DISCUSSION PAPER SERIES

No. 3404

WHY VOTE FOR LOSERS?

Micael Castanheira

PUBLIC POLICY

Centre for Economic Policy Research

www.cepr.org

Available online at: **www.cepr.org/pubs/dps/DP3404.asp**

ISSN 0265-8003

WHY VOTE FOR LOSERS?

Micael Castanheira, ECARES, Université Libre de Bruxelles and CEPR

Discussion Paper No. 3404
June 2002

Centre for Economic Policy Research
90–98 Goswell Rd, London EC1V 7RR, UK
Tel: (44 20) 7878 2900, Fax: (44 20) 7878 2999
Email: cepr@cepr.org, Website: www.cepr.org

CEPR Discussion Paper No. 3404

June 2002

ABSTRACT

Why Vote for Losers?*

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). Such predictions are not always met in practice, however. 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: D72 and D81
Keywords: Duverger's law, poisson games, signalling and voting

Micael Castanheira
ECARES
ULB CP 114
50 Avenue Roosevelt
1050 Bruxelles
BELGIUM
Tel: (32 2) 650 4467
Fax: (32 2) 650 3369
Email: mcasta@ulb.ac.be

For further Discussion Papers by this author see:
www.cepr.org/pubs/new-dps/dplist.asp?authorid=126467

* 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, three anonymous referees and the editor. I also greatly benefited from discussions with Patrick Bolton, Laura Bottazzi, Isabelle Brocas, Timothy Feddersen, John Fingleton, Steve Lawford, Mikko Leppamaki, 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. Remaining errors are (of course) my own.

Submitted 17 May 2002

1 Introduction

History abounds with examples of candidates who receive a substantial vote share, even though they are certain to lose the elections. Arguably, this may happen because of dull political campaigns or because main candidates propose uninteresting policies (see *e.g.* Cox 1997, p.76). However, this cannot be the sole reason why they get votes: disinterest should increase abstention, whereas –in the U.S. at least– the share of votes cast on such "losers" is empirically correlated with *lower* abstention rates.[1] Another argument could be that some fraction of the electorate "enjoys" voting for small parties. In this case, however, the vote share of those parties should remain almost constant over time. Instead, it varies substantially.[2] It might also be that voters are not able to coordinate on *who* the losers are. Losers may thus get votes "by mistake." Yet, in many instances, parties receive votes even though they are **clear** losers (in the U.S. 1992 presidential election for instance, Ross Perot received 19.7 million votes, compared to George Bush, who lost with 39.2 million votes). Can we really expect so many voters to be consistently mistaken? Or is it that we still do not understand why some voters sometimes vote for mainstream parties and, at other times, decide to vote for a loser? What motivates this voting behaviour? How does it affect implemented policies?

In this paper, we show that voters can use "losers" to influence the political stance of mainstream parties. Voting-for-losers allows voters to "voice" their political preferences: when the vote share obtained by a loser becomes large, mainstream parties know they should incorporate some elements of his political programme into their own platform. Thus, voting-for-losers becomes valuable because it affects platforms differently from a vote for one of the mainstream parties. To show this, we use a model of repeated first-past-the-post elections with opportunistic parties. The paper demonstrates that, in any "stable" equilibrium, a strictly positive share of the electorate must vote for losers, as long as there is enough uncertainty about the preferences of the electorate. In reaction, parties adopt a somewhat extremist stance, to prevent losers from receiving too many votes.

By contrast, existing theory would predict that rational voters should only allocate their ballots to two candidates (Duverger (1954), Palfrey (1989), Cox (1994), Myerson (1999)). This result is known as *Duverger's Law*, and derives from a very simple –yet extremely powerful– force. Consider a rational voter (she) for whom only the *outcome* of the election matters. If two candidates lead the polls, she knows that only one of them can win the election. Therefore, to have any effect on the outcome, she must cast her ballot on one of these two candidates. Voting for another candidate is useless, since he will never be elected. In this case, only two parties compete for election, and

[1]For U.S. Presidential elections between 1960 and 2000, and including a time control, a one point increase in the vote share of the losers appears to increases turnout by 21 percentage points (p-value: 6.4%).

[2]In the U.S. for instance, it increased from 1% in 1988 to 19.5% in 1992.

they should adopt centrist platforms if they want to maximize their probability of being elected (Hotelling (1929), Downs (1957), Black (1958)).

One could use various arguments to explain deviations from the above predictions: first, parties' extremism might be ideologically-motivated (see Witmann (1983), Calvert (1985), Alesina (1988), Alesina and Rosenthal (1995), or Roemer (1997)). Yet, the results of this paper suggest that opportunistic parties may need to adopt an extremist stance in order to capture more votes. Second, voters might not be fully rational or not solely concerned about the outcome of the election. Yet, we show that rational voters may need to vote for losers in order to influence election outcomes. In other words, this paper shows that straightforward extensions of the "standard" modelling of elections (i.e. when parties are opportunistic and voters only care about outcomes) can explain both the behaviour of the parties and that of the voters. To this end, it is sufficient to consider two features of the democratic process: first, the complete distribution of preferences in the electorate is generally not known before the election. Second, elections follow one another. We thus consider two-period repeated game, such that parties learn about the preferences of the electorate from first-election results, and can use this information to select their platforms in the second election. Then, we show that a vote-for-losers carries a specific type of signal to the parties, which is why the share of votes obtained by the losers affects second-period platforms differently from the share of the main parties. As a consequence, mainstream parties do not only compete against one another in equilibrium. It also becomes important to reduce the losers' vote share, which can only be achieved by adopting sufficiently different platforms.[3] In other words, voters' extremism alone can induce purely opportunistic parties to adopt a partisan stance.

Note that the role of information in elections and its influence on voting behavior has already been studied, although in different frameworks. Lohmann (1994) showed 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,[4] 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

[3] Although the model does not address the runoff system, the first round of the French Presidential election in 2002 can be viewed as a striking example of this result: prior to the election, opinion polls revealed that a large fraction of the French electorate could not tell the difference between the programmes of Lionel Jospin and Jacques Chirac. As a result, Jospin lost the first round against the "outsider" Le Pen, and these three candidates taken together only collected 52.76% of the votes.

[4] 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.

