WHY VOTE FOR LOSERS?

by

Micael Castanheira[*]
ECARES (ULB), FNRS, and CEPR
email: mcasta@ulb.ac.be

Abstract

Voting Theory generally concludes that, in first-past-the-post elections, all votes should go to effective candidates (Duverger's Law), and parties should adopt a similar platform (Median Voter Theorem). However, such predictions are not always met in practice. We show why divergence and vote dispersion is a natural outcome when (i) parties are opportunistic, (ii) there is uncertainty on the position of the median voter and (iii) elections are repeated. "Voting for losers" increases the informational content of elections, and may induce mainstream parties to relocate towards extremists. As a result, to maximize their probability of being elected, they do not adopt median platforms, but instead diverge to a certain extent.

JEL Classification Numbers: D72, D81

Keywords: Voting, Duverger's Law, Signalling, Poisson games.

[*]I wish to express my gratitude to Juan Carrillo, Mathias Dewatripont, Denis Gromb, and Gérard Roland for their support since the first stages of this project. The present version of the paper owes a lot to comments and suggestions by Estelle Cantillon, Alain Desdoigts, Abdul Noury, Nicolas Sahuguet, two anonymous referees and the editor. I also greatly benefitted from discussions with Patrick Bolton, Laura Bottazzi, Isabelle Brocas, Timothy Feddersen, John Fingleton, Steve Lawford, Mikko Leppämäki, James Mirrlees, Roger Myerson, Michele Polo, Anouk Rivière, Howard Rosenthal, Georges Siotis, Guido Tabellini, Jean Tirole, and seminar participants at ULB (Brussels), CentER (Tilburg), ESF (Florence), IGIER (Milan), SSE (Stockholm), Universidad Carlos III (Madrid), and the Helsinki University. The financial support from the FNRS is gratefully acknowledged. The usual disclaimer applies.

1 Introduction

Standard voting theory predicts that rational voters should concentrate all their votes on parties that have a serious chance of winning the election (Duverger's Law). This prediction proves to be extremely strong theoretically[1] but is often far off the mark in practice. Historical events abound with examples of candidates who receive a substantial vote share, even though they are certain to lose the election. To explain this deviation from Duverger's Law, one needs to consider more recent developments in the literature. Myerson (1998) and others have shown that voters can use elections to pool private pieces of information. In that setup, one can then show that voters may want to "communicate" in elections by supporting trailing parties, if they have a chance of winning in the future (Piketty 2000).

Still, two problems remain unaddressed. First, the interactions between the strategy of the voters and that of the parties has never been analyzed in such a setup. Which kind of political stance should small parties adopt? How does their presence affect mainstream parties' strategy? Second, there are cases in which voters consistently support pure "losers," that is parties who do not appear to have *any* chance of winning, either in present or in future elections. Why voting for pure losers?

To address these questions, we develop a two-period model of repeated elections, in which the distribution of preferences in the electorate is random. Like in the Downs-Hotelling framework, voters are purely outcome-minded and rational, and parties maximize their probability of winning. In that setup, we show why small parties receive a strictly positive vote share even if they have no chance of winning either present or future elections, and why *both* "losers" and mainstream parties may select non-centrist platforms in the first election. Put differently, we show that,

[1] See for instance Duverger (1954), Palfrey (1989), Cox (1994), or Myerson (1999).

in contrast with the predictions of the Median Voter Theorem (Hotelling (1929), Downs (1957), Black (1958)), the presence of extremist voters is sufficient to induce platform divergence. Opportunistic parties thus mimic the behaviour of ideologically-motivated parties (see Witmann (1983), Calvert (1985), Alesina (1988), Alesina and Rosenthal (1995), or Roemer (1997)).

The rationale behind these results is the following. By definition, when the distribution of preferences is uncertain, the exact position of the median voter is unknown. If this uncertainty is strong enough, observing the vote results of only two parties is not sufficient to learn where the median voter stands. The vote share of losers thus reveals *additional* information. For instance, the vote results of the mainstream "left-wing" and "right-wing" parties may reveal whether the median voter is left- or right-leaning, but does not tell "how far" the median voter is. Observing additional information (the vote shares of extremist parties) instead provides this information. Then, mainstream parties have an incentive to adopt extreme positions in the second election if they observe that extremist parties collect a substantial vote share. In other words, by voting for losers, extremist voters can attract mainstream parties towards their favorite position. Extremist platforms are implemented with positive probability, even if extremist parties never get elected. Anticipating this reaction, extremist voters value positively a vote for a pure loser. How does this influence the strategy of the parties? We show that, since these small parties only collect votes from extremist voters, they endogenously choose to adopt extremist platforms in order to make themselves most "attractive" to their core electorate, . Similarly, mainstream parties must also make themselves look "attractive" to their extremist electorate to increase their probability of election. For this reason, mainstream parties may also be constrained to adopt extreme platforms in the first election.

Note that the role of information in elections and its influence on voting behaviour has already been studied, although in different frameworks. Lohmann (1994) shows how *pre-election* political actions (*e.g.* strikes) generates information that can improve – or worsen – the ability of voters to select the best platform at the election stage. More recently, Myerson (1998) generalized the Condorcet Jury Theorem,[2] to show that elections aggregate information efficiently. In his setup, voters share a common valuation about the outcome of the vote: they all prefer to elect a "good" party rather than a "bad" one. That is, voters are a priori indifferent between two parties. Then, each voter receives an imperfect signal about the value of the parties, which tilts their preferences. In this setup, Myerson shows that there exists an "informative" equilibrium, in which the "good" candidate is elected with a probability approaching one as the size of the electorate increases.

Feddersen and Pesendorfer (1997) extend that framework to an electorate with heterogenous preferences. They show that voters who are "almost" indifferent between the parties vote informatively, and hence that the election still selects the good party with probability one.[3] Piketty (1999) surveys this literature in more detail.

Piketty's (2000) analysis is most closely related to the present one. Like him, we analyze a repeated-election framework: two separate elections take place at two different dates, and the first election generates information, which is exploited in the second one. His setting extends Myerson's results to three-party competition: a majority has imperfect signals about which of two candidates they should support, whereas the minority supports a third candidate. He shows that the majority will use the first election to "communicate," and thereby uncover which party they must vote for in the second election.

Our model departs from Condorcet's approach in two respects. First, we jointly

[2]Condorcet (1785) demonstrated that, when information is dispersed across individuals, the decision is more likely to be appropriate if it is made by a Jury rather than by a single Judge.

[3]However, Caillaud and Tirole (1997) show that aggregate uncertainty weakens this result.

analyze the strategy of the voters and that of the parties. This extends existing results to endogenize the strategy of the parties,[4] whereas platforms are assumed exogenous in the above-mentioned papers. Second, voters know ex ante, both which is their preferred platform and who will be the main challengers in the second election. Like in the standard Downs-Hotelling model, voters have different (private) valuations of the possible election outcomes, by opposition to Feddersen and Pesendorfer (1997) and Myerson (1998). However, they do not know whether or not they are a majority, and they cannot condition their second-period strategy on the information revealed in the first election, in contrast with Piketty (2000).

The remainder of the paper is organized as follows. Section 2 lays out the model. Section 3 describes the equilibrium behaviour of the voters and Section 4 that of the parties. Section 5 analyzes some extensions and Section 6 concludes. Most proofs are relegated to the Appendix.

2 The Model

We focus on first-past-the-post elections and consider four purely opportunistic parties that compete for a plurality of the votes. Voters behave strategically and their utility only depends on the platform that is implemented. To facilitate comparisons with the literature, we assume that the policy space is one-dimensional (platforms and preferences are represented by points on the real line), and that voters have single-peaked preferences. The strategy space of the parties is represented by the set of platform positions they can adopt, and the strategy space of the voters is the set

[4] Two exceptions are Razin (2000) and Shotts (2000). Shotts provides a rationale for abstention in a two-period game. He shows that abstention carries a moderation signal to the parties, who use it to determine their second-period platforms. Razin analyzes the interaction between the margin of victory and parties' platforms. His approach is however quite different: first, he considers the case of a single election with two ideologically-minded parties. Second, the platforms preferred by the parties *and* by the voters vary according to the state of nature.

of parties for which they can vote. In this way, the model follows the "standard" assumptions of the Downs-Hotelling model with strategic voting (see Section 2.1).

The model extends this classical Downs-Hotelling framework in two directions. Firstly, we assume that both voters and parties are uncertain about the distribution of preferences in the electorate. In particular, even the *shape* of this distribution is unknown: there is "aggregate uncertainty" about the preferences of the electorate (see Section 2.2). Secondly, we consider a *dynamic* election game, *i.e.* there are two elections, at two different dates (see Section 2.3).

2.1 Players

Parties are indexed by p, and P denotes the set of parties, with $P = \{\alpha, \beta, \gamma, \delta\}$. Throughout the paper, α and β denote the "leaders" (who could be seen as "traditional" parties) and γ and δ represent the "losers" (who could be seen as "challengers," or newly created fringe parties).

All parties are assumed to be purely opportunistic. That is, from the perspective of period 1, the intertemporal utility of party p is:

$$\Pi_p\left(\mathsf{x}^1, \mathsf{x}^2\right) = \Pr(W^1 = p|\mathsf{x}^1) + \Pr\left(W^2 = p|\mathsf{x}^2\right), \tag{1}$$

where $\Pr(W^t = p|\mathsf{x}^t)$ denotes the probability that party p wins election t given the set of platforms $\mathsf{x}^t \equiv \{x^t_\alpha, x^t_\beta, x^t_\gamma, x^t_\delta\}$.[5] Parties are free to select any platform on the real line. At each election date $t = 1, 2$, each party p will choose the platform $x^t_p \in \mathcal{R}$ that maximizes his probability of winning.

Voters are indexed by v and only care about the platform implemented after each

[5]The discount factor is set equal to 1, both for the parties and the voters. This is without loss of generality, and all results directly extend to any strictly positive –but finite– discount factors, which need not be the same for voters and parties.

election. For simplicity, assume there are four types of voters: $v \in V = \{eL, L, R, eR\}$, where eL stands for extreme left, L for left, R for right, and eR for extreme right. Ranking these preferences along the real line, we set $eL < L < 0 < R < eR$, and assume symmetry, *i.e.* $eL = -eR$, $L = -R$, and $eR - R \simeq R - L$. In Section 3.2, we briefly discuss how results would be affected if there were only three types of voters.

The voters' instantaneous utility function $u(\cdot)$ is single-peaked, symmetric, and not too concave.[6] As of period 1, the expected intertemporal utility of a voter with type v is given by:

$$\mathsf{E}U_v = \mathsf{E}_{W^1}\left[u\left(|x_{W^1} - v|\right)\right] + \mathsf{E}_{W^2}\left[u\left(|x_{W^2} - v|\right)\right], \text{ with } -\infty < u' < 0.$$

With this utility function, voters would be exactly indifferent between parties that propose a same platform. To alleviate technical problems linked to exact indifference, assume that, in case several parties propose a same platform, each voter uses some other criterium that allows her to rank those parties from "most preferred" to "least preferred" (lexicographic preferences). Moreover, each ranking is equivalently likely, and the intensity of this preference is arbitrarily small.

The only action voters can take is to vote for one of the parties (abstention is discussed in Section 5). Their action space is thus given by $P = \{\alpha, \beta, \gamma, \delta\}$. Accordingly, for any voter v, the space of mixed strategies is the 3-dimensional simplex:

$$\Phi_v = \left\{ \phi_{v,p} \colon \phi_{v,p} \geq 0, \forall p \text{ and } \sum_{p \in P} \phi_{v,p} = 1 \right\}, \quad \forall v \in V,$$

where $\phi_{v,p}$ denotes the probability that a voter of type v casts a ballot for party p. Let also Φ denote the matrix of all voters' strategies.

Importantly, voters have exact knowledge of their own type and bliss point, which implies that the "state of nature" (see below) never affects their preferences. By

[6] In particular, the standard utility function $u(x - v) = -|x - v|$ would be sufficient for our results to hold.

contrast, no player is informed about the actual distribution of preferences in the electorate, which depends on the state of nature.

