Political Advocacy with Collective Decision Making ⋆

Sophie Bade and Andrew Rice

Penn State University[1]

Abstract

The model presented in this paper captures some of the effects of a pre-electoral debate on the incentives for information acquisition of voters that belong to different ideological strands. We introduce the option to publicly share information into a fairly standard model of information aggregation through an election with costly information acquisition. We find that this option dramatically changes the incentives to acquire information. Without the option to share one's signal no extremist has any incentive to acquire information. With this option present the extremists' incentive to acquire information is even stronger than the independents' incentive. In equilibrium this extra incentive leads the extremists acquire more information than the independents. We use this to explain the empirically observed correlation between extremism and information.

JEL Classification: D72; D82; D83

Key words: Information acquisition; Debates; Persuasion; Condorcet; Extremism

⋆ We thank Kalyan Chatterjee, Ed Green, Yao Huang, Jim Jordan, Vijay Krishna, Joris Pinkse, and seminar participants at the economic theory seminar at Cal Tech, the political science seminar at NYU, and the political economy seminar at Northwestern (MEDS) for their insightful comments. Sophie Bade wishes to thank Elmar Wolfstetter and Deutsche Forschungsgemeinschaft (DFG), SFB Transregio 15, "Governance and Efficiency of Economic Systems.", TP A7 for their hospitality while working on this project. This paper was originally titled "Pre-Electoral Debate: The Case of a Large Election".

[1] Department of Economics, Penn State University, 518 Kern Graduate Building, University Park, PA 16802-3306, USA. Phone: (814) 865 8871, Fax: (814) 863 - 4775 `sub18@psu.edu` and `arice@psu.edu`

1 Introduction

In most economic models of elections voters act as if they are in isolation; they might obtain some possibly costly information about the candidates, they digest this information, and then vote. Typically it is assumed that all information is private and is only aggregated through the vote. In this context a voter that has extreme preferences about the outcome of an election may never want to buy any information since it will have little or no influence on their voting decision. This can lead to a prediction that more extreme voters will be less informed.

In contrast, it turns out that there is strong evidence that voters with more extreme views actually tend to be better informed. This empirical result has been shown in "The Relationship between Information, Ideology, and Voting Behavior" by Palfrey and Poole (1987) as well as a host of papers in the literature on political psychology (see for example Sidanius (1988) and the references cited therein). In the model presented here we attempt to explain this correlation by allowing voters to endogenously acquire costly information and then partake in a pre-electoral debate.

for the most part, pre-electoral debates have been overlooked in the political economics literature; yet it seems that they play an important role in the democratic process. Debates and discussions about the running candidates are among the most common features of political life as we know it. In fact, according to The American National Election Studies (ANES) poll in 2004, 34% of respondents reported yes to the question, "During the campaign, did you talk to any people and try to show them why they should vote for or against one of the parties or candidates?". This desire to persuade could lead to the desire to be more informed, and as such a model of information acquisition before an election without some form of communication between voters seems incomplete. We show that the option to share one's information can change the incentives to acquire information such that extremists will acquire more information than independents.

The driving force behind this result is that extremists are much less concerned about spreading information that leads to a wrong choice. Our model builds upon a standard Condorcetian model with two candidates and two states. In each state all voters weakly prefer one of the two candidates; we call this candidate the "correct" choice. The other candidate is called the "incorrect" choice. Extremists are distinguished from independents in that extremists suffer much less if their preferred candidate is chosen wrongly. Independents, however, suffer equally as much from the incorrect choice of either candidate. Given that the broadcast of positive information about a candidate increases the probability of that candidate winning the election, extremists will send any information in favor of their preferred candidate and will withhold any other information. In case that this candidate is the correct choice the extremist has much to gain, in case this candidate is the incorrect choice the extremist faces only a small loss. The independents face a more complicated decision. If an independent sends her signal she will gain no more than an extremist in the event that the signal helps the correct candidate win, however she could potentially cause herself pain by helping the incorrect candidate win. Since this is the case, an independent faces a significant implicit cost that an extremist need not consider. In short, independents worry more about disseminating a signal that could lead to a wrong choice and so their incentive to acquire signals ends up being lower than the care free extremists' since the independents face both the explicit signal cost as well as an implicit utility penalty.

We complement this main result by some observations for the case of large electorates. We find that when the expected number of voters grows to infinity, information is acquired only by a shrinking number of extremists and total information converges to a strictly positive amount. Interestingly, in the limit the total amount of signals acquired and broadcast when there are no extremists in the electorate is strictly less than the amount that would be broadcast when extremists exist.

1.1 Review of the Literature

In this section we relate our work to two main strands in the literature: endogenous information acquisition before elections, related models of information exchange. We also discuss some models that combine endogenous information acquisition and informational exchange in the context of politics.

Two main questions in the literature on endogenous information acquisition before an election are: "who acquires information" and "what is the probability that a large electorate of endogenously informed voters elects the better candidate?" In response to the first question Oliveros (2006) and Degan (2006) both find that more extreme voters will acquire less information. To understand the stark difference between their results and ours, observe that information cannot be transmitted in either of these models. In such a model information is only useful if it could lead an informed voter to vote differently from an uninformed voter. Independents and extremists differ insofar as that information has to be stronger to sway an extremist. Given a fixed price and strength of information the most extreme voters would never acquire information, since they know that they would not base their vote on this information.

We also contribute to the literature on information aggregation with endogenous information. Martinelli (2006) pioneered a formal model to analyze the question how much costly information a large electorate would acquire and aggregate. Like Martinelli (2006) we find that as an electorate grows large the amount of information per voter converges to zero. However, unlike Martinelli (2006) in our model the magnitude of total information acquired converges to a strictly positive finite amount that does not guarantee a correct choice being made. We concur with Martinelli (2007)'s result that there can exist equilibria with positive information acquisition for any size of electorate, however we do not require, as Martinelli does, the support of the marginal cost of signal acquisition to include zero. In our case people are willing to acquire information even when incurring a significant cost. This highlights the importance of the option to convince others in voting models.

Economic theory offers two main approaches to model informational exchange: "cheap talk" and "games of persuasion". It is assumed in either approach that information can be transmitted without cost; the difference between the two is that only true messages can be sent in a persuasion game, whereas models of cheap talk assume that the recipients of messages have no means to ascertain the truth value of a message. Coughlan (2000) noted the introduction of a cheap talk stage into a model of collective decision making can have a significant impact on the equilibrium outcome of such a model. This however, depends on the assumption that the voters or agents have similar enough preferences. It is well known that the effectiveness of cheap talk depends on the similarity of preferences among the speakers and listeners. This was already observed by Crawford and Sobel (1982) and has been extended to the context of pre-electoral deliberation by Austen-Smith and Feddersen (2005) and Austen-Smith and Feddersen (2006). The preferences of voters in our model are different enough such as to prevent any informative information transmission by extremists under the cheap talk assumption. We therefore follow the alternative route and assume that all messages are provable. This places our work in the literature on persuasion games. [2]

[2] Real life debates are certainly best described by an intermediate assumption between that of cheap talk or persuasion games. Mathis (2006) shows that Austen-Smith and Feddersen (2006)'s unfavorable judgment of the unanimity rule depends on the cheap talk assumption. He shows that their results can be mitigated when assuming at least partial provability of signals. While we agree that real life debates

Within this literature on persuasion games our result is strongly related to Shin (1994)'s result that an impartial arbitrator should place a higher burden of proof on the better informed advocate to extract the optimal amount of information from two opposing advocates. We find in analogy that independent voters would require more right signals to be convinced to vote for the right wing candidate if they knew that the right wing extremists are better informed than the left wing extremists. While Shin (1994) only derives a qualitative result we are able to fully characterize the "burden of proof" in our context as our assumptions on the information technology are more explicit. We will get back to this discussion once we derived the "burden of proof" for our context (see footnote 5 for this discussion).