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.[5] (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 then used in the second one. His setting extends Myerson's results to three-party competition: a majority has imperfect signals about which of two candidates would maximize their utility, 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, voters will be assumed to be perfectly informed about their preferred platform. That is, voters have different (private) valuations of the possible election outcomes, like in the standard Downs-Hotelling model. This paper is thus based on a mechanism different from that of Condorcet's Jury.[6] Second, we jointly analyze the strategy of the voters and that of the parties. This extends existing results to endogenize the platforms of the parties.[7] In the above-mentioned papers, platforms are exogenous. Information gathering is then essentially a coordination game amongst (almost) indifferent voters. By contrast, we show that, when platforms are endogenized, extremist voters may also have an incentive to vote for small extremist parties. The predictions of the two approaches are thus quite different. According to the above literature, voters with strong political preferences never "communicate." Communication focuses on moderate parties who have a chance of being elected, either in "today's" or in "tomorrow's" election. Here, parties need not be eligible to gather ballots, and they endogenously choose not to be too moderate.

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

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

[6] Whereas the Condorcet Jury approach is related to common value auctions (see also Feddersen and Pesendorfer (1996)), the present approach is closer to that of private value auctions.

[7] One exception is Razin (2000), who also analyzes the interactions between vote results 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. Hence, the share of votes going to the two parties will determine which platform they implement eventually. In equilibrium, voters must pay attention to the "mandate" they give to the parties, and the information generated by this mandate allows them to fine-tune the implemented policy.

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 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:

$$\Pi_p(x) = \Pr(X^1 = x_p^1) + \Pr\left(X^2 = x_p^2\right), \tag{1}$$

where X^t is the platform implemented at time t, and $\Pr(X^t = x_p^t)$ denotes the probability that party p wins the election.[8] 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 $x_p^t \in \mathcal{R}$ that maximizes his probability of winning.

Voters are indexed by v and only care about the platform implemented after each election. For simplicity, assume there are four types of voters: $v \in V = \{eL, L, R, eR\}$,

[8]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.

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 have $eL < L < 0 < R < eR$, and we assume symmetry, *i.e.* $eL = -eR$, and $L = -R$. 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.[9] As of period 1, the expected intertemporal utility of a voter with type v is given by:

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

With this utility function, any voter must be exactly indifferent between parties that propose a same platform. To alleviate technical problems linked to perfect 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 defined as the 3-dimensional simplex:

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

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 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 states of nature, $\omega \in \Omega = \{\omega_{eL}, \omega_L, \omega_R, \omega_{eR}\}$, and the position of the median voter (μ^{ω}) is different in each of them: $\mu^{\omega_v} = v$.

[9] We impose $u(eL - eL) - u(L - eL) \geq \frac{1}{3}\left|u(eL - v) - u(L - v)\right|, \forall v \in V$, to ensure that voters give enough weight to attracting platforms toward their own position.

To capture this aggregate uncertainty, it proves convenient to adopt the following specification: in state of nature ω, the probability r_v^ω 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 $r_L = r_R$.

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:

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

where $\Pr(\omega) = 1/4$ is the probability that nature chooses ω. Hence, $q(\omega|v)$ 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:

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

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:

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

where $\tilde{N}$ 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 $\Pr(v|\omega) = r_v^\omega$ (see (2) above). This implies that the number of voters with type v follows a distribution $\mathcal{P}\left(r_v^\omega \lambda\right)$ in state ω.

From Myerson's work, we know several properties of such a random population game. 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. This probability is called the "pivot probability" in the literature.

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 for which party they want to vote. 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 across the two elections, 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 4 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 3 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 $s^{\omega}_p(\Phi) \equiv \sum_v r^{\omega}_v\, \phi_{v,p}$ denotes the vote *share* of party p in state ω. The argument of the exponential, $-\theta^{\omega}_{pp'}(\Phi)$, is called the *magnitude* of this 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':

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

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 $s^{\omega}_{p} < \min\left(s^{\omega}_{\alpha}, s^{\omega}_{\beta}\right), \forall\, \omega \in \Omega$, 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 3 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. Using this Property, we find:

Lemma 1 *In the second election, if parties γ and δ are perceived as losers then the PBE is unique and such that* i) *γ and δ receive no vote ($s^{\omega}_{\gamma} = s^{\omega}_{\delta} = 0$, $\forall\omega$),* 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.[10] Lemma 1 thus extends to aggregate uncertainty some results that are already present in the literature: 1) absent communicational motives, strategic voting generates Duvergerian Outcomes; 2) rankings are self-fulfilling.[11]

Result *ii* in Lemma 1 is also reminiscent of already existing results: parties maximize their probability of winning by locating at the position of the median voter.[12] Note that when more than two parties compete against one another, such a result

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

[11] This implies that if, say, α and β are perceived as losers (instead of γ and δ), then all voters coordinate their votes on γ and δ only. That is, there exist different stable coordination equilibria, that all share the property of being Duvergerian.

[12] This result hinges in part on the assumption that voters have lexicographic preferences. If voters

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 the strategy of the voters and that of the parties 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 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 vote 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 these parties always locate between L and R in the first election. For the sake of concreteness, consider the situation in which 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, all 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*

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.

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 determine second-period platforms. But what is the probability that a vote-for-losers affects those platforms? Defining $K \equiv \log\left[s_{p1}^{\omega_v}/s_{p1}^{\omega_w}\right] / \log\left[s_{p2}^{\omega_v}/s_{p2}^{\omega_w}\right]$, we find:

Lemma 2 *If two types of voters, v and w, vote for a same party, p_1, and types v also vote for another party p_2, a vote for p_2 can move platforms from w to v in the second period. Technically, for $\phi_{v,p_1} + \phi_{v,p_2} = 1$, $\phi_{w,p_1} = 1$ and $\phi_{v',p_1} = \phi_{v',p_2} = 0$, $\forall v' \neq v$, w, we have:*

$$\Pr(com_{vw}) \propto \exp\left[-\chi\left(\phi_{v,p_2}\right)\lambda\right]/\sqrt{\lambda}, \;\; \text{with } \chi\left(\phi_{v,p_2}\right) = \frac{\left(K\, s_{p1}^{\omega_v}(\Phi) + s_{p2}^{\omega_v}\left(\phi_{v,p_2}\right)\right)^2}{K^2\, s_{p1}^{\omega_v}(\Phi) + s_{p2}^{\omega_v}\left(\phi_{v,p_2}\right)}$$

where '$\propto$' means 'proportional to' and $\chi\left(\phi_{v,p_2}\right)$ has the following properties:

$$\lim_{\phi_{v_1,p_2} \to 0} \chi\left(\phi_{v,p_2}\right) = 0 \quad \text{and} \quad \chi\left(\phi_{v,p_2}\right) > 0, \; \forall \phi_{v,p_2} > 0.$$

Moreover the probability of being communication-pivotal between any other pair of positions is infinitely smaller than $\Pr(com_{vw})$.