2.2 Aggregate uncertainty and pivot probabilities

2.2.1 Aggregate uncertainty

Aggregate uncertainty is introduced by letting "Nature" select the distribution of preferences in the electorate, leaving the position of the median voter −which we will refer to as the *state of nature*− unknown to all players in the game. There are four, equally likely,[7] states of nature, $\omega \in \Omega = \{\omega_{eL}, \omega_L, \omega_R, \omega_{eR}\}$, and the position of the median voter ($\mu(\omega)$) is different in each of them: $\mu(\omega_v) = v$.

To capture this aggregate uncertainty, it proves convenient to adopt the following specification: in state of nature ω, the probability $r_v(\omega)$ that a randomly picked voter has type v is equal to:

$$r_v(\omega) \equiv \Pr(v|\omega) = \begin{cases} r_v & (< 1/2), \text{ if } \mu^\omega \neq v \\ r_v + \eta & (> 1/2), \text{ if } \mu^\omega = v, \end{cases} \tag{2}$$

subject to $r_v > 0$, $\forall v$ and $\sum_{v \in V} r_v = 1 - \eta$. That is to say, a fraction $1 - \eta$ of the population ($0 < \eta < 1$), is attributed a type independently of the state of nature. Conversely, nature increases the share of voters with the "median voter type" by η, and the restrictions on r_v and $r_v + \eta$ ensure that there are voters of every type in all states of nature, but that they cannot represent a majority unless natures decides so.

Finally, to maintain as much symmetry as possible between the "left" and the "right," set $r_{eL} = r_{eR}$, and $r_L = r_R$.

[7] Section 5 discusses deviations from this symmetry assumption.

2.2.2 Prior Beliefs

Voters will use all available information to assess the probability of being in each state of nature. Assuming that each state of nature is equally likely *ex ante*, voters use their – privately observed – type to shape their beliefs through Bayesian updating:

$$b(\omega|v) = \frac{\Pr(\omega)\ r_v(\omega)}{\sum_{\tilde{\omega}\in\Omega} \Pr(\tilde{\omega})\ \ r_v(\tilde{\omega})}, \tag{3}$$

where $\Pr(\omega) = 1/4$ is the probability that nature chooses ω. Hence, $b(\omega|v)$ denotes voters' beliefs about the state of nature ω *prior* to the first election.

2.2.3 Winning Probabilities and Poisson Games

In first-past-the-post elections, the party collecting the highest number of votes is elected. However, some rule must be devised in case there is a tie. The classical assumption is that ties are resolved by the toss of a fair coin. However, and without losing generality, computations can be simplified by assuming that the party ranked alphabetically first is elected in case of ties. For instance, if α and β receive the same number of votes, α gets elected. Letting $\tilde{z}_p^t$ denote the (random) **number** of votes for party p at time t, we have:

$$\begin{aligned} \Pr\left(W^t = \alpha\right) &= \Pr\left(\tilde{z}_\alpha^t \geq \tilde{z}_p^t,\ \forall\, p \in P\setminus\alpha\right), \\ \Pr\left(W^t = \beta\right) &= \Pr\left(\tilde{z}_\beta^t > \tilde{z}_\alpha^t \text{ and } \tilde{z}_\beta^t \geq \tilde{z}_p^t,\ p \in \{\gamma, \delta\}\right), \end{aligned}$$

and likewise for the other parties.

Thus, to derive the probability that one party is elected, the distribution of the number of votes for each party must be characterized. Clearly, this distribution is determined by the number of voters who participate in the election and on their voting strategy. The equilibrium strategy of the voters will be determined in Section 3. It remains to characterize the distribution of the number of voters.

Myerson (1997, 1998, 2000) shows that if the electorate is large, and if each voter faces some exogenous (and not-too-small) probability of not showing up at the election, then the total number of voters who participate in the election follows a Poisson Distribution:

$$\tilde{Z} \sim \mathcal{P}(\lambda),$$

where $\tilde{Z}$ is the (random) total number of voters, $\mathcal{P}$ denotes the Poisson distribution, and the parameter λ represents the expected total number of votes. Following his approach, we assume that the total number of voters is distributed according to this Poisson law and that, for a given number of voters and a given state of nature, preferences are attributed by independent and identically distributed (i.i.d.) draws, with probability $r_v(\omega)$ (see (2) above). This implies that the number of voters with type v follows a distribution $\mathcal{P}(r_v(\omega)\ \lambda)$ in state ω.

From Myerson's work, we know several properties of games. These are summarized in Appendix 1 and will be introduced in the text when needed. Among other things, they allow us to characterize the probability that a given vote is pivotal in determining the outcome of the election (the "pivot probability").

2.3 Timing

The election game is repeated twice. At time 0, Nature chooses a state of nature ω, which prevails for the whole game. At time 1, the first election is held: parties propose a platform, voters' types are attributed according to the Poisson distribution, and voters cast their ballot. The party receiving the largest number of votes is elected, and payoffs are realized. Beliefs about the different states of nature are updated and, at time 2, there is a second election: parties select a new platform, voters' types are attributed, ballots are cast, a party is elected, and time-2 payoffs are realized. The game then ends.

Summing up, the model considers four types of voters and four opportunistic parties who face two subsequent elections. Nature decides the distribution of preferences in the electorate, and hence where the median voter will be located. Then, parties are free to select any location x_p^t on the real line so as to maximize their probability of election, and voters observe these locations before deciding which party they want to vote for. Finally, since both the voters and the parties are uncertain about the state of nature, they will use first-election results to update their second-period beliefs about the relative likelihood of the different states of nature.

3 Voting for losers and its effect on platforms

To solve for the equilibrium strategy of the voters and that of the parties, we look for Perfect Bayesian Equilibria (PBE) of this game, and this for the limiting case of infinitely large populations: $\lambda \to \infty$. Solving for these equilibria by backward induction, we first focus on the second period of the game.

3.1 Second-period equilibrium

In the second period, two elements determine voting behaviour: *i)* the *positions* of the parties, and *ii)* the *pivot probability* of electing a party over another. Hence, to determine the second-period equilibrium of the game, we have to characterize those pivot probabilities.

One ballot is pivotal in electing party p over p' (we denote this event by $piv_{pp'}$) if two conditions are met: first, in the absence of this ballot (if the voter abstained), party p' must be elected. By Property A3 in Appendix 1, a necessary condition for this to happen is that the vote share of p' is at least as large as the vote shares of the parties other than p and p'. To capture this condition, define the indicator function $\mathcal{I}_{p,p'}(\Phi, \omega)$, which takes value zero if at least one of the parties other than p and p' has a larger vote share than p' in the state of nature ω, value one if the vote share of p' is strictly larger than that of these parties, and value $1/n$ if n parties have the same vote share. The second condition is that, by casting her ballot for p, the voter modifies the outcome of the election. Clearly, this can only happen if p is trailing behind p' by at most one vote. By Property A2 in Appendix 1, we have:

$$\begin{aligned} \Pr\left(z_{p'} = z_p | \Phi, \omega\right) &= \frac{\exp\left[\theta^{\omega}_{pp'}(\Phi)\ \lambda\right]}{2\sqrt{\pi}\,\lambda\,(s^{\omega}_p\, s^{\omega}_{p'})^{1/4}}, \\ \Pr\left(z_{p'} = z_p + 1 | \Phi, \omega\right) &= \sqrt{s^{\omega}_{p'}/s^{\omega}_p}\ \Pr\left(z_{p'} = z_p | \Phi, \omega\right), \end{aligned} \tag{4}$$

where $\theta^{\omega}_{pp'} = -\left(\sqrt{s^{\omega}_p(\Phi)} - \sqrt{s^{\omega}_{p'}(\Phi)}\right)^2$, and $s^{\omega}_p(\Phi) \equiv \sum_v r_v(\omega)\ \phi_{v,p}$ denotes the *expected vote share* of party p in state ω. The argument of the exponential, $\theta^{\omega}_{pp'}(\Phi)$, is called the *magnitude* of this probability, and the smaller (the more negative) is this magnitude, the smaller is the pivot probability.

Taking these two conditions together, and summing across states of nature, we thus find the probability that a vote is pivotal between parties p and p':

$$\begin{aligned} \Pr\left(piv_{pp'} | \Phi, v\right) &= \sum_{\omega \in \Omega} \Pr(\omega | \mathsf{I}_v)\ \mathcal{I}_{p,p'}(\Phi, \omega) \times \Pr\left(z_{p'} = z_p | \Phi, \omega\right), \text{ if } p \succ p' \\ &= \sum_{\omega \in \Omega} \Pr(\omega | \mathsf{I}_v)\ \mathcal{I}_{p,p'}(\Phi, \omega) \times \Pr\left(z_{p'} = z_p + 1 | \Phi, \omega\right), \text{ if } p \prec p', \end{aligned} \tag{5}$$

where $p \succ p'$ means "p is alphabetically after p'," and I_v represents all the information available to the voter at time $t = 2$.

Given (5) and the definition of $\theta^{\omega}_{pp'}$, it follows immediately that, if two parties lead the election (that is, if there exist two parties, say α and β, such that $\min[s^{\omega}_{\alpha}, s^{\omega}_{\beta}] > \max\left[s^{\omega}_{\gamma}, s^{\omega}_{\delta}\right]$, $\forall \omega \in \Omega$), we have:

Property 1 (Myerson (2000), Corollary 1)

Consider any party $p \in \{\gamma, \delta\}$. *If* $s_p^\omega < \min(s_\alpha^\omega, s_\beta^\omega), \forall\, \omega \in \Omega$, *then:*

$$\lim_{\lambda \to \infty} \frac{\Pr(piv_{p\alpha}|\Phi, v)}{\Pr(piv_{\beta\alpha}|\Phi, v)} = \lim_{\lambda \to \infty} \frac{\Pr(piv_{p\beta}|\Phi, v)}{\Pr(piv_{\alpha\beta}|\Phi, v)} = 0, \forall v.$$

Proof. Immediate from Property A2 in Appendix 1. ■

In words, any voter who expects two parties (α and β) to lead the election realizes that her only chance to affect the outcome of the election will be to also vote for one of these two parties. Technically, this happens because the magnitude of the probability of electing a "serious contender" (α or β) is larger than that of electing a "loser" (γ or δ). Using this property, we find:

Lemma 1 *In the second election, if parties* γ *and* δ *are not perceived as serious contenders, then the PBE is unique and such that* i) γ *and* δ *receive no vote* ($s_\gamma^\omega = s_\delta^\omega = 0$, $\forall \omega$), ii) α *and* β *locate at the perceived position of the median voter.*

Proof. See Appendix A.2.1. ■

Result i in Lemma 1 is exactly the game-theoretic interpretation of Duverger's Law proposed by Palfrey (1989) and Cox (1994, 1997): "*Some voter, whose favorite candidate has a poor chance of winning, notices that she has a preference between the top two candidates; she then rationally decides to vote for the most preferred of these top two competitors rather than for her overall favorite, because the latter vote has a much smaller chance of actually affecting the outcome*" (Cox (1997), p.71).

Accordingly, define a *Duvergerian Outcome* as an equilibrium in which only two parties gather a positive number of votes.[8] Lemma 1 thus extends to aggregate

[8]Cox also shows that non-Duvergerian equilibria can occur, but only if parties γ or δ have exactly the same vote share as min $\left[s_\alpha^\omega, s_\beta^\omega\right]$. See Fey (1997, Theorems 3 and 4), for a proof that such equilibria are "globally expectationally unstable".

uncertainty some results that are already present in the literature: 1) absent communicational motives, strategic voting generates Duvergerian Outcomes; 2) rankings are self-fulfilling.[9]

Result *ii* in Lemma 1 is also reminiscent of existing results: parties maximize their probability of winning by locating at the position of the median voter.[10] Note that when more than two parties compete against one another, such a result needs not always hold (see *e.g.* Palfrey (1984) and Myerson and Weber (1993)). The reason why the Median Voter Theorem holds here is that, in equilibrium, only two parties receive votes. The second election thus reduces to a duel between the top two candidates, in spite of the candidacy of the two losers.