The result that partisans have stronger incentives to acquire information also appears in Dewatripont and Tirole (1999). However, they derive this finding from a technology of information exchange that can neither be described as a game of cheap talk nor as a persuasion game. The informational setup could be described as a persuasion game in which two opposing signals are observationally equivalent to no signal. Consequently information is not fully contractible in their model, an independent who obtained two opposing signals receives the same reward as an independent that did not receive a signal. This renders the value of information non-monotonic for independents. It turns out to be more costly to induce such independents to acquire two signals than it is to induce two opposing advocates to acquire two signals. While we obtain a similar result: extremists are more motivated to acquire information the intuition behind our result is different. Independents in our model view information as less beneficial since they are afraid of any wrong decision, conversely extremists do not care if the their preferred candidate is chosen wrongly.

As far as we could discern the only work that allows for endogenous information acquisition as well as pre-electoral communication is Gerardi and Yariv (2006). This work differs from ours in their goal (they are interested in optimal committee design whereas we are interested in informational structure of an electorate) as well as their communication structure (cheap talk). Gerardi and Yariv (2006) show that incentives for information acquisition have to be traded off against efficient information aggregation when designing an optimal voting rule, when a cheap talk stage among all voters precedes the vote.

Hafer and Landa (2006) must also be mentioned when discussing the intersec-

---

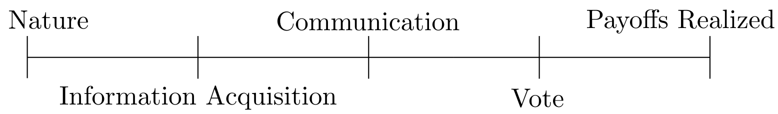
are certainly best described by an assumption of partial provability. Real life turns out to be a little bit too rich for our theoretical model. As a first pass we content ourselves with the assumption that all messages are provable.

tion of the literature on endogenous information acquisition before an election and informational exchange. Strictly speaking their work does not belong to either literature. Hafer and Landa (2006) do not analyze the politics of collective decision making but "private political" decisions. In the case of such decisions some citizens use the public or political sphere to influence the private decisions of others. The decision whether or not to have an abortion is an example of such a "private political" decision. The focus on such "private political" decisions allows Hafer and Landa (2006) to analyze a rich model of debates without being encumbered by usual complications in the analysis of games of collective decision making. The model of deliberation proposed by Hafer and Landa (2006) does neither belong to the family of cheap talk models nor is it a persuasion game. In just one sentence their non-Bayesian approach might best be described as a persuasion game without agreement on truth. So while the our model differs from Hafer and Landa (2006)'s in a variety of respects we share their conclusion that only the most extreme members of a society choose to broadcast information.

Finally it needs to be pointed out that we model the game played by the voters as a Poisson game following Myerson (1998) and Myerson (2000). It turns out that the analysis of the present model is considerably simplified by the setup as a Poisson game. This will be described in the section on the solution of this game.

2 The Model

Following Myerson (1998), we define $\{\Omega, \alpha, T, n, p, C, U\}$ as an extended Poisson game with the state space $\Omega = \{l, r\}$ and a common prior of $\alpha = \frac{1}{2}$ that the state is l. The parameter T denotes the set of voter types $T = T_1 \times T_2$ where $t_1 \in T_1 = \{l, i, r\}$ denotes the preference type of a voter and $t_2 \in T_2 = \{l, r\}$ denotes the informational type of a voter. The size of the electorate is a Poisson random variable with mean n. For each state of the world, $\omega \in \{l, r\}$, the preference and information type of a voter are drawn from independent distributions $p_1(\cdot|\omega)$ and $p_2(\cdot|\omega)$ respectively. We assume that the preference type of a voter is independent of the state of the world, and that any voter is equally likely to be or preference type l or r. We let $p_1(t_1 = l|\omega) = p_1(t_1 = r|\omega) = \eta < \frac{1}{2}$ for $\omega = \{l, r\}$. We assume that $n(1 - 2\eta) > 1$ which amounts to a very weak statement of the requirement that the electorate is "large". So that the informational type t_2 of a voter depends on the state of the world, we assume that $p_2(t_2 = l|\omega = l) = p_2(t_2 = r|\omega = r) = p > \frac{1}{2}$. Both the strategy set, C, and the preference set, U, will be explained in detail after the stages of the game are described.



The game proceeds according to the above time-line: First Nature draws a state of the world and an electorate. Next voters choose whether or not to acquire information. Voters then choose whether to publish their information or not. After the publication stage voters then learn the value of all the information that has been published. The Supreme Court throws a fair coin which can come up R or L, and lastly voters vote for L or R after which utility is realized.

The information acquisition stage is modeled very simply. A voter has to decide whether to learn his informational type t_2 or not. If he chooses to learn his type he will incur a fixed cost $c > 0$. We denote a voter's decision to become informed by $x_1 = 1$, otherwise we write $x_1 = 0$. In essence a voter's informational type is an informative signal on the state of the world. This is due to the assumption that $p_2(t_2 = l|\omega = l) = p_2(t_2 = r|\omega = r) = p > \frac{1}{2}$. From this point on we refer to t_2 as the informed voter's signal.

The exchange of information is modeled as a persuasion game following Glaeser and Rubinstein (2001), and Shin (1994). An informed voter can decide to either share his signal t_2 with the electorate as a whole, or stay silent. He cannot lie, reveal partial truths, or communicate with subsets of the entire electorate. The electorate learns the total number of l and r signals that have been published, denoted by the vector $\boldsymbol{s} = (s_r, s_l)$, and nothing more. Voters do not learn the total number of voters, the number of voters in any one camp, or how many signals have been acquired.

In the voting stage voters have to pick between R and L.[3] The vote is decided by majority rule. In case of a tie the Supreme Court picks on behalf of independent voters, meaning that the Court picks the candidate whom an independent would like better given all the information contained in the public signal vector $\boldsymbol{s}$. If this does not yield a clear cut result then the Supreme Court picks the candidate that was determined by a fair coin toss in the fifth stage.

[3] We will show in section 7 that not considering abstention is without loss of generality in the context of the present model, in the sense that we would obtain the same equilibrium result if we where to allow voter's to abstain.

The utility of a voter depends on the outcome of the vote $W = L$ or R (which in turn depends on the votes of all voters and in case of a tie on $\boldsymbol{s}$), his decision whether to acquire information or not $x_1 \in \{0, 1\}$, the state of the world and his preference type. We can now define: $U : \{L, R\} \times \{0, 1\} \times \{l, i, r\} \times \{l, r\} \to \mathbb{R}$:

$$U(W, x_1, t_1, \omega) = \begin{cases} -\delta(t_1) - x_1 c & \text{if } \omega = r \text{ and } W = L \\ -(1 - \delta(t_1)) - x_1 c & \text{if } \omega = l \text{ and } W = R \\ -x_1 c & \text{otherwise} \end{cases}$$

The preference type of a voter, t_1, determines the disutility a voter incurs when candidate X is picked in state y. For leftist voters we assume that $\delta(l) = 0$. A leftist does not receive any disutility from an incorrect choice of candidate L, however he receives the maximal disutility when candidate R wins the election in state l. Conversely we assume $\delta(r) = 1$, so a rightist's utility is minimized when L wins in state r. Finally we assume that independents suffer an equal amount of disutility from either mistake, $\delta(i) = \frac{1}{2}$.