Proof. See Appendix A.2.3. ■

In words, if types v vote (with positive probability) for party p_2, a ballot for p_2 may induce parties to locate in v instead of some other position w in the second period. For instance, assume that types eL vote for both $p_1 = \alpha$ and $p_2 = \gamma$, and types L only vote for $p_1 = \alpha$. 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), 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 will thus depend on the relative probability of determining the outcome of the first election (outcome-pivotability) compared to that of influencing platforms in the second election (communication-pivotability). That is, it will depend on the relative magnitudes of the two events. Our first proposition shows that:

Proposition 1 *For $x_p^1 \in [eL, eR]$, $\forall p$; $x_\alpha^1 \simeq -x_\beta^1$, $x_\gamma^1 \simeq -x_\delta^1$, 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 value of r_{eL} and r_L, either i) *two parties remain 'losers,' and voting strategies are then independent of their platforms, or* ii) *each party has a probability of being elected equal to 1/4.*
• ***Type-II equilibria: three parties receive a strictly positive vote share.***
A Type-II equilibrium always exists if $\eta < 1/3$. Under this condition, only two parties get elected with strictly positive probability.
• ***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 p_2, 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 p_2? Since parties know we vote for p_2, 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 p_2 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 p_2 with very small probability, we have $\Pr(com_{vw}) / \max_{p,p'} [\Pr(piv_{pp'})] \to \infty$, which implies that such a strategy cannot be an equilibrium. Put differently, the fact that parties interpret a vote for p_2 as a signal in favour of some state ω_v triggers a voting strategy under which types v must vote for p_2 with strictly positive probability.

The above gives the general intuition regarding the trade-offs 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}. For some parameter values, the equilibrium 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, the equilibrium vote shares of γ and δ may become "large," which implies that they can get elected, and thus that 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 $\eta = 1 - 2(r_L + r_{eL})$ small enough, whereas the latter case arises for η large enough. In Example 1 below, we set $\eta = 0.4$ 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 $r_{eL} = r_L = 0.15$; $x^1_\alpha = L$, $x^1_\beta = R$, $x^1_\gamma = eL$, and $x^1_\delta = eR$.

Type-I equilibria. Four parties get votes in such an equilibrium. For instance,[13] if types eL vote for γ and types eR vote for δ, the equilibrium is characterized by $\phi^*_{eL,\gamma} = \phi^*_{eR,\delta} = 0.35$. Under this strategy, the probability of being outcome-pivotal (when voting for α or β) and of being communication-pivotal (when voting for γ or δ) have the same magnitude. The equilibrium vote shares of the parties are:

Table 1: Equilibrium vote shares when extremist types vote for losers.[14]

	State of nature			
	ω_{eL}	ω_L	ω_R	ω_{eR}
s^ω_γ	*0.19*	0.05	0.05	0.05
s^ω_α	**0.51**	**0.65**	**0.25**	**0.25**
s^ω_β	**0.25**	**0.25**	**0.65**	**0.51**
s^ω_δ	0.05	0.05	0.05	*0.19*

[13]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.

[14]Bold-faced figures represent the largest two vote shares in each state of nature.

Type-II equilibria. Three parties get votes in such an equilibrium. For instance, there is an equilibrium in which types eL vote for γ with probability 0.29 and for α with probability $1 - 0.29 = 0.71$, whereas all other types only vote for α or β. 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^{ω}_{γ}	*0.16*	0.04	0.04	0.04
s^{ω}_{α}	**0.54**	**0.66**	**0.26**	**0.26**
s^{ω}_{β}	**0.30**	**0.30**	**0.70**	**0.70**
s^{ω}_{δ}	0	0	0	0

Type-III equilibria. If only α and β collect votes, α wins against β by a 70%-30% margin in states ω_{eL} and ω_L, and loses by a 30%-70% 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: for any $0 < \phi_{eL,\gamma} < 0.32$, a vote is infinitely more likely to be communication-pivotal than outcome-pivotal. In $\phi_{eL,\gamma} = 0.32$, the two probabilities are equal. However, a vote for γ is most likely to be outcome-pivotal against α under this strategy. 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, for any $\phi_{eL,\gamma} \in (0, 1)$: 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.

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:

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 overall vote shares of γ and δ in state ω are thus:

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

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.[15] Proposition 2 instead demonstrates that this presumed stability hinges on the perceived role of the election. Piketty (2000) already showed that non-Duvergerian outcomes can be stable. However, according to his results, this can only happen when some voters are uncertain about their preferences. Instead, Propositions 1 and 2 show that his conclusion overlooks the impact of election results on the parties' strategy. As soon as parties' future (re)actions are taken into account, voting-for-losers has positive value, even for voters with fixed preferences. By voicing their ideological differences, voters threaten main parties with loss of the election. If there are sufficiently many voters who convey such a message, the main parties will have to concede, and 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.

[15]In the words of Fey (1997), this means that, when the value of information is introduced in the model, equilibria with voting-for-losers become "expectationally stable," whereas Duvergerian outcomes are not.

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 share to that of the second leader (the "SF-ratio").[16] 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.

Our model predicts that, in the former case, 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, our model predicts that losers should obtain a larger vote share than when the election is close. With an SF-ratio that is most likely to lie somewhere 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.

[16] 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 behavior, it does not provide any valuation of the welfare costs entailed by vote dispersion. Note however that, by definition, losers get elected with probability 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 position of the* ***actual*** *median voter.*

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.[17] In our framework, this problem can be simplified by focusing on the set of parameters for which two parties remain losers, independently of their platform positions (as in Example 1 above). Under such parameter values, losers *never* get elected, unless leaders were adopting "wrong" platforms.

Leaders, on the other hand, compete with one another for election. However, a part of their electorate also votes for losers. Hence, the vote shares of the losers also matter to determine their probability of getting elected.

