WHY VOTE FOR LOSERS ?[1]

Micael Castanheira
IGIER
Via Salasco, 5, 20136 Milan
E-Mail: mcasta@uni-bocconi.it

First draft: 7/30/96.
This draft: 30 March 1998.

[1]This paper is part of a larger research project jointly led with Denis Gromb (MIT Sloan Shool of Management). I am particularly indebted to him and this work incorporates conclusions of lengthy discussions we had. I would also like to thank Patrick Bolton, Juan Carrillo, Mathias Dewatripont, Timothy Feddersen, John Fingleton, Mikko Leppämäki, Jan Mahrt-Smith, James Mirrlees, Roger Myerson, Michele Polo, Anouk Rivière, Gérard Roland, Howard Rosenthal, Jean-Francois Ruhashyankiko, Georges Siotis, Guido Tabellini and Jean Tirole; seminar participants at ULB (Brussels), CentER (Tilburg), ESF (Florence) and IGIER (Milan) for helpful (and sometimes very long) discussions.

Abstract

Voting Theory generally concludes that –in first-past-the-post elections– 1) All votes should go to effective candidates (Duverger's Law); 2) Parties' platforms should converge (Median Voter Theorem). Observations, though, suggest that such predictions are not met in practice. We show that divergence and dispersion of votes is a natural election outcome when there is uncertainty and repetition of elections. "Voting for losers" increases the informational content of elections, and forces main parties to relocate towards extremists. As a result, they maximize their probability of being elected, not by converging to the median but by diverging to a certain extent. Ideological behavior results then from optimizing considerations alone.

JEL Classification Numbers: D72, D81

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

1 Introduction

Standard voting theory predicts that, in a first-past-the-post election, rational voters should allocate their votes to only two candidates (Duverger 1954, Cox 1994). Indeed, if voters care only about which candidate/platform wins the election, they should not waste their vote on "losers". Voting for a loser means that fewer votes are given to a second best option that could possibly win otherwise. This result, known as "*Duverger's Law*", is quite often far off the mark in practice. For example, the U.S. presidential elections showed that a third candidate could receive a substantial share of votes even without a chance of having any elected representative. In Europe, the U.K. elections display the same feature (see also Piketty 1995), and in the 1995 French presidential elections, around 40% of the votes were cast on small –mainly extremist– parties which had no chance of winning[1,2].

While they conflict with a literal interpretation of Duverger's law, such observations do not imply its irrelevance as a benchmark. It could simply be that one or more of the assumptions necessary for the result to hold are violated. A tempting explanation is thus that some voters are not rational, or that they do not only care about who is elected. Yet, casual observation suggests that such vote for losers influences mainstream parties' platforms. This contradicts the "median voter theorem", which states that opportunistic parties should position themselves at the median voter, not incorporate extremists' views. If these votes were sincere, such party behavior would itself be irrational.

This paper explores another alley that sustains the assumption of rational voters who care only about the implemented policy, and purely opportunistic parties, but introduces information concerns in elections. We show that in the case of repeated elections and incomplete information, some voters (mainly extremists) may vote for sure "losers" in order to communicate information about the distribution of voters and, in particular, about the

[1] Duverger's Law holds only for the first-past-the-post system. Cox (1994) proved a generalized version of this Law, which is now known as the "$n+1$" Law. That is, if n candidates have to be elected, the appointed ones being the n candidates receiving the highest shares of votes, then only $n+1$ candidates should receive votes at all. For a broader discussion, see also Cox (1997).

[2] The French electoral system is different, as the president is elected in two rounds. Nevertheless, insofar as the results in Cox (1994) extend to a two-round election, where the first round can be seen as a separate election meant to elect two candidates by plurality rule, only three candidates should receive votes in the first round, by the $n+1$ Law.

position of the median voter. Extremists have an incentive to vote for losers if more such votes mean that the median is closer to the voted loser. In this case, parties would adapt their platforms in future elections and come closer to the voted extremist party. Were party positions exogenous, voting for losers would never be a rational option for extremist voters. As a consequence, the very uncertainty on the distribution of voters provides a rationale for a violation of Duverger's Law.

If parties rationally select their platforms, they will try and reduce the tendency to vote for losers. If not, "small" parties could become too important and become a political threat. We show that, to this end, main parties need to take some distance with the median and propose somewhat extremist platforms. This shows that convergence of platforms to the median *cannot* occur in equilibrium, in opposition to the median voter theorem (Hotelling 1929, Black 1958). This means that party' partisanship is not needed to explain imperfect convergence of political platforms (as e.g. Calvert 1985 or Wittman 1977, 1983 would suggest). Divergence of platforms may be due to the voters' own dispersion of preferences.

Recent research stressed the informative role of elections. In that regard, Piketty (1995) is the paper that is closest to ours. It shows that, in repeated elections with uncertainty on the preferences of voters, the latter can communicate in the first election in order to aggregate the information that is dispersed among the electorate. This communication is initiated by voters who know they share a common preference for a candidate, but have imperfect knowledge about the identity of the latter. That is, they are "moderate" voters whose preferences for one or another party vary with the "state of the world". In equilibrium, communication can then induce a violation of Duverger's Law, because *voters hesitate*. In our model, voters perfectly know their preferences but try to modify the platform of the main candidates. This means that with Piketty (1995)'s assumption of fixed platform, our setup would not generate any violation of Duverger's Law. Piketty (1995) analyzes a vote for potential winners in the second election, while we want to focus specifically on the role of extremist voting, and on the influence of extremist parties on the platforms of moderate parties. The shared feature of the two models is that Duverger's Law is violated for communicational reasons only. That is, both models require repeated elections and an informational link between these two elections for communication to occur.

Feddersen and Pesendorfer (1997) show that voters with strong preferences over the alternatives will not use their private information in the vote. Only moderate voters, whose preferred party depends on the state of the world, will make use of their piece of information (with the share of such voters converging to zero as the size of the electorate increases). Such a behavior, or abstentionism (Feddersen and Pesendorfer, 1996a,b) increases the efficiency of the election, as the selected platform will depend on the true state of the world. Caillaud and Tirole (1997) show however that the joint presence of strong heterogeneity among voters and of any (aggregate) uncertainty about the distribution of preferences weakens these results.

Yet, again, these papers do not explain party's behavior. Were platforms a choice variable, would parties propose such distinct platforms that dissatisfy a strict majority of voters? Why does a "vote for losers" take this form of "protest" voting for small extremist parties? It is common knowledge that such parties will never attract a majority of votes. In addition, such "protest" vote seems to be mainly cast by extremist voters who have a clear and strong preference. Feddersen and Pesendorfer or Piketty explain the behavior of moderate and undecided voters, not that of extremists.

Banerjee and Somanathan (1997) focus on information aggregation through "voice" in large populations, not particularly in elections. Their results in the case of private information on "voters' type"[3] give however a good hint of the mechanism also at work in our model. They show that people with more extremist views are more likely to send relevant signals because they gain more from altering the decisions of the leader. We show here that elections are particularly appropriate to aggregate the information dispersed among the electorate, as they endogenously generate a cost of sending a message that permits efficient "voicing". This is achieved by parties when they select their platforms.

One of the main features of this paper is precisely that we endogenize parties' platforms in

[3] Given that Banerjee and Somanathan analyze a broader variety of communication problems, they compare different information cases. They do not actually consider elections, as there is only one leader/dictator, but people can try and influence the leader's decisions by voicing him/her some piece of information. This may make comparisons a bit thorny, but their case with unverifiable preferences of the sender (the voter?) and unverifiable information is nevertheless close enough to the framework of an election for this comparison to be meaningful.

a world of "informative elections". In that case, parties' positions must also be used to prevent communication. That is, parties can partly control how much of the votes they attract on themselves. Attracting the votes of extremists may prove too difficult and, in equilibrium, parties make the cost of protest voting exactly such that "voice" efficiently generates enough information about the distribution of voters, without making extremist parties potential winners of the elections.

Myerson and Weber (1993) analyze strategic interactions between numerous parties and voters, but in a framework where no aggregate information is to be communicated. Opportunistic parties generally converge to the median voter's position and Duverger's law should apply, except for peculiar beliefs. Our framework combines these two approaches and shows their complementarinesses.

We present the model in the next section. In section 3, we describe the equilibria that arise in two "intermediate" cases. First, we consider the full information case and then incomplete information, but with non-secret vote. This can be seen as a vote by raising hand. This allows us to shed light on the importance of the assumptions on information. In full information, we reproduce the standard results of convergence of parties and no vote for losers. In the case of observable ballot, we show that these two results cannot hold simultaneously. Information incompleteness implies a trade-off between convergence of parties and concentration of votes on a small number of parties.

In section 4, we present the central results of the paper. In the case of secret ballot (that is when only aggregate vote results are observed), there cannot be convergence to the median in equilibrium and, in addition, there is generally a dispersion of votes on losers. This occurs because anonymity of elections allows some voters to garble their informational content. When choosing platforms to increase their probability of being elected, parties must trade-off their desire to propose the median voter's preferred platform and the need for divergence to prevent excessive voting for losers. In equilibrium, parties must not converge, as the cost of vote dispersion exists only if losing elections imposes a relevant loss in terms of implemented platforms. In addition, parties usually do not want to capture the vote of extremists, which implies the presence of some voting for losers in equilibrium. Section 5 presents some comparative statics. Section 6 uses Piketty (1995)'s insights to show that the mechanisms described in the paper with a small population extend to Poisson games with a large number of voters. Section 7 concludes.

2 Model

Until section 5, we consider a small electorate faced with two sequential elections, in which four political parties are in competition. In section 6 we show this framework extends to a large number of voters, even if this is achieved at the expense of more intricate computations.

For voters, these elections also embrace a signaling game, as parties are unsure about the actual distribution of the electorate. Voters must provide an accurate signal about their preferences at the same time as they elect their representatives. Voters and parties are described below.

2.1 Voters

A finite number of voters, n, must elect a party in two distinct, successive, elections. Throughout the paper, we study different cases where the number of voters, n, varies. In the next two sections, $n = 5$. In each period, a voter i has a utility that is symmetric with respect to her bliss point, θ_i. The overall utility is function of x_1 and x_2, the platforms implemented in each of the two elections:

$$\begin{aligned} U\left(\theta_i, x_1, x_2\right) &= u_1(\theta_i, x_1) + \rho\ u_2\left(\theta_i, x_2\right) \\ &= -\left(\theta_i - x_1\right)^2 - \rho\ \left(\theta_i - x_2\right)^2, \end{aligned} \tag{1}$$

where ρ is a positive weighting (or discount) factor, denoting the relative importance of the two elections. ρ needs not be smaller than one.

Voters must cast their ballot so as to maximize their expected utility. Their vote influences the probability of having one platform elected over another. The probabilities of victory will be spelled out when needed.

The bliss point of any voter i can take only six different values, symmetrically disposed around zero, such discretization of the preference space being meant to simplify the model. Possible bliss points are shown in figure 1.

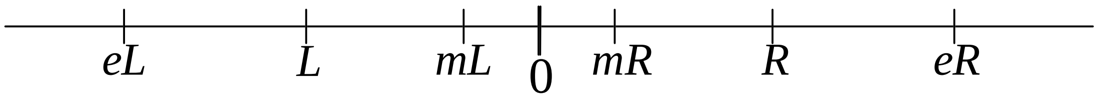

Figure 1: Possible Voters' Bliss Points. $L<0<R$, $e>1$, $0<m<1$.

