_h)$. Finally, since low popularity parties capture less rents and have a lower vote share than in a symmetric equilibrium, then $P^*(t_l, t_h)$ is lower than the number of entrants in a symmetric equilibrium. $\square$

We now may approach the case where both the fraction of high-popularity parties and that of low-popularity parties are endogenous. We denote P_l and P_h the number of parties with a low and a high reputation respectively. The equilibrium probability of winning are given by:

$$\begin{aligned} G_h(t_l, t_h, P_l, P_h) &= \frac{t_h^K}{P_h t_h^K + P_l t_l^K Z(\chi_h, \chi_l)} \quad \text{and} \\ G_l(t_l, t_h, P_l, P_h) &= \frac{t_l^K}{P_l t_l^K + P_h t_h^K / Z(\chi_h, \chi_l)}, \end{aligned}$$

with

$$Z(\chi_h, \chi_l) = \left(\frac{1-\chi_l}{1-\chi_h}\right)^{(1-\epsilon)\theta K}. \tag{B.46}$$

Since parties maximize their expected rents, by analogy with the previous application, the optimal levels of extraction solve the following system of first-order conditions:

$$1 - \frac{\chi_h}{(1-\chi_h)} K(1-\epsilon)\theta \left(1 - \frac{t_h^K}{P_h t_h^K + P_l t_l^K Z(\chi_h, \chi_l)}\right) = 0 \tag{B.47}$$

for the P_h high reputation parties, and

$$1 - \frac{\chi_l}{(1-\chi_l)} K(1-\epsilon)\theta \left(1 - \frac{t_l^K}{P_l t_l^K + P_h t_h^K / Z(\chi_h, \chi_l)}\right) = 0 \tag{B.48}$$

for their P_l low reputation challengers.

Following the steps of the previous section, we can still deduce two loci $\chi_h(\chi_l, t_l, t_h, P_l, P_h)$ and $\chi_l(\chi_h, t_l, t_h, P_l, P_h)$ from the first-order conditions. Furthermore, by combining (B.47) and (B.48), we can establish a third relationship between χ_h and χ_l that proves the unicity of the intersection of the two loci (the determination of the intersection of the two loci resembles that represented in figure 4). We establish the following result below, which is the formal version of the proposition and the subsequent discussion provided in the main text.

Proposition 7.

- *There exists a unique electoral equilibrium where the parties with popularity t_l (resp. t_h) extract a level of rents $\chi_l^*(t_l, t_h, P_l, P_h)$ (resp. $\chi_h^*(t_l, t_h, P_l, P_h)$), with $\chi_l^*(t_l, t_h, P_l, P_h) < \chi_h^*(t_l, t_h, P_l, P_h)$. $\chi_l^*(t_l, t_h, P_l, P_h)$ increases with t_l and decreases with t_h, while $\chi_h^*(t_l, t_h, P_l, P_h)$ decreases with t_l and increases with t_h.*

- *Both $\chi_l^*(t_l, t_h, P_l, P_h)$ and $\chi_h^*(t_l, t_h, P_l, P_h)$ decrease with P_l and P_h. Consequently, the system*

$$\begin{cases} y\chi_l^*(t_l, t_h, P_l, P_h) G_h(t_l, t_h, P_l, P_h) = c_l \\ y\chi_h^*(t_l, t_h, P_l, P_h) G_l(t_l, t_h, P_l, P_h) = c_h \end{cases} \tag{B.49}$$

  admits a unique solution $(P_l^*(t_l, t_h, c_l, c_h), P_h^*(t_l, t_h, c_l, c_h))$. The aggregate number of parties is such that $P_h^S < P_l^* + P_h^* < P_l^S$, with P_l^S (resp. P_h^S) the number of parties in a symmetric equilibrium where there are only low (resp. high) types running for the election.