[17] 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.

For the sake of concreteness, this section focuses on the type of voting equilibrium presented in Table 1 above: types eL vote for α and γ, types eR vote for β and δ, and the two moderate types L and R only vote for the mainstream parties α and β. In this case, α competes for the votes of types eL on the left and, to capture the votes of R-voters, it would need to get strictly closer to R than x^1_β.

When extremist voters support γ and δ, mainstream parties must trade off the tension between moderation and extremism: if α moves closer to β, it loses interest in the eyes of eL-voters (γ's vote share increases and α's vote share falls). Conversely, if α moves left (towards eL), it gains interest in the eyes of eL-voters, but loses support in the center (L-voters may find β more attractive). 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 *The equilibrium locational strategy of the parties in the first period 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. ■

On the side of losers, the intuition for these results is straightforward: 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 (and hence their probability of election).[18] For them, locating in eL and eR is a strictly dominant strategy.

The intuition behind the locational strategy of the main two parties (α and β) is more intricate. Extremist voters always value positively a vote-for-losers, independently of their platforms. By contrast, the value of a vote for one of the main parties depends on the distance between their platforms. If the two platforms are identical, losing the election is virtually costless.[19] Protest voting on extremist parties then becomes close to a cheap-talk game, and this induces excessive communication (γ and δ must be potential winners of the election in this case). Instead, maintaining enough distance between platforms ensures that, in the eyes of an extremist voter, losing the election does have a cost. This reduces the propensity to vote for losers, and in turn increases the probability that both α and β are elected.

This shows that main parties want to maintain "some" distance between their

[18] One could object that losers face an election probability of zero. However, this is only true in the limit, for $\lambda = \infty$.

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

platforms. How much distance? This depends on the shape of the utility function. By moving towards eL, α reduces both eL voters' and eR voters' propensity to vote for losers. Therefore, both the share of α and of β increase. If eL-voters are more sensitive to this move than types eR, then α's share increases more than β's. This in turn increases the probability that α is elected. Conversely, if eR-voters react more strongly, α's probability of election is reduced. Depending on the case, the main parties thus maintain more or less distance between their platforms.

These interactions were never analyzed in the literature before. The Condorcet Jury literature, for instance, always considers parties as passive players of the electoral game. According to that literature, communication happens in election because voters privately receive information that tells them one party seems to be "better" than the other. Hence, if their political preferences are similar, voters use elections to pool their information. By contrast, extremists would never communicate in such a way, since information generated by moderate voters would never affect their political opinion. Yet, in today's elections, extremists seem to have an important weight, which cannot be rationalized by these models. The results in this paper fill this gap: by analyzing the linkages between election results and parties' platforms, the model allows us to better understand why extremist voters are willing to signal, and why extremists can weigh heavily on moderate parties' policies.

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 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, parties adopt more extreme positions, even though the median voter may actually be moderate. If equilibrium platform locations are close to eL and eR, the platform implemented at the end of the first period is thus necessarily extreme, even though the median voter might be moderate. That is, 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 [20]

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

[20] I thank an anonymous referee for stressing the issues raised in this section.

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 the outcome of the election in her favour.[21] 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 in the framework of this model. 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 policy platforms.

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 equivalently likely. Nonetheless, all the proofs for the results of Section 3 were carried out assuming any set of (strictly positive) prior probabilities. This shows that, as long as the probability of all states of nature is strictly positive, 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 election, *in addition to* voting-for-losers. Put differently, these two signalling devices are complementary, rather than substitutes.[22]

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).

[21] 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 own preferred platform.

[22] Lohmann (1993, 1994, and 2000) analyzes the problems of collective action through such "alternative" communication methods (as well as their effects on political choices by the parties) in much greater detail.

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 can 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.

This shows that elections also have the power of eliciting accurate information about the complete distribution of preferences in the electorate. This, in turn, 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 thus 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, as sufficiently extreme platforms increase the parties' probability of being elected.[23]

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.[24] Yet,

[23] 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.

[24] Note that Mueller (1989) and Myerson (1999) compare the properties of the different systems. However, they abstract from information aggregation considerations. Piketty (2000) discusses how

generalizing the model in this 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, A. (1988), "Credibility and Policy Convergence in a Two-Party System with Rational Voters", *American Economic Review*, Vol. **78**, 796-805.

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

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

CAILLAUD, B. and TIROLE, J. (1997), "Voters homogeneity and aggregation of information" (Unpublished, CERAS and IDEI).

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

CASTANHEIRA, M. (1998), "Why Vote for Losers?" (IGIER Working Paper No125).

COX, G. (1994), "Strategic Voting Equilibria Under the Single Non-Transferable Vote", *American Political Science Review*, **88**, 608-621.

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

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

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

FEDDERSEN, T. and PESENDORFER, W. (1996), "The Swing Voter's Curse", *American Economic Review*, **86**, 408-424.

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

FEY, M. (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, H. (1929), "Stability in Competition", *Economic Journal*, **39**, 41-57.

JOHNSON, N. and KOTZ, S. (1969) *Discrete Distributions* (New York: John Wiley & Sons).

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, his model also focuses on first-past-the-post elections.

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

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

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

MUELLER, D. (1989), *Public Choice II* (Cambridge: Cambridge University Press).

MYERSON, R. (1997), "Population Uncertainty and Poisson Games" (Northwestern University Discussion Paper No1102R).

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

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

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

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

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

PALFREY, T. (1989), "A Mathematical Proof of Duverger's Law", in P. Ordeshook (ed), *Models of Strategic Choice in Politics* (Ann Arbor: University of Michigan Press), 69-91.

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

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

RAZIN, R. (2000), "The effect of Vote Shares on Policy and Information Aggregation" (Unpublished, New York University).

RIVIÈRE, A. (2000), "Citizen Candidacy, Party Formation and Duverger's Law" (Unpublished, Université Libre de Bruxelles).