$R>0$ denotes a right-wing preference, $L<0$, a left-wing preference, with $L=-R$. $e>1$ is a parameter of "extremization", $0<m<1$ is a parameter of moderation.

2.2 Voters' Distributions

The actual distribution of voters is initially unknown. There are four equiprobable[4] states of the world $\mathcal{EL}$ (Extreme Left), $\mathcal{ML}$ (Moderate Left), $\mathcal{MR}$ (Moderate Right), and $\mathcal{ER}$ (Extreme Right), corresponding to different distributions of voters and thus to different positions of the median voter. Let μ^S denote the median voter in state S. We have

$$\mu^{\mathcal{EL}}=L \quad < \quad \mu^{\mathcal{ML}}=m.L \quad < \quad \mu^{\mathcal{MR}}=m.R \quad < \quad \mu^{\mathcal{ER}}=R. \tag{2}$$

In the case of 5 voters, those can be distributed in the way described in figure 2.

The heart of the model lies here. The expected median is 0 on average, but the uncertainty on the distribution is sufficiently complex for making the process of learning the exact position of the median difficult. In the two left-wing (resp. right-wing) states of the world, there are 3 voters to the left (resp. right) of zero. That is, the quantile in 0 is the same for $\mathcal{ML}$ and $\mathcal{EL}$ or for $\mathcal{MR}$ and $\mathcal{ER}$, but the position of the median voter is different in each of these states of the world. The difference between a "moderate" and an "extremist" state of the world is that the presence of an extremist voter moves the position of the median of the distribution farther away from zero.

This justifies why we use six positions for the voters and not fewer: were m converging to 1, there would be no difference between the position of the median voter in extremist and moderate states of the world. Were e be **equal** to one, there would be no difference between an extremist and a "core" voter in L or R (Still, our results are valid for any e arbitrarily close to –but different from– one). Voters in L and R should be seen as the "solid bastions" of

[4]Equiprobability makes the set-up completely symmetric around zero. This will enable us to satisfy conditions for left-wing and right-wing voters and parties at the same time. When probabilities are modified, the conditions must be checked separately for left-wing and right-wing players. This is done in section 5.

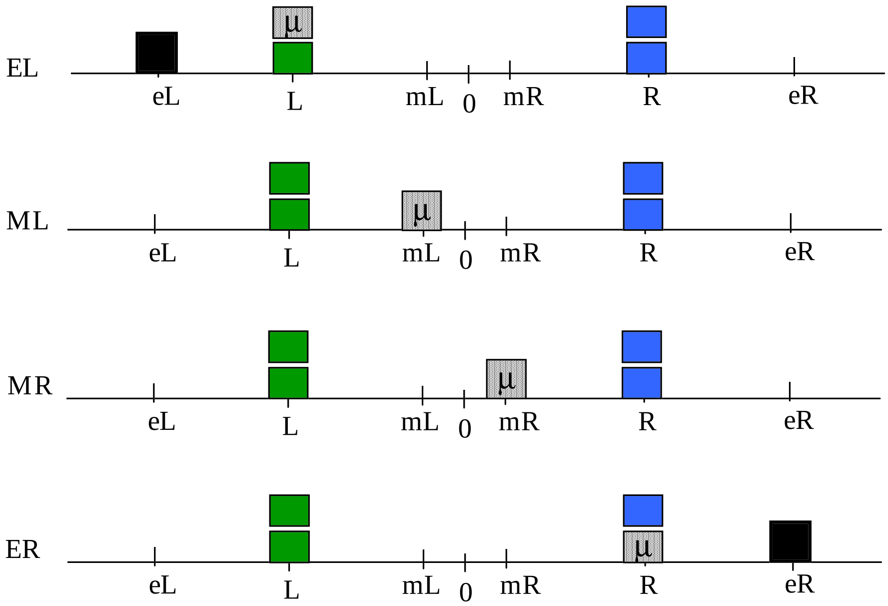

Figure 2: Potential Distributions of Voters. The figure displays the four possible distributions of voters in the case of five voters. μ identifies the median voter. Each box represents the presence of one voter.

main parties: they are not extremist, and have a strong preference for a left- (or right-)wing policy. Moderate voters are easier to swing, as they have weaker political preference.

2.3 Parties

Parties are separated in two classes. Two parties, Γ and Δ , are purely ideological and have "extremist" preferences −defined as a bliss policy that is distant from the bliss point of the median voter−. These parties will always propose a fixed platform that is equal to their bliss policy:

$$\pi_k^1 = \pi_k^2 = \theta_k,$$

where θ_k is the bliss policy of party $k \in \{\Gamma, \Delta\}$ and π_k^t is the platform proposed by party k in election t. For clarity of exposition, let us assume $\theta_\Gamma < L < R < \theta_\Delta$. Given their extremism, we assume that **these parties are secondary, they are not "focal". That is, they are a priori assumed to be losers of the elections**, as long as such beliefs are not inconsistent. This simply assumes away beliefs where moderate parties are felt to be less important than extremists who would attract a majority of the votes (Cox 1994, Myerson-Weber 1993).

This also prevents extremist parties from being considered as "potential entrants" who can threaten main parties (Palfrey 1984). Extremist parties are losers. Only a wrong action by main, opportunistic, parties could enable their victory, as it will appear in section 4.

The other two parties, A and B, are purely opportunistic. They only care about winning the elections. Elections select a winner by plurality rule (first-past-the-post) with a coin toss in case of a tie. This means that expected payoffs for these parties are

$$E\left[V(\theta_j, \pi_j^1, \pi_j^2)\right] = \sum_{t=1}^{2} \left[Prob\left(n_j^t > n_{-j}^t\right) + \frac{1}{2} Prob\left(n_j^t = n_{-j}^t\right)\right],$$

where n_j^t is the number of votes cast on party $j \in \{A, B\}$ at election t and n_{-j}^t is the highest number of votes cast on a party other than j in election t. Introducing discounting in the pay-off function of parties is of no influence, as parties' information is symmetric.

A priori, these parties are the likely winners of the election. That is, we assume that beliefs focus the attention of voters on these two parties as running for victory. This assumption of a dichotomy between the two sets of parties (2 "main parties" versus 2 "Losers") will be maintained throughout the paper and allows us to contrast with Myerson and Weber (1993): without this assumption, we should accept beliefs where main parties receive no votes, which increases the number of equilibria. To repeat, we want to dismiss this and focus on the effect of small parties (losers) on elections.

2.4 Timing

At date 0, Nature draws the state of the world, that is: a distribution of voters. In period 1, parties simultaneously select a platform, followed by voters who cast their ballot on one of the parties. The party getting the highest number of votes wins the elections (with a coin toss in case of a tie). Its platform is implemented and first-period payoffs are realized. Election results modify the information set of parties who compute posterior beliefs by Bayesian updating at the beginning of period 2, followed by a repetition of the period-1 game. The game ends at the end of period 2.

Period 0:

Nature selects state of the world.

Period 1:

Parties select a platform, π_j^1

Voters cast their ballot, v_i^1

Winning platform is implemented

Payoffs are realized.

Period 2:

Parties select a platform, π_j^2

Voters cast their ballot, v_i^2

Winning platform is implemented

Payoffs are realized.

3 Strategies and Equilibrium

Solving this game requires us to look for a Perfect Bayesian Equilibrium. The first logical step for the resolution consists of solving the equilibrium for the second period.

Lemma 1 *In a one shot election, parties' platforms converge to the expected median and all votes are cast on main parties in the Nash Equilibrium. As a result, both Duverger's Law and the median voter theorem always hold in the second period.*

Proof. Following Myerson and Weber (1993), and specifying Γ and Δ as losers a priori, only the pivot probability p_{AB} is positive[5]. This means that no vote will go to Γ and Δ in equilibrium (Myerson and Weber, pp103-104). As Γ and Δ are felt to be losers whatever main parties' platforms, the latter always converge to the position of the expected median (This is simply an application of Myerson and Weber's Theorem 3, where the probability vector in the set of all pairs of candidates always converges to 1 for the pair A, B only. Hence, among the several equilibria in Myerson and Weber, only one is made sustainable). ■

Lemma 1 implies that the implemented platform in the second period will be the bliss point of the expected median voter. If a voter can influence posterior beliefs about the position of the expected median voter, she can also modify the implemented policy in the second period.

[5] The pivot probability p_{kl} is the perceived probability that parties k and l are tied for the first place. In that event, one vote can change the election result.

As in Piketty (1995), the first election will be used for two purposes: information revelation and election of a platform. As we shall see, these two objectives conflict when first-period platforms do not converge fully.

We are looking for the Perfect Bayesian Equilibrium of this game, when voters' types are private information and only aggregate vote results can be observed. However, the mechanisms in the game are not straightforward, because of the multiple interactions between players. For the sake of clarity, the analysis will then proceed as follows: in the next section, we study the full information and an "intermediate" incomplete information cases, before going to the more realistic incomplete information, unobservable moves case in section 4.

3.1 Intermediate Information Cases

From now on, we shall concentrate mainly on the first period of the game and on the behavior of opportunistic parties. Indeed, extremist parties always propose a platform that is equal to their bliss point θ_j, $j{=}\Gamma,\Delta$ and –to repeat– are considered as losers. By lemma 1, the second-period equilibrium displays convergence to the expected median, votes are only cast on opportunistic parties (Duverger's Law applies) and each of the latter has a probability of victory equal to $\frac{1}{2}$.

3.1.1 Full information

Under **full information**, each voter's type is observed, so that the position of the median voter is common knowledge.

Proposition 1 *Under full information, opportunistic parties locate at the position of the median voter in both periods and voters cast their ballot on the closest opportunistic party (the distance between the parties being positive but arbitrarily close to zero). Both the median voter theorem and Duverger's law hold.*

Proof. The subgame perfect equilibrium of the repeated game is the repetition of the equilibrium of the one shot game. ■

This shows that the model encompasses the "standard" case where no vote goes to losers and parties locate at the position of the median voter. But comparing this full information case with incomplete information cases will nevertheless show that full information is a pre-condition for these results to hold.

3.1.2 Intermediate Incomplete Information Case

This case isolates the role of the election as an information aggregator. Let us temporarily assume that a) every cast vote is observed separately (as in a raising hand ballot) and b) only one of the voters' type is private information to her. That is, core voters (with $\theta_i = L, R$) are known perfectly, while the fifth voter, z, has an unobservable type $\theta_z \in \{eL, mL, mR, eR\}$.

Banerjee and Somanathan (1997) offers a more detailed analysis of the interactions between a sender and a receiver in a communication game of this type. In contrast to them, again, our model rather focuses on a case where elections aggregate information, which puts more restrictions on acceptable information setups. In that respect, we must stress that such a case cannot pretend to proxy secret ballot elections, as voters' ballots are not disclosed in the latter case. It nevertheless serves as a useful benchmark if one accepts that some districts (or socio-economic groups, ...) are solid bastions who clearly stand for a given party and cannot "cheat" on that regard. Beside these bastions, there are some other groups whose political feelings are unclear (unemployed, declining industrialized regions, etc...) which are closely watched at.

In this setup, as the type of core voters is perfectly observed, their vote can carry no signal. Hence, they will always vote for the closest opportunistic party (as in Myerson and Weber 1993), which implies that both A and B are sure to get at least two votes.