3 Strategies and Equilibrium

The strategy set of a voter, $C = C_1 \times C_2 \times C_3$, is composed of an *information acquisition* rule $C_1 = \{0, 1\}$ a *broadcasting* rule $C_2 = (\mathcal{L} \times \mathcal{R})$ and a *voting* rule $C_3 = \{f : \mathbb{N}_0 \times \mathbb{N}_0 \times T \times \{L, R\} \to \{L, R\}\}$.[4] The broadcasting rule consists of two elements: $\mathcal{L} = \{\emptyset, l\}$ is the choice to broadcast a signal with value l, and $\mathcal{R} = \{\emptyset, r\}$ is the decision to broadcast a signal with value r, in each case $\emptyset$ stands for the suppression of the signal. C_3 is the set of all voting rules given vector of broadcast signals $\boldsymbol{s}$, a voter's preference type t_1, his informational type t_2 if he chose to acquire information, and the result of the court's coin. A typical pure strategy is denoted as a vector x with $x_1 \in \{0, 1\}$ as the voter's decision whether to acquire information or not, (x_2, x_3) as the voter's broadcasting strategy and x_4 as the voter's voting rule. The vector $x = (1, \emptyset, r, f)$, for instance, denotes the strategy in which the voter acquires information, only broadcasts r signals and follows a voting rule f.

We denote a mixed strategy of a voter of type t_1 by τ_{t_1}. We define $\tau = (\tau_l, \tau_i, \tau_r)$ as a mixed strategy profile for the game. For example, $\tau_l(0, \emptyset, \emptyset, f)$ denotes

[4] Here we follow the convention that $\mathbb{N}_0 := \mathbb{N} \bigcup \{0\}$

the probability that a leftist does not acquire information, never shares his signal, and follows the voting rule f. We define $\tau_{t_1}(\cdot, l, r, f)$ as the probability that a voter of type t_1 passes on any signal and plays voting strategy f. Formally $\tau_{t_1}(\cdot, l, r, f) = \tau_{t_1}(0, l, r, f)+\tau_{t_1}(1, l, r, f)$; the expressions $\tau_{t_1}(\cdot, \cdot, r, f)$, $\tau_{t_1}(\cdot, l, r, \cdot)$ and so forth are defined analogously. We define $EU_{t_1}(\tau, x)$ as the expected utility of a voter of preference type t_1 when this voter uses the pure strategy x while all other voters follow the profile τ. The probability that the state is r and candidate L wins when all other voters follow strategy τ and the voter under consideration follows strategy x is denoted by $Pr(L, r|\tau, x)$. Conversely the probability that the state is l and R wins when the voter uses the pure strategy x and all other voters follow the strategy profile τ is denoted by $Pr(R, l|\tau, x)$. Given these definitions we can express the expected utility of a voter of type t_1 as follows:

$$EU_{t_1}(\tau, x) = -\delta(t_1)Pr(L, r|\tau, x) - (1 - \delta(t_1))Pr(R, l|\tau, x) - cx_1. \quad (1)$$

Given this utility form, it is possible to define our equilibrium concept. We refine the Myerson's Poisson game equilibrium to look only at the equilibria with the following features.

Definition 1 *A strategy profile τ is an equilibrium if*

(1) $EU_{t_1}(\tau, x) \geq EU_{t_1}(\tau, y)$ *for all* $x \in C : \tau_{t_1}(x) > 0,\ y \in C,\ t_1 \in T_1$.
(2) The strategy profile is symmetric in the sense that extremists of both camps are equally likely not to acquire any information, we have that $\tau_l(0, \cdot, \cdot, \cdot) = \tau_r(0, \cdot, \cdot, \cdot)$.
(3) The independents will reveal any signal that they acquire, we have that $\tau_i(1, l, \emptyset, f) = \tau_i(1, \emptyset, r, f) = \tau_i(1, \emptyset, \emptyset, f) = 0$ *for all* $f \in C_2$.
(4) $E[U(L, x_1, t_1, \omega)|\boldsymbol{s}, x_1, t_1] \geq E[U(R, x_1, t_1, \omega)|\boldsymbol{s}, x_1, t_1] \Rightarrow f(\boldsymbol{s}, t_1) = l$, *i.e. voters vote sincerely.*

The first condition says that any strategy that a voter plays with a positive probability in τ has to be a best response to the strategy profile τ. This is Roger Myerson's definition of an equilibrium in an extended Poisson game (Myerson 1998). This definition alone does not rule out certain odd behaviors. A profile τ that prescribes that all voters vote L if more than 2 signals have been sent can be an equilibrium profile. In this case voters can infer that there are more than 2 voters in the electorate. So no voter has a chance to change the outcome of the vote. Unfortunately we cannot use weak domination to address this issue for reasons described in section 7.

To deal with the above issues we require conditions 2, 3 and 4. An important question arises: is the set of all 4 conditions compatible? For example, it is well known that sincere and strategic voting need not coincide. It turns out that it is straightforward to show that the second requirement is compatible with the first. This cannot be said of the latter two requirements: some of the most involved proofs in this paper are devoted to showing that the third and fourth requirements are compatible with the first as shown in sections 7 and 8.

4 Voting Strategies

4.1 Sincere Voting

Sincere voting requires a voter to vote for the candidate whom he would choose if he alone had to determine the winner based on all information available to him. We define $Pr[\boldsymbol{s}, x_1t_2|\omega]$ as the probability that the signal vector and the voters own information is $\boldsymbol{s}$ and x_1t_2 in state ω. If the voter has no private information beyond the public vector of signals ($x_1 = 0$), then x_1t_2 is equal to 0, on the other hand x_1t_2 equals t_2 for the case that the voter has private information i.e., $x_1 = 1$. We calculate a voters expected utility of the two outcomes L and R, when the information available to the voter is the signal vector, $\boldsymbol{s}$, and the voter's own private information, x_1t_2, as

$$-\delta(t_1)\frac{Pr[\boldsymbol{s}, x_1t_2|r]}{Pr[\boldsymbol{s}, x_1t_2|l] + Pr[\boldsymbol{s}, x_1t_2|r]} \tag{2}$$

$$-(1-\delta(t_1))\frac{Pr[\boldsymbol{s}, x_1t_2|l]}{Pr[\boldsymbol{s}, x_1t_2|l] + Pr[\boldsymbol{s}, x_1t_2|r]}. \tag{3}$$

for L and R respectively. If candidate L is chosen only one type of mistake matters: namely the case that L is chosen in state r. To obtain a voter's expected utility the probability of this mistake has to be multiplied with the disutility that this voter receives if L is chosen in state r, this disutility is $-\delta(t_1)$. Analogously we obtain the expected utility of candidate R given $\boldsymbol{s},\ x_1t_2,$ and t_1.

Voting sincerely requires a voter to vote for candidate L if expression 2 is larger than expression 3. If the opposite inequality holds true a voter that votes sincerely has to vote for candidate R. If the two expressions are equal the voter is indifferent between the two candidates. We assume that indifferent voters base their vote on the coin thrown by the court, if the coin comes up on X they vote for X.