A noteworthy aspect of this result is that parties locate where they *think* the median voter stands. Remember that the median voter can take one in four positions, and we assumed that each of these positions is equally likely *ex ante*. However, at the time of the second election, parties have had the opportunity to observe the results of the first election and to update their beliefs accordingly. As is shown below, this strategic use of information by the parties will be crucial to determine equilibrium strategies in the first election.

3.2 Voting for losers in the first election

Now, let us turn our attention to the voting stage of the first election. Since first-period platforms have already been chosen at this stage, vote results in the first

[9]If, say, α and β are perceived as losers, then all voters coordinate their votes on γ and δ only. That is, there exist different stable coordination equilibria, which all are Duvergerian.

[10]This result hinges in part on the assumption that voters have lexicographic preferences. If voters were exactly indifferent between α and β when $x_\alpha^2 = x_\beta^2$, the value of a ballot for α or β would drop to zero instead of being equal to $\pm\varepsilon \to 0$. Hence, the voters' optimal strategy might be discontinuous in $x_\alpha^2 = x_\beta^2$. Introducing lexicographic preferences instead shows that this discontinuity is not robust to the introduction of even a second-order differentiation of the parties.

election can only achieve two purposes: 1) they determine which of the proposed platforms is implemented at the end of the first period. 2) they determine the updating of beliefs at time 2, and thereby platforms positions in the second period.

To assess the value of her ballot, the voter must take these two effects into account. On the one hand, her ballot might be *outcome-pivotal*, *i.e.* pivotal in affecting the outcome of the *first* election. On the other hand, it might prove *communication-pivotal*. That is, her ballot may induce parties to change their second-period perception about the location of the median voter.

To understand the link between first-period results and second-period platforms, imagine for a moment that voters only vote for α and β. Given this strategy, it is easy to check that parties always locate between L and R in the first election. For the sake of concreteness, assume that all left-wing voters prefer the platform of α, and right-wing voters prefer that of β. In this case, if α wins the first election, parties (and voters) will infer that the median voter is "left-wing." Conversely, if β wins the first election, all players will infer that the median voter is "right-wing."

But how far is the median voter? Is she extreme or moderate? From our definition of aggregate uncertainty (See (2) in Section 2.2.1), we have:

Property 2 *The aggregate share of left-wing and right-wing voters is the same in the states of nature ω_{eL} and ω_L, and in the states of nature ω_R and ω_{eR} :*

$$\begin{cases} r_{eL}(\omega_{eL}) + r_L(\omega_{eL}) = r_{eL}(\omega_L) + r_L(\omega_L) \quad (> 1/2) \\ r_{eL}(\omega_R) + r_L(\omega_R) = r_{eL}(\omega_{eR}) + r_L(\omega_{eR}) \quad (< 1/2) \end{cases} \tag{6}$$

Therefore, although observing the vote shares of α and β reveals whether left-wing states of nature (ω_{eL} and ω_L) are more or less likely than right-wing states (ω_R and ω_{eR}), observing only two vote results does not discriminate between extreme and moderate states of nature. That is, if α wins the first election, both α and β will locate in L to maximize their probability of winning the second election (see Lemma 1 above and Lemma 3 in Appendix A.2.2). Conversely, if β wins the first election, α and β locate in R in the second election.

What happens if instead voters adopt a different strategy, and vote for γ and/or δ with strictly positive probability? In this case, parties will use the vote results of the *four* parties to update their beliefs. Thus, the vote results of the losers will also influence second-period platforms. Consider the probabilities faced, *e.g.* by types eL if they vote for both α and γ with strictly positive probability (the results extend to other types and/or parties). Defining $K \equiv \log\left[s_{\alpha}^{\omega_L}/s_{\alpha}^{\omega_{eL}}\right]/\log\left[s_{\gamma}^{\omega_{eL}}/s_{\gamma}^{\omega_L}\right]$, we find:

Lemma 2 *Consider two types of voters, say eL and L. If types L only vote for α, whereas types eL vote both for parties α and γ with strictly positive probability, a vote for γ is* communication-pivotal*, and moves second-period platforms from L to eL, with a probability* $\Pr(com_{eL,L})$ *such that:*

$$\lim_{\lambda\to\infty}\log\left[\Pr(com_{eL,L})\right]/\lambda = \chi\left(\phi_{eL,\gamma}\right) = \sum_{p\in P}\chi(s_p^{\omega},\sigma(p)),\ \omega\in\{\omega_L,\omega_{eL}\},$$

$$\begin{aligned} \textit{with}\quad & \chi\left(s_p^{\omega},\sigma(p)\right) = \sigma(p)\ \left[1-\log\left(\sigma(p)/s_p^{\omega}\right)\right] - s_p^{\omega}, \\ & \sigma(\gamma)/\sigma(\alpha) = K;\ \ \sigma(\beta)/\sigma(\delta) = s_{\beta}^{\omega}/s_{\delta}^{\omega}, \end{aligned}$$

where $\chi\left(\phi_{eL,\gamma}\right)$ *is the* magnitude *of this probability and has the properties:*

$$\lim_{\phi_{eL,\gamma}\to 0}\chi\left(\phi_{eL,\gamma}\right) = 0 \quad \textit{and} \quad \partial\chi\left(\phi_{eL,\gamma}\right)/\partial\phi_{eL,\gamma} < 0,\ \forall\phi_{eL,\gamma} > 0.$$

Proof. See Appendix A.2.3. ■

In words, if types eL vote (with positive probability) for some party γ, a ballot for γ may induce parties to locate in eL instead of some other position (L in this example) in the second period. Moreover, the probability that this vote be communication-pivotal is decreasing in the fraction of eL-voters who vote for γ. The rationale for this result is the same as above: if α wins the election, parties learn that left-wing states of nature are most likely. But, is the actual state eL or L? Simply, if the share of γ is "large" (how large it should be depends on the strategy adopted by types eL, *i.e.* on $\phi_{eL,\gamma}$), parties learn that the proportion of eL's in the electorate is also large, and hence locate in eL in the second period. If instead the share of γ is "small," parties learn that it is the proportion of L's which is large, and locate in L.

The best response of a voter thus depends on the relative probability of determining the outcome of the first election (outcome-pivotability) compared to that of influencing second-period platforms (communication-pivotability). That is, it depends on the relative magnitudes of the two events. Our first proposition shows that:

Proposition 1 *For* $x_p^1 \in [eL, eR]$, $\forall p$*;* $x_\alpha^1 \simeq -x_\beta^1 \neq 0$*,* $x_\gamma^1 \simeq -x_\delta^1 \neq 0$*, and* $r_{eL} \leq r_L$*, three types of Perfect Bayesian voting equilibria can arise in the first period:*

• ***Type-I equilibria**: **four parties receive a strictly positive vote share**.*

In such an equilibrium, depending on the values of r_{eL} *and* r_L *and on platform positions, either* i) *there exist equilibria in which two parties remain 'losers,' and voting strategies are then independent of their platforms, or* ii) *each party is elected with probability 1/4.*

• ***Type-II equilibria**: **three parties receive a strictly positive vote share**.*

A Type-II equilibrium always exists if $\eta \leq 1/2$*. Under this condition, only two parties get elected with strictly positive probability, and two parties remain losers.*

• ***Type-III equilibria**: **only two parties receive a strictly positive vote share (Duvergerian outcomes).***

Proof. See Appendix A.2.4. ■

From the properties of the two pivot probabilities and of Perfect Bayesian Equilibria, this result is rather intuitive. However, it hides some effects that deserve detailed attention. Put yourself in the mind of a voter, and consider the following reasoning:

> How should I vote if all players expect that voters with *my* type do not vote for losers? If parties do not expect us to vote for losers, they will consider my voting for a loser as "babbling" and not use my vote to update their beliefs. Hence, voting for losers has zero-value, and is thus a dominated action.

This reasoning explains why there always exist at least one equilibrium in which some type(s) do(es) not vote for losers (Type-II and Type-III equilibria). However, if types v are expected to vote for some loser o, then it is always in the interest of any type-v voter to vote for that loser with strictly positive probability. The intuition behind this result is linked to the results of Lemma 2:

> How should I vote if all players expect that voters with *my* type are voting for loser o? Since parties know we vote for o, they will adopt "my" preferred platform if his share is large enough. Moreover, the probability that my ballot is pivotal in determining their future platform is "high" if the expected vote share of o is small. Hence, if I expect this share to be small, the value of my ballot will be maximized if I vote for my loser.

Technically, if types v vote for o with very small probability, we have $\Pr(com_{vw}) / \max_{p,p'} \left[\Pr\left(piv_{pp'}\right)\right] \rightarrow \infty$, which implies that such a strategy cannot be an equilibrium. Put differently, the fact that parties interpret a vote for o as a signal in favor of some state ω_v triggers a voting strategy under which types v vote for o with strictly positive probability.

The above gives the general intuition regarding the trade-off that the voter is facing. Yet, this trade-off opens the way for different types of equilibria and pay-off structures. There actually exist two types of equilibria under voting-for-losers, and the occurrence of one or the other depends on the parameters r_L and r_{eL} and on platform positions. For some parameter values, there always exists a Type-I equilibrium such that the shares of γ and δ are always lower than those of α and β. By Property 1, a vote for γ or δ cannot be outcome-pivotal in this case, and the equilibrium strategy of the voters is thus independent of their platforms. For other parameter values instead, such a Type-I equilibrium does not exist. The equilibrium vote shares of γ and δ become "large," which implies that they become serious contenders in the election. In that case, the value of a vote for γ or δ also depends on their platform.

By means of numerical simulations, we observe that the former case typically arises for $\eta (\equiv 1 - 2(r_L + r_{eL}))$ small enough, whereas the latter case arises for η large enough. In Example 1 below, we set $\eta = 1/3$ to illustrate the former case, and in Example 2, we set $\eta = 0.6$ to illustrate the latter:

Example 1: two parties are losers. Let $r_{eL} = r_L = 1/6$; $x^1_\alpha = L$, $x^1_\beta = R$, $x^1_\gamma = eL$, and $x^1_\delta = eR$.

Type-I equilibria. Four parties get votes in such an equilibrium. For instance,[11] if types eL vote for γ and types eR vote for δ, the equilibrium with γ and δ being losers[12] is characterized by $\phi^* = \phi^*_{eL,\gamma} = \phi^*_{eR,\delta} = 0.44$. For $\phi < \phi^*$, the probability of being communication-pivotal (when voting for γ or δ) is infinitely larger than the probability of being outcome-pivotal (when voting for α or β). Thus, extremist voters never want to vote-for-losers with a probability below ϕ^*. For $1/2 > \phi > \phi^*$, the probability of being outcome-pivotal between α and β is infinitely larger than that of

[11]Under Type-I and Type-II equilibria, different signalling structures can emerge. For instance, types eL may vote for δ and types eR for γ. Similarly, there may exist equilibria in which moderate voters vote for losers to signal their moderation. This multiplicity is however of little interest, since each equilibrium amounts to a relabeling of the players' names.

[12]There also exists one equilibrium in which each type of voter votes for her preferred party with probability 1, and there is no loser.

being communication-pivotal. These values of ϕ thus cannot either be an equilibrium. Under the strategy ϕ^* instead, the two probabilities have the same magnitude, and extremists are indifferent between voting for their loser or their preferred mainstream party. In ϕ^*, equilibrium vote shares are:

Table 1: Equilibrium vote shares when extremist types vote for losers.

	State of nature			
	ω_{eL}	ω_L	ω_R	ω_{eR}
s_γ^ω	*0.22*	0.07	0.07	0.07
s_α^ω	**0.45**	**0.60**	**0.26**	**0.26**
s_β^ω	**0.26**	**0.26**	**0.60**	**0.45**
s_δ^ω	0.07	0.07	0.07	*0.22*

Type-II equilibria. Three parties get votes in such an equilibrium. Following the same intuition as above, there is an equilibrium in which types eL vote for γ with probability 0.33 and for α with probability $1 - 0.33 = 0.67$, whereas all other types only vote for α or β. Under this strategy, the magnitude of the two probabilities are equalized, and equilibrium vote shares become:

Table 2: Equilibrium vote shares when types eL vote for losers.