In contrast to the ballot of these four voters, the ballot of the unidentified voter can carry a signal. This voter can have 4 possible types: eL, mL, mR, or eR, that are all different from the type of core voters. Given the distribution of voters, identifying the type of z allows all players to pinpoint exactly the state of the world.

We see then that this intermediate case is actually a standard signalling game where the type of one player is unknown but can be inferred from her behavior in a separating equilibrium. The problem remains to see whether such separating equilibrium exists. Clearly, z whatever her type, has an incentive to communicate her preferences, as platforms will get closer to her in the second period. What are the available signals at hand?

The only signal a voter can use is her ballot[6]. We must then associate each party with the message it can carry. With four possible types and four parties, we have the choice among $4! = 24$ message structures. Nevertheless, only one of them seems acceptable in elections: a more left-wing party entails a more left-wing message. That is, the communication structure in elections satisfies beliefs monotonicity. Hence, eL has the choice to communicate (and vote for Γ) or to elect her preferred main party, A. eR faces the same trade-off between Δ and B. Moderate voters, by contrast, face no trade-off when parties' platforms are symmetrically positioned around 0: electing and communicating require the same ballot. Their problem becomes more complicated if main parties (A and B) have asymmetric positions. Anticipating on the results, moderate voters will not face any trade-off in equilibrium, but may face one when parties deviate from the equilibrium.

By backward induction, we must first study the equilibrium behavior of voters in the continuation game where opportunistic parties' platforms are already set, and study the optimal positioning of parties in a second step.

If main parties' platforms are identical, the first-period pay-offs are independent of z's vote ($\pi_A = \pi_B$), and the election becomes a cheap-talk game. As the preferences of the sender and the receiver of the message are perfectly aligned, we know that the separating equilibrium exists: eL votes for the extreme-left party, Γ, mL for the moderate left party, A, etc... We do not study pooling (babbling) equilibria here, as they are of limited interest for our purpose.

What happens when first-period platforms differ? Maintaining our message structure, extremist voters (eL or eR) face a trade-off: either they do not signal their type, but ensure the victory of their favored opportunistic party or they "protest" (vote for an extremist party), at the expense of an uncertain first-period payoff (each main party wins with a probability a half). To decide whether they vote "protest" (higher payoff in the second period, but higher risk in the first one) or "elect" (lower risk, but lower future payoff), they must compute the

[6] This abstracts volontarily from other communication tools as strikes, demontrations, terrorism, etc... that are hard to consider in such a model, but also from abstentionism that raises different issues (cost of voting, ideology of abstentionism, etc...).

expected utility of each strategy. This leads us to the following lemma:

Lemma 2 *The closer main parties' platforms in the first-period, the more extremist voters are prone to vote for a loser. (Proof in the appendix)*

When opportunistic parties select a platform, they will take this into account. Differentiating sufficiently one's platform allows to catch the vote of extremists, converging towards the median voter attracts moderates. This obtains:

Proposition 2 *With observable ballot, there exist two sets of equilibria. In the first set, there is divergence of platforms and no vote is cast on losers, which implies that the true state of the world is not learned perfectly. The second set displays a separating equilibrium, where extremists vote for losers and the true state of the world is learned perfectly. In this set, platforms converge.[proof in the appendix]*

This result comes from the trade-off spelled out in lemma2: parties know that, by proposing themselves a more extremist platform, they can prevent extremist voting. If, by doing so, they reduce their probability of winning, moderate platforms are an equilibrium. $\pi_A^1 = \pi_B^1 = 0$ is proven to be an equilibrium when m, the parameter of moderation, is small enough[7] (a small m implies that moderate voters swing to the other opportunistic party when one of the parties become more extremist).

Another candidate equilibrium is the one where parties share symmetric but sufficiently distinct positions, so that extremist voters do not vote for losers. This set of platforms is an equilibrium if running on moderate platforms does not increase their probability of winning. When m is large enough, moderate voters are difficult to swing as they pay a lot of attention to signalling their difference with moderates of the other side. Hence, moderating platforms means that the party loses the vote by extremist voters, without getting extra votes from centrists.

[7] Any symmetric pair of platforms in the vicinity of {0,0} is also an equilibrium as the space of voters' preferences is discrete. Nevertheless, if we were considering a continuous space, where $\theta_{mL} \sim \mathcal{N}(mL, \sigma_{mL})$ with $lim\, \sigma_{mL} \rightarrow 0$, then {0,0} would be the unique pair of platforms candidate for this equilibrium.

Again, proposition 2 hinges on the assumption that 4 voters are perfectly identified and cannot communicate. It nevertheless sheds light on the trade-off between convergence of platforms and the application of Duverger's Law. It also shows that the forces towards convergence dominate when there are strong tensions within a political "camp" (i.e. when extremists and moderate voters strongly oppose each other, parties must moderate themselves). The latter equilibrium is obtained when tensions are stronger between the left and the right than within a political camp (if moderates and extremists are close enough, parties must become more extreme).

Such results do not appear in Piketty (1995) because the type of uncertainty and the motivation of voters is different in our model. Here, there is a tension between extremist voters and moderate voters. An extremist is sure to prefer an extremist policy, whatever the state of the world. The state of the world determines if such extremism is relevant for parties or not. In Piketty (1995), unsure voters jointly want to learn whether party A or party B best represents their interest. In our model, one could see his case as moderate voters who do not know whether they are left-wing or right-wing because of uncertainty on the state of the world.

To summarize, two equilibria are possible. When the parameter of moderation, m, is small, parties converge to the expected median in period one (the median voter theorem applies), but this induces a vote for losers by extremist voters (Duverger's Law does not). If m is large enough, this convergence to the median is no longer an equilibrium. Parties' dominant strategy is to impose a cost to "protest" vote by diverging (Median voter theorem is violated, but Duverger's Law holds). For intermediate values of m, usually both candidates are an equilibrium. The equilibrium strategy of parties is then to mimic the other party. That is, be extremist if the other party is expected to play the extremist strategy and converge to the median if the other party is expected to converge to the median.

The latter result however does not apply for any set of parameter values. There exist narrow ranges of parameters (e very close to 1, $\rho \simeq 1$, and a narrow range of m) such that no equilibrium exists in pure strategy. Some weak restrictions on parameter values would ensure the existence of at least one equilibrium (e.g. e sufficiently larger than 1). Figure 3 displays the possible equilibria in function of the parameters.

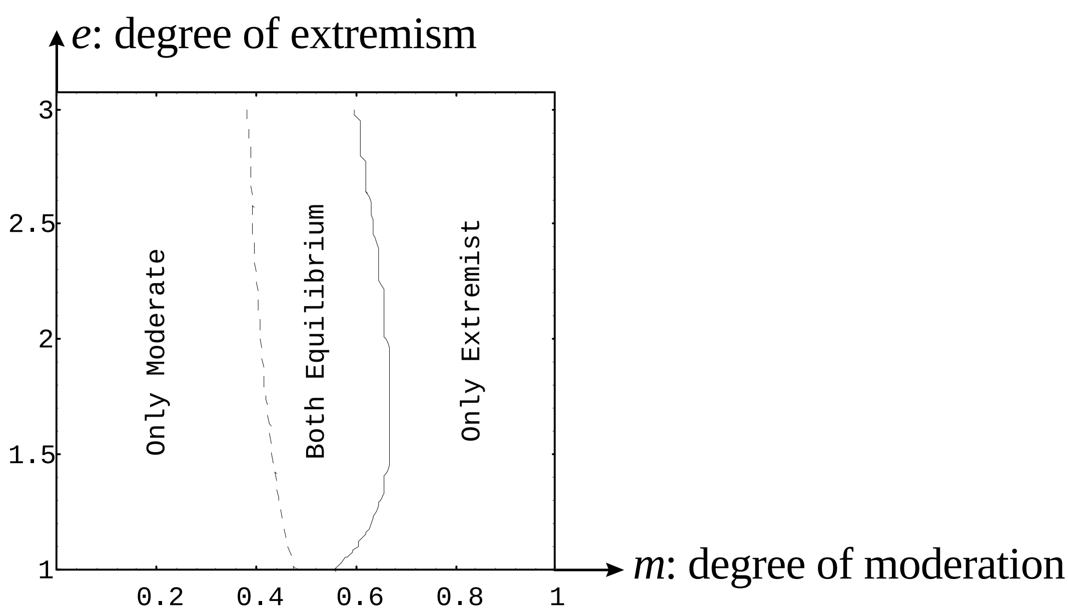

Figure 3: possible equilibria when ballots are observable. Horiz. Axis: m / Vert. Axis: e. ($\rho = 2.5$, $R = 1$).

This case of observable ballots remains partly unsatisfactory insofar as it does not capture the informational interactions between voters in L and R and the others. This leads us to the following case.

4 Incomplete Information: Secret Ballot

In "anonymous" elections, ballots are secret and only aggregate election results can be observed. In this case, voters in L and R can also influence beliefs as their ballot needs not being distinguishable from that of other types. The previous case was built so as to analyze a "standard" signalling game between one voter (the sender) and parties (the receivers of the signal) while abstracting from adverse selection issues. This is no longer the case in secret ballot elections: core voters (L and R) may be tempted to mimic extremist voters (and vote for extremist parties) in order to alter beliefs. This means there is a coordination problem to resolve: if the true state of the world is, say, extreme left, a second vote for a loser increases the number of votes Γ gets, but does not alter beliefs. However, if the true state of the world is, say, moderate left, then if an L-voter casts an extreme left ballot, she could be mistaken for the extreme left voter (eL).

Empirically, a lot of questions are raised about "who is the extremist voter?". In France, the National Front can get a majority in municipal elections, without convincing the public opinion that a majority of voters are Fascists. Political scientists have troubles in identifying "the" extreme right voter. Our model sheds some light on this point: even "not-so-extreme" voters —voters who would refuse to elect an extremist party— vote for it because of the signal this vote carries. Hence, there is no "typical" extremist voter. The electorate of these parties are composite.

In the case analyzed in the previous section, vote results are (numbers represent the vector of vote results, $(n^1_\Gamma, n^1_A, n^1_B, n^1_\Delta)$.)

State	Vote results:	Potential Winners
$\mathcal{EL}$	(1,2,2,0)	A, B
$\mathcal{ML}$	(0,3,2,0)	A

While, if a voter in L deviates and votes for Γ, these vote results become

State	Vote results:	Potential Winners
$\mathcal{EL}$	(2,1,2,0)	Γ, B
$\mathcal{ML}$	(1,2,2,0)	A, B

and state $\mathcal{ML}$ is mistaken for state $\mathcal{EL}$ in posterior beliefs of parties if the deviation by L was not expected.

But one sees that, in state $\mathcal{EL}$ in the second table, the loser becomes a potential winner. To simplify the analysis, let us expand the size of the electorate and add one voter in L and one voter in R. In this case, even when a voter deviates, she cannot make the extremist party a potential winner[8].

In this set-up, there are 7 voters, all of them being able to hide their type. Still, there always are six "core" voters (3 in L and 3 in R) while the seventh voter has a type that is specific to the state of the world, as before.