4.2 Extremists

Observe that $Pr[\boldsymbol{s}, x_1t_2|\omega] \neq 0$ for all possible signal profiles $(\boldsymbol{s}, x_1t_2)$ and both states $\omega = l, r$. Our assumption that extremists vote sincerely implies the following Lemma:

Lemma 1 *In equilibrium, for all signal vectors, all leftists vote for candidate L and rightists vote for candidate R.*

Proof Expression 2 equals to 0 if $\delta(t_1) = 0$. On the other hand for no vector $(\boldsymbol{s}, x_1t_2)$ does the probability $Pr(\boldsymbol{s}, x_1t_2|l)$ equal 0. So expression 3 is always negative for a leftist. Consequently expression 2 is larger than expression 3 for any $(\boldsymbol{s}, x_1t_2)$. A leftist votes sincerely if he votes for L. The same argument holds mutatis mutandum for for a rightist. □

This can also be thought of as an issue of weak domination. When $\omega = r$ a leftist will have a zero utility no matter the outcome of the vote. In this case the extremist is indifferent between voting L and R. When $\omega = l$ a leftist will have a zero payoff if the electorate chooses L and a payoff of -1 if the electorate chooses R. The leftist has an incentive to vote left in some cases and is indifferent in every other case, thus has a weakly dominant strategy to vote left.

4.3 Independents

We show next that in equilibrium the independent voters vote according to a simple cutoff rule.

Lemma 2 *There exists a cutoff $g(n, \eta, p, \tau)$ such that an independent voter votes for L if $s_l > s_r + g(n, \eta, p, \tau)$ and votes for R if $s_l < s_r + g(n, \eta, p, \tau)$, otherwise the independent is indifferent.*

Proof We need to show that

$$Pr[\boldsymbol{s}, x_1t_2|l] > Pr[\boldsymbol{s}, x_1t_2|r] \tag{4}$$

if and only if $s_l > s_r + g(n, \eta, p, \tau)$ for some expression $g(n, \eta, p, \tau)$.

If the voter has not invested in information acquisition, that is if $x_1 t_2 = 0$, then

$$Pr[\boldsymbol{s}, tx_2|l] = \frac{e^{-npS_l}[npS_l]^{s_l}}{s_l!} \cdot \frac{e^{-n(1-p)S_r}[n(1-p)S_r]^{s_r}}{s_r!}$$
$$Pr[\boldsymbol{s}, tx_2|r] = \frac{e^{-n(1-p)S_l}[n(1-p)S_l]^{s_l}}{s_l!} \cdot \frac{e^{-npS_r}[npS_r]^{s_r}}{s_r!}$$

for

$$S_l = \eta[\tau_l(1, l, \emptyset, \cdot) + \tau_l(1, l, r, \cdot) + \tau_r(1, l, \emptyset, \cdot) + \tau_r(1, l, r, \cdot)] \\ +(1-2\eta)[\tau_i(1, l, \emptyset, \cdot) + \tau_i(1, l, r, \cdot)]$$

and

$$S_r = \eta[\tau_l(1, \emptyset, r, \cdot) + \tau_l(1, l, r, \cdot) + \tau_r(1, \emptyset, r, \cdot) + \tau_r(1, l, r, \cdot)] \\ +(1-2\eta)[\tau_i(1, \emptyset, r, \cdot) + \tau_i(1, l, r, \cdot)]$$

Observe that expression 2 is larger than expression 3 if and only if

$$\frac{Pr[\boldsymbol{s}|l]}{Pr[\boldsymbol{s}|r]} = (\frac{p}{1-p})^{[s_l - s_r]} \cdot e^{[n(1-2p)(S_l - S_r)]} < 1 \tag{5}$$

we can rearrange to find the cutoff condition

$$s_l < \frac{n(2p-1)}{\ln(\frac{p}{1-p})}(S_l - S_r) + s_r \tag{6}$$

which is always well defined in the given environement. If equation 6 holds with equality then the voter is indifferent.

We next need to show that an independent voter that has observed a signal follows the same rule when voting sincerely. To see this assume that the voter has observed an l-signal. So the public signal count without his signal then becomes $s_l - 1$ and s_r. In this case the voter estimates the probabilities $Pr[\boldsymbol{s}, tx_2|l]$ and $Pr[\boldsymbol{s}, tx_2|r]$ as

$$Pr[\boldsymbol{s}, tx_2|l] = p\frac{e^{-npS_l}[npS_l]^{(s_l-1)}}{(s_l-1)!} \cdot \frac{e^{-n(1-p)S_r}[n(1-p)S_r]^{s_r}}{s_r!}$$
$$Pr[\boldsymbol{s}, tx_2|r] = (1-p)\frac{e^{-n(1-p)S_l}[n(1-p)S_l]^{(s_l-1)}}{(s_l-1)!} \cdot \frac{e^{-npS_r}[npS_r]^{s_r}}{s_r!}$$

which yields the exact same condition on the voter's behavior. So we find that the statement of the Lemma holds true for $g(n, \eta, p, \tau) = \frac{n(2p-1)}{\ln(\frac{p}{1-p})}(S_l - S_r)$ □

In the next section on optimal communication we will formally show that no extremists will ever send evidence that favors the opposing candidates or $\tau_l(1,\cdot,r,\cdot) = \tau_r(1,l,\cdot,\cdot) = 0$. Taking this result for granted, we can rewrite $S_l - S_r$ as $\tau_l(1,l,0,\cdot) - \tau_r(1,0,r,\cdot)$. We can now see that this term acts as a counterbalance to the biased information that extremists are broadcasting. If for some reason the rightists acquire and send signals with higher probability than the leftists; the independent will discount these signals by a factor depending on the discrepancy between the strategies of the both camps of extremists. If the signal strength is exceptionally strong ($p \to 1$) this bias is dampened since $\frac{n(2p-1)}{\ln(\frac{p}{1-p})} \to 0$. Intuitively in this situation it becomes incredibly unlikely that a signal opposite the true state is obtained and broadcast even if there are large differences in extremist strategies. As such signal count becomes more important than biased information concerns. [5]

By the symmetry of our equilibrium concept extremist strategies are equal and the model assumption is that the groups are on average the same size so $S_l - S_r = 0$. Since this is the case we have that $g(n,p,S_r,S_l) = 0$ and an independent voter is indifferent between the two candidates if equally many signals have been sent for both candidates. This reduces the voting rule for independents to voting for the candidate that has a majority of signals. Note that the reduction of the voting rule to a simple majority of signals is an artifact of the symmetry of the setup. As described above, independents are indeed considering that some of the information is biased, however in the special symmetric case the bias of the left signals cancels out the bias of the right ones.

It is important to note that the simplicity of the above argument comes from the Poisson structure. Without it calculating the voting rules would be a much more complicated issue, and it is not clear if this exercise would benefit the exposition of the ideas set forth in this paper. The reason for this is that under the assumption of the Poisson distribution a voter's own own signal carries as much information as anyone else's. For other assumptions the case in which the independent has not acquired any information differs significantly from the alternative case in which he has acquired information.

[5] An interesting comparison with Shin (1994) arises. The expression $g(n,\eta,p,\tau)$ can be interpreted as Shin's burden of proof: if g is large the leftists face a large burden of proof, in the sense that they need to provide more signals for any given amount of right signals to convince an independent that L is the better candidate. This form of the burden of proof has the advantage that it can be completely characterized in terms of $n,\ \eta,\ p$ and τ.