	State of nature			
	ω_{eL}	ω_L	ω_R	ω_{eR}
s_γ^ω	*0.16*	0.05	0.05	0.05
s_α^ω	**0.51**	**0.62**	**0.28**	**0.28**
s_β^ω	0**.33**	**0.33**	**0.67**	**0.67**
s_δ^ω	0	0	0	0

Type-III equilibria. If only α and β collect votes, α wins against β by a 67 to 33 percent margin in states ω_{eL} and ω_L, and loses by a 33 to 67 percent margin in the other states.

Example 2: all parties face a positive probability of election. We focus here on Type-I equilibria. Let $r_{eL} = r_L = 0.10$, and platforms are the same as in Example 1. With this change in parameter values, there is no Type-I equilibrium in which γ and δ are losers: consider the problem of an eL voter (by symmetry, the same reasoning holds for eR voters). For any $0 < \phi_{eL,\gamma} < 0.36$, a vote is infinitely more likely to be communication-pivotal than outcome-pivotal. In $\phi_{eL,\gamma} = 0.36$, the two probabilities have the same magnitude. In that point, however, γ is not a loser anymore and a vote for γ can be pivotal against α. Since eL voters' preferred platform is precisely that of γ, the value of a vote for γ is thus always higher than the value of a vote for α: voting for α is a dominated action for eL voters. Hence, in equilibrium, eL-voters will only vote for γ, L-voters for α, etc. In this case, α gets elected in state ω_L, β in state ω_R, γ in state ω_{eL} and δ in state ω_{eR}, each time with a $0.7 (= r_{eL} + \eta)$ vote share. Note however that such equilibria are not the primary focus of this paper, since there are no losers for such parameter values.[13]

Equilibrium Selection. Given the multiplicity of equilibria presented in Proposition 1, the model seems to have little predictive power. Multiplicity of equilibria is a typical outcome in signalling games, and this model clearly makes no exception. However, we can check that only Type-I equilibria are robust. In other words, even though the model cannot always predict which type of voters should vote for which loser (if any) in equilibrium, Proposition 2 allows us to predict that, in any robust equilibrium, all four parties should receive votes:

[13] There also exists a counterintuitive equilibrium with losers, in which eL-voters are expected to vote for δ, and eR-voters to vote for γ. In this case, $\phi_{eL,\delta} = \phi_{eR,\gamma} = 0.36$ is an equilibrium. The reason for this result is the following: under that strategy, eL voters are indifferent between being outcome-pivotal in favour of α, against δ (in which case, $x^1 = L$ instead of eR) and being communication-pivotal (in which case, $x^2 = eL$ instead of L). Instead, for $\phi_{eL,\delta} > 0.36$, the magnitude of being communication-pivotal is lower, and the value of a vote for δ is thus lower than that of a vote for α. Therefore, $\phi_{eL,\delta}$ (and by symmetry $\phi_{eR,\gamma}$)> 0.36 cannot be an equilibrium. The same reasoning applies if eL-voters vote for γ but $x^1_\gamma << eL$, and types eR vote for δ but $x^1_\delta >> eR$.

Proposition 2 *Type-II and Type-III equilibria are not robust to the introduction of an infinitesimal fraction of ideological voters.*

Proof. Consider that some (small) fraction of the electorate always votes for some party. For instance, a fraction $\varepsilon \to 0$ of types eL always votes for γ, of types L for α, of types R for β and of types eR for δ. The vote shares of γ and δ in state ω are thus:

$$s_\gamma^\omega = \varepsilon\, r_{eL}^\omega + \sum_{v\in V} \phi_{v,\gamma}\, \left(r_v\left(\omega\right) - \varepsilon\right) \quad \text{and} \quad s_\delta^\omega = \varepsilon\, r_{eR}^\omega + \sum_{v\in V} \phi_{v,\delta}\, \left(r_v\left(\omega\right) - \varepsilon\right).$$

Can the strategy $\phi_{v,\gamma} = \phi_{v,\delta} = 0$, $\forall v \in V$ be an equilibrium? By Lemma 2 and the pivot probability (5), we have:

$$\phi_{eL,\gamma} = 0 \Rightarrow \exists w \neq eL \text{ s.t. } \frac{\Pr\left(com_{eL,w}\right)}{\max_{p,p'}\left[\Pr\left(piv_{pp'}\right)\right]} = \infty,\ \forall p,p' \in P,\ p \neq p',$$

and similarly for types eR. Hence, $\phi_{eL,\gamma} = 0$ is a strictly dominated strategy. Clearly, this reasoning holds for any $\varepsilon < \phi_{eL,\gamma}^*$, where $\phi_{eL,\gamma}^*$ denotes the probability that types eL vote for γ in a Type-I equilibrium when ideological voters are absent. ■

Proposition 2, in a similar fashion to Piketty (2000), thus shows that Type-II and Type-III equilibria are "knife-edge" or "unstable." Such equilibria only exist because a vote-for-loser has no value if parties initially expect no vote-for-losers at all. By contrast, the presence of ideological voters ensures that the losers' vote shares can never be exactly zero, in which case the logic behind Type-II and Type-III equilibria ceases to hold. Hence, the losers' vote shares must always be bounded above zero in the presence of even an infinitesimal fraction of ideological voters.

This conclusion is antagonistic to the "standard" prediction that only Duvergerian outcomes can be "stable," while non-Duvergerian outcomes would be "unstable" in nature.[14] Proposition 2 instead demonstrates that this presumed stability

[14] In the words of Fey (1997), this means that, when the value of information is introduced in

hinges on the perceived role of the election. Piketty (2000) already showed that non-Duvergerian outcomes can be stable, and Propositions 1 and 2 extend his result to the case of endogenous platforms and more parties. As soon as parties' future (re)actions are taken into account, voting-for-losers has positive value: by voicing their ideological preferences, voters threaten main parties with loss of the election. If there are sufficiently many voters who convey such a message, it becomes a dominant strategy for mainstream parties to adapt their platforms in the second election.

Interestingly, this result is also robust to a change in the preferences of the electorate. Reducing the number of voter types down to three would imply that Type-I equilibria do not exist. Yet, applying the proofs for Propositions 1 and 2 to the three-type case, one can check that Types II and III equilibria would still coexist, and only Type-II equilibria would be robust.

Theoretical predictions and empirical regularities. Cox (1994, Theorem 1) predicts that, because of Duverger's Law, only Duvergerian outcomes should be observed in equilibrium (and Lemma 1 above reproduces this result). Under Duvergerian Outcomes, losers can only receive votes because of "noise" in voters' behaviour (voters may fail to coordinate on two parties, or only some voters vote strategically).

Instead, Proposition 2 predicts that Duvergerian Outcomes should *not* be observed in equilibrium, even when only two parties receive "elective ballots," and Proposition 1 rationalizes both the concentration of votes on two leaders and the remaining dispersion across losers.

It is thus interesting to confront our predictions with Cox's (1997) detailed evidence on voting behaviour in British elections between 1983 and 1992. To measure the political strength of the losers, he computes the ratio of the main loser's vote

the model, equilibria with voting-for-losers become "expectationally stable," whereas Duvergerian outcomes are not.

share to that of the second leader (the "SF-ratio").[15] Across all elections, this ratio is shown to be consistently and substantially larger than zero, which confirms that Duvergerian outcomes are quite uncommon in practice (in all elections covered by Cox, the SF-ratio fell below 0.1 only three times −see Cox (1997, Figure 4.1, page 87)). Furthermore, by splitting the sample into two subsets, Cox contrasts voting behaviour between those elections in which the two leaders are "close," with voting behaviour in elections where one party has a clear and strong lead.

In close elections, the probability of being outcome-pivotal is quite large. This, in turn, should reduce the amount of voting-for-losers. Such a prediction is corroborated by Cox's figures: when the expected lead is small, the SF-ratio is more likely to be lower than 0.3 than between 0.3 and 0.7. By any measure, this starkly contrasts with the "Duvergerian Prediction" that the SF-ratio should be very close to zero most of the time. Next, when one party has a clear-and-strong lead, the probability of being outcome-pivotal is drastically reduced. In this case, losers should obtain a larger vote share than when the election is close. With an SF-ratio that is most likely to lie between 0.3 and 0.7 in such elections, Cox's evidence again confirms our predictions.

Cox too observes that vote results are closer to Duverger's predictions when the threat of losing the elections is more evident (*i.e.* when the election is close). However, in the strict sense, his model cannot be reconciled with SF-ratios that are consistently and substantially different from 0 or 1. Instead, our results not only explain why this SF-ratio can −and should− be strictly between 0 and 1; they also explain why the distribution of SF-ratios is so strikingly different in close and non-close elections.

[15] Note that we define "losers" and "leaders" differently from Cox (1997). According to his definitions, the party ranked second is the "first loser," and the candidate ranked third, the "second loser". His SF–ratio is given by the "Second-to-First" losers' votes totals: $z^t_{p_3}/z^t_{p_2}$, where p_r is the r^{th}-ranked party.

Longer horizons. Lemma 1 and Proposition 1 show that voting-for-losers can occur in the first election, but not in the second. However, our results directly extend to longer horizons if we assume a Markovian process in which the state of nature at time $t+1$ remains the same as at time t with some probability larger than 1/2 but smaller than 1. In this case, learning that the median voter is in v at date t would imply that she is still most likely to be in v at date $t+1$. Thus, parties will have an incentive to locate close to v in $t+1$. Yet, with some probability (smaller than 1/2), the position of the median voter will have changed, which implies that some voters still have an incentive to vote for losers in $t+1$. In such a framework, equilibria with voting-for-losers thus exist in all elections except the last.

Welfare implications. While Proposition 1 characterizes equilibrium voting behaviour, it does not provide any valuation of the welfare costs entailed by vote dispersion. Note however that, by definition, losers get elected with probability close to zero. Hence, for any **given** set of platform positions, there is no welfare cost associated with vote dispersion in the first election. Instead, the following corollary, which is still an asymptotic property of the voting game, shows that it entails some benefits:

Corollary 1 *Under Type-I equilibria, there is perfect learning about the state of nature. As a result, by Lemma 1, leading parties will always locate at the* ***actual*** *position of the median voter in the second election.*

Holding platforms positions fixed, voting-for-losers thus generates additional information at no aggregate cost, since "extremist" candidates are not elected. However, such a result is only valid in the case of fixed platforms. Hence, we must also analyze the first-period locational strategy of the parties in order to fully assess those costs.

4 Positioning strategy in the first election

Platform selection is typically a complex problem when more than two parties gather votes.[16] Yet, there is a natural asymmetry between "mainstream" parties and losers (or "challengers") in our setup, which helps simplifying the analysis. To maintain coherence with the previous section, call α and β the two mainstream parties. Conversely, γ and δ are the challengers, and only collect votes from eL- and eR-voters respectively.

To give weight to the asymmetry between leaders and losers we focus on a subset of Type-I equilibria. Namely, if different Type-I equilibria exist, we assume that voters always select the one in which γ and δ receive the lowest possible number of votes. That is if, like in Example 1 above, there exists a Type-I equilibrium in which these two parties remain losers, that equilibrium is selected. This defines the leadership of α and β over the other two parties.

To further simplify the analysis, we make two (more restrictive) assumptions: first, α is constrained to be a "left-wing" party, and can only select a platform $x^1_\alpha \leq 0$. By symmetry, β is a "right-wing" party and is constrained to select a platform $x^1_\beta \geq 0$. Second, we restrict the space of policy preferences in the electorate to be such that $eL \leq 2L$ and $eR \leq 2R$. Taken together, these two assumptions ensure that, for any shape of the voters' utility function, a pure-strategy locational equilibrium exists.[17]

When such an equilibrium with losers exists (i.e. if η is small enough –see Section 3.2), the asymmetry between losers and leaders is such that only the two leaders compete with one another for election, whereas γ and δ face a probability of election

[16] For instance, Myerson and Weber (1993) show that if parties are presumed to have a lead only for some ranges of platforms, they will locate within this range. Palfrey (1984) shows that parties may maintain some distance between their platforms to prevent entry by a third party.