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

WITTMAN, D. (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 3 (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*

$$\lim_{\lambda\to\infty} \Pr\left(\tilde{z}_p = \tilde{z}_q + c|\omega, s_p^{\omega}, s_q^{\omega}\right) = \left(s_p^{\omega}/s_q^{\omega}\right)^{\frac{c}{2}} \frac{\exp\left[-\left(\sqrt{s_p^{\omega}}-\sqrt{s_q^{\omega}}\right)^{2}\lambda\right]}{2\sqrt{\pi\lambda}\,\left(s_p^{\omega}\, s_q^{\omega}\right)^{1/4}},\ p,q \in P,\ \omega \in \Omega. \tag{8}$$

Proof. From (7), we have:

$$\Pr\left(\tilde{z}_p = \tilde{z}_q + c|\omega, s_p^{\omega}, s_q^{\omega}\right) = \sum_{k=0}^{\infty} \frac{e^{-\left(s_p^{\omega}+s_q^{\omega}\right)\lambda} \times \left(s_p^{\omega}\lambda\right)^{k+c} \times \left(s_q^{\omega}\lambda\right)^{k}}{(k+c)!\ k!} \tag{9}$$

In addition, the definition of the modified Bessel function I of degree c is given by:

$$I_c(2\lambda) = \sum_{k=0}^{\infty} \frac{\lambda^{2k+c}}{(k+c)!\ k!},\ \text{with } \underset{\lambda\longrightarrow\infty}{lim}\ I_c(\lambda) = \frac{e^{\lambda}}{\sqrt{2\pi\lambda}} \tag{10}$$

and, after some manipulations, substituting (10) into (9) yields (8). ■

Property 4 below is a direct corollary of Property 3:

Property 4 *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 (8) , 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 3 and 4. 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:

$$W(v,p)=\sum_{\tilde{p}\in P\backslash p}\Pr(piv_{p\tilde{p}})\cdot\left[u\left(\left|x_p^2-v\right|\right)-u\left(\left|x_{\tilde{p}}^2-v\right|\right)\right], \tag{11}$$

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 $\lambda\to\infty$:

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

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 expected median voter. Denote by $\Pr(\mu^\omega=v)=\Pr(\omega|\mathsf{I}_p)$ the parties' perceived probability that the median voter has type v, given all the information available to them, including first election results. Denote also the *perceived position of the median voter* by μ^*, defined as:

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

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

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

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* $\mathsf{z}^1=\left\{z_\alpha^1,z_\beta^1,z_\gamma^1,z_\delta^1\right\}$, *the ratio of posterior beliefs about two given states of nature,* ω *and* $\tilde{\omega}$, *is given by:*

$$\frac{\Pr\left(\omega|\mathsf{z}^1\right)}{\Pr\left(\tilde{\omega}|\mathsf{z}^1\right)}=\frac{\Pr(\omega)}{\Pr(\tilde{\omega})}\prod_{p\in P}\left(\frac{s_p^\omega}{s_p^{\tilde{\omega}}}\right)^{z_p^1}. \tag{12}$$

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

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

Bayesian updating also tells us that:

$$\Pr\left(\omega|\mathsf{z}^1\right)=\frac{\Pr\left(\mathsf{z}^1|\omega\right)\cdot\Pr(\omega)}{\Pr\left(\mathsf{z}^1\right)},$$

which directly leads to (12). ■

A.2.3. Proof of Lemma 2. For the sake of concreteness, consider one particular case (the reasoning clearly extends to any other voter v and any other party): 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, $\Pr\left(\omega_{eL}|\mathsf{z}^1\right) / \Pr\left(\omega_L|\mathsf{z}^1\right)$ is strictly increasing in z_{γ}^1, and a vote for γ can thus move second-period platforms from L to eL. Arguably, this vote could also be communication-pivotal between ω_{eL} and ω_R or between ω_{eL} and ω_{eR}. In these two cases, however, the ratios of posterior beliefs involve at least three vote shares and vote results, and it is straightforward to show that, for λ "large," the probability of being communication-pivotal becomes infinitely smaller in those cases than between ω_{eL} and ω_L.

Next, affecting the parties' posterior beliefs about the relative likelihood to be in states ω_{eL} or ω_L does **not** influence second-period platforms if $\Pr\left(\omega_{eL}|\mathsf{z}^1\right) + \Pr(\omega_L|\mathsf{z}^1) < \Pr\left(\omega_R|\mathsf{z}^1\right) + \Pr(\omega_{eR}|\mathsf{z}^1)$. By Property 4 (see Appendix 1), for $\lambda \to \infty$, the probability that this happens converges to one (respectively: zero) for $\omega \in \{\omega_R, \omega_{eR}\}$ (resp.: $\omega \in \{\omega_{eL}, \omega_L\}$): a vote for α or γ can only be communication-pivotal if the true state of nature is ω_{eL} or ω_L.

When can a ballot for γ move second-period platforms from L to eL? This can only happen if, *without* this ballot, we have:

$$\frac{q\left(\omega_{eL}|\mathsf{z}^1\right)}{q\left(\omega_L|\mathsf{z}^1\right)} < 1, \tag{13}$$

while with one additional vote for γ we have:

$$\frac{q\left(\omega_{eL}|\mathsf{z}^1\right)}{q\left(\omega_L|\mathsf{z}^1\right)} > 1. \tag{14}$$

That is, one vote for γ increases the ratio (12) from a value below 1 to a value above 1 iff:

$$\frac{q\left(\omega_{eL}|\mathsf{z}^1\right)}{q\left(\omega_L|\mathsf{z}^1\right)} = \left(\frac{s_{\alpha}^{\omega_{eL}}}{s_{\alpha}^{\omega_L}}\right)^{z_{\alpha}^1} \left(\frac{s_{\gamma}^{\omega_{eL}}}{s_{\gamma}^{\omega_L}}\right)^{z_{\gamma}^1} < 1 < \left(\frac{s_{\alpha}^{\omega_{eL}}}{s_{\alpha}^{\omega_L}}\right)^{z_{\alpha}^1} \left(\frac{s_{\gamma}^{\omega_{eL}}}{s_{\gamma}^{\omega_L}}\right)^{z_{\gamma}^1+1}. \tag{15}$$

Taking logarithms, this condition becomes:

$$-1 < y \equiv z_{\alpha}^1 \frac{\log\left(s_{\alpha}^{\omega_{eL}}/s_{\alpha}^{\omega_L}\right)}{\log\left(s_{\gamma}^{\omega_{eL}}/s_{\gamma}^{\omega_L}\right)} + z_{\gamma}^1 < 0, \tag{16}$$

and we still need to compute the probability that this condition is met.