Sincere voting does not imply anything for the case that $s_l = s_r + g(n, \eta, p, \tau)$. We assume that independent voters follow the coin throw of the Supreme Court in this particular case.

We conclude this section by defining the voting rules $f_l, f_i, f_r \in C_3$ by $f_l(\boldsymbol{s}, X) = L$, $f_r(\boldsymbol{s}, X) = R$ and $f_i(\boldsymbol{s}, X) = L$ if $s_l > s_r$, $f_i(\boldsymbol{s}, X) = R$ if $s_l > s_r$ and finally if $s_l = s_r$ then $f_i(x) = R$ if and only if the publicly thrown coin came up $X = R$. In the two preceding Lemmata 1 and 2 we have shown that in any equilibrium profile τ the voters will only follow these strategies. In short, we have shown that $\tau_{t_1}(\cdot, \cdot, \cdot, f_{t_1}) = 1$ for $t_1 \in \{l, i, r\}$.

5 Optimal Communication

In this section we show that in equilibrium extremists will pass on signals in their favor and suppress signals in favor of the other candidate. The broadcasting behavior of the independents is determined by equilibrium condition 3: in equilibrium independents broadcast all signals. We will show in section 8 that this assumption is consistent with strategic behavior. We start by showing that extremists will never acquire information without any plans to broadcast it.

Lemma 3 *In equilibrium we have that $\tau_l(1, \emptyset, \emptyset, f) = \tau_r(1, \emptyset, \emptyset, f) = 0$ for all voting strategies f.*

Proof Suppose in equilibrium some $x = (1, \emptyset, \emptyset, f)$ was played with positive probability by the left wing extremists. As f_l is a weakly dominant voting strategy for left wing extremists we have that $EU_l(\tau, x) \leq EU_l(\tau, x^*)$ for $x^* = (1, \emptyset, \emptyset, f_l)$. Now consider the alternative strategy $x' = (0, \emptyset, \emptyset, f_l)$, which differs from x^* only insofar as that the voter does not acquire information. Remember that a voter's expected utility can be expressed as

$$EU_{t_1}(\tau, y) = -\delta(t_1)Pr(L, r|\tau, y) - (1 - \delta(t_1))Pr(R, l|\tau, y) - cy_1.$$

for $t_1 \in \{l, r, i\}$. The signal acquired by the extremist is only known to him, since this signal will never sway his vote we have that $Pr(Y, \omega|\tau, y)$ does not depend on this signal. Thus $EU_l(\tau, x^*) - EU_l(\tau, x') = -c$. Consequently the voter strictly prefers x' to x^* and x, we must have that $\tau_l(1, \emptyset, \emptyset, f) = 0$ for all voting strategies f. An analogous argument holds for right wing extremists.
□

Note that this proof cannot be used to show that $\tau_i(1, \emptyset, \emptyset, f) = 0$ in equilibrium as the independents do not have a dominant voting strategy.

To analyze the optimal broadcasting strategies we need to find out how the broadcasts of l and r signals change the winning probabilities of the two candidates in the two states. We define $Pr(E|\tau, \omega)$ as the probability that event E occurs in state ω when all voters follow strategy τ. Let us assume the stance of one particular voter, and define the events R and L such that in these events candidates R and L win the election, when the voter under consideration does not send a signal. Additionally, we define the events $X + r$ $(X + l)$ as the event that the candidate X wins the election given that the voter sent an r (l) signal .

The effect of an additional signal l or r being sent on the expected utility comes through the effect of that additional signal on the winning probabilities of the two candidates in the two states. We define $\Delta^z_{Y,\omega}$ as the change in probability that candidate Y is being chosen in state ω when another z signal is being sent. Formally

$$\Delta^z_{Y,\omega} := Pr(Y + z|\tau, \omega) - Pr(Y|\tau, \omega).$$

Observe that $\Delta^r_{R,\omega} > 0$ and $\Delta^l_{L,\omega} > 0$, i.e. in either state the probability that candidate Y wins is increasing in the number of signals sent in his favor. It is also true that a right (left) wing extremist would never pay to acquire a signal and then send a left (right) signal. Lemma 4 formalizes this idea.

Lemma 4 *In equilibrium extremists will never send a signal counter to their bias, i.e.* $\tau_l(1, \cdot, r, f_l) = \tau_r(1, l, \cdot, f_r) = 0$. *Moreover once a signal is obtained that lies with their bias, an extremist will always broadcast it.*

Proof Conditional on having received an l-signal the following holds true for a left wing extremist:

$$\begin{aligned} EU_l(\tau, x) - EU_l(\tau, x') = p(Pr(L + l|\tau, l) - Pr(L|\tau, l)) = p\Delta^l_{L,l} > 0 \\ \text{for } x = (1, l, \cdot, f_l) \text{ and } x' = (1, \emptyset, \cdot, f_l) \end{aligned}$$

To see this observe that the utility of a left wing extremist is always 0 in state r, so utility changes only happen in state l. Given that the extremist has observed an l signal he thinks that state l is the true state with probability p. So any left wing extremist strictly prefers sending an l signal to being silent as $\Delta^l_{L,l} > 0$. Let us check next that no leftist extremist will send an r signal in equilibrium. To see this observe that, conditional on having received an r we have that

$EU_l(\tau, x) - EU_l(\tau, x') = -(1-p)((Pr(R+r|\tau) - Pr(R|\tau, l)) = -(1-p)\Delta^r_{R,l} < 0$ for $x = (1, \cdot, r, f_l)$ and $x' = (1, \cdot, 0, f_l)$, so a left wing extremist strictly prefers to stay silent. Analogous arguments hold for right wing extremists. $\square$

Lemma 4 along with equilibrium assumption 2 imply that $\tau_l(1, l, \emptyset, f_l) = \tau_r(1, \emptyset, r, f_r)$. [6] To save on notation we say from now on that λ is the probability that an extremist acquires information in the equilibrium strategy profile τ. Lemmata 3 and 4 imply that the extremists always pass on information that supports their case.

Lemma 4 together with the symmetry assumption also implies that $S_l = S_r$ and, as previewed in the previous section, $g(n, \eta, p, \tau) = 0$ so that the independents equilibrium rule for sincere voting becomes vote for the candidate for whom more signals have been broadcast. The symmetry of the equilibrium also implies that $\Delta^r_{R,r} = \Delta^l_{L,l}$ and $\Delta^l_{L,r} = \Delta^r_{R,l}$. To save on notation we define $\Delta^r_{L,r} = \Delta^l_{R,l} := \Delta^+$ as the increase in probability of a correct choice Y in state $\omega = y$ given that another signal in favor of Y is being sent. Analogously we define $\Delta^l_{L,r} = \Delta^r_{R,l} := \Delta^-$ as the increase in the probability of the incorrect choice of candidate Y in state $\omega \neq y$ given that another signal in Y's favor is being sent.

Economizing further on notation we say that π is the probability that an independent acquires information in the equilibrium strategy profile τ. Our assumption that independents will share any signal they acquired implies that $\tau_i(1, l, r, f_i) = \pi$.

To summarize what is known so far, observe that in any equilibrium τ left wing extremists vote for the left candidate and right wing extremists vote for the right candidate. Independents vote according to a very simple cutoff rule: if more l-signals have been published they vote for candidate L, conversely if more r-signals have been published they vote for candidate R, otherwise they base their vote on the coin of the Supreme Court. We also know that extremists would only broadcast signals in their favor. We show that our assumption that independents broadcast all their signals is consistent with equilibrium in section 8. We are now ready to prove our main result, namely the fact that in equilibrium that extremists never acquire less information than the independents.