All non moderate voters have an incentive to vote for a loser. We know that moderate

[8]E.g., the vector of vote results becomes (1,3,3,0) in state $\mathcal{EL}$ without deviation and (2,2,3,0) with deviation.

voters do not have an incentive to do so, but are not yet sure about which, from the extremist or the core voter, has the highest payoff from sending a signal at the expense of an increased risk in the first period. The following lemma answers this question:

Lemma 3 *The tendency to vote for a loser increases in the extremism of the voter. [Proof in the appendix]*

A priori, one could have expected the contrary: by the concavity of the utility function, more extreme voters experience a higher utility gain from electing their preferred main party. But this is actually matched by a larger second-period gain when platforms are attracted towards the extreme. The remaining difference between voters is then their degree of absolute risk aversion, as is directly relates to the risk-premium of losing elections in the first period. As more extremist voters have a lower absolute risk aversion, eL and eR voters are always more prone to vote "protest" than voters in L and R. But, as a consequence:

Proposition 3 *: In anonymous elections, there is no equilibrium with full convergence of platforms. [Proof in the appendix].*

This result comes directly from the standard result where separating equilibria do not exist in such a cheap-talk game. As core voters are always tempted to mimic extremist voters, with full convergence, the only cost that may prevent core voters from always voting for extremist parties is the fear that the latter may actually be elected. In other words, full convergence of parties implies that extremist parties become potential winners. As probabilities of winning must sum up to one, this is equivalent to saying that the probability of winning is less than $\frac{1}{2}$ for each of the opportunistic parties. We show in the appendix that a party increases its probability of winning by proposing more extremist platforms. This ensures that convergence of platforms is never an equilibrium.

Hence, taken jointly, uncertainty about the distribution of voters and anonymity of elections always induce partisan platforms even when parties are non-partisan. In the literature, most of the justifications for platforms divergence involve parties' partisanship (Calvert 1985, Wittman 1977, 1983) or beliefs that explicitly prevent parties from converging (Myerson and

Weber 1993). Proposition 3 shows that when the shape of voters' distribution is not known, non-convergence of platforms is a natural outcome because **voters** are partisan. In the intermediate case, such equilibrium with convergence was possible because core voters did not have the possibility to influence beliefs. Vote secrecy restores this possibility and renders convergence unsustainable.

Proposition 4 *In anonymous elections,*

- If m is small, there exists an equilibrium where parties' platforms are selected in pure strategy. These platforms imply there is voting for losers and parties do not converge in equilibrium.

- If m is too large, parties select platforms in mixed strategies. [Proof in the appendix].

When m is small, parties locate symmetrically around the median such that $-\pi_A^1 = \pi_B^1 = \frac{(1-m)^2 \rho R}{6} > 0$. This ensures that voters in L and R do not vote for losers, but extremist voters do[9]. Note that the distance between platforms is increasing in ρ, the relative importance of the second election. This is because the payoff of garbling information takes place in the second election, which is weighed at rate ρ. One should then observe more extreme platforms in less important elections that precede more important ones.

Proving this result requires to combine the mechanisms used in propositions 2 and 3.

From proposition 3, we know that parties cannot converge in equilibrium. From Proposition 2 and lemmas 2 and 3, there exists a pair of platforms where core (L and R) voters do not vote for losers, unless a party moves towards the center. This is the case in $-\pi_A^1 = \pi_B^1 = \pi^* = \frac{(1-m)^2 \rho R}{6}$, where any deviation towards the center reduces the probability of winning. Any "more extreme" platform is dominated as it allows the competing party to move and increase its probability of winning. Hence, the choice (in pure strategy) of any pair of platforms more extreme than $\pm\pi^*$ is dominated.

Still, parties may have an incentive not to stay in $\pm\pi^*$ if deviating towards more extreme positions can swing the vote of extremist voters while keeping that of the moderate one.

[9] Again, extremist voters are more extreme than core voters. But this does not mean that extremist voters would like the extremist party to be elected. They vote for the extreme party only because this vote carries a signal. For instance, if e=1.1 and $\pi_\Gamma = 3.L$, the voter in eL prefers π_A to π_Γ but votes Γ anyway.

From proposition 2, we know this is the case if m is large enough. As all states of the world are equiprobable, such deviation does not increase the probability of winning strictly[10]. But for any slight difference between the probabilities of each extreme state of the world, $\pm\pi^*$ becomes strictly dominated for one of the parties. In that case, there is no pure strategy equilibrium if m is too large.

The first graph below displays the platforms B (top curve) and A (bottom dashed curve) propose in equilibrium, provided this one exists. The bottom, continuous, curve displays the position A must take in order to swing the vote of eL. The second graph plots the values of the parameters where $-\pi_A^1 = \pi_B^1 = \frac{(1-m)^2\rho R}{6}$ is an equilibrium (white zone) and the values for which there is no equilibrium in pure strategies (black zone).

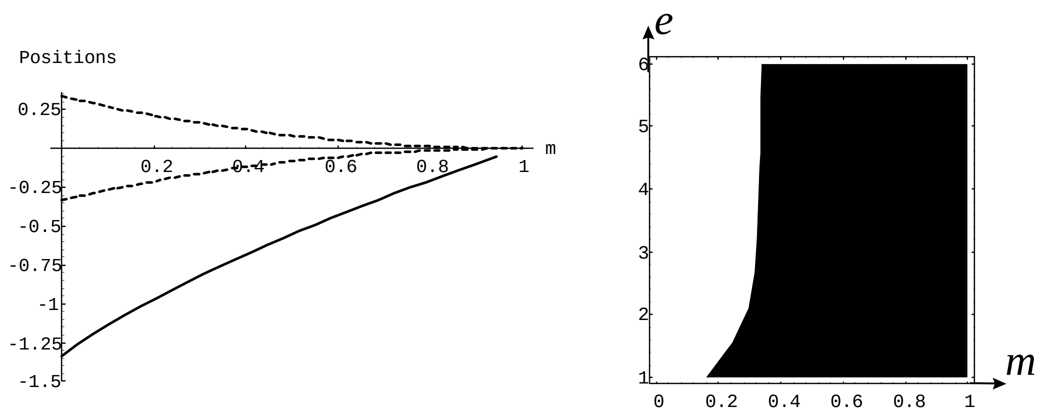

Figure 4: Equilibria in anonymous elections Left graph: dashed lines represent the plaforms of A and B that are a candidate equilibrium, in function of m, the parameter of moderation. The solid line displays the value of π_A^1 that swings the vote of the extreme left voter. Right graph: The white zone represents the combination of parameters that ensures that positions on the dashed lines are an equilibrium. In the black zone, parties deviate outwards (e.g., A proposed the platform represented by the solid line).

This result shows that anonymous (secret ballot) elections make the dispersion of votes on more parties than Duverger's Law would suggest a natural outcome. This is the case because there is a tension for parties between going to the median and creating a cost to vote dispersion. As long as the shape of voters' distribution is unknown, such a tension cannot be completely resolved.

[10] Without deviation: $p(\pi_A^1|\mathcal{EL}) = p(\pi_A^1|\mathcal{ER}) = 0.5$. With deviation, $p(\pi_A^1|\mathcal{EL}) = 1$, but $p(\pi_A^1|\mathcal{ER}) = 0$.

Proposition 4 implies that spreading votes on "too many" parties can be a rational strategy, even for voters who experience no direct payoffs from expressing an opinion (sincere voting) nor desire to elect the small party. Spreading is used as a signalling tool, as voting for a sure loser makes sure that the motivation of such vote is not the election of the party.

Let us make another few remarks before pursuing our analysis. First, we assume non observable votes only because this is a clear feature of secret ballots. Technically, however, private information on the type of all voters is sufficient to drive our results. Second, we also implicitly assumed that extremist parties are so extreme that π^* is always smaller than the position of extremist parties ($\pi^* << |\pi_\Gamma|, \pi_\Delta$). If this were not the case, main parties could become more extreme than extremists, and voters may like to change their coordination in favor of Γ and Δ. Analyzing this properly would nevertheless bring us too far away from the focus of this paper.

5 Comparative Statics

Until now, we have always assumed equiprobable states of the world for simplicity. Running some comparative statics around this simple case nevertheless brings some interesting results.

Let us abstain from specifying the probability of each state of the world. In this case, the expected utility gain of voting for Γ instead of A for a voter in L is

$$(p(\mathcal{EL}) + p(\mathcal{ML}) + p(\mathcal{ER})) . \tfrac{u_1(L,\pi_B^1) - u_1(L,\pi_A^1)}{2} + \rho . p(\mathcal{ML}) \left(u_2(L,L) - u_2(L,mL) \right) \quad (3)$$

which must be negative in equilibrium. By Taylor expansion and substituting for r_a, the degree of absolute risk aversion, this negativity condition becomes

$$\pi_B^1 - \pi_A^1 + \frac{\left(\pi_B^1\right)^2 - \left(\pi_A^1\right)^2}{2R} > \frac{p(\mathcal{ML})}{1 - p(\mathcal{MR})} \rho R (1-m)^2 = \lambda_L . K, \quad (4)$$

where $\lambda_L = \frac{p(\mathcal{ML})}{1-p(\mathcal{MR})}$ and $K = \rho R(1-m)^2$. Computing the same condition for a voter in R obtains

$$\pi_B^1 - \pi_A^1 + \frac{\left(\pi_A^1\right)^2 - \left(\pi_B^1\right)^2}{2R} > \frac{p(\mathcal{MR})}{1 - p(\mathcal{ML})} K = \lambda_R . K, \quad (5)$$

where $\lambda_R = \frac{p(\mathcal{MR})}{1-p(\mathcal{ML})}$. By binding these two conditions, we can find the frontier of indifference for L and R voters. This leads to the following proposition:

Proposition 5 *Probabilities of extreme states do not influence platforms that are candidate equilibria in anonymous elections. However, a movement of the expected median to the* ***left*** *(resp. the right) due to a variation in the probabilities of states $\mathcal{ML}$ and $\mathcal{MR}$, moves to the* ***right*** *(resp. left) the platforms of both A and B in the first election.*
[Proof in the appendix].

A twofold mechanism generates this result. On the side of voters, a higher probability of a moderate state of the world means that the probability that information garbling actually works has increased for core voters. If $p(\mathcal{ML})$ increases and $p(\mathcal{MR})$ decreases, L-voters want to protest more and R-voters want to protest less. This means that, to achieve an equilibrium, the cost of protesting must be higher for L than for R. On the side of parties, the "average voter" has also moved to the right: the left-wing party, A, can count less on the L voters and more on the mL voter. The right wing party counts less on mR and more on R. This altogether explains why, when the median voter moves left in expected terms, parties should move right. This mechanism is displayed in figure 4.

Checking whether the displayed positions in figure 4 are an equilibrium requires point to point calculations, but in the vicinity of $p(\mathcal{ML})$=$p(\mathcal{MR})$, the mechanism of the proofs remains unchanged, and hence the results. Nevertheless, if $p(\mathcal{ML})$ increases relative to $p(\mathcal{MR})$, A wins more often. This means that B would have preferred to be assigned the left-wing identity in the first place. But conditional on A being more left-wing than B, these curves still display the only possible candidates for a pure strategy equilibrium.