- *Allowing political parties to make high campaign investments decreases party fragmentation and increases rent extraction from popular parties. Furthermore, P_l^* increases with t_l and c_h while it decreases with t_h and c_l. Alternatively, P_h^* decreases with t_l and c_h while it increases with t_h and c_l.*

Proof. The first-step for determining the electoral equilibrium consists in establishing the monotonicity of the two loci $\chi_h(\chi_l, t_l, t_h, P_l, P_h)$ and $\chi_l(\chi_h, t_l, t_h, P_l, P_h)$ from the first-order conditions.

By differentiating the first-order condition of a high popularity party (B.47) with respect to χ_l, we find that

$$W(\chi_h, \chi_l)\frac{\partial \chi_h}{\partial \chi_l} + \frac{\chi_h}{1-\chi_h}K\theta\frac{t_h\theta P_l t_l}{1-\chi_h}(\frac{1-\chi_l}{1-\chi_h})^{K\theta-1}\frac{1}{(P_l t_l Z + P_h t_h)^2} = 0, \tag{B.50}$$

with $W(\chi_h, \chi_l)$ the differential of the LHS of (B.47) with respect to χ_h, which is negative (note that $W(\chi_h, \chi_l)$ is not the second-order condition, since we derive with respect to χ_h, not with respect to the rents extracted by a single high reputation party).

Since the second term in the LHS of (B.50) is positive, it is direct that $\partial\chi_h/\partial\chi_l > 0$, meaning that along the equilibrium path, there is a complementarity between the rents extracted by the low and high reputation parties. Similarly, we can establish with the first-order condition of a low popularity party (B.48) that $\partial\chi_l/\partial\chi_h > 0$ as well. Thus, the unicity of the electoral equilibrium is not straightforward from the first-order conditions considered separately.

The second step consists in combining the two first-order conditions in order to establish a third relationship that in turn proves the unicity of the electoral equilibrium. To this aim, we will express Z from (B.47) and substitute its expression in (B.48). From (B.47),

$$\frac{1-\chi_h}{K\theta\chi_h} = 1 - \frac{t_h^K}{P_l t_l^K Z + P_h t_h^K}, \tag{B.51}$$

so

$$P_l t_l^K Z + P_h t_h^K = \frac{t_h^K \theta\chi_h}{\theta\chi_h - (1-\chi_h)} \tag{B.52}$$

and we deduce that

$$P_l t_l^K Z = \frac{-t_h^K \theta K\chi_h(P-h-1) + P_h t_h^K(1-\chi_h)}{\theta K\chi_h - (1-\chi_h)}, \tag{B.53}$$

As (B.48) rewrites

$$1 = \frac{\chi_l}{1-\chi_l}\theta K[1 - \frac{Z t_l^K}{P_l t_l^K Z + P_h t_h^K}], \tag{B.54}$$

we deduce that

$$1 = \frac{\chi_l}{1-\chi_l}\theta K[1 - \frac{-\theta K\chi_h(P_h-1) + P_h(1-\chi_h)}{P_l\theta K\chi_h}], \tag{B.55}$$

from which we establish that

$$\chi_l = \frac{\chi_h P_l}{\chi_h((1+\theta K)(P_l+P_h) - \theta K) - P_h}. \tag{B.56}$$

In turn, (B.56) provide a negative relationship between χ_l and χ_h, which allows to determine the unicity of the Nash equilibrium. Furthermore, notice that the relationship between χ_l and χ_h in (B.56) is independent from the popularity parameters. This is why it is direct that χ_l^* increases with t_l and decreases with t_h, while χ_h^* decreases with t_l and increases with t_h. This, in turn, implies that $\chi_l^*(t_l, t_h, P_l, P_h) < \chi_h^*(t_l, t_h, P_l, P_h)$.

Furthermore, by differentiating the FOCs with respect to P_l or P_h, we find that $\chi_l(\chi_h, t_l, t_h, P_l, P_h)$ (resp. $\chi_h(\chi_l, t_l, t_h, P_l, P_h)$) decreases with P_l and P_h for a given value of χ_h (resp. χ_l). This implies that when P_l or P_h increases, then the two loci intersect for strictly lower values of χ_l and χ_h. Both $\chi_l^*(t_l, t_h, P_l, P_h)$ and $\chi_h^*(t_l, t_h, P_l, P_h)$ decrease with P_l and P_h.