Considering jointly condition (16) and that ω must belong to $\{\omega_{eL}, \omega_L\}$, the probability that one vote for γ is communication-pivotal reduces to:

$$\Pr\left(com_{eL,L}\right) = \Pr\left(\omega_{eL}\right) \cdot \Pr\left(-1 < \tilde{y} < 0|\omega_{eL}\right) + \Pr\left(\omega_L\right) \cdot \Pr\left(-1 < \tilde{y} < 0|\omega_L\right),$$

in which we defined $\tilde{y} \equiv \tilde{z}_{\alpha}^1 \frac{\log\left(s_{\alpha}^{\omega_{eL}}/s_{\alpha}^{\omega_L}\right)}{\log\left(s_{\gamma}^{\omega_{eL}}/s_{\gamma}^{\omega_L}\right)} + \tilde{z}_{\gamma}^1$. It thus remains to determine the probability that condition (16) is met for $\omega = \omega_{eL}$ or $\omega = \omega_L$.

By the law of large numbers, large values of λ ensure that a Poisson distribution $\mathcal{P}(\lambda)$ is very well approximated by a Normal $\mathcal{N}(\lambda, \lambda)$ (Johnson and Kotz (1969, p. 99)). Defining $K \equiv \log\left(s_{\alpha}^{\omega_{eL}}/s_{\alpha}^{\omega_L}\right) / \log\left(s_{\gamma}^{\omega_{eL}}/s_{\gamma}^{\omega_L}\right)$; and remembering that $\tilde{z}_{\alpha}^1 \sim \mathcal{P}\left(s_{\alpha}^{\omega}\lambda\right)$ and $\tilde{z}_{\gamma}^1 \sim \mathcal{P}\left(s_{\gamma}^{\omega}\lambda\right)$, we find that, for $\lambda \to \infty$:

$$\tilde{y} \sim \mathcal{N}\left(\left(K\, s_{\alpha}^{\omega} + s_{\gamma}^{\omega}\right)\lambda, \left(K^2\, s_{\alpha}^{\omega} + s_{\gamma}^{\omega}\right)\lambda\right).$$

Then, using the properties of Normal distributions, we have:

$$\begin{aligned}\Pr\left(-1<\tilde{y}<0|\omega\right) &\simeq \kappa\left(\phi_{eL,\gamma}\right)\cdot\exp\left[-\chi\left(\phi_{eL,\gamma}\right)\lambda\right]/\sqrt{\lambda},\text{ in which} \\ \chi\left(\phi_{eL,\gamma}\right) &= \frac{\left(K\,s_\alpha^\omega\left(\phi_{eL,\gamma}\right)+s_\gamma^\omega\left(\phi_{eL,\gamma}\right)\right)^2}{K^2\,s_\alpha^\omega\left(\phi_{eL,\gamma}\right)+s_\gamma^\omega\left(\phi_{eL,\gamma}\right)},\text{ and} \\ \kappa\left(\phi_{eL,\gamma}\right) &= \left(1-\exp\left[-\chi\left(\phi_{eL,\gamma}\right)\right]\right)/\chi\left(\phi_{eL,\gamma}\right).\end{aligned} \tag{17}$$

By l'Hospital's rule, one can check that $\chi\left(\phi_{eL,\gamma}=0\right)=0$, and it is easy to verify that $\chi\left(\phi_{eL,\gamma}\right)>0$, for any $\phi_{eL,\gamma}>0$ (By means of numerical simulations, we could also verify that $\partial\chi/\partial\phi>0$ for all admissible ranges of parameters).

Also, the probability that a vote for α is communication-pivotal between L and eL is proportional to $\Pr\left(0<\tilde{y}<1|\omega\right)$ and has thus the same magnitude $\chi\left(\phi_{eL,\gamma}\right)$ as in (17).

This proof directly extends to cases in which two types of voters support a same leader (*e.g.* if both L and R vote for α). Hence, the behaviour of $\chi\left(\phi_{v,p}\right)$ derived here holds for any non-degenerate voting strategy (*i.e.* if at least one type of voter votes for a loser). ■

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:

$$\begin{aligned}W\left(v,p\right) = \sum_{\omega\in\Omega} q\left(\omega|v\right)\Bigg[&\sum_{p'\in P\setminus p}\Pr\left(piv_{pp'}|\omega\right)\,\left(u\left(x_p^1-v\right)-u\left(x_{p'}^1-v\right)\right)+ \\ &\sum_{v',w\in V}\Pr\left(com_{v',w}|p,\omega\right)\,\left(u\left(v'-v\right)-u\left(w-v\right)\right)\Bigg].\end{aligned}$$

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?

From Lemmas 2 and 3, if the vote share of p is exactly zero, 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 $\max_{p\in\{\gamma,\delta\}}\left[W\left(v,p\right)\right]<\max_{\tilde{p}\in\{\alpha,\beta\}}\left[W\left(v,\tilde{p}\right)\right]$, 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: $\phi_{eL,\gamma}+\phi_{eL,\alpha}=1$, $\phi_{eL,\gamma},\phi_{eL,\alpha}>0$; $\phi_{L,\alpha}=1$; $\phi_{R,\beta}=\phi_{eR,\beta}=1$.

To ensure that left-wing voters prefer x_α^1 to x_β^1 and conversely for right-wing voters, set $-x_\alpha^1\simeq x_\beta^1>0$ (in the opposite case, the identity of α and β must be reversed). By Lemmas 2, a vote for γ can be communication-pivotal between eL and L, and a vote for α between L and eL. By Property 1, for $\phi_{eL,\gamma}$ close to zero, we have:

$$\begin{aligned}W\left(v,\gamma\right) &\simeq \sum_{\omega\in\Omega} q\left(\omega|v\right)\Pr\left(com_{eL,L}\,|\omega\right)\,\left(u\left(eL-v\right)-u\left(L-v\right)\right) \\ W\left(v,\alpha\right) &\simeq \sum_{\omega\in\Omega} q\left(\omega|v\right)\left[\Pr\left(piv_{\alpha\beta}\,|\omega\right)\,\left(u\left(x_\alpha^1-v\right)-u\left(x_\beta^1-v\right)\right)+\right. \\ &\qquad \left.\Pr\left(com_{L,eL}\,|\omega\right)\,\left(u\left(L-v\right)-u\left(eL-v\right)\right)\right].\end{aligned} \tag{18}$$