6 Extension to a Large Electorate

A lot of research remains to be done in order to fully generalize the model and pinpoint the general conditions on the type of uncertainty that could drive vote for losers to be a "normal" result. As the purpose of this paper is not to characterize such general conditions, we shall restrict ourselves here to an extension of our 6-type-setup to a large population of random

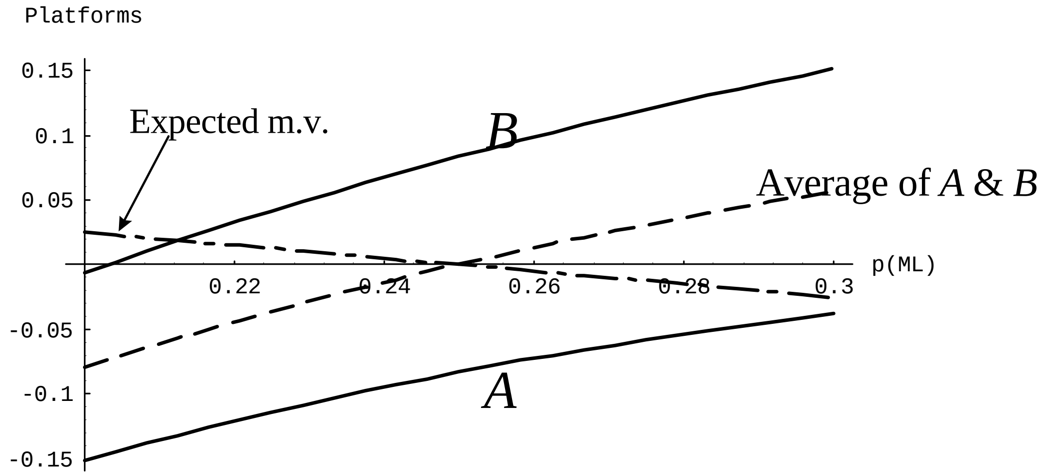

Figure 5: Comparative statics. Comparing party positions for different probabilities of state ML. Top curve: positions for B. Bottom curve: A; dashed line: average of the two positions. The "expected m.v. line represents the position of the expected median voter.

size.

Following Myerson (1994), we want to analyze a Poisson Game[11] where (on expectation) a share $0 < \alpha < 1$ of the population has a known type (these can be core, moderate or extremist voters) and a fraction $(1-\alpha)$ has an uncertain type. Among the α, a fraction q_i has type $i \in \{eL, L, mL, mR, R, eR\}$, with $\sum_i q_i = 1$. Hence, independently from the state of the world, at least a fraction αq_i is expected to have type i. For symmetry purposes, let us assume that $q_L = q_R$, $q_{eL} = q_{eR}$ and $q_{mL} = q_{mR}$.

The type of the remaining fraction $(1-\alpha)$ of voters is contingent on the realized state of the world. That is, they are of type eL in state $\mathcal{EL}$; of type mL in state $\mathcal{ML}$, and so forth, as in the 7-voter case. Each state of the world is equiprobable.

For this extension to be meaningful, the set of parameters must obtain

$$\begin{array}{lll} (1-\alpha) + \alpha\, q_{eL} < .5 & (1-\alpha) + \alpha\,(q_{eL} + q_L) > .5 & \Rightarrow \quad \mu^{\mathcal{EL}} = L \\ \alpha/2 < .5 \text{ (Automatic)} & (1-\alpha) + \alpha\,(q_{eL} + q_L + q_{mL}) > .5 \text{ (Auto.)} & \Rightarrow \quad \mu^{\mathcal{ML}} = mL \\ \alpha\,(q_{eL} + q_L + 2q_{mL}) < .5 & (1-\alpha) + \alpha\,(q_{eL} + q_L + 2q_{mL}) > .5 \text{ (Auto.)} & \Rightarrow \quad \mu^{\mathcal{MR}} = mR \\ \alpha\,(q_{eL} + q_L + 2q_{mL}) < .5 & \alpha\,(q_{eL} + 2q_L + 2q_{mL}) > .5 & \Rightarrow \quad \mu^{\mathcal{ER}} = R \end{array}$$

[11] That is, the aggregate size of the population, $\tilde{n}$ is random and follows a Poisson distribution of parameter, say, n. Each voter is then attributed a type independently from the realization of $\tilde{n}$ and from the type of the other voters. In that case, the number of voters having type i follows a Poisson distribution of parameter $\lambda_i n$ where λ_i is the probability of being of type i. For more details, see Myerson (1994) or Feddersen and Pesendorfer (1996b), or Castanheira (1998).

such that the median voter is different in each state of the world.

In this framework, and considering Poisson games of sufficiently large populations:

Lemma 4 *Extremist states of the world bear the highest pivot probabilities for electing a party.*

Proof. Start from the assumption that there is some protest voting by extremist voters only. Let $1 > \gamma_{eR} = \gamma_{eL} > 0$ be the probability with which they vote for the extremist party. In state $\mathcal{EL}$, the share of votes going to party A and B, $\lambda_A^{\mathcal{EL}}$ and $\lambda_B^{\mathcal{EL}}$ are respectively:

$$\lambda_A^{\mathcal{EL}} = (1-\gamma_{eL})\,(1-\alpha+q_{eL}) + \alpha\,(q_L+q_{mL}) \tag{6}$$

$$\lambda_B^{\mathcal{EL}} = (1-\gamma_{eL})q_{eL} + \alpha\,(q_L+q_{mL}) \tag{7}$$

In state $\mathcal{ML}$, these shares are

$$\lambda_A^{\mathcal{ML}} = (1-\gamma_{eL})q_{eL} + 1-\alpha + \alpha\,(q_L+q_{mL}) \tag{8}$$

$$\lambda_B^{\mathcal{ML}} = (1-\gamma_{eL})q_{eL} + \alpha\,(q_L+q_{mL}) \tag{9}$$

It follows that $\lambda_A^{\mathcal{ML}} > \lambda_A^{\mathcal{EL}} > \lambda_B^{\mathcal{EL}} = \lambda_B^{\mathcal{ML}}$. It is straightforward to see that the same holds for right-wing states of the world: $\lambda_B^{\mathcal{MR}} > \lambda_B^{\mathcal{ER}} > \lambda_A^{\mathcal{ER}} = \lambda_A^{\mathcal{MR}}$. Knowing that the pivot probability of having a tie between main parties is approximately

$$Prob\,(\tilde{n}_A = \tilde{n}_B) \simeq e^{-\left(\sqrt{\lambda_A}-\sqrt{\lambda_B}\right)^2 n}\,\left(2\sqrt{\pi n}\sqrt[4]{\lambda_A\lambda_B}\right)^{-1}, \tag{10}$$

it is clear that the closer the margin of victory, the highest the pivot probability. This proves lemma 4.

■

By the properties of Poisson games, the probability of a state of the world conditional on being pivotal converges to one if this state of the world bears the highest pivot probability and to zero if it does not (Myerson 1994a). Hence, voters will behave only in accordance to the potential result of their vote in extremist states of the world. We can then discard moderate states of the world to compute the equilibrium voting behavior. It remains to check whether the above assumption that only extremist voters cast ballots for losers is always verified. Before doing this, we need another lemma:

Lemma 5 *When extremist voters' (mixed) strategy is non-degenerate, they are the only ones to vote for extremist parties.*

Proof. If extremist voters use a mixed strategy, and using the same notation as in Piketty (1994), we know that

$$p_{com}.\Delta U_{com}^{eL} = p_{dm}.\Delta U_{dm}^{eL}, \tag{11}$$

where p_{com} is the probability of being pivotal for altering beliefs (being "communicative" in the words of Piketty) and p_{dm} is the probability of being "decision-making", that is of being pivotal for electing A or B. ΔU_{com}^{eL} is the utility gain for a voter in eL when she changes beliefs and ΔU_{dm}^{eL} is that of changing the result of the election.

By lemma 3, we immediately deduce that, for L-voters

$$p_{com}.\Delta U_{com}^{L} < p_{dm}.\Delta U_{dm}^{L} \tag{12}$$

as each ballot must have the same probability to influence the vote outcome. Hence core voters vote only for main parties. As moderate voters never want to vote for extremist parties, lemma 5 is proved. ■

It remains to check whether lemmas 4 and 5 are applicable in the equilibrium. This leads to the following proposition:

Proposition 6 *With large populations of Poisson random size, there is always vote for losers in equilibrium, but only by extremist voters.*

Proof. To check this, let us start from a mixed strategy where $\gamma_{eL} \underset{>}{\rightarrow} 0$. By lemma 5, we do not have to care about core voters in such a case. As usual, we look only at left-wing voters, knowing that the same reasoning applies to right-wing voters.

Being communicative consists of increasing the number of votes for Γ, the extreme-left party, to a value that is high enough for players to update their beliefs and think that the state of the world is $\mathcal{EL}$.

Following Piketty (1995), the probability that the number of votes for the extremist party be k in state of the world S is

$$Prob\left(\tilde{n}_{\Gamma}=k|S\right)=\frac{e^{-\lambda_{\Gamma}^{S}n}\left(\lambda_{\Gamma}^{S}n\right)^{k}}{k!} \tag{13}$$

where λ_{Γ}^{S} is the expected share of votes for Γ in S. We know that $\lambda_{\Gamma}^{\mathcal{EL}}>\lambda_{\Gamma}^{\mathcal{ML}}$. There is only one value of k that is equiprobable in both the states of the world $\mathcal{EL}$ and $\mathcal{ML}$. Let this value be $(\lambda_{com}n)$. If the realized number of these votes is higher (lower) than this threshold, state $\mathcal{EL}$ (resp. $\mathcal{ML}$) becomes most likely in the posterior beliefs. As the expected size of the population, n, increases, these posterior beliefs will converge to certainty. Hence, any single voter can be pivotal for beliefs: if the realization of $\tilde{n}_{\Gamma}$ is $\tilde{n}_{\Gamma}=\lfloor\lambda_{com}n\rfloor$, one extra vote changes posterior beliefs from $Prob\left(S=\mathcal{EL}|n_{\Gamma}\right)=0$ to $Prob\left(S=\mathcal{EL}|n_{\Gamma}\right)=1$.

The share realized votes that is equiprobable under the two states of the world is then

$$\lambda_{\Gamma}^{\mathcal{EL}}>\lambda_{com}=\frac{\lambda_{\Gamma}^{\mathcal{EL}}-\lambda_{\Gamma}^{\mathcal{ML}}}{log\left[\lambda_{\Gamma}^{\mathcal{EL}}/\lambda_{\Gamma}^{\mathcal{ML}}\right]}=\frac{\gamma_{eL}\left(1-\alpha\right)}{log\left[1+\frac{1-\alpha}{\alpha q_{el}}\right]}=\phi\,\gamma_{eL}\,\alpha\,q_{eL}>\lambda_{\Gamma}^{\mathcal{ML}} \tag{14}$$

where ϕ is a value larger than 1 that converges to 1 as $\lambda_{\Gamma}^{\mathcal{EL}}\rightarrow\lambda_{\Gamma}^{\mathcal{ML}}$.