To prove that (B.49) admits a unique solution, consider first the equation

$$y\chi_l^*(t_l, t_h, P_l, P_h)vs_l(t_l, t_h, P_l, P_h) = c_l. \tag{B.57}$$

We know that $\chi_l^*(t_l, t_h, P_l, P_h)$ decreases with P_l and P_h. Notice that in equilibrium, the first-order condition (B.47) rewrites

$$1 - \frac{\chi_l}{1-\chi_l}\theta K(1 - vs_l) = 0, \tag{B.58}$$

so we can simply express the vote share of a party with low reputation as a function of χ_l,

$$vs_l(t_l, t_h, P_l, P_h) = 1 - \frac{1 - \chi_l(t_l, t_h, P_l, P_h)}{\theta K \chi_l(t_l, t_h, P_l, P_h)}. \tag{B.59}$$

From this point, it is direct that vs_l increases with χ_l, so it decreases with P_l and P_h. Consequently, the LHS of (B.57) is a decreasing function of both P_l and P_h. This is why (B.57) admits a unique solution and defines a locus $P_l(P_h)$, which is decreasing in P_h. By the same kind of reasoning, we can establish that

$$y\chi_h^*(t_l, t_h, P_l, P_h)vs_h(t_l, t_h, P_l, P_h) = c_h \tag{B.60}$$

admits a unique solution and defines a locus $P_h(P_l)$ that is decreasing in P_l. It is then direct that the loci $P_l(P_h)$ and $P_h(P_l)$ intersect only once.

Finally, for a given value of $P_h \geq 1$, then low-popularity parties have a lower incentive to enter the race relative to a case where there is no high reputation parties because (i) they capture less rents and (ii) they get a lower vote share. Thus, the locus $P_l(P_h)$ is below the line $P_h + P_l = P_l^S$, where P_l^S is the number of low popularity parties that enter the race in a symmetric equilibrium.

By a similar token, when $P_l \geq 1$, then high reputation parties have a higher incentive to enter the race relative to a case where they only face high reputation challengers because (i) they can capture more rents and (ii) they get a higher vote share. The locus $P_h(P_l)$ is above the line $P_h + P_l = P_h^S$, where P_h^S is the number of high popularity parties that enter the race in a symmetric equilibrium.

Consequently, the intersection of the two loci necessarily occurs on the subspace delimitated by the two lines $P_h + P_l = P_l^S$ and $P_h + P_l = P_h^S$, which implies that $P_h^S < P_l^* + P_h^* < P_l^S$. This concludes the proof of the third point of the proposition. The fourth point is a direct consequence of the third. □

Finally, consider case b) where the parties can invest in order to increase their expected turn-out. Then the equilibrium probability of elections are:

$$\begin{aligned} G_h(t_l, t_h, P_l, P_h) &= \frac{\alpha_h t_h^K}{P_h \alpha_h t_h^K + P_l \alpha_l t_l^K Z(\chi_h, \chi_l)} \quad \text{and} \\ G_l(t_l, t_h, P_l, P_h) &= \frac{\alpha_l t_l^K}{P_l \alpha_l t_l^K + P_h \alpha_h t_h^K / Z(\chi_h, \chi_l)}, \end{aligned}$$

with

$$Z(\chi_h, \chi_l) = (\frac{1-\chi_l}{1-\chi_h})^{(1-\epsilon)\theta K}. \tag{B.61}$$

This is precisely the same formulas as above, with $\alpha_i t_i^K$ instead of t_i^K. The formal development leading to the second point of the proposition are then similar to those established above. **QED.**

B.6 Proof of Proposition 4

The objective of party i rewrites

$$vs_i(\chi_i y - rP(1/2 - vs_i)) + (1 - vs_i)\frac{P/2-1}{P-1} rPvs_i, \tag{B.62}$$

so

$$vs_i(\chi_i y + 1/2rP) + vs_i(1 - vs_i)\frac{-P/2}{P-1} rP. \tag{B.63}$$

The first-order equation then writes

$$-1/Pn_s + \frac{\partial vs_i}{\partial \tau_s}\{\chi_i y + 1/2rP + (1-2vs_i)\frac{-P/2}{P-1} rP\} = 0. \tag{B.64}$$