[17] If one of these two assumptions were violated, and if voters had a convex utility function, the equilibrium would be in mixed strategy for α and β.

close to zero. Yet, a fraction of the extremist voters also votes for losers. Hence, the vote shares of the losers influences election probabilities. The two leaders, α and β, thus face a trade-off between moderation and extremism: adopting a moderate position increases their attractiveness to moderate voters but decreases their appeal to extremist voters, and conversely. Being outcome-pivotal in the first election influences voters' payoffs only insofar as it affects the platform implemented in the first period. If the two leaders adopt similar platforms, being outcome-pivotal has little value, and extremist voters may then prefer to vote for losers. Conversely, if the two leaders adopt an extreme position, extremist voters become increasingly willing to support them. Moderate voters instead may prefer another party, if it proposes a moderate platform.

Altogether, Proposition 3 determines the equilibrium location of the four parties when this trade-off is present, that is when γ and δ are losers and gather votes from eL- and eR-voters respectively:

Proposition 3 *If γ and δ are losers, the first-period equilibrium locational strategy of the parties is such that: i) γ and δ select the locations $x^1_\gamma = eL$ and $x^1_\delta = eR$; ii) α and β always adopt different platforms (the Median Voter Theorem does not apply); and*

iii) $x^1_\alpha \in [eL, L)$ *and* $x^1_\beta \in (R, eR]$ *if* $u'(|eL - L|) < u'(|eR - L|)$, *or*
$x^1_\alpha \in [eL, 0)$ *and* $x^1_\beta \in (0, eR]$ *if* $u'(|eL - L|) \geq u'(|eR - L|)$.

Proof. See Appendix 3. ■

The intuition for these results is straightforward for the losers: by assumption, only extremist voters cast their ballots on losers. It is therefore in their interest to locate as close as possible to extremists to maximize their vote share.[18] For them,

[18] One could object that losers face an election probability of zero. However, this is only true in the limit, for $\lambda = \infty$. For a finite λ, this probability is *close* to zero, and increasing in their vote share.

locating in eL and eR is a dominant strategy.[19] In other words, if some type of voter v is expected to vote for some loser p, that party endogenously chooses to locate close to v in the first period. The "counterintuitive equilibrium" discussed in footnote 13 thus only holds in the subgame where platforms have already been chosen.

The intuition behind the locational strategy of the main two parties (α and β) is more intricate. Extremist voters value positively the signal carried by an extremist ballot. They only vote for a mainstream party to avoid losing the first election. Thus, the value of a vote-for-leaders mainly depends on the distance between their platforms. The more distant the platforms of the leaders (the larger is $|x^1_\alpha - x^1_\beta|$), the higher is the value of being outcome-pivotal in the first election. Conversely, if the two platforms are identical, the utility of electing either leader is virtually the same; being outcome-pivotal has almost no value.[20] Protest voting on extremist parties then becomes close to a cheap-talk game, and this induces excessive communication; γ and/or δ might win the election. Instead, maintaining enough distance between platforms ensures that, in the eyes of an extremist voter, being outcome-pivotal does have a "significant" value. This reduces the propensity to vote for losers, and increases the probability that α or β is elected.

This shows that main parties want to maintain "some" distance between their platforms. How much distance? This depends on the shape of the utility function. By moving towards eR, β increases his vote share, because eR voters have a lower propensity to vote for losers. However, the same effect holds for eL voters, whose propensity to vote for losers is also reduced. Therefore, the vote shares of both α

[19] Ideologically-minded extremist parties might of course select platforms that are even more extreme than eL or eR. If they are "too" extreme, their vote share will still be positive, but they would face a probability of election close to zero for any value of r_{eL} and r_L (See Proposition 1 and example 2, footnote 13).

[20] If the two parties adopted the same platform, and if the magnitude of the probabilities of being outcome- and communication-pivotal had the same magnitude, we would have $U(eL, \alpha) = -\Pr(com_{L,eL}) \times (u(0) - u(eL - L)) + \Pr(piv_{\alpha\beta}) \times \varepsilon < 0$

and β increase. If eR-voters are more sensitive to this move than types eL, then β's share increases more than α's, which increases the probability that β is elected. Conversely, if eL-voters react more strongly, β's probability of election is reduced.[21] Depending on the case, the two leaders thus maintain more or less distance between their platforms: if the voters' utility function is concave, both leaders adopt similar, centrist, platforms. If it is convex, α locates close to eL and β close to eR. In the more general case where the voters' utility function is nor concave nor convex everywhere, leaders should adopt platforms that are strictly inside that interval.[22]

Proposition 3 applies to the case in which γ and δ are losers. What happens if they are not? Simply, if γ and δ are "serious contenders" in the first election, α and β never locate close to 0. If they were doing so, L-voters would prefer to vote for γ rather than α and R-voters for δ instead of β. Thus, if the vote share of extremist parties becomes "significant," leading parties never adopt centrist positions. In other words, the fact that leaders may adopt centrist positions results from the initial assumption we made, *i.e.* that voters select the Type-I equilibrium in which γ and δ have the lowest possible vote share. If the leadership of α and β were less marked, they would be constrained to adopt non-centrist positions for any shape of the voters' utility function.

These interactions have never been analyzed in the literature. The Condorcet Jury literature, for instance, always considers parties as passive players of the electoral game. The results in this paper fill this gap: by analyzing the linkages between election results and parties' platforms, we show why voting for a "pure loser" may

[21] Note that, the larger is electorate size, the less platform positions influence the voters' equilibrium mixed strategy. However, the sign of $d\mathsf{E}\left(z_\alpha^1 - z_\beta^1\right)/dx_p^1$, $p \in \{\alpha, \beta\}$ is independent of electorate size.

[22] Utility functions with a second derivative that changes sign are seldom used, because of their lack of tractability. This does not imply they are unrealistic: if, say, an extreme-left person strongly values the differences between a Maoist, a Leninist and a Socialist party, whereas he or she views "all extreme right parties as equally bad," we have an indication that the utility function should be concave close to $|x - v| = 0$ and convex (to become almost flat) in $|x - v| = eL - eR$.

buy off an extremist platform, even though the party initially defending that platform is never elected. In addition, our results explain why, under population uncertainty, extremist voters weigh heavily on mainstream parties' policies, even though they might be a minority and extremist parties never get elected in equilibrium.

Welfare implications. We now have a better grasp of the welfare costs induced by the presence of losers. In the absence of voting-for-losers, mainstream parties always adopt moderate positions, which reduces the variance of the voters' payoffs in the first period. However, some information is lost, and parties may thus fail to locate at the position of the actual median voter in the second period.

Under voting-for-losers equilibria, benefits and costs are reversed: from Corollary 1, additional information is generated for the second period. In the first period, however, all parties may adopt extreme positions. If equilibrium platform locations are close to eL and eR, the platform implemented at the end of the first period is necessarily extreme, even though the median voter might be moderate. Voters must choose among extremist platforms just because of the *possibility* that the median voter is extreme. Voting-for-losers thus imposes a short-run cost (political extremism) and a long-run benefit (improved learning).

5 Pre-Play Communication and Abstention

One may wonder whether the equilibria with voting-for-losers vanish when other means of transmitting information to the parties, such as abstention, pre-election polls or mass-demonstrations, are available.

In itself, abstention would be a dominated move in any of the equilibria presented in Proposition 1, since it would prevent the voter from affecting election outcomes in her favor.[23] Still, abstention can be introduced in the model differently, by posing that either γ or δ represents abstention instead of a party. Put in this way, voting for a loser becomes formally equivalent to abstaining. Yet, this does not mean that voting-for-losers should disappear in equilibrium: if extreme-left voters signal their preference by abstaining (playing γ with some probability), then it is in the best interest of extreme-right voters to use another signal (namely: vote for δ), rather than abstain. We may thus conjecture that voting-for-losers has at least one advantage over abstention: it carries a clearer signal about the desired direction towards which voters want to move platforms.[24]

In contrast with abstention, mass demonstrations and opinion polls require an extension of the model. Mass demonstrations can be considered as a pre-play signalling device that affects the prior beliefs of the players regarding states of nature. In the presentation of the model, we assumed that each state was equally likely. Nonetheless, all the proofs for the results of Section 3 were carried out assuming some set of (strictly positive) prior probabilities. This shows that, as long as the probability of all states of nature are not too different from one another, equilibria with voting-for-losers still exist. Still, manipulating priors may affect which equilibrium survives, and which platforms parties adopt in the first election. (In a different setup, we analyzed how changes in prior beliefs affect those platforms, see Castanheira, 1998). Therefore, mass demonstrations may be used as a tool to influence the outcome of the

[23] Feddersen and Pesendorfer's (1996) "Swing voter's curse" motivates abstention in a different way: if only a fraction of the voters are aware of the relative merits of the main candidates, then uninformed and indifferent voters rationally decide to abstain, to let informed voters alone select the best candidate. In our setup, however, this rationale for abstention is absent, since all voters have exact knowledge of their preferred platform.

[24] Shotts (2000) analyzes abstention in a setup that is close to the one in this paper. In his model, abstention proves to be a signal that moderate voters want to send to candidates, in order to moderate second-period platforms. The same result would hold in our model if we consider three types of voters, and an equilibrium in which moderates want to communicate, either by voting for a moderate loser or, equivalently, by voting "abstain".

election, *in addition to* voting-for-losers. Put differently, these two signalling devices are complementary, rather than substitutes.[25]

Opinion polls play yet a different role in the model. Polls reveal the vote shares of the different parties *during* the election campaign. That is, they facilitate voters' coordination between the platform selection stage and the voting stage of the game. Observing a sequence of polls thus reveals how voters adapt their strategy as additional information becomes available. If polls were perfectly accurate; *i.e.*, if they could perfectly reveal the actual state of nature, then voting-for-losers would become pointless at the voting stage. However, polls are known to generate noisy signals, which means that voting-for-losers should still arise in equilibrium (again, this hinges upon a strictly positive prior probability of being in each state of nature).

6 Conclusions

This paper develops a model of repeated elections when the distribution of preferences in the electorate is uncertain. We show that, in this context, voters can elicit additional information about their preferences by voting for small parties. By taking such action, they influence the platforms of mainstream parties and attract them towards "extreme" platforms, which would never be implemented in the absence of "votes for losers." Central to the analysis is the interplay between the actions of the voters and the reactions of the parties. If parties were passive actors in the election game, *i.e.* if their platforms were exogenous, voters would not have any incentive to vote for small parties. Conversely, if parties select the platform that maximizes their probability of being elected, they must exploit all the information available, and the vote share of those "losers" becomes relevant in determining their political stance.

[25] Lohmann (1993, 1994, and 2000) analyzes the problems of collective action through such "alternative" communication methods in much greater detail.

This shows that elections have the power of eliciting accurate information about the complete distribution of preferences in the electorate. In turn, information elicitation reinforces the dynamic efficiency of the electoral system, allowing for quicker and more substantial adaptation of parties' platforms.