From equation (13), and applying Stirling's formula[12], we see that

$$\begin{aligned} plim_{com}=&\lim_{n\rightarrow\infty}\ \frac{1}{n}\,log\left(p_{com}\right)=-\lambda_{\Gamma}^{\mathcal{ML}}+\lambda_{com}\ \left[log\left(\frac{\lambda_{\Gamma}^{\mathcal{ML}}}{\lambda_{com}}\right)+1\right]-\frac{1}{2n}\,log\left(2\pi\,\lambda_{com}n\right)\\ &\underset{n\rightarrow\infty}{\longrightarrow}\left(\phi-1\right)\lambda_{\Gamma}^{\mathcal{ML}}-\lambda_{com}\,log\left(\phi\right) \end{aligned} \tag{15}$$

We can now see how $plim_{com}$ varies with γ_{eL} :

$$\lim_{\gamma_{eL}\rightarrow 0}\ plim_{com}=0$$

Compare it with the pivot probability p_{dm} in state $\mathcal{EL}$:

$$\begin{aligned} plim_{dm}=&\lim_{n\rightarrow\infty}\ \frac{1}{n}\,log\left(p_{dm}\right)=-\left(\sqrt{\lambda_{A}^{\mathcal{EL}}}-\sqrt{\lambda_{B}^{\mathcal{EL}}}\right)^{2}-\frac{1}{2n}\,log\left(2\pi\sqrt{\lambda_{A}^{\mathcal{EL}}\,\lambda_{B}^{\mathcal{EL}}}n\right)\\ &\underset{n\rightarrow\infty}{\longrightarrow}-\left(\sqrt{\lambda_{A}^{\mathcal{EL}}}-\sqrt{\lambda_{B}^{\mathcal{EL}}}\right)^{2} \end{aligned} \tag{16}$$

which negative for any $\lambda_{A}^{\mathcal{EL}}\neq\lambda_{B}^{\mathcal{EL}}$. As $\left(\sqrt{\lambda_{A}^{\mathcal{EL}}}-\sqrt{\lambda_{B}^{\mathcal{EL}}}\right)^{2}$ is maximal in $\gamma_{eL}=0$ and equal to zero when $\gamma_{eL}=1$, $plim_{dm}$ must be strictly negative when $\gamma_{eL}=0$. As the actual probability

[12] $n!\simeq(n/e)^{n}\sqrt{2\pi n+\pi/3}$

is $e^{plim_{dm} \cdot n}$, we know that, for $\gamma_{eL} \to 0$

$$\frac{p_{com}}{p_{dm}} \underset{n \to \infty}{\to} \infty$$

which ensures that voters in eL and L always want to vote for Γ and $\gamma_{eL} > 0$ in equilibrium (unless the distance between parties' platforms is infinite). This proves that all equilibria involve a positive amount of voting for losers by extremist voters.

By computing $plim_{com}$ and $plim_{dm}$ for $\gamma_{eL} \to 1$, one can easily check that all equilibrium strategies are non degenerate, that is $\gamma_{eL} < 1$ in equilibrium.

■

This shows that all the results obtained in previous sections generalize to more complicated cases of uncertainty and large populations, at least on the side of voters' behavior. One could have thought that vote for losers becomes worthless in larger electorates, but the point is precisely that voting for main parties also has less effect when population size increases. If there were no vote for losers at all (but votes still interpreted as signals), the probability of being pivotal for the election of one of the main parties would fall very fast, while that of being "communicative" would have remained high. This ensures that some vote for losers must take place in equilibrium. Of course, there always exists an (unstable) equilibrium where there is no vote for losers because signals are not taken into account at all, as in Piketty (1995). But it is also clear that with the slightest probability that a signal may be interpreted, this equilibrium is not sustainable.

A direct corollary from the last proposition will also ensure that our results concerning parties' behavior also generalize to large electorates:

Corollary 1 *The closer main parties' platforms, the more weight is put on protest voting.*

Proof. From equation (11), in equilibrium,

$$\frac{p_{com}}{p_{dm}} = \frac{\left(\pi_B^1 - \pi_A^1\right)/\rho}{(1-m)\,R - r_a\,(eL)\,\frac{(1-m^2)R^2}{2}} = \sqrt{\frac{2\sqrt{\lambda_A^{\mathcal{EL}}\,\lambda_B^{\mathcal{EL}}}}{\phi\gamma_{eL}\,\alpha\,q_{eL}}} \tag{17}$$

which implies that, if $\left(\pi_B^1 - \pi_A^1\right)$ decreases, λ_A^S and λ_B^S must increase (remember that $\lambda_A^{\mathcal{EL}}\,\lambda_B^{\mathcal{EL}}$ also decreases in γ_{eL}). ■

And hence:

Proposition 7 *In equilibrium, main parties' platforms always diverge sufficiently to render extremist parties sure losers.*

Proof. If platforms converge fully, $p_{com}.\Delta U_{com}^{j} > p_{dm}.\Delta U_{dm}^{j} = p_{dm}.0$, $\forall j \in \{eL, L, R, eR\}$ and all voters want to communicate, until "losers" become potential winners. Hence the probability of victory of main parties must have decreased. An ε outward-deviation ($\varepsilon > 0$, $\varepsilon \to 0$) must then increase the probability of winning. Unfolding this reasoning, any candidate equilibrium must ensure that $p(\pi_A^1) + p(\pi_B^1) = 1$. ■

From these characteristic of the equilibrium, some other results follow:

Corollary 2 *The state of the world is learnt with a probability converging to one as the size of the population goes to infinity (similar to Piketty 1995)*

Proof. Straightforward from the fact that voting shares are different in each state of the world. ■

Corollary 3 *The main party favored by a majority of the population is elected with a probability converging to one as population size increases to infinity, but with a margin of victory that is lower than without protest voting (similar to Piketty 1995).*

Proof. Straightforward from the computation of voting shares in equilibrium. ■

Note that, as extremists are the voters that carry information, heterogeneity of preferences increases the share of protest voting and hence the aggregate efficiency of elections, in contrast to Feddersen and Pesendorfer (1997) and Caillaud and Tirole (1997). This ensures both the efficiency of information aggregation and the election of the median voter's most favored party in the first election as well as the implementation of her bliss policy in the second election.

7 Conclusion

This paper has shown in a simple framework that far sighted voters can use elections as a signalling device when the distribution of voters is not perfectly known. Such signalling requires extremist voters to vote for extremist parties even though the latter are clear losers. This implies that, in equilibrium, votes will be dispersed across more parties than standard theory predicts. In other words, signalling in equilibrium implies the presence of "protest" voting and the violation of Duverger's Law.

In reaction to this phenomenon of vote dispersion, purely opportunistic parties will not converge fully and will propose somewhat extremist platforms. Platform divergence is required to prevent extremist parties from becoming potential winners in these elections. If platforms are too close to each other, voting for losers has too low a cost: even if the disliked party is elected, it will anyway implement a similar policy. In such a case, only the threat of seeing extremists being elected prevents voters from voting more and more for extremist parties.

When there is enough distance between platforms, losing elections is perceived as a substantial cost for voters. In equilibrium, parties will locate so as to prevent their core electorate from voting for losers, but will never try to capture the vote of all extremists, unless more information is available (e.g. in "raising hand votes"). This shows that purely opportunistic parties should not converge fully in equilibrium. The informational content of elections more than counterbalances the forces driving the median voter theorem, which must be violated in equilibrium.

The driving force behind these interactions is that, as in all games with incomplete information, information elicited in a first election will be rationally used by parties to relocate efficiently in the next election. If there is uncertainty about the distribution of voters, extra pieces of information allow parties to relocate more accurately. If the exact *shape* of the distribution is unknown, information about the tails of the distribution is relevant and motivates "protest" voting. Hence, the more important the next election –in the mind of voters–, the more parties will have to diverge in equilibrium to prevent excessive voting for losers.

This has important implications with respect to the interpretation of protest voting. Some democratic parties obviously feel puzzled by the size of such a phenomenon. Faced with a large share of votes for the extreme right, parties react by warning voters of the dangers of a fascist come back. It is likely that some of these voters want extreme right parties to be elected and should be warned. But it is also likely that extremist voting is also used as a credible signal to warn main parties. Some people who only want the latter to cure themselves from corruption and/or too abstract politics would stop voting for extremists as soon as they become "too" important. This also explains there is no "typical" extremist voter. All national front voters are not pure fascists. Another question is to know the real proportion of those.

This paper does not address the reasons why extremist parties form. However some clues can be inferred from our analysis. We assumed that extremist parties where a priori extremist. This assumption has no other role than making clear that voting for one such party carries an extremist signal. If one takes the thread of this argument from the other end, we can deduce that forming a new party that has a clear ideological content, mainly if it is non-median and points to clear directions for new policies, may attract more votes than forming a middle-of-the-road party. This points to two potential rationales for forming such a party: either a group of voters is not interested in winning elections, but forms a party in order to influence main parties; or a group of opportunistic politicians may think this is an easier way to become famous and moderate one's positions afterwards. Some "extreme" parties like the Greens in Germany succeeded becoming "important" in this way. The French experience also shows that main parties may be induced to react swiftly enough to try and eat up the vote shares of these extremists.

These conjectures also stress some of the loopholes of this paper, as well as interesting paths for future research: there is a need for making this analysis more dynamic and endogenizing extremist parties' behavior to see how small parties can become big. Another alley for future research is to extend the analysis to more general distributions of voters, in order to allow more direct comparisons of its predictions with reality. We also assumed that voter's preferred policy is given; a strong concern is that extremist parties' opinions may contaminate the electorate. If this is the case, voters' bliss points should also be made partly endogenous.

References

[1] Alesina, Alberto and Howard Rosenthal (1995) "Partisan Politics, Divided Government, and the Economy", Cambridge University Press.

[2] Banerjee, Abhijit and Rohini Somanathan (1997) "A simple Model of Voice" Mimeo, May

[3] Black, D. (1958) "The Theory of Committees and Elections", Cambridge University Press

[4] Caillaud, Bernard and Jean Tirole (1997) "Voters homogeneity and aggregation of information" Mimeo, May

[5] Calvert, R. (1985) "Robustness of the Multidimensional Voting Model: Candidates's Motivations, Uncertainty and Convergence", American Political Science Review, Vol 29, No1

[6] Cox, Gary (1994) "Strategic Voting Equilibria Under the Single Non-Transferable Vote", American Political Science Review, Vol 88, No3, Sept. 1994

[7] Cox, Gary W. (1997) "Making Votes Count" Cambridge University Press.

[8] Delwit, Pascal (1995) "Des profils incertains", in Le Paige (1995), pp81-102.

[9] Downs (1957) "An Economic theory of Democracy". Harper and Row

[10] Duverger, Maurice (1954) "Political Parties", Wiley, New York

[11] Feddersen and Pesendorfer (1997) "Voting Behavior And Information Aggregation in Elections with Private Information", Econometrica, Vol. 65, No5 and (1994) Northwestern University Discussion Paper No1117

[12] Feddersen and Pesendorfer (1996a) "The Swing Voter's Curse", American Economic Review

[13] Feddersen and Pesendorfer (1996b) "Abstention and Common Values", Northwestern University Mimeo

[14] Fudenberg, Drew and Jean Tirole (1991) "Game Theory" MIT Press

[15] Hotelling, H. (1929) "Stability in Competition", Economic Journal, vol 39

[16] Gibbons, Robert (1992) "A Primer in Game Theory", Harvester Wheatsheaf

[17] Laffont, Jean-Jacques (1989) "The economics of Uncertainty and Information", MIT

Press

[18] Le Paige, Hugues et al. (1995) "Le Désarroi Démocratique. L'extrême droite en Belgique.", Editions Labor, Brussels.

[19] Morton, Rebecca (1993) "Incomplete Information and Ideological Explanation of Platform Divergence", American Political Science Review, Vol 87, No2, June

[20] Mueller, Dennis (1989) "Public Choice II", Cambridge University Press

[21] Myerson, Roger (1994a) "Population Uncertainty and Poisson Games", Northwestern University Discussion Paper No1102

[22] Myerson, Roger (1994b) "Extended Poisson Games and the Condorcet Jury Theorem", Northwestern University Discussion Paper No1103

[23] Myerson, Roger and Robert Weber (1993) "A theory of voting equilibria" American Political Science Review, Vol 87, N01, March

[24] Palfrey, Thomas (1984) "Spatial Equilibrium with Entry", Review of Economic Studies, pp 139-156

[25] Piketty, Thomas (1995) "Voting as Communicating", mimeo

[26] Wittman, D. (1977) "Candidates with policy preferences: A dynamic Model", Journal of economic theory, Vol 14.