In a symmetric equilibrium, this simplifies to the condition given in the main text,

$$-\frac{1}{P} n_s + (1-\epsilon)\frac{n_s}{n}\frac{\theta_s}{c_s}\{\chi_i y + \frac{1}{2} r \frac{P}{P-1}\}\frac{P-1}{P^2} = 0. \tag{B.65}$$

It is then direct that $c_s/\theta_s = c_k/\theta_k$ for any pair $s, k \in S$. Following the steps of the first application, this implies that

$$c_s = \frac{\theta_s}{n\overline{\theta}}(y(1-\chi_i) - r(\frac{P}{2} - 1)). \tag{B.66}$$

Replacing θ_s/c_s in the first-order condition then, we find that

$$-\frac{1}{P} + (1-\epsilon)\overline{\theta}\frac{1}{y(1-\chi) - r(P/2-1)}\{\chi_i y + \frac{1}{2}r\frac{P}{P-1}\}\frac{P-1}{P^2} = 0, \tag{B.67}$$

from which we deduce that

$$y\chi = \frac{y - r(P/2 - 1 + (1-\epsilon)\overline{\theta}/2)}{1 + (1-\epsilon)\overline{\theta}(P-1)/P}. \tag{B.68}$$

Observe that if $r = 0$ we find the result of the first application. It is interesting to notice that in a symmetric equilibrium, the cost of forming a winning coalition is precisely equal to the expected benefit from participating to governing coalitions without being the formateur. Indeed, the expected utility of party i simplifies to

$$W = \frac{y\chi}{P}. \tag{B.69}$$

Since χ is decreasing in the number of parties, and denoting c the cost of party formation, we find that the equation $W = (1/P) \cdot y\chi = c$ admits a unique solution and deduce the formula of P_c^* given in the main text.

QED.

B.7 Proof of Lemma 1

With the notations of the canonical model of Section 3, i and j are ranked first by the citizenry when

$$\min(\eta(i, P)vs_i^t, \eta(j, P)vs_j^t) > \eta(k, P)vs_k^t \text{ for any } k \in P \setminus \{i, j\}. \tag{B.70}$$

Given that

$$\eta(i, P) = \frac{\mu_i}{\sum_{k \in P} \mu_k}, \tag{B.71}$$

with μ_i distributed according to a Fréchet distribution of cdf $F(\mu) = exp(-\mu^{-K})$, we deduce i and j are ranked first when

$$\min(\mu_i vs_i^t, \mu_j vs_j^t) > \mu_k vs_k^t \text{ for any } k \in P \setminus \{i, j\}. \tag{B.72}$$

It is easy to show that the distribution of $\min(\mu_i vs_i^t, \mu_j vs_j^t)$ is given by

$$F_{ij}(\epsilon) = \exp(-vs_i^K \epsilon^{-K}) + \exp(-vs_j^K \epsilon^{-K}) - \exp(-[vs_i^K + vs_j^K]\epsilon^{-K}). \tag{B.73}$$

Indeed,

$$\Pr(\min(\mu_i vs_i^t, \mu_j vs_j^t) \geq \epsilon) = \Pr(\mu_i vs_i^t \geq \epsilon)\Pr(\mu_j vs_j^t \geq \epsilon), \tag{B.74}$$

so

$$\Pr(\min(\mu_i vs_i^t, \mu_j vs_j^t) \geq \epsilon) = 1 - \exp(-vs_i^K \epsilon^{-K}) - \exp(-vs_j^K \epsilon^{-K}) + \exp(-[vs_i^K + vs_j^K]\epsilon^{-K}), \tag{B.75}$$

from which we deduce $F_{ij,s}(\epsilon)$. The probability that the pair i, j is chosen by an individual with attribute s can then be expressed as

$$ij(q_i, q_{-i})) = \int_0^{\infty} \prod_{k \in P \setminus \{i,j\}} F(\epsilon/vs_k) dF_{(ij)}(\epsilon). \tag{B.76}$$

By substituting $F_{ij}(.)$ and $F(.)$, we find the formula given in the main text.