However, voting-for-losers also generates welfare costs. When losers receive ballots, even purely opportunistic parties may need to adopt an "extremist" stance. Through this link between voting behaviour and parties' platforms, extremist voters impose a cost upon moderate voters, even when the majority of the electorate is actually moderate. Interestingly, this means that there is a causality link between the preferences of the electorate and the apparent preferences of the parties. This contrasts with Wittman (1983), Calvert (1985) or Alesina (1988) who *hypothesize* parties with extremist preferences. Following their premises, one could have legitimately wondered why "natural selection" always makes extreme candidates emerge as "natural leaders;" why couldn't moderate politicians play this role? Our results instead justify the Calvert-Wittman assumption: in equilibrium, even purely opportunistic candidates have an incentive to mimic the behaviour of an extremist candidate, since sufficiently extreme platforms increase the parties' probability of being elected.[26]

Even though some of the results are specific to the model, the insights it provides are more general. In particular, the model should be extended to different types of elections, such as the proportional system, run-off elections, and approval voting. Extending the analysis in that direction is however not as straightforward as it seems: in run-off elections, for instance, the first round can be used to signal preferences. However, only the *second* round determines which party is elected. The costs of losing the first round are thus less clear than in a first-past-the-post setting.[27] Yet,

[26] Another interesting way to explain parties' ideological bias is provided by Rivière (2000), who argues that parties are endogenously created by a non-median fringe of the electorate.

[27] Note that Myerson (1999) compares the properties of the different systems. However, he ab-

generalizing the model in that direction would also be desirable from the point of view of testing predictions empirically. We showed that Cox's (1997) evidence on first-past-the-post systems tends to confirm its predictions, but the analysis will not be complete until we can also compare different systems.

References

ALESINA, Alberto (1988). "Credibility and Policy Convergence in a Two-Party System with Rational Voters." *American Economic Review*, Vol. **78**, 796-805.

ALESINA, Alberto and Howard ROSENTHAL (1995). *Partisan Politics, Divided Government, and the Economy.* Cambridge: Cambridge University Press.

BLACK, Duncan (1958). *The Theory of Committees and Elections.* Cambridge: Cambridge University Press.

CAILLAUD, Bernard and Jean TIROLE. (1997). "Voters homogeneity and aggregation of information." Mimeo. Toulouse: IDEI.

CALVERT, Randall (1985). "Robustness of the Multidimensional Voting Model: Candidates's Motivations, Uncertainty and Convergence." *American Journal of Political Science*, **29**, 69-95.

CASTANHEIRA, Micael (1998). "Why Vote for Losers?" IGIER Working Paper, 125.

CONDORCET, Marie Jean, Marquis de (1785). "Essai sur l'application de l'analyse à la probabilité des décisions rendues à la pluralité des voix." Paris

COX, Gary (1994). "Strategic Voting Equilibria Under the Single Non-Transferable

stracts from information aggregation considerations. Piketty (2000) discusses how his model could be interpreted as a run-off election. In his model, however, pay-offs directly depend on the first election results. That is, *stricto sensu*, his model also focuses on first-past-the-post elections.

Vote." *American Political Science Review*, **88**, 608-621.

COX, Gary (1997). *Making Votes Count.* Cambridge: Cambridge University Press.

DOWNS, Anthony (1957). *An Economic theory of Democracy.* New York: Harper and Row.

DUVERGER, Maurice (1954). *Political Parties.* New York: Wiley.

FEDDERSEN, Timothy and Wolfang PESENDORFER (1996). "The Swing Voter's Curse." *American Economic Review*, **86**, 408-424.

FEDDERSEN, Timothy and Wolfang PESENDORFER (1997). "Voting Behavior and Information Aggregation in Elections with Private Information." *Econometrica*, **65**, 1029-58.

FEY, Mark (1997). "Stability and Coordination in Duverger's Law: A Formal Model of Preelection Polls and Strategic Voting." *American Political Science Review*, **91**, 135-147.

HOTELLING, Harold (1929). "Stability in Competition." *Economic Journal*, **39**, 41-57.

LOHMANN, Susanne (1993). "A Signaling Model of Informative and Manipulative Political Action." *American Political Science Review*, **87**, 319-333.

LOHMANN, Susanne (1994). "Information Aggregation Through Costly Political Action." *American Economic Review*, **84**, 518-530.

LOHMANN, Susanne (2000). "Collective Action Cascades: an Informational Rationale for the Power in Numbers." *Journal of Economic Surveys*, **14**, 655-684.

MYERSON, Roger (1997). "Population Uncertainty and Poisson Games." Northwestern University Discussion Paper. 1102R.

MYERSON, Roger (1998). "Extended Poisson Games and the Condorcet Jury Theorem." *Games and Economic Behavior*, **25**, 111-131.

MYERSON, Roger (1999). "Theoretical Comparisons of Electoral Systems." *European Economic Review*, **43**, 671-697.

MYERSON, Roger (2000). "Large Poisson Games." *Journal of Economic Theory*, **94**, 7-45.

MYERSON, Roger and Robert WEBER (1993). "A Theory of Voting Equilibria." *American Political Science Review*, **87**, 102-114.

PALFREY, Thomas (1984). "Spatial Equilibrium with Entry." *Review of Economic Studies*, **51**, 139-156.

PALFREY, Thomas (1989). "A Mathematical Proof of Duverger's Law." In *Models of Strategic Choice in Politics,* edited by P. Ordeshook. Ann Arbor: University of Michigan Press.

PIKETTY, Thomas (1999). "The information-aggregation approach to political institutions." *European Economic Review*, **43**, 791-800.

PIKETTY, Thomas (2000). "Voting as Communicating." *Review of Economic Studies*, **67**, 169-191.

RAZIN, Ronny (2000). "Signaling and Electing Motivations in a Voting Model with Common Values and Responsive Candidates." *Econometrica*, forthcoming.

RIVIÈRE, Anouk (2000). "Citizen Candidacy, Party Formation and Duverger's Law." Mimeo. Brussels: Université Libre de Bruxelles.

ROEMER, John (1997). "Political-Economic Equilibrium when Parties Represent Constituents: The Unidimensional Case." *Social Choice and Welfare*, **14**, 479-502.

SHOTTS, Ken (2000). "A Signalling Model of Repeated Elections." Mimeo. Evanston: Northwestern University

WITTMAN, Donald (1983). "Candidate Motivation: a Synthesis of Alternatives." *American Political Science Review*, **77**, 142-157.

Appendix

Appendix 1 summarizes some properties of Poisson Games already proven by Myerson (1997, 2000) or that are straightforward extensions from his work. Appendices 2 and 3 demonstrate the claims made in Sections 3 and 4 respectively.

Appendix 1: Some Properties of Poisson Games

By the definition of Poisson distributions,

$$\text{if } \tilde{\zeta} \sim \mathcal{P}(\lambda), \text{ then } \Pr\left(\tilde{\zeta} = k\right) = \frac{e^{-\lambda}\,\lambda^k}{k!},\ \lambda,\, k \in \mathcal{N} \tag{7}$$

Among other things, Myerson proved that, for any value of λ: 1) In the state of nature ω, the number of voters with type v is distributed according to a Poisson Law $\mathcal{P}(r_v(\omega)\,\lambda)$; 2) By the *independent actions property*, the number of votes for a party, z_p, is also distributed according to a Poisson distribution: $\tilde{z}_p^\omega \sim \mathcal{P}\left(\lambda_p^\omega\right)$, where $\lambda_p^\omega = \lambda \cdot \sum_{v \in V} \phi_{v,p}\, r_v(\omega)$ and $\tilde{z}_p^\omega$ is independent of $\tilde{z}_q^\omega$, $\forall p \neq q$; 3) There always exists at least one equilibrium in such a voting game; 4) The property of *environmental equivalence* holds, *i.e.* a voter can compute the value of her vote by taking into account population distribution alone.

From those properties, it follows that:

Property A1 (Myerson, 2000, Theorem 1) *The probability that* ***realized*** *vote shares are* $\{\sigma(p)\}$, *subject to* $\sum_p \sigma(p) = 1$, *when* ***expected*** *vote shares are* $\left\{s_p^\omega\right\}$, *with* $\sum_p s_p^\omega = 1$, *is:*

$$\Pr\left(\sigma(p)\,|s_p^\omega\right) \underset{\lambda\to\infty}{\longrightarrow} \max_{\sigma(p)} \prod_{p \in P} \exp\left[\chi\left(s_p^\omega,\, \sigma(p)\right)\,\lambda\right] / \sqrt{2\pi\,\lambda\,\sigma(p) + \pi/3}, \tag{8}$$

$$\text{with: } \chi\left(s_p^\omega,\, \sigma(p)\right) = \sigma(p)\left[1 - \log\left(\sigma(p)/s_p^\omega\right)\right] - s_p^\omega\ (\leq 0)$$

Property A2 (Myerson (2000, Theorem 2) and Feddersen and Pesendorfer (1996)) *The probability that two parties receive a number of votes that differs by a constant c*

$(c << \lambda)$ is:

$$\lim_{\lambda\to\infty} \Pr\left(\tilde{z}_p = \tilde{z}_q + c|\omega, s_p^\omega, s_q^\omega\right) = \left(s_p^\omega/s_q^\omega\right)^{\frac{c}{2}} \frac{\exp\left[-\left(\sqrt{s_p^\omega}-\sqrt{s_q^\omega}\right)^2\lambda\right]}{2\sqrt{\pi\lambda}\left(s_p^\omega\, s_q^\omega\right)^{1/4}},\ p,q \in P,\ \omega \in \Omega. \tag{9}$$

Property A3 below is a direct corollary of Property A2:

Property A3 *The probability that the realized ordering of parties is consistent with initial beliefs converges to one as the expected population size increases:*

$$\lim_{\lambda\to\infty} \Pr(\tilde{z}_q \geq \tilde{z}_p|\omega, s_q^\omega > s_p^\omega) = 1,\ p,q \in P,\ \omega \in \Omega.$$

Proof. From (9), for $\lambda \to \infty$:

$$\frac{\Pr\left(\tilde{z}_p \geq \tilde{z}_q|\omega\right)}{\Pr\left(\tilde{z}_p \leq \tilde{z}_q|\omega\right)} = \frac{\sum_{c=0}^{\infty}\left(s_p^\omega/s_q^\omega\right)^{\frac{c}{2}}}{\sum_{c=0}^{\infty}\left(s_p^\omega/s_q^\omega\right)^{-\frac{c}{2}}} = \begin{cases} 0 \text{ iff } s_q^\omega > s_p^\omega \\ \infty \text{ iff } s_q^\omega < s_p^\omega \end{cases}$$

■

Property 1 in Section 3 follows immediately from Properties A2 and A3. For instance, $\Pr\left(piv_{\alpha\beta}\right) = \Pr\left(\tilde{z}_\beta = \tilde{z}_\alpha + 1\right) \cdot \Pr\left(\tilde{z}_\beta \geq \max\left[\tilde{z}_\gamma, \tilde{z}_\delta\right]\right)$.

Appendix 2: Proofs for Section 3

A.2.1. **Proof of Lemma 1.** In the eyes of a voter v, the value of a ballot for party p is:

$$U(v,p) = \sum_{\tilde{p}\in P\backslash p} \Pr(piv_{p\tilde{p}}) \cdot \left[u\left(\left|x_p^2 - v\right|\right) - u\left(\left|x_{\tilde{p}}^2 - v\right|\right)\right], \tag{10}$$

in which $piv_{p\tilde{p}}$ denotes the event that the voter is pivotal in electing p over $\tilde{p}$.

From Property 1, it is immediate that, for $\lambda \to \infty$:

$$\max_{\bar{p}\in\{\alpha,\beta\}} U(v,\bar{p}) > \max_{p\in\{\gamma,\delta\}} U(v,p),\ \ \forall v \in V;\ \forall\left|x_p^2\right|,\ \left|x_{\bar{p}}^2\right| < \infty,$$

and hence that voting for γ or δ is a dominated action.