[27] Wittman (1983) "Candidate Motivation: A Synthesis of Alternatives" American Political Science Review, vol 77 pp142-157.

8 Technical Appendix

Computation of the value of a vote for losers ($d\mu^S$ and $d^2\mu^S$).

We only look at eL-voters. By symmetry, the argument applies immediately to eR-voters. eL, by voting for Γ sends a signal different from that of any other voter. By the assumptions on possible signals, this leads to posterior beliefs $E\left[\mu^S|v_{eL}=\Gamma\right]=L$. This defines posterior beliefs in case eL votes for Γ. But the *value* of this vote depends on the alternative beliefs, in case eL were voting for A.

If eL is **expected to vote for** Γ, but deviates and votes for A, $E\left[\mu^S|v_{eL}=A\right]=mL$. In this case, the second period platform differential is $d\mu^S=(1-m)L$. The second order effect is $d^2\mu^S=L^2-(mL)^2=\left(1-m^2\right)L^2$.

If eL **and** mL **are expected to vote for** A, we have a semi-pooling equilibrium: $E\left[\mu^S|v_z=A\right]=\frac{1+m}{2}L$, $z\in\{eL,mL\}$. This implies $d\mu^S=L-\frac{1+m}{2}L=\frac{1-m}{2}L$ and $d^2\mu^S=\left(1-\left(\frac{1+m}{2}\right)^2\right)L^2$.

Proof of Lemma 2:

Second-period platforms are the expected position of the median voter. Hence voter i's ex-ante payoff of casting a ballot v_i is

$$E\left[U\left(\theta_i, x_1, x_2\right)|v_i\right] = -E\left[\left(\theta_i - x_1\right)^2 |v_i\right] - \rho\ \left(\theta_i - E\left[\mu^S|I^1\right]\right)^2 \tag{18}$$

where I^1 is the information set obtained after the first election. With $\pi_B^1 = -\pi_A^1 > 0$, if eL does not cast a ballot, $n_A^1 = n_B^1 = 2$ and $p\left(\pi_A^1\right) = p\left(\pi_B^1\right) = \frac{1}{2}$. eL may consider voting for A or for Γ (showing that voting for B or Δ is a dominated strategy given the assumed beliefs is easy). If $v_z^1 = A$, $n_A^1 > n_B^1$ and $x_1 = \pi_A^1$ for sure. If $v_z^1 = \Gamma$, the tie between A and B is maintained. Taking the difference between the payoffs of these two strategies, when eL is expected to vote for Γ, and taking second order Taylor expansion in zero obtains —As $u_1(.,.)$ and $u_2(.,.)$ have the same shape, $\frac{\partial u_1(\theta_i,0)}{\partial x_1} = \frac{\partial u_2(\theta_i,0)}{\partial x_2}$. We denote these partial derivatives $u'(\theta_i, 0)$—:

$$\begin{aligned} & E\left[U\left(\theta_{eL}, x_1, x_2\right)|v_{eL}^1 = \Gamma\right] - E\left[U\left(\theta_{eL}, x_1, x_2\right)|v_{eL}^1 = A\right] = \qquad (19) \\ & u'(\theta_{eL}, 0).\left[\frac{\pi_B^1 - \pi_A^1}{2} + r_a(\theta_{eL}).\frac{\left(\pi_B^1\right)^2 - \left(\pi_A^1\right)^2}{4} + \rho\ \left(d\mu^S + r_a(\theta_{eL}).\frac{d^2\mu^S}{2}\right)\right] \\ \propto & -\frac{\pi_B^1 - \pi_A^1}{2} + \rho\ \left(\left|d\mu^S\right| - r_a(\theta_{eL}).\frac{d^2\mu^S}{2}\right) \qquad (20) \end{aligned}$$

where $\propto$ stands for "proportional to" and $r_a(\theta_{eL}) = \left|\frac{u_1''(\theta_{eL},0)}{u_1'(\theta_{eL},0)}\right| = \left|\frac{1}{\theta_{eL}}\right|$ denotes the degree of absolute risk aversion in 0 (Notation borrowed from Laffont (1989), but with an important difference: $u' < 0$ for $\theta_z < 0$). With quadratic utility, Taylor expansion is not an approximation, but allows to rewrite (19) differently.

If (20) is positive, eL votes for Γ. Isolating first-period platforms, we get

$$-\pi_A^1 = \pi_B^1 < \rho\left(-d\mu^S - r_a(\theta_{eL}).\frac{d^2\mu^S}{2}\right) \tag{21}$$

as a condition for a voter in θ_{eL} to vote for Γ. Hence, the larger the distance between first period platforms, the lower the surplus (the larger the loss) of protest voting.

It is straightforward to show that this mechanism extends to L-voters if the latter have private information on their type or to right-wing voters, with $\theta_z > mR$, if one compares the value of voting for B and Δ. This proves lemma 2. ■

Proof of proposition 2.

By lemma 2, there are two possible types of equilibria: either parties diverge sufficiently to "kill" vote for losers (VfL from now on), or they don't. Let us analyze the two types of candidate equilibria sequentially.

Case 1: No VfL. If there is no VfL, this means that

$$E\left[U(eL, x_1, x_2)|v_{eL}^1 = \Gamma\right] < E\left[U(eL, x_1, x_2)|v_{eL}^1 = A\right] \tag{22}$$

$$E\left[U(eR, x_1, x_2)|v_{eR}^1 = \Delta\right] < E\left[U(eR, x_1, x_2)|v_{eR}^1 = B\right] \tag{23}$$

and extremist voters are *expected* to vote for main parties. By our computation of $d\mu^S$ and $d^2\mu^S$ and by lemma 2, this requires

$$\frac{\pi_B^1 - \pi_A^1}{2} + r_a(eL).\frac{\left(\pi_B^1\right)^2 - \left(\pi_A^1\right)^2}{4} > -\frac{1-m}{2}\rho\left(1 - r_a(eL).\frac{3+m}{4}L\right)L$$

Simplifying this obtains

$$-\pi_A^1 = \pi_B^1 > \phi/2, \tag{24}$$

where $\phi = \rho(1-m)R\left(1 - \frac{3+m}{4e}\right)$.

It remains to compute when such platforms are an equilibrium. Clearly neither A nor B have an incentive to deviate outwards. They do have an incentive to deviate inwards if they can add the moderate voter of the other camp to their electorate without losing their own moderate voter. Looking at A, mR votes for A when $\left|\pi_A^1 - mR\right| < \left|\pi_B^1 - mR\right|$ and

$$\pi_B^1 - \pi_A^1 - \frac{\left(\pi_B^1\right)^2 - \left(\pi_A^1\right)^2}{2mR} < -\rho\frac{mR}{2} \tag{25}$$

Similarly, mL votes for A as long as $\left|\pi_A^1 - mL\right| < \left|\pi_B^1 - mL\right|$ and

$$\pi_B^1 - \pi_A^1 + \frac{\left(\pi_B^1\right)^2 - \left(\pi_A^1\right)^2}{2mR} > -\rho\frac{mR}{2} \tag{26}$$

As the minimum of the left hand side in (25) is reached in $\pi_A^1 = mR$, a necessary and sufficient condition for $(-\phi/2, \phi/2)$ to be a Nash equilibrium is either that $\phi/2 < mR$ or that equation (25) is violated in $\pi_A^1 = mR$ (Note that (25) is a tighter constraint than (26)). By equation (25), if $\phi/2 > \left(1 + \sqrt{\rho}\right) mR$, both parties have an incentive to deviate towards the median. This is the case when

$$2m < \frac{\phi/R}{1+\sqrt{\rho}} \tag{27}$$

that is, when m is sufficiently low.

Case 2: VfL and convergence. From lemma 2, eL votes for Γ as long as

$$\left|\pi_A^1\right| - \frac{\left(\pi_A^1\right)^2}{2eR} < -\left(\pi_B^1 + \frac{\left(\pi_B^1\right)^2}{2eR}\right) + \rho\left(2\left|d\mu^S\right| - \frac{d^2\mu^S}{eR}\right) \tag{28}$$

Hence, B renders the deviation hardest by positioning in 0. From there, to swing the vote of eL, A must select a platform π_A^1 such that

$$\pi_A^1 + \frac{\left(\pi_A^1\right)^2}{2eR} \leq -\rho R\left(2(1-m) - \frac{1-m^2}{e}\right) \tag{29}$$

i.e. a platform such that 28 be violated.

A necessary but not sufficient condition for $(0,0)$ to be an equilibrium is that no value of π_A^1 allows to satisfy this inequality. Indeed, eR votes for B when $\pi_A^1 - \frac{\left(\pi_A^1\right)^2}{2eR} < \rho L\left(2(1-m) - \frac{1-m^2}{e}\right)$, which holds automatically if condition (29) does. (29) is then a worthwhile deviation only if mL does not swing to B. If mL does not swing, probability of victory does not change. Any preference for a higher average number of votes, or any change in the relative probabilities of $\mathcal{EL}$ and $\mathcal{ER}$ makes the deviation strictly worthwhile for at least one of the parties. If mL does swing, the probability of victory falls, which ensures that $(0,0)$ is strictly an equilibrium.

The condition for $(0,0)$ to be a strong Nash positional equilibrium is then that the conditions such that eL and mL vote for A never hold together. When A deviates, it selects $\pi_A^{1*} = -eR + R\sqrt{e^2 + 2\rho(1-m)(1+m-2e)}$, provided it is real. mL votes for A when $2mR\pi_A^1 + \left(\pi_A^1\right)^2 < \rho f(m) R^2$, where $f(0) = 0$ and $f'(m) > 0$. One sees this condition is violated for very small $\left|\pi_A^1\right|$ when m is close to zero. As m closer to zero also implies that π_A^{1*} is more negative, $(0,0)$ is a Nash positional equilibrium when m is small enough.

To conclude, when m is small (sufficiently close to zero), there is no equilibrium with divergence, but there is one with convergence. When m is large (sufficiently close to one), there is no (strictly preferred) equilibrium with convergence, but there is one with divergence. By virtue of continuity, a pair of platforms in the vicinity of those proposed in the proof may also be an equilibrium, but the type of sustainable equilibrium remains unchanged. ■

Proof of lemma 3: From lemma 2, it is enough to rewrite (20) to see that a voter in $z < mL$ votes for Γ as soon as

$$r_a(\theta_z) < \frac{1}{d^2\mu^S}\left(-\frac{\pi_B^1 - \pi_A^1}{\rho} + 2\,d\mu^S\right). \tag{30}$$

As $r_a(\theta_z) = |\theta_z|^{-1}$, the more extremist the voter, the more easily the condition is met. ■

Proof of proposition 3.