QED.

B.8 Proof of Proposition 5

In a symmetric equilibrium, the first-order condition with respect to the transfers $\tau_{s,i}$ to group s simplifies to

$$-\frac{n_s}{P} + K\chi(1-\epsilon)\frac{\theta_s}{c_s}\frac{n_s}{n}[\frac{1}{2}\frac{(P-2)(2P-1)}{P^2(P-1)} + \frac{1}{2}\frac{1}{P}] = 0 \tag{B.77}$$

when the solution is interior. This implies that $\theta_s/c_s = \theta_k/c_k$ for any pair $s, k \in S$. From the budget constraint we still deduce that

$$c_s = \frac{\theta_s}{n\bar{\theta}} y(1-\chi). \tag{B.78}$$

We can substitute θ_s/c_s in the first-order condition in order to determine the optimal level of extraction χ_m in runoff elections.

QED.

References

Acemoglu, Daron, James A. Robinson, and Thierry Verdier (2004), "Kleptocracy and divide-and-rule: A model of personal rule." *Journal of the European Economic Association*, 2, 162–192.

Ames, B. (2001), *The Deadlock of Democracy in Brazil*. Interests, Identities, And Institutions In Comparative Politics, University of Michigan Press.

Amorim Neto, O (2002), "Legislative politics in latin america." In *Electoral laws and their political consequencesn* (S. Morgenstern and B. Nacif, eds.), 48–78, Cambridge: Cambridge University Press.

Amorim Neto, Octavio (2006), "The presidential calculus: Executive policy making and cabinet formation in the americas." *Comparative Political Studies*, 39, 415–440.

Banks, Jeffrey S. and John Duggan (2005), "Probabilistic voting in the spatial model of elections: The theory of office-motivated candidates." In *Social Choice and Strategic Decisions: Essays in Honor of Jeffrey S. Banks* (David Austen-Smith and John Duggan, eds.), 15–56, Springer Berlin Heidelberg, Berlin, Heidelberg.

Baron, David P. (1993), "Government formation and endogenous parties." *The American Political Science Review*, 87, 34–47.

Baron, David P. and Daniel Diermeier (2001), "Elections, governments, and parliaments in proportional representation systems." *The Quarterly Journal of Economics*, 116, 933–967.

Becker, Gordon M., Morris H. Degroot, and Jacob Marschak (1963), "Stochastic models of choice behavior." *Behavioral Science*, 8, 105–142.

Besley, Timothy and Stephen Coate (1997), "An economic model of representative democracy." *The Quarterly Journal of Economics*, 112, 85–114.

Bordignon, Massimo, Tommaso Nannicini, and Guido Tabellini (2016), "Moderating political extremism: Single round versus runoff elections under plurality rule." *American Economic Review*, 106, 2349–70.

Bordignon, Massimo, Tommaso Nannicini, and Guido Tabellini (2017), "Single round vs. runoff elections under plurality rule: A theoretical analysis." *European Journal of Political Economy*, 49, 123 – 133.

Bouton, Laurent (2013), "A theory of strategic voting in runoff elections." *American Economic Review*, 103, 1248–88.

Bouton, Laurent and Micael Castanheira (2012), "One person, many votes: Divided majority and information aggregation." *Econometrica*, 80, 43–87.

Bouton, Laurent, Micael Castanheira, and Aniol Llorente-Saguer (2016), "Divided majority and information aggregation: Theory and experiment." *Journal of Public Economics*, 134, 114 – 128.

Bouton, Laurent, Micael Castanheira, and Aniol Llorente-Saguer (2017), "Multicandidate elections: Aggregate uncertainty in the laboratory." *Games and Economic Behavior*, 101, 132 – 150. Special Issue in Honor of John O. Ledyard.

Bouton, Laurent and Gabriele Gratton (2015), "Majority runoff elections: Strategic voting and duverger's hypothesis." *Theoretical Economics*, 10, 283–314.