It now remains to show that α and β always locate close to the median voter. Denote by $\Pr\left(\mu^\omega = v\right) = \Pr\left(\omega|\mathsf{I}_p\right)$ the parties' perceived probability that the median voter has type v, given all the information available to them, I_p. Denote also the *perceived position of the median voter* by μ^*, defined as:

$$\mu^* \equiv x \; s.t. \; \Pr(\mu^\omega < x) < 1/2 \text{ and } \Pr(\mu^\omega \leq x) \geq 1/2.$$

From the first part of the Lemma, we know that voters only vote for α and β. Hence, we focus on the locational strategy of α and β alone. Given any platform x_β^2, we have:

$$\Pr\left(W^2 = \alpha | \; |x_\alpha^2 - \mu^*| > |x_\beta^2 - \mu^*|\right) < \Pr\left(W^2 = \alpha | \; |x_\alpha^2 - \mu^*| \leq |x_\beta^2 - \mu^*|\right).$$

That is, for any x_β^2, α can increase his probability of winning by locating closer to μ^* than β. Similarly, β increases his probability of winning by locating closer to μ^* than α. Hence, $x_\alpha^2 = x_\beta^2 = \mu^*$ is the unique locational equilibrium for α and β. ■

A.2.2. Lemma 3 and proof

Lemma 3 *For given first-election results* $\mathsf{z}^1 = \{z_\alpha^1, z_\beta^1, z_\gamma^1, z_\delta^1\}$, *the ratio of posterior beliefs about two given states of nature,* ω *and* $\tilde{\omega}$, *is given by:*

$$\frac{\Pr\left(\omega|\mathsf{z}^1\right)}{\Pr\left(\tilde{\omega}|\mathsf{z}^1\right)} = \frac{\Pr\left(\omega\right)}{\Pr\left(\tilde{\omega}\right)} \prod_{p \in P} \left(\frac{s_p^\omega}{s_p^{\tilde{\omega}}}\right)^{z_p^1}. \tag{11}$$

Proof. From (7), the probability of observing the vector of vote results z^1 in the state of nature ω is given by:

$$\Pr\left(\mathsf{z}^1|\omega\right) = \frac{(s_\alpha^\omega \lambda)^{z_\alpha^1} \cdot \left(s_\beta^\omega \lambda\right)^{z_\beta^1} \cdot \left(s_\gamma^\omega \lambda\right)^{z_\gamma^1} \cdot \left(s_\delta^\omega \lambda\right)^{z_\delta^1}}{z_\alpha^1! \; z_\beta^1! \; z_\gamma^1! \; z_\delta^1! \; e^\lambda}.$$

Then, using Bayes' rule yields (11). ■

A.2.3. Proof of Lemma 2.

For the sake of concreteness, consider one particular case (the reasoning clearly extends to any voters and parties): assume that $\phi_{eL,\alpha} + \phi_{eL,\gamma} = 1$, with $\phi_{eL,\alpha}$, $\phi_{eL,\gamma} > 0$; that $\phi_{L,\alpha} = 1$; and that the other voters only vote for β and/or δ. Under this strategy, we have $s_\gamma^{\omega_{eL}} > s_\gamma^{\omega_L}$; $s_\alpha^{\omega_{eL}} < s_\alpha^{\omega_L}$; $s_\beta^{\omega_{eL}} = s_\beta^{\omega_L}$; and $s_\delta^{\omega_{eL}} = s_\delta^{\omega_L}$. Hence, by Lemma 3, the ratio $\Pr\left(\omega_{eL}|\mathsf{z}^1\right)/\Pr\left(\omega_L|\mathsf{z}^1\right)$ is strictly increasing in z_γ^1, and is independent of z_β^1 and z_δ^1.

By Property A1, this ratio is equal to one in:

$$\frac{z_\gamma^1}{z_\alpha^1} = \frac{\sigma(\gamma)}{\sigma(\alpha)} = K = \frac{\log\left(s_\alpha^{\omega_L}/s_\alpha^{\omega_{eL}}\right)}{\log\left(s_\gamma^{\omega_{eL}}/s_\gamma^{\omega_L}\right)}. \tag{12}$$

Then, by Property A1, the probability that $\sigma(\gamma)/\sigma(\alpha) \simeq K$ is given by:

$$\begin{aligned}\lim_{\lambda\to\infty}\Pr(com_{eL,L}) &= \lim_{\lambda\to\infty}\Pr\left(\{\sigma(p)\}\,|\,\{s_p^\omega\}\right)\\ &= \max_{\sigma(p)}\prod_{p\in P}\exp\left[\lambda\times\chi\left(s_p^\omega,\sigma(p)\right)\right]/\sqrt{2\pi\,\lambda\,\sigma(p)+\pi/3},\\ &\qquad\text{subject to (12) and } \textstyle\sum_p\sigma(p)=1.\end{aligned}$$

By Lagrangean analysis, this yields the implicit relationships:

$$\begin{aligned}\sigma(\gamma) &= \left(s_\gamma^{\omega_{eL}}\right)^{c_1}\left(s_\gamma^{\omega_L}\right)^{1-c_1}c_2,\\ \sigma(\alpha) &= \left(s_\alpha^{\omega_{eL}}\right)^{c_1}\left(s_\alpha^{\omega_L}\right)^{1-c_1}c_2,\\ \sigma(\beta) &= s_\beta^\omega c_2;\;\; \sigma(\delta) = s_\delta^\omega c_2,\ \omega\in\{\omega_{eL},\omega_L\},\end{aligned}$$

where $c_1 \in (0,1)$ is the value that solves $\sigma(\gamma)/\sigma(\alpha) = K$ and c_2 is the value that solves $\sum_p \sigma(p) = 1$. Solving for these values then obtains:

$$\begin{cases}\sigma(\alpha) = [1+K+K^{K/(1+K)}\dfrac{r_{eL}+r_L}{\left(s_\alpha^\omega\left(s_\gamma^\omega\right)^K\right)^{1/(1+K)}}]^{-1}\\ \sigma(\gamma) = K\,\sigma(\alpha)\\ \dfrac{\sigma(\beta)}{s_\beta^\omega} = \dfrac{\sigma(\delta)}{s_\delta^\omega} = K^{K/(1+K)}\left(s_\alpha^\omega\left(s_\gamma^\omega\right)^K\right)^{-1/(1+K)}\sigma(\alpha).\end{cases} \tag{13}$$

Now, define $\chi\left(\phi_{eL,\gamma}\right) \equiv \sum_p \chi\left(s_p^\omega,\sigma(p)\,|\phi_{eL,\gamma}\right)$. Then, since $\partial\chi\left(s_p^\omega,\sigma(p)\right)/\partial\left(s_p^\omega-\sigma(p)\right)\times\left(s_p^\omega-\sigma(p)\right) < 0$, $\forall\left(s_p^\omega-\sigma(p)\right)\neq 0$, and since $\partial\left|s_p^\omega-\sigma(p)\right|/\partial\phi_{eL,\gamma} > 0$, for $p\in\{\alpha,\gamma\}$, we have $\partial\chi\left(\phi_{eL,\gamma}\right)/\partial\phi_{eL,\gamma} < 0$, $\forall\phi_{eL,\gamma} > 0$.

Arguably, a vote for γ could also be communication-pivotal against other states of nature. However, one can check that $\Pr(com_{eL,L})/\Pr(com_{eL,v}) \to \infty$, $\forall\phi, v \neq L$.

On the other hand, one can verify that, for $r_{eL} \leq r_L$, there exists $\bar{\phi}$ close or equal to one such that $\Pr\left(com_{eL,L}|\bar{\phi}\right) / \Pr\left(com_{L,R}|\bar{\phi}\right) = 1$. ■

A.2.4. Proof of Proposition 1.

Given $\mathsf{E}U_v$, a voter with type v values of a ballot for party p in the first election as:

$$U(v,p) = \sum_{\omega\in\Omega} b(\omega|v) \left[\sum_{p'\in P\backslash p} \Pr\left(piv_{pp'}|\omega\right) \left(u\left(x_p^1 - v\right) - u\left(x_{p'}^1 - v\right)\right) + ... \right.$$
$$\left. ... + \sum_{v',w\in V} \Pr\left(com_{v',w}|p,\omega\right) \left(u\left(v'-v\right) - u\left(w-v\right)\right)\right]. \tag{14}$$

Existence of Type-III equilibria. What is the best action a voter with type v can take if only two parties (say α and β) receive votes?

By Lemmas 2 and 3, if $s_p^\omega = 0$, $\forall p$, a vote for p cannot be communication-pivotal. Moreover, by Property 1, a vote for γ or δ is also infinitely less likely to be outcome-pivotal than a vote for α or β, which implies $\max_{p\in\{\gamma,\delta\}}\left[U(v,p)\right] < \max_{\tilde{p}\in\{\alpha,\beta\}}\left[U(v,\tilde{p})\right]$, and hence that Type-III equilibria must always exist for $x_\alpha^1 \simeq -x_\beta^1$.

Existence of Type-II equilibria. First, we show that, if it is common knowledge that $\phi_{v,p} > 0$ for some type v, and $\phi_{v',p} = 0$, $\forall v' \neq v$, then $\phi_{v,p} \to 0$ cannot be part of the equilibrium. To verify this, consider the following case: $\phi_{eL,\gamma} + \phi_{eL,\alpha} = 1$, $\phi_{eL,\gamma}, \phi_{eL,\alpha} > 0$; $\phi_{L,\alpha} = 1$; $\phi_{R,\beta} = \phi_{eR,\beta} = 1$. By Lemma 2, a vote for γ can thus be communication-pivotal between eL and L, and a vote for α between L and eL and, By Property 1, for $\phi_{eL,\gamma}$ close to zero, we have:

$$U(v,\gamma) \simeq \Pr\left(com_{eL,L}\right)\left(u\left(eL-v\right) - u\left(L-v\right)\right) \tag{15}$$

$$U(v,\alpha) \simeq \Pr\left(piv_{\alpha\beta}\right)\left(u\left(x_\alpha^1 - v\right) - u\left(x_\beta^1 - v\right)\right) + \Pr\left(com_{L,eL}\right)\left(u\left(L-v\right) - u\left(eL-v\right)\right). \tag{16}$$

Lemma 2 shows that $\Pr\left(com_{eL,L}\right) / \Pr\left(piv_{\alpha\beta}\right) = \infty$ for $\phi_{eL,\gamma} \to 0$, and hence that $U(eL,\gamma) > 0 > U(eL,\alpha)$. This shows by contradiction that $\phi_{eL,\gamma}$ close to zero cannot be part of a Type-II equilibrium.

Now, we show that a Type-II equilibrium with $1 > \phi^*_{eL,\gamma} = 1 - \phi^*_{eL,\alpha} > 0$ and $\Pr\left(W^1 = \gamma\right) \to 0$ always exists for $\eta \equiv 1 - 2\left(r_{eL} + r_L\right) \leq 1/3$.

The vote shares of α, β and γ in each state of nature are respectively:

$$s^\omega_\alpha = r^\omega_L + \phi_{eL,\alpha}\, r^\omega_{eL};\ s^\omega_\beta = r^\omega_R + r^\omega_{eR};\ s^\omega_\gamma = \phi_{eL,\gamma}\, r^\omega_{eL}.$$

It follows that, under the strategy $\bar{\phi}_{eL,\alpha} = r_{eL}/\left(r_{eL} + \eta\right)$, we have: $s^{\omega_{eL}}_\alpha = s^{\omega_{eL}}_\beta \geq s^{\omega_{eL}}_\gamma$. That is, $\bar{\phi}_{eL,\alpha}$ implies that $\Pr\left(com_{eL,L}\right)/\Pr\left(piv_{\alpha\beta}\right) = 0$, and therefore $U\left(eL,\gamma\right) < U\left(eL,\alpha\right)$ in $\bar{\phi}_{eL,\alpha}$, which can thus not either be part of an equilibrium. By continuity, there must exist an "intermediate" strategy Φ^*, with $\phi^*_{eL,\alpha} \in \left(\bar{\phi}_{eL,\alpha},\, 1\right)$ for which:

$$\chi\left(\phi^*_{eL,\gamma}\right) \simeq \max_\omega \left[-\left(\sqrt{s^\omega_\alpha\left(\phi^*_{eL,\alpha}\right)} - \sqrt{s^\omega_\beta}\right)^2\right],$$

and that ensures $U\left(eL,\gamma\right) = U\left(eL,\alpha\right) > U\left(eL,\beta\right)$. Is the strategy Φ^* an equilibrium for all voters? It is straightforward to check that types L do not want to deviate from $\phi_{L,\alpha} = 1$. However, right-wing voters face a more complex trade-off. By voting for β, they increase the probability that $W^1 = \beta$. By voting for α, they could moderate second-period platforms. However, since $\Pr\left(com_{eL,L}\right) \simeq \Pr\left(com_{L,eL}\right)$ and $\Pr\left(piv_{\alpha\beta}\right) \simeq \Pr\left(piv_{\beta\alpha}\right)$, and since eL voters adopt a mixed strategy (which implies that $U\left(eL,\gamma\right) = U\left(eL,\alpha\right)$), we find that $U\left(v,\alpha\right) < U\left(v,\beta\right)$, for $v \in \{R, eR\}$ if the utility function $u\left(\cdot\right)$ is not too concave. In particular, $u\left(0\right) - u\left(eL - L\right) > \left(u\left(R - eL\right) - u\left(eR - eL\right)\right)/4$ is a sufficient condition to ensure that Φ^* is an equilibrium, and the utility function $u\left(x\right) = -\left|x\right|$ always satisfies this condition.