[6] Since $\tau_l(0, \cdot, \cdot, f_l) = 1 - \tau_l(1, l, \emptyset, f_l)$ and $\tau_r(0, \cdot, \cdot, f_r) = 1 - \tau_r(1, \emptyset, r, f_r)$

6 Extremists Are More Informed

Theorem 1 *In an equilibrium the utility gain from acquiring information for an extremist is always greater than that of an independent.*

The intuition for this result goes as follows. In any equilibrium τ the extremists only broadcast information in their favor. For example, a left wing extremist would only broadcast an l signal. This reduces the probability that candidate R is chosen in state l; at the same time this increases the probability of the alternative mistake that candidate L is chosen in state r. A left wing extremist does not care about the increase in the probability of the second mistake. An independent also broadcasts an l signal if he receives one, however, unlike the extremists both effects are felt by the independent. The independent appreciates the fact that the probability of an erroneous choice of R is being reduced. At the same time an independent suffers from the fact that the additional l signal increases the probability that L is chosen in state r.

Proof We compare the expected utility of right wing extremist voter for the pure strategies $x_r = (1, \emptyset, r, f_r)$ and $x'_r = (0, \emptyset, \emptyset, f_r)$ to each other. From the proof of Lemma 4 we know that a right wing extremist values a right wing signal at $p\Delta^+$. Given that the signal costs c, and that right wing signals are obtained with probability $\frac{1}{2}$, we conclude that:

$$EU_r(\tau, x_r) - EU_r(\tau, x'_r) = \frac{p}{2}\Delta^+ - c.$$

Analogously we have for a left wing extremist, where $x_l = (1, l, \emptyset, f_l)$ and $x'_l = (0, \emptyset, \emptyset, f_l)$:

$$EU_l(\tau, x_l) - EU_l(\tau, x'_l) = \frac{p}{2}\Delta^+ - c.$$

To investigate the utility difference for an independent, let $x_i = (1, l, r, f_i)$ and $x'_i = (0, \emptyset, \emptyset, f_i)$, so that

$$
\begin{aligned}
EU_i(\tau, x_i) = & -\frac{1}{4}pPr(L+r|\tau, r) - \frac{1}{4}(1-p)Pr(R+r|\tau, l) \\
& -\frac{1}{4}(1-p)Pr(L+l|\tau, r) - \frac{1}{4}pPr(R+l|\tau, l) - c \\
EU_i(\tau, x_i') = & -\frac{1}{4}pPr(L|\tau, r) - \frac{1}{4}(1-p)Pr(R|\tau, l) \\
& -\frac{1}{4}(1-p)Pr(L|\tau, r) - \frac{1}{4}pPr(R|\tau, l)
\end{aligned}
$$

The expected utility difference between acquiring and not acquiring then becomes:

$$
\begin{aligned}
EU_i(\tau, x_i) - EU_i(\tau, x_i') &= \frac{p}{2}\Delta^+ - \frac{1-p}{2}\Delta^- - c \\
&< \frac{p}{2}\Delta^+ - c = EU_l(\tau, x_l) - EU_l(\tau, x_l') = EU_r(\tau, x_r) - EU_r(\tau, x_r').
\end{aligned}
$$

Where the inequality follows from the observation that $\Delta^- > 0$. We conclude that for any equilibrium strategy profile τ information is more valuable for extremists than for independents. $\square$

We observe in passing that the proof of Theorem 1 implies that the left and right wing extremists benefit equally much from the acquisition of information. Consequently the second assumption in our equilibrium concept, that all extremists are equally likely not to acquire information, is consistent with our assumption that all players are best responding.

Our model generates a positive correlation between information and extremism. We state this result as a simple corollary of Theorem 1.

Corollary 1 *If the equilibrium probability that independents acquire information is positive then all extremists acquire information in equilibrium. If the equilibrium probability that an extremist acquires no information is positive, then no independent acquires any information in equilibrium.*

Proof We can express Corollary 1 as:

$$\pi > 0 \Longrightarrow \lambda = 1 \text{ and } \lambda < 1 \Longrightarrow \pi = 0.$$

Using the results derived in Theorem 1:

$$
\begin{aligned}
\pi > 0 &\Rightarrow EU_i(\tau, x_i) - EU_i(\tau, x_i') = 0 \\
&\Rightarrow \frac{p}{2}\Delta^+ - \frac{1-p}{2}\Delta^- = c \\
&\Rightarrow \frac{p}{2}\Delta^+ > c \\
&\Rightarrow EU_l(\tau, x_l) - EU_l(\tau, x_l') > c \Rightarrow \lambda = 1
\end{aligned}
$$

and

$$
\begin{aligned}
\lambda < 1 &\Rightarrow EU_l(\tau, x_l) - EU_l(\tau, x_l') = 0 \\
&\Rightarrow \frac{p}{2}\Delta^+ = c \\
&\Rightarrow \frac{p}{2}\Delta^+ - \frac{1-p}{2}\Delta^- < c \\
&\Rightarrow EU_i(\tau, x_i) - EU_i(\tau, x_i') < 0 \Rightarrow \pi = 0
\end{aligned}
$$

□

We conclude that information and extremism is indeed positively correlated in equilibrium, if such an equilibrium exists. The next three sections are devoted to showing that such an equilibrium indeed exists. We need to show in particular that the assumptions of sincere voting and transmission of all signals by the independents do not stand in contradiction with strategic voting.

7 Sincere Voting is Strategic Voting: The Swing Voter's Boon

Equilibrium requirement 4 imposes that all voters vote sincerely in any equilibrium τ. In section 4 we describe the sincere voting strategies of the three voter types. It remains to be shown that no voter has an incentive to deviate from the sincere strategies given that all other voters vote sincerely. In other words, the question whether strategic behavior (Equilibrium requirement 1) is consistent with sincere voting (Equilibrium requirement 4)is still open. We need to inquire whether a sincere voter for L would like to vote for L if he knew that he was the pivotal voter. For the extremists this is easy to see, as a left wing extremists strictly prefers voting for L to voting for R in any case in which the state of the world might (with some positive probability be l.

It is somewhat more difficult to see that an independent would like to vote sincerely. Consider the case of a signal structure $\boldsymbol{s}$ such that the equilibrium profile τ prescribes that independents vote L. If the expected utility of R is higher than the expected utility of L given $\boldsymbol{s}$ *and given that the voter is pivotal* then a strategic independent should vote R instead of L. The crucial question is whether an independent who knows that he is pivotal thinks that the state r is more likely than the state l. We will show that the information contained in the pivotality event generally strengthens the information contained in the public signal $\boldsymbol{s}$.

The intuition goes as follows: Suppose we have that $s_l \geq s_r$, and that the coin of the Supreme Court came up L (which is relevant only in the case that $s_l = s_r$). According to the strategy profile τ all independents will vote L. An independent would be pivotal if L and R receive equally as many votes. The fact that all independents vote for L implies that there must be more right wing extremists in the electorate than there are left wing extremists. Now consider the fact that more l signals have been sent. Taken together this implies it is likely that more right wing extremists decided to hide an l signal than there are a left wing extremist who hid an r signal.