Brams, Steven J. and Michael K. O'Leary (1970), "An axiomatic model of voting bodies." *American Political Science Review*, 64, 449–470.

Caillaud, Bernard and Jean Tirole (2002), "Parties as political intermediaries." *The Quarterly Journal of Economics*, 117, 1453–1489.

Callander, Steven (2005), "Duverger's hypothesis, the run-off rule, and electoral competition." *Political Analysis*, 13, 209–232.

Chamon, Marcos, Sergio Firpo, Joao M. P. de Mello, and Renan Pieri (2018), "Electoral rules, political competition and fiscal expenditures: Regression discontinuity evidence from brazilian municipalities." *The Journal of Development Studies*, 0, 1–20.

Cohen, Ronald N. and Kenneth A. Shepsle (1990), "Multiparty competition, entry, and entry deterrence in spatial models of elections." In *Advances in the Spatial Theory of Voting* (James M. Enelow and Melvin J. Hinich, eds.), Cambridge University Press.

Coughlin, Peter and Shmuel Nitzan (1981), "Directional and local electoral equilibria with probabilistic voting." *Journal of Economic Theory*, 24, 226 – 239.

Coughlin, P.J. (1992), *Probabilistic Voting Theory*. Probabilistic Voting Theory, Cambridge University Press.

Cox, G.W. (1997), *Making Votes Count: Strategic Coordination in the World's Electoral Systems*. Political Economy of Instituti, Cambridge University Press.

Dickson, Eric S. and Kenneth Scheve (2010), "Social identity, electoral institutions and the number of candidates." *British Journal of Political Science*, 40, 349?375.

Diermeier, Daniel, Hulya Eraslan, and Antonio Merlo (2002), "Coalition governments and comparative constitutional design." *European Economic Review*, 46, 893–907.

Diermeier, Daniel and Antonio Merlo (2004), "An empirical investigation of coalitional bargaining procedures." *Journal of Public Economics*, 88, 783–797.

Dixit, Avinash and John Londregan (1996), "The determinants of success of special interests in redistributive politics." *The Journal of Politics*, 58, 1132–1155.

Dow, Jay K. and James W. Endersby (2004), "Multinomial probit and multinomial logit: a comparison of choice models for voting research." *Electoral Studies*, 23, 107 – 122.

Duverger, M. (1964), *Political Parties: Their Organization and Activity in the Modern State*, 3rd edition. London: Methuen.

Eaton, Jonathan and Samuel Kortum (2002), "Technology, Geography, and Trade." *Econometrica*, 70, 1741–1779.

Figueiredo, Argelina Cheibub and Fernando Limongi (2000), "Presidential power, legislative organization, and party behavior in brazil." *Comparative Politics*, 32, 151–170.

Fudenberg, Drew and Jean Tirole (1991), *Game Theory*. MIT Press, Cambridge, MA. Translated into Chinesse by Renin University Press, Bejing: China.

Fujiwara, Thomas (2011), "A regression discontinuity test of strategic voting and duverger's law." *Quarterly Journal of Political Science*, 6, 197–233.

Hinich, Melvin J. (1977), "Equilibrium in spatial voting: The median voter result is an artifact." *Journal of Economic Theory*, 16, 208 – 219.

Hinich, Melvin J., John O. Ledyard, and Peter C. Ordeshook (1972), "Nonvoting and the existence of equilibrium under majority rule." *Journal of Economic Theory*, 4, 144 – 153.

Junior, Arnaldo Mauerberg, Carlos Pereira, and Ciro Biderman (2015), "The evolution of theories about the brazilian multiparty presidential system." *Journal of Politics in Latin America*, 7, 143–161.

Lin, Tse-Min, James M. Enelow, and Han Dorussen (1999), "Equilibrium in multicandidate probabilistic spatial voting." *Public Choice*, 98, 59–82.

Lindbeck, Assar and Jorgen Weibull (1987), "Balanced-budget redistribution as the outcome of political competition." *Public Choice*, 52, 273–297.