Finally, since $\eta \leq 1/3$ is sufficient to ensure the existence of such an equilibrium, there must exist a threshold $\bar{\eta} > 1/3$, such that, $\forall \eta < \bar{\eta}$, $s^{\omega_{eL}}_\gamma$ remains strictly lower than $s^{\omega_{eL}}_\beta$ in Φ^*, which warranties the existence of such an equilibrium. By means of numerical simulations, one can check that $\bar{\eta} \simeq 1/2$. For instance, in $\eta = 0.4$, and $r_{eL} = r_L = 0.15$, we have: $\phi^*_{eL,\gamma} = 0.35$, $s^{\omega_{eL}}_\gamma = 0.19$ and $s^{\omega_{eL}}_\beta = 0.3$.

Type-I equilibria. The procedure to demonstrate the existence of Type-I equilibria is similar: set $\phi_{eL,\gamma} > 0$, $\phi_{v,\gamma} = 0$, $\forall v \neq eL$ and $\phi_{eR,\delta} > 0$, $\phi_{v,\delta} = 0$, $\forall v \neq eR$. In this case, following the same reasoning as for Type-II equilibria, neither $\phi_{eL,\gamma} \to 0$ nor $\phi_{eR,\delta} \to 0$ can be part of the equilibrium. Thus, both γ and δ must receive a strictly positive vote share. Still, the exact shape of this Type-I equilibrium depends on parameter values:

Case 1. γ and δ are losers. If there exists a strategy profile Φ^*, such that:

$$\phi^*_{eL,\gamma} + \phi^*_{eL,\alpha} = 1, \;\; \phi^*_{L,\alpha} = 1, \;\; \phi^*_{R,\beta} = 1, \;\; \phi^*_{eR,\beta} + \phi^*_{eR,\delta} = 1 \tag{17}$$

$$\chi\left(\phi^*_{eL,\gamma}\right) \simeq \max_{\omega}\left[-\left(\sqrt{s^{\omega}_{\alpha}\left(\phi^*_{eL,\alpha}\right)} - \sqrt{s^{\omega}_{\beta}\left(\phi^*_{eR,\beta}\right)}\right)^2\right] \tag{18}$$

$$\chi\left(\phi^*_{eR,\delta}\right) \simeq \max_{\omega}\left[-\left(\sqrt{s^{\omega}_{\alpha}\left(\phi^*_{eL,\alpha}\right)} - \sqrt{s^{\omega}_{\beta}\left(\phi^*_{eR,\beta}\right)}\right)^2\right] \tag{19}$$

$$\max\left[s^{\omega}_{\gamma}\left(\phi^*_{eL,\gamma}\right), s^{\omega}_{\delta}\left(\phi^*_{eL,\delta}\right)\right] < \min\left[s^{\omega}_{\alpha}\left(\phi^*_{eL,\alpha}\right), s^{\omega}_{\beta}\left(\phi^*_{eL,\beta}\right)\right], \; \forall \omega \in \Omega, \tag{20}$$

a vote for γ or δ is infinitely less likely to be outcome-pivotal than communication-pivotal. Hence, by Property 1, the platforms of these parties do not matter in determining voters' payoffs (like in (15) above). Moreover, for $u(0) - u(eL - L) > (u(R - eL) - u(eR - eL))/3$, no voter wants to deviate from Φ^*, which is thus an equilibrium. (Clearly other Type-I equilibria exist since one can re-label the name of parties and voters.)

Case 2. γ and δ are not losers. If (20) is always violated when (17)-(19) are fulfilled, then a vote for α (resp. β) becomes more likely to be pivotal against γ than against β (resp. δ and α) in Φ^*. Thus, the platforms of γ and δ also matter in determining voters' payoffs.

Call γ the party closest to eL, α the party closest to L, β the party closest to R and δ the party closest to eR. In this case, if:

$$\chi\left(\phi_{eL,\gamma} = 1\right) < -\left(\sqrt{r_{eL} + \eta} - \sqrt{r_L}\right)^2, \tag{21}$$

the strategy profile Φ^{**}, such that $\phi^{**}_{eL,\gamma} = \phi^{**}_{L,\alpha} = \phi^{**}_{R,\beta} = \phi^{**}_{eR,\delta} = 1$ is an equilibrium: if (21) holds, the value of a ballot only depends on outcome-pivotability and, for $u\left(x^1_\gamma - eL\right) - u\left(x^1_\alpha - eL\right) \geq \left(u\left(x^1_\beta - eL\right) - u\left(x^1_\delta - eL\right)\right)/3$, one can verify that no voter can increase her expected utility by deviating from Φ^{**}. With symmetric platforms, the utility function $u(x) = -|x|$ ensures that this condition is satisfied. Moreover, if eL (resp. eR) voters are expected to vote for δ (resp. γ), they would always adopt a strategy that prevents them from being elected, since $u\left(x^1_\delta - eL\right) - u\left(x^1_\alpha - eL\right) < 0$.

If (21) is violated, the value of a ballot only depends on communication-pivotability. However, by Lemma 2, there exists a value $\bar{\phi}$ close or equal to one, for which $\lim_{\lambda\to\infty} \Pr\left(com_{eL,L}\right) / \Pr\left(com_{L,R}\right) = 1$ and, following the same methodology as above, one can easily verify that $\phi^{**}_{eL,\gamma} = \phi^{**}_{eR,\delta} \simeq \bar{\phi}$ is an equilibrium. ∎

Appendix 3: Proof of Proposition 3

As a prerequisite, note that the extremist voters only adopt a mixed strategy if:

$$U(eL,\gamma) = U(eL,\alpha) \quad \text{and} \quad U(eR,\delta) = U(eR,\beta). \tag{22}$$

This implies:

$$\frac{\Pr\left(com_{eR,R}\right)}{\Pr\left(piv_{\beta\alpha}\right)} \simeq \frac{u\left(x^1_\beta - eR\right) - u\left(x^1_\alpha - eR\right)}{2(u(0) - u(R-eR))} - \frac{\Pr\left(piv_{\delta\beta}\right)}{\Pr\left(piv_{\beta\alpha}\right)} \frac{\left(u\left(x^1_\delta - eR\right) - u\left(x^1_\beta - eR\right)\right)}{2(u(0) - u(R-eR))}, \tag{23}$$

and a similar condition holds for eL voters. The right-hand side of this expression is clearly increasing in $x^1_\beta - x^1_\alpha$ and in $\left|x^1_\beta - eR\right|$, and decreasing in $\left|x^1_\delta - eR\right|$. In turn, $\Pr\left(com_{eR,R}\right)$ is decreasing in the vote share of δ, by Lemma 2. Moreover, in ϕ^* (the lowest values of $\phi_{eL,\gamma}$ and $\phi_{eR,\delta}$ for which the magnitudes of the two probabilities on the left-hand side of (23) are equal), the ratio $\Pr\left(com_{eR,R}\right) / \Pr\left(piv_{\beta\alpha}\right)$ is decreasing in the vote share of δ. Finally, if δ is a loser, the magnitude of $\Pr\left(piv_{\delta\beta}\right)$ is smaller than that of $\Pr\left(piv_{\beta\alpha}\right)$.

Now, we are ready to demonstrate Proposition 3, which is done in two steps:

Step 1: optimal location of γ and δ.

γ only collects votes from eL, and δ only from eR. By (23), the closer x_δ^1 is to eR, the smaller the ratio $\Pr(com_{eR,R}) / \Pr(piv_{\beta\alpha})$ must be. Since this ratio is decreasing in the vote share of δ, locating x_δ^1 closer to eR can only increase δ's vote share, and thereby its probability of election. It is thus a dominant strategy for δ to locate in eR. By symmetry, γ locates in eL.

Step 2: optimal positioning of α and β.

First, we show that $x_\alpha^1 = x_\beta^1 = 0$ is a dominated strategy for α and β: from the first step, eL-voters prefer the platform of γ and eR-voters that of δ. If γ and δ are sure losers, the value of a vote for, say, α in the eyes of an eL-voter is asymptotically given by (16), where the first term is arbitrarily close to zero in $x_\alpha^1 = x_\beta^1$, and the second term is negative, which implies $U(eL, \alpha|\cdot) < 0$.

Therefore, $x_\alpha^1 = x_\beta^1$ ensures that α will not collect votes from eL-voters (and β will not collect votes from types eR by symmetry). As a result, γ and δ must be eligible with probabilities $\Pr(\omega_{eL})$ and $\Pr(\omega_{eR})$, whereas α and β are elected with probabilities $\Pr(\omega_L)$ and $\Pr(\omega_{eR})$ respectively. Conversely, $x_\alpha^1 < x_\beta^1$ increases the value of $U(eL, \alpha)$: by selecting a platform that is different from β, α increases his vote share, and hence his probability of election up to $\Pr(\omega_{eL}) + \Pr(\omega_L)$. The same applies for β, which shows that $x_\alpha^1 = x_\beta^1 = 0$ is a dominated strategy.

Now, we solve for the optimal positions of α and β. Start from a set of initial platforms $x_\alpha^1 = L = -x_\beta^1$, such that $\Pr(piv_{\delta\beta}) / \Pr(piv_{\beta\alpha})$ is close to zero. Do parties prefer to adopt platforms that are closer or more distant from one another? Consider potential deviations for β (they are symmetric for α). As long as types R vote for β in pure strategy, only eR's behaviour determines β's vote share. Since the voting strategy of types eR solves (23), locating x_β^1 closer to eR must increase $\phi_{eR,\beta}$, and thus β's vote share.

However, moving x^1_β towards eR also increases eL-voters' valuation of a vote for α. Thus, if:

$$\partial\phi_{eR,\beta}/\partial x^1_\beta > \partial\phi_{eL,\alpha}/\partial x^1_\beta, \tag{24}$$

locating closer to eR increases both s^ω_β and $(s^\omega_\beta - s^\omega_\alpha)$ in all states of nature, and as a result β's probability of election. By contrast, if (24) is violated, $(s^\omega_\beta - s^\omega_\alpha)$ decreases, as well as β's probability of winning. Using (22), we find:

$$\partial\phi_{eR,\beta}/\partial x^1_\beta \gtreqless \partial\phi_{eL,\alpha}/\partial x^1_\beta \Leftrightarrow \left|u'\left(eR - x^1_\beta\right)\right| \gtreqless \left|u'\left(eL - x^1_\beta\right)\right|.$$

Hence, if the utility function is concave (resp. convex) everywhere, platforms will be close to each other (resp. close to eL and eR) in equilibrium. If instead the second derivative of the utility function changes sign, the equilibrium can be strictly within these boundaries. Moreover, the assumption that $eR \leq 2L$ ensures that even if α adopts the platform $x^1_\alpha = 0$, types R still prefer the platform of β to that of α in $x^1_\beta = eR$, which ensures that $x^1_\alpha = 0$ is a dominated Nash response to $x^1_\beta = eR$. Thus, for $u(.)$ convex everywhere, $x^1_\alpha \to eL$ and $x^1_\beta \to eR$ is an equilibrium. This yields part iii) of Proposition 3.

■