The following Lemma will prove useful in this context:

Lemma 5 *Suppose that $Pr(l|\boldsymbol{n}, \boldsymbol{s}) \geq Pr(r|\boldsymbol{n}, \boldsymbol{s})$. Let $\boldsymbol{n'}, \boldsymbol{s'}$ be such that $n'_r \geq n_r$, $n'_l \leq n_l$, $n'_i = n_i$, $s'_r \leq s_r$, $s'_l \geq s_l$ and either $\boldsymbol{n'} \neq \boldsymbol{n}$ or $\boldsymbol{s'} \neq \boldsymbol{s}$ or both, then we have that $Pr(l|\boldsymbol{n'}, \boldsymbol{s'}) > Pr(r|\boldsymbol{n'}, \boldsymbol{s'})$.*

Proof Observe that $Pr(l|\boldsymbol{n}, \boldsymbol{s}) \geq Pr(r|\boldsymbol{n}, \boldsymbol{s})$ holds if and only if $Pr(\boldsymbol{n}, \boldsymbol{s}|l) \geq Pr(\boldsymbol{n}, \boldsymbol{s}|r)$. We proceed by distinguishing two cases: $\pi = 0$ and $\pi > 0$.

In case that $\pi = 0$ the signals s_l, s_r are drawn from binomial distributions $B(n_l, \lambda p), B(n_r, 1 - \lambda p)$ in state l and $B(n_l, 1 - \lambda p), B(n_r, \lambda p)$ in state r. The inequality $Pr(\boldsymbol{n}, \boldsymbol{s}|l) \geq Pr(\boldsymbol{n}, \boldsymbol{s}|r)$ holds if and only if:

$$\binom{n_l}{s_l}(\lambda p)^{s_l}(1-\lambda p)^{n_l - s_l}\binom{n_r}{s_r}(\lambda(1-p))^{s_r}(1-\lambda(1-p))^{n_r - s_r} \geq$$
$$\binom{n_l}{s_l}(\lambda(1-p))^{s_l}(1-\lambda(1-p))^{n_l - s_l}\binom{n_r}{s_r}(\lambda p)^{s_r}(1-\lambda p)^{n_r - s_r} \Leftrightarrow$$
$$\left(\frac{p}{1-p}\right)^{s_l - s_r} \geq \left(\frac{1-\lambda p}{1-\lambda(1-p)}\right)^{n_r - n_l + s_l - s_r}$$

As $1 - \lambda p > 1 - \lambda(1 - p)$ and $p > (1 - p)$ we have that

$$\left(\frac{p}{1-p}\right)^{(s'_l - s'_r)} > \left(\frac{1-\lambda p}{1-\lambda(1-p)}\right)^{(n'_r - n'_l)+(s'_l - s'_r)}.$$

for $s'_l - s'_r \geq s_l - s_r$ and $n'_r - n'_l \geq n_r - n_l$ with at least one of the inequalities strict. We conclude that $Pr(l|\boldsymbol{n'}, \boldsymbol{s'}) > Pr(r|\boldsymbol{n'}, \boldsymbol{s'})$ for $\boldsymbol{s'}, \boldsymbol{n'}$ described in the statement of the Lemma.

When $\pi > 0$ the state l is more likely if more signals have been sent in its favor. The independent needs to compare the total number of signals, revealed or hidden, in favor of l (call this s_l^*), with the total number of signals in favor of r, (call this s_r^*). Not knowing these numbers the independent can calculate the expected difference between the two numbers for a fixed $\boldsymbol{s}, \boldsymbol{n}$ as $E(s_l^* - s_r^*|\boldsymbol{s}, \boldsymbol{n})$; $Pr(l|\boldsymbol{n}, \boldsymbol{s}) \geq Pr(r|\boldsymbol{n}, \boldsymbol{s})$ holds if and only if $E(s_l^* - s_r^*|\boldsymbol{s}, \boldsymbol{n}) \geq 0$.

Define s_l^i, s_r^i as the number of l and r signals sent by the independents. Since $\pi > 0$, we have by Corollary 1 that all extremists acquire information. Consequently every silent extremist hides a signal in favor of the opposite candidate. We can calculate s_l^*, s_r^* as:

$$s_l^* = s_l + n_r - (s_r - s_r^i) \qquad s_r^* = s_r + n_l - (s_l - s_l^i)$$

The expected difference becomes:

$$\begin{aligned} E(s_l^* - s_r^*|\boldsymbol{s}, \boldsymbol{n}) = E(2(s_l - s_r) + (n_r - n_l) + (s_l^i - s_r^i)|\boldsymbol{s}, \boldsymbol{n}) = \\ 2(s_l - s_r) + (n_r - n_l) + E(s_r^i - s_l^i|\boldsymbol{s}, n_i) \end{aligned}$$

The expression $E(s_l^* - s_r^*|\boldsymbol{s}, \boldsymbol{n})$ is increasing in $n_r - n_l$. The expression is also increasing in $s_l - s_r$ as an increase of $s_l - s_r$ by 1 increases $s_l - s_r$ by 2 while it decreases $E(s_r^i - s_l^i|\boldsymbol{s}, \boldsymbol{n})$ by at most 1. □

Theorem 2 *(The Swing Voter's Boon) Sincere voting does not conflict with strategic voting in equilibrium, formally requirement 4 and 1 are consistent.*

Proof Suppose that $\boldsymbol{s}$ is such that τ prescribes for an independent to vote L. Would this independent want to vote L if he knew that he was pivotal? In other words, if T is the event that i is pivotal i.e. $T = \{\boldsymbol{n}|n_r = n_l + n_i, n_r = n_l + n_i + 1\}$, is it true that $Pr(l|\boldsymbol{s}, T) \geq Pr(r|\boldsymbol{s}, T)$. We know from Lemma 5 that $Pr(l|\boldsymbol{s}, T) \geq Pr(r|\boldsymbol{s}, T)$ holds for all cases in which independents are supposed to pick L if it holds for the case in which $s_l - s_r$ and $n_r - n_l$ are being minimized given $s_l \geq s_r$ and the independent is pivotal, namely $s_l = s_r$ and $n_r - n_l = n_i$. In the sequel we will only investigate this case and show that the independent has an incentive to vote for L even in this worst case scenario.

When $\pi = 0$ we can use the argument given in the proof of Lemma 5 to show that $Pr(l|\boldsymbol{n}, \boldsymbol{s}) \geq Pr(r|\boldsymbol{n}, \boldsymbol{s})$ if and only if

$$\left(\frac{p}{1-p}\right)^{s_l - s_r} \geq \left(\frac{1-\lambda p}{1-\lambda(1-p)}\right)^{n_r - n_l + s_l - s_r}$$

Observe that for $s_l = s_r$ and $n_r - n_l = n_i \geq 0$ the above in equality always holds so we are done.

When $\pi > 0$ Lemma 5 implies that we need to show $E(s_l^* - s_r^* | s_r = s_l, n_r - n_l = n_i) \geq 0$. Following the arguments given in the proof of Lemma 5 we can calculate $E(s_l^* - s_r^* | s_r = s_l, n_r - n_l = n_i)$ as $E(n_i + s_r^i - s_l^i | s_r = s_l, n_r - n_l = n_i)$.