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

Lizzeri, Alessandro and Nicola Persico (2005), "A drawback of electoral competition." *Journal of the European Economic Association*, 3, 1318–1348.

Luce, R.D. (1959), *Individual Choice Behavior: A Theoretical Analysis*. Wiley.

Mainwaring, S. (1999), *Rethinking Party Systems in the Third Wave of Democratization: The Case of Brazil*. Political Science, Stanford University Press.

Mainwaring, S. and M.S. Shugart (1997), *Presidentialism and Democracy in Latin America*. Cambridge Studies in Comparative Politics, Cambridge University Press.

Martinez-Gallardo, C (2005), *Presidents, Posts, and Policy: Ministerial Appointments and Political Strategy in Presidential Regimes*. Ph.D. Dissertation, Columbia University.

McFadden, Daniel (1974), "Conditional Logit Analysis of Qualitative Choice Behavior." In *Frontiers in Econometrics* (P. Zarembka, ed.), 105–142, Academic Press, New York.

Melo, Karine (2017), "Political reform approved in brazil, some rules apply for 2018 elections." *Agencia Brasil*, URL `http://agenciabrasil.ebc.com.br/en/politica/noticia/2017-10/political-reform-approved-brazil-some-rules-apply-2018-elections`.

Morelli, Massimo (2004), "Party formation and policy outcomes under different electoral systems." *The Review of Economic Studies*, 71, 829–853.

Myerson, Roger B. (1993), "Incentives to cultivate favored minorities under alternative electoral systems." *The American Political Science Review*, 87, 856–869.

Osborne, Martin (1995), "Spatial models of political competition under plurality rule: A survey of some explanations of the number of candidates and the positions they take." *Canadian Journal of Economics*, 28, 261–301.

Osborne, Martin J. and Al Slivinski (1996), "A model of political competition with citizen-candidates*." *The Quarterly Journal of Economics*, 111, 65–96.

Padró i Miquel, Gerard (2007), "The control of politicians in divided societies: The politics of fear." *The Review of Economic Studies*, 74, 1259–1274.

Persson, T. and G.E. Tabellini (2002), *Political Economics: Explaining Economic Policy*. Zeuthen lecture book series, MIT Press.

Raile, Eric D., Carlos Pereira, and Timothy J. Power (2011), "The executive toolbox: Building legislative support in a multiparty presidential regime." *Political Research Quarterly*, 64, 323–334.

Riker, William H. (1982), "The two-party system and duverger's law: An essay on the history of political science." *The American Political Science Review*, 76, 753–766.

Sartori, G (1986), "The influence of electoral systems: Faulty laws or faulty method?" In *Electoral laws and their political consequencesn* (B. Grofman and A. Lijphart, eds.), 43–68, New York: Agathon Press.

Scherlis, Gerardo (2014), "Political legitimacy, fragmentation and the rise of party-formation costs in contemporary latin america." *International Political Science Review*, 35, 307–323.

Schofield, Norman (1993), "Political competitition and multiparty coalition governments." *European Journal of Political Research*, 23, 1–33.

Schofield, Norman (1997), "Political competition." *European Journal of Political Research*, 31, 179–192.

Schofield, Norman (2007), "The mean voter theorem: Necessary and sufficient conditions for convergent equilibrium." *The Review of Economic Studies*, 74, 965–980.

Schofield, Norman, Itai Sened, and David Nixon (1998), "Nash equilibrium in multiparty competitionwith "stochastic" voters." *Annals of Operations Research*, 84, 3–27.

Shepsle, K.A. (1991), *Models of Multiparty Electoral Competition*. Fundamentals of pure and applied economics, Harwood Academic Publishers.

Snyder, James M. and Michael Ting (2002), "An informational rationale for political parties." *American Journal of Political Science*, 46, 90–110.

The Economist (2017), "Elections on a shoestring: Brazil's congress starts to reform itself." *The Economist*.

Vaz de Melo, Pedro O. S. (2015), "How many political parties should brazil have? a data-driven method to assess and reduce fragmentation in multi-party political systems." *PLOS ONE*, 10, 1–24.