Lemma 2 shows that $\max_\omega\left[\Pr\left(com_{eL,L}|\omega\right)\right]/\Pr\left(piv_{\alpha\beta}\right)=\infty$ for $\phi_{eL,\gamma}\to 0$, and hence that $W\left(eL,\gamma\right)>0>W\left(eL,\alpha\right)$. This shows by contradiction that $\phi_{eL,\gamma}$ close to zero cannot be

part of a Type-II equilibrium.
Now let us show that a Type-II equilibrium with $1 > \phi^*_{eL,\gamma} = 1 - \phi^*_{eL,\alpha} > 0$ always exists for $\eta \equiv 1 - 2(r_{eL} + r_L) < 1/3$.
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 $\bar{\phi}_{eL,\alpha} = r_{eL}/(r_{eL} + \eta)$, we have: $s^{\omega_{eL}}_\alpha = s^{\omega_{eL}}_\beta > s^{\omega_{eL}}_\gamma$. That is, $\bar{\phi}_{eL,\alpha}$ implies that $\max_\omega [\Pr(com_{eL,L}|\omega)] / \Pr(piv_{\alpha\beta}) = 0$, and therefore $W(eL,\gamma) < W(eL,\alpha)$ in $\bar{\phi}_{eL,\alpha}$, which can thus not either be part of an equilibrium. However, since $\phi_{eL,\gamma} \to 0$ implies $W(eL,\gamma) > W(eL,\alpha)$ and $\phi_{eL,\alpha} \to \bar{\phi}_{eL,\alpha}$ implies $W(eL,\gamma) < W(eL,\alpha)$, by continuity there must exist an "intermediate" strategy Φ^*, with $\phi^*_{eL,\alpha} \in (\bar{\phi}_{eL,\alpha}, 1)$ for which:

$$\chi(\phi^*_{eL,\gamma}) \simeq \min_\omega(\theta^\omega_{\alpha\beta}) = \min_\omega\left(\sqrt{s^\omega_\alpha(\phi^*_{eL,\alpha})} - \sqrt{s^\omega_\beta}\right),$$

and that ensures $W(eL,\gamma) = W(eL,\alpha) > W(eL,\beta)$. Is the strategy Φ^* an equilibrium for all voters? Types-L clearly do not want to deviate from $\phi_{L,\alpha} = 1$: by voting for α, they both signal their type (they can be communication-pivotal against eL) and ensure that their preferred platform is more likely to be implemented in the first period. Right-wing voters face a more complex trade-off. By voting for β, they increase the probability of x^1_β being implemented. However, by voting for α, they could moderate second-period platforms. However, using the following properties of pivot probabilities:

$$\Pr(com_{eL,L}|\omega) \simeq \Pr(com_{L,eL}|\omega); \quad \Pr(piv_{\alpha\beta}|\omega) \simeq \Pr(piv_{\beta\alpha}|\omega),$$

together with the equilibrium condition $W(eL,\gamma) = W(eL,\alpha)$, it is straightforward to verify that $W(v,\alpha) < W(v,\beta)$, for $v \in \{R, eR\}$ if the utility function $u(\cdot)$ is not too concave. In particular, $u(0) - u(eL - L) > (u(R - eL) - u(eR - eL))/3$ is a sufficient condition to ensure that Φ^* is an equilibrium.

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:

$$\phi^*_{eL,\gamma} + \phi^*_{eL,\alpha} = 1,\ \ \phi^*_{L,\alpha} = 1,\ \ \phi^*_{R,\beta} = 1,\ \ \phi^*_{eR,\beta} + \phi^*_{eR,\delta} = 1 \tag{19}$$

$$\chi(\phi^*_{eL,\gamma}) \simeq \min_\omega\left(\sqrt{s^\omega_\alpha(\phi^*_{eL,\alpha})} - \sqrt{s^\omega_\beta(\phi^*_{eR,\beta})}\right) \tag{20}$$

$$\chi(\phi^*_{eR,\delta}) \simeq \min_\omega\left(\sqrt{s^\omega_\alpha(\phi^*_{eL,\alpha})} - \sqrt{s^\omega_\beta(\phi^*_{eR,\beta})}\right) \tag{21}$$

$$\max\left[s^\omega_\gamma(\phi^*_{eL,\gamma}), s^\omega_\delta(\phi^*_{eL,\delta})\right] < \min\left[s^\omega_\alpha(\phi^*_{eL,\alpha}), s^\omega_\beta(\phi^*_{eL,\beta})\right],\ \forall\omega \in \Omega, \tag{22}$$

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 (18) above). Moreover, under the same condition on the concavity of $u(.)$, no voter wants to deviate from Φ^*, which is thus an equilibrium. (Clearly other Type-I equilibria exist if γ and δ are losers. For instance, types eL may vote for δ instead of γ and types eR for

γ instead of δ. However, the properties of this alternate equilibrium are identical: although γ and δ remain losers, they receive a strictly positive vote share.)

Case 2. γ and δ are not losers. If (22) is always violated when (19)-(21) are fulfilled, then a vote for α becomes more likely to be pivotal against γ than against β. By symmetry, a vote for β is also more likely to be pivotal against δ than against α. 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, the strategy profile Φ^{**}, such that $\phi^{**}_{eL,\gamma} = \phi^{**}_{L,\alpha} = \phi^{**}_{R,\beta} = \phi^{**}_{eR,\delta} = 1$ must be an equilibrium: under that strategy, $r_{eL} \leq r_L$ ensures that:

$$\chi\left(\phi_{eL,\gamma} = 1\right) = \frac{\left(K\, r_L + r_{eL} + \eta\right)^2}{K^2\, r_L + r_{eL} + \eta} > \min_{\omega\in\Omega}\left[\theta\left(\Phi^{**}\right)\right] = \left(\sqrt{r_{eL}+\eta} - \sqrt{r_L}\right)^2.$$

Hence, the value of a ballot only depends on outcome-pivotability, and, under the same conditions on the concavity of $u\,(.)$, one can verify that no voter can increase her expected utility by deviating from Φ^{**}.

■

Appendix 3: Proof of Proposition 3

We demonstrate Proposition 3 in two steps: first, we analyze the losers' strategy. Second, we analyze the leaders' strategy.

Step 1: optimal location of γ and δ.