Let us start the proof of this proposition with the following lemma:

Lemma 6 *With full convergence of main parties' platforms, there cannot exist sure losers.*

Proof. From lemma 2, with full convergence of platforms and as long as extremist parties have no chance of winning the elections, there is no cost to voting for losers, while there is an informational benefit. Hence, voters in L and R will vote for losers. But this implies that extremist parties become potential winners, inducing a contradiction. Hence, L and R must be voting in mixed strategy when A and B fully converge, that is

$$p_{A,B}\,\Delta U_{A,B}^L + p_{A,\Delta}\,\Delta U_{A,\Delta}^L + p_{A,\Gamma}\,\Delta U_{A,\Gamma}^L + p_{\Gamma,\Delta}\,\Delta U_{\Gamma,\Delta}^L = p_{com}\,\Delta U_{com}^L \tag{31}$$

$$p_{B,A}\,\Delta U_{B,A}^R + p_{B,\Gamma}\,\Delta U_{B,\Gamma}^R + p_{B,\Delta}\,\Delta U_{B,\Delta}^R + p_{\Delta,\Gamma}\,\Delta U_{\Delta,\Gamma}^R = p_{com}\,\Delta U_{com}^R \tag{32}$$

where $p_{j,k}$ is the *differential* of pivot probabilities between party j and k when voting "centrist" instead of "extremist"; $p_{com}\,\Delta U_{com}^z$ is the expected utility of altering beliefs for the second elections and ΔU_{event}^z is the utility differential for the voter z of being pivotal for *event* (communication or election).

With full convergence, $\Delta U_{A,B}^L = \Delta U_{B,A}^R = 0 < p_{com}\,\Delta U_{com}^L$ or $p_{com}\,\Delta U_{com}^R$. Hence, extremist parties may be enabled to win elections for the equalities to be restored (in which case $p_{A,\Delta}$, $p_{A,\Gamma}$, $p_{\Gamma,\Delta}$, $p_{B,\Gamma}$, $p_{B,\Delta}$ and $p_{\Delta,\Gamma}$ become strictly positive). ■

Proof of proposition 3: PARTIES. From lemma 6, an equilibrium with full convergence would give to main parties a lower probability of victory than another (hypothetical) equilibrium with enough divergence. The question remains to check whether A or B can increase his probability of winning by deviating away from the median.

For sufficiently extreme losers ($\Delta U^L_{A,\Gamma}$ and $\Delta U^R_{B,\Delta}$ sufficiently big), we can be sure that even with full convergence, A and B win more often than Γ and Δ (and hence we can maintain our primary assumption that main parties remain more focal than extremist parties). If a party (say A) deviates to the left by ε (very small), $\Delta U^L_{A,B}$ and $\Delta U^R_{B,A}$ increase, necessitating a new mixture by voters to restore the two above equalities. As
a) left-wing voters can influence the relative probability of A or Γ being elected against B or Δ, but cannot influence the relative probability of having B or Δ.
b) symmetrically, voters in R must take as given the distribution of votes for A and Γ but can influence that of B and Δ,
we can deduce that the deviation by A implies L and R-voters must for sure vote more for A and B respectively. With a small enough deviation[13], we ensure that $p_{com}\,\Delta U_{com}$ does not change and that the effect is symmetric for L and R-voters. Hence, a small deviation towards outside must increase the probability of being elected for both A and B and will be implemented. It is easy to see that, insofar as we look at symmetric cases, this argument applies to any pair of platforms that is not sufficiently divergent, that is to any pair where extremists are potential winners. Hence, any stable equilibrium must have divergence and extremist parties being sure losers. ■

Illustration of proposition 3: a particular case

If $\rho\,(1-m)^2 \leq \frac{2}{3}$, the optimal deviation from (0,0) is simple to compute: A voter in L votes for A in pure strategy if

$$\left|\pi^1_A\right| - \frac{\left(\pi^1_A\right)^2}{2R} > \frac{\rho R}{3}\,(1-m)^2 \tag{33}$$

Solving the symmetric condition for voters in R obtains

$$\left|\pi^1_A\right| + \frac{\left(\pi^1_A\right)^2}{2R} > \frac{\rho R}{3}\,(1-m)^2 \tag{34}$$

Condition 33 is tighter than 34. This implies that when π^1_A swings L-wing voters to A (who then vote for A in pure strategy), right-wing voters are automatically swung to B. This, in turn ensures that extremist parties have indeed become pure losers. This necessarily increases the joint probability of victory of the two main parties. If the voter in mL starts to vote for B because of the deviation from 0, the result is not altered. To the contrary, L-voters are totally disciplined as they cannot alter the information set in state $\mathcal{ML}$, while R-voters experience a higher pay-off from garbling information (the utility differential has increased and they can garble information in both states $\mathcal{ML}$ and $\mathcal{MR}$). This implies that the new condition for R becomes

$$\left|\pi^1_A\right| + \frac{\left(\pi^1_A\right)^2}{2R} > \rho R. \tag{35}$$

So, when π^1_A is chosen so as not to bind (35) while mL votes for B, A increases his probability of winning (All L voters vote for A, while R voters coordinate so as to have one

[13] For a larger deviation, there is a change in the distribution of votes which can affect both types of voters dissymmetrically. We illustrate this in the particular case below.

vote who votes for Δ in pure strategy and the other two vote for B in pure strategy to avoid that Δ becomes a potential winner).

Proof of proposition 4.

L votes for A and R votes for B respectively if

$$3\left(\pi_B^1-\pi_A^1+\frac{\left(\pi_B^1\right)^2-\left(\pi_A^1\right)^2}{2R}\right)-\rho(1-m)^2R>0 \tag{36}$$

$$\text{and}\quad 3\left(\pi_B^1-\pi_A^1-\frac{\left(\pi_B^1\right)^2-\left(\pi_A^1\right)^2}{2R}\right)-\rho(1-m)^2R>0 \tag{37}$$

This means that, starting from symmetric positions,

$$\frac{dCond^nL}{d\pi_A^1}=-1-\frac{\pi_A^1}{R}=-1+\left|\frac{\pi_A^1}{R}\right| \quad (>-1) \tag{38}$$

$$\frac{dCond^nR}{d\pi_A^1}=-1+\frac{\pi_A^1}{R}=-1-\left|\frac{\pi_A^1}{eR}\right| \quad (<-1) \tag{39}$$

implying that from any pair of positions $\pi_A^1=-\pi_B^1$, satisfying (36) and (37) (that is, the *strict* inequalities), a party can increase his probability of winning by moving towards the center. Indeed, this increases the VfL at a higher rate for voters of the opposite party. The only case where it does not work is when (36) and (37) are replaced by their equivalent equalities. That is, the only candidate equilibrium is

$$\pi^*=-\pi_A^1=\pi_B^1=(1-m)^2\rho\frac{R}{6} \tag{40}$$

Can dissymmetric positions be an equilibrium ? Certainly not, as the two conditions for L and R then cannot bind together, inducing the move inwards by one of the parties.

To check whether $\pm\pi^*$ is an equilibrium, one needs to see under which condition parties do not want to move outwards. By the same argument as in proposition 2, this is met when eL and mL never vote for A at the same time. From proposition 2, we know this is the case for m sufficiently small. If m is not small enough, this pair of platforms is not an equilibrium. But for any other pair of positions, parties want to move inwards to as to swing the vote of the opponent's voters. It then comes out that the equilibrium will necessarily be in mixed strategies if m is not sufficiently small. ■

Proof of proposition 5.

To begin with, let us simply check how (40) changes when the probabilities of the states of the world are modified. The computation for L is in the body of the text. For R,

$$\begin{aligned}
&p(\mathcal{ER}).\left(u_1(R,\pi_A^1)-\frac{u_1(R,\pi_B^1)+u_1(R,\pi_A^1)}{2}\right)+\\
&p(\mathcal{MR}).\left(\frac{u_1(R,\pi_B^1)+u_1(R,\pi_A^1)}{2}-u_1(R,\pi_B^1)+\rho.\left[u_2(R,R)-u_2(R,mR)\right]\right)+\\
&p(\mathcal{ML}).\ 0+p(\mathcal{EL})\left(u_1(R,\pi_A^1)-\frac{u_1(R,\pi_B^1)+u_1(R,\pi_A^1)}{2}\right)<0
\end{aligned} \tag{41}$$

Simplifying this obtains equation (5). By binding conditions (4) and (5), and summing these gives the required distance between the two parties to avoid that protest vote is cheap talk:

$$\pi_B^1 - \pi_A^1 = \frac{(\lambda_L + \lambda_R)}{2} K. \tag{42}$$

Taking the difference between these two equations allows to compute the average of the positions of A and B:

$$\begin{aligned} \left(\pi_B^1\right)^2 - \left(\pi_A^1\right)^2 &= (B + \pi_A^1)(B - \pi_A^1) = (\lambda_L - \lambda_R).K.R \\ \frac{\pi_B^1 + \pi_A^1}{2} &= \frac{(\lambda_L - \lambda_R).K.R}{2(\pi_B^1 - \pi_A^1)} = \frac{\lambda_L - \lambda_R}{\lambda_L + \lambda_R} R \end{aligned} \tag{43}$$

This implies that with $p(\mathcal{ML}) = p(\mathcal{MR})$, the average of the two positions is zero, as in the text. If $p(\mathcal{ML}) > p(\mathcal{MR})$, the average of the positions becomes positive. That is, if the state of the world is more likely to be left-wing than right-wing, parties need to move *right* to prevent protest voting by typical left-wing and right-wing voters.

Increasing the value of R, and maintaining K constant, increases the amplitude of the curves' movement (the absolute value of the average increases) without influencing the distance between the two parties. Increasing the value of K without changing R (by increasing ρ or decreasing m) increases the required distance between the two parties, without modifying the average of the two positions.

This exercise gives us parties positions that make L and R voters indifferent. But it is important to note that, now, the behavior of moderate voters is not granted anymore. If parties move too much e.g. to the right, even the moderate right voter could prefer to vote for A in the first election: mL votes A iff

$$E(U(mL, x_1, x_2) | v_{mL}^1 = A) > E(U(mL, x_1, x_2) | v_{mL}^1 = B) \tag{44}$$

and mR votes B iff

$$E(U(mR, x_1, x_2) | v_{mR}^1 = B) > E(U(mR, x_1, x_2) | v_{mR}^1 = A). \tag{45}$$

Expanding these conditions:

$$\begin{aligned} u_1(mL, \pi_A^1) + \rho u_2(mL, mL) &> u_1(mL, \pi_B^1) + \rho u_2(mL, 0) \\ \pi_A^1 - \pi_B^1 + r_a(mL)\frac{\left(\pi_A^1\right)^2 - \left(\pi_B^1\right)^2}{2} &< -\rho\left(mL + r_a(mL)\frac{(mL)^2}{2}\right) \\ -(\lambda_L + \lambda_R) - \frac{(\lambda_L - \lambda_R)}{m} &< -\rho\frac{mL}{K} = \rho\frac{mR}{K} \quad (>0) \end{aligned} \tag{46}$$

while mR votes B when

$$\pi_B^1 - \pi_A^1 - \frac{\left(\pi_B^1\right)^2 - \left(\pi_A^1\right)^2}{2mR} > -\rho\frac{mR}{2} \tag{47}$$

$$-(\lambda_L + \lambda_R) + \frac{(\lambda_L - \lambda_R)}{m} < \rho\frac{mR}{K} \tag{48}$$

So, λ_L and λ_R must satisfy these conditions for the threshold values of π_A^1 and π_B^1 computed above to hold. When λ_L and λ_R do not satisfy these inequalities, other thresholds must be computed, but the mechanism remains the same.

This proves proposition 5 $\blacksquare$.