γ only collects votes from eL, and δ only from eR. Taking the case of eR, the value of a vote for δ is given by:

$$\begin{aligned} W(eR,\delta) \quad = \quad & \sum_{p\in P\backslash\delta} \Pr\left(piv_{\delta p}|\phi_{eL,\gamma},\, \phi_{eR,\delta}\right)\cdot\left(u\left(\left|x^1_\delta - eR\right|\right) - u\left(\left|x^1_p - eR\right|\right)\right) \\ & + \Pr\left(com_{eR,R}|\phi_{eR,\delta}\right)\,\cdot\,\left(u\left(|eR - eR|\right) - u\left(|R - eR|\right)\right), \end{aligned}$$

where $\Pr\left(piv_{\delta p}|\phi_{eL,\gamma},\phi_{eR,\delta}\right) = \Pr\left(\tilde{z}^1_\delta = \tilde{z}^1_p\right)\cdot\Pr\left(\tilde{z}^1_p \geq \max_{p'\in P}\left(\tilde{z}^1_{p'}\right)\right)$. Clearly, $W(eR,\delta)$ is monotonically decreasing in $\left|x^1_\delta - eR\right|$ and, therefore, locating x^1_δ closer to eR can only increase δ's vote share, and hence his probability of victory (the same applies for γ when locating closer to eL). It is therefore a dominant strategy for γ to locate in eL and for δ to locate in eR.

Step 2: optimal positioning of α and β.

First, we show that $x^1_\alpha = x^1_\beta$ is a dominated strategy for α and β: from the first step, we know that eL-voters prefer the platform of γ and eR-voters that of δ. Moreover, if γ and δ are sure losers, then the value of a vote for, say, α in the eyes of an eL-voter is given by:

$$\begin{aligned} W\left(eL,\alpha|\gamma \text{ and } \delta \text{ are losers}\right) \quad = \quad & \Pr\left(piv_{\alpha\beta}|\phi_{eL,\gamma},\phi_{eR,\delta}\right)\cdot\left(u\left(\left|x^1_\alpha - eL\right|\right) - u\left(\left|x^1_\beta - eL\right|\right)\right) \\ & - \Pr\left(com_{L,eL}|\phi_{eL,\gamma}\right)\,\cdot\,\left(u\left(|0|\right) - u\left(|eL - L|\right)\right), \end{aligned}$$

where the first term is close to zero if $x^1_\alpha = x^1_\beta$, and the second term is negative, which implies $W\left(eL,\alpha|\cdot\right) < 0$. In turn, when γ is not a loser, and one can check that a vote for α is most likely to be pivotal against γ. In this case, $W\left(eL,\alpha\right)$ becomes even more negative:

$$\begin{aligned} W\left(eL,\alpha|\gamma \text{ is not a loser}\right) \quad = \quad & \Pr\left(piv_{\alpha\gamma}|\phi_{eL,\gamma},\phi_{eR,\delta}\right)\cdot\left(u\left(\left|x^1_\alpha - eL\right|\right) - u\left(|0|\right)\right) \\ & - \Pr\left(com|\phi_{eL,\gamma}\right)\,\cdot\,\left(u\left(|0|\right) - u\left(|eL - L|\right)\right) < 0. \end{aligned}$$

Therefore, $x_\alpha^1 = x_\beta^1$ ensures that α will not collect votes from eL-voters (and β will not collect votes from eR by symmetry). As a result, γ (and, by symmetry, δ) must be eligible with a probability $\Pr(\omega_{eL}) = 1/4$ (by symmetry, δ wins the elections with a probability $\Pr(\omega_{eR}) = 1/4$). As the probabilities of election must sum to 1, probability of election will also be 1/4 for both α and β. Conversely, $x_\alpha^1 < x_\beta^1$ increases the value of $W(eL, \alpha)$: by selecting a platform that is even marginally different from β, α can increase his vote share, and hence his probability of election. The same reasoning applies for β, which shows that $x_\alpha^1 = x_\beta^1$ is a dominated strategy.

Now, we solve for the optimal position of α and β. Consider a set of initial platforms $x_\alpha^1 = -x_\beta^1$. Do parties prefer to adopt platforms that are closer or more distant one from another? As initial platforms are symmetric around zero, we know that voting strategies will also be symmetric for any population size. Let us consider potential deviations for α (they are symmetric for β). As long as L's vote for α in pure strategy, only eL's behaviour matters to determine α's vote share. eL's value a vote for α as

$$\begin{aligned} W(eL,\alpha) &= \sum_{p\in-\alpha}\left[\Pr(piv_{\alpha p})\cdot\left(u\left(\left|x_\alpha^1 - eL\right|\right) - u\left(\left|x_p^1 - eL\right|\right)\right)\right] \\ &\qquad - \Pr(com_{L,eL}|\cdot)\cdot\left(u(0) - u(|L - eL|)\right) \end{aligned}$$

Clearly, for any given voting strategy, $\partial W(eL,\alpha)/\partial|x_\alpha - eL| < 0$. Therefore, locating closer to eL increases $\phi_{eL,\alpha}$.

However, moving x_α^1 towards eL also increases eR-voters' valuation of a vote for β. Therefore, if

$$\left|\frac{\partial\phi_{eL,\alpha}}{\partial x_\alpha^1}\right| > \left|\frac{\partial\phi_{eR,\beta}}{\partial x_\alpha^1}\right|,$$

locating closer to eL increases both s_α^ω and $\left(s_\alpha^\omega - s_\beta^\omega\right)$, in all states of nature, and as a result α's probability of winning also increases. By contrast, if

$$\left|\frac{\partial\phi_{eL,\alpha}}{\partial x_\alpha^1}\right| < \left|\frac{\partial\phi_{eR,\beta}}{\partial x_\alpha^1}\right|,$$

α's vote share increases, but $\left(s_\alpha^\omega - s_\beta^\omega\right)$ decreases, as well as α's probability of winning. As the equilibrium is defined by

$$W(eL,\gamma) = W(eL,\alpha) \text{ and } W(eR,\delta) = W(eR,\beta),$$

we find

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

Therefore, 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.

Beyond these marginal changes in position, a party could "leap" to the position most preferred by the electorate of the other leader. For instance, if $x_\alpha^1 << L$, then β could select $x_\beta^1 = L$ in order to induce a deviation from L voters. If this can happen, then α is constrained to locate "not too far" from L, but platforms still belong to the set (eL, eR) in equilibrium.

■