If k out of the n_i independents acquired information then we can calculate a lower bound on $E(s_r^i - s_l^i) = (1-p)k - pk - 1 = (1-2p)k - 1$ (assuming that the state is l, the independent under consideration observed an l signal, and $s_l = s_r > 2k$ which in turn implies that the constraint of the total number of signals does not bind). This expression is bounded from below by $(1-2p)n_i - 1$ as $k = 0, ..., n_i$. So a lower bound for $E(n_i + s_r^i - s_l^i | s_r = s_l, n_r - n_l = n_i)$ is $E(n_i - n_i(1-2p) - 1) = n(1-2\eta)2p - 1$. This expression is non-negative as $n(1-2\eta) \geq \frac{1}{2p} \geq 1$ as was assumed in the setup of the model. □

Theorem 2 establishes the presence of a swing voter's boon, in the sense that the information contained in the pivotality event reaffirms any decision by the independent. As a consequence of the swing voter's boon we can without loss of generality incorporate abstention into our model. To see this take a game $G = \{\Omega, \alpha, T, n, p, C, U\}$ as described in section 2, let τ be an equilibrium for this game. Now consider an alternative game G' that differs from G only insofar as that voters have the option to abstain ($C' \neq C$). We claim that τ is an equilibrium in G'[7]. Since τ is an equilibrium in G we know that no voter can be made better off by a deviation to a strategy that is already contained in C. We only have to consider deviations to strategies in C'/C, i.e. we have to only consider strategies that involve abstention. Theorem 2 tell's us that abstention can never be a better response than the voting strategies prescribed by τ. On the equilibrium path, a voter that conditions on the signal vector and on the event of being pivotal prefers to vote according to the prescription by the voting rule. We conclude that our assumption that voters cannot abstain is without loss of generality in the present context.

[7] This statement is, of course, somewhat sloppy as τ is not even a strategy profile in G'. Strictly speaking we should define a strategy profile τ' for G' such that $\tau_i'(x) = \tau_i(x)$ whenever $\tau_i(x) > 0$.

To see that sincere voting is not the only voting rule that is consistent with strategic voting consider the following case: Let each of the three camps of voters have 1000 members on average, that is $\eta n = (1-2\eta)n = 1000$. Consider a strategy profile in which no independents but all extremists acquire information, that is $\pi = 0$ and $\lambda = 1$. Consider a voting rule that prescribes an R vote for all independents when $s_r = 0$ and $s_l = 2$. Clearly this is a deviation from sincere voting, based on the signals alone the voter should prefer L to R. We show next that a strategic independent has no incentive to deviate from this voting rule. To see this observe that the independent is pivotal if and only if either $n_i + n_r = n_l$ or $n_i + n_r = n_l + 1$. In either case the independent expects that there are about 1000 more left wingers than there are right wingers. Given that the strategy profile prescribes that all voters acquire information, this implies that there are about 1000 more hidden r-signals than there are hidden l-signals. The information contained in the pivotality event easily outweighs the information contained in the signal vector. We conclude that if we where to replace the requirement that voters vote sincerely by the weaker and more commonly used requirement that no player uses a weakly dominated strategy, we would be left with a much larger equilibrium set.

8 Independents never want to acquire signals for private use only

Equilibrium requirement 3 says that any independent that acquires a signal would broadcast it, formally $\tau_i(1, l, \emptyset, f) = \tau_i(1, \emptyset, r, f) = \tau_i(1, \emptyset, \emptyset, f) = 0$ for all $f \in C_2$. Our solution is based on this requirement. We, therefore, need to show that no strategic voter would like to deviate from this behavior. We need to show that $EU_i(\tau, (1, \emptyset, \cdot, \cdot)) \leq EU_i(\tau)$.

Lemma 6 *The disclosure of all signals by the independents does not conflict with strategic behavior in equilibrium; formally, equilibrium requirement 1 and 3 are consistent.*

Proof We know from the prior section that f_i is the best voting strategy for an independent given that he either acquired and sent a signal or that he did not acquire information. We will show in the following proof that a preference of acquiring and hiding a signal over not acquiring a signal implies that sending the signal is even better than hiding it. Formally we will show that if there exists a voting rule f such that:

$$EU_i(\tau, (1, \emptyset, \cdot, f)) \geq EU_i(\tau, (0, \cdot, \cdot, f_i))$$

then it must be that:

$$EU_i(\tau, (1, l, ., f_i)) > EU_i(\tau, (1, \emptyset, ., f))$$

This in turn implies that $(1, \emptyset, ., .)$ is never a best reply for an independent. For this proof let us assume the role of an informed independent, i^*, that has received an l signal. Observe that there are $n_i - 1$ other independents in the electorate. As before we assume that $\boldsymbol{s}^* = (s_l^*, s_r^*)$ is the vector of all public signals when i^* does send his signal. For the context of the section we also define $\boldsymbol{s}^* - l$ as the set of all public signals when voter i^* decides not to disclose his signal. Observe that f_i prescribes to vote L if $s_l^* > s_r^*$ when i^* does send his signal and $s_l^* - 1 > s_r^*$ when i^* does not send his signal. We assume that f^* maximizes $EU_i(\tau, (1, \emptyset, \cdot, f))$, so f^* is the optimal voting rule for the independent that received an l signal but did not broadcast this signal.

For $EU_i(\tau, (1, \emptyset, ., f^*)) \geq EU_i(\tau, (0, ., ., f_i))$ we must have that $f^* \neq f_i$. If these two would coincide then the independent would be strictly better off not to incur any cost for the private signal which he would never use.[8] We know from Theorem 2 that sincere voting is strategic voting: If $\boldsymbol{s}^*$ and $\boldsymbol{s}^* - l$ yield the same voting recommendation under f_i, then it must be true that $f_i(\boldsymbol{s}^*) = f^*(\boldsymbol{s}^* - l)$. So we need to only consider the cases in which the signal of the independent could possibly sway the vote of all independents. This happens if either $s_l^* = s_r^*$ and $X = L$, or $s_l^* = s_r^* + 1$ and $X = R$ where X represents the coin of the Court.

In the case that $s_l^* = s_r^*$ and $X = L$ all independents except for i^* will vote R if i^* does not share his signal ($f_i(\boldsymbol{s}^* - l) = R$). Voter i^* does not have an incentive to do any different from all the other independents. If the Court would have recommended R for the case that all signals have been sent Theorem 2 implies an independent prefers to vote R. However, the independent has the same amount of information as if he did not send the signal, so the independent would prefer to vote R in that case as well. We conclude that $f_i(\boldsymbol{s}^*, L) = f^*(\boldsymbol{s}^* - l, L)$. If i^* has an incentive to vote according to his private signal this must happen when $s_l^* = s_r^* + 1$ and $X = R$.

[8] We used this argument in Lemma 3 to show that the extremists would never buy a signal without any plans to send it.

Focusing on this case; assume that i^* withheld his signal. [9] Since this is our last chance to find a $\boldsymbol{s}^*$ such that $f_i(\boldsymbol{s}^*, R) \neq f^*(\boldsymbol{s}^* - l, R)$ it must be true that $f^*(\boldsymbol{s}^* - l, R) = L$ for $s_l^* = s_r^*$.

To fully understand the incentives of the privately informed voter we need to find out when this voter would be pivotal. In the case under study all independents except for i^* ($n_i - 1$ independents) vote R. Without i^* 's vote R receives $n_r + n_i - 1$ votes whereas l receives n_l votes. Voter i^* is pivotal in exactly two cases namely $n_r + n_i = n_l$ and $n_r + n_i = n_l + 1$ (Remember that we are only looking at the case in which the Court's coin came up R, so R wins the election in case of a tie). So we have found a necessary condition for $EU_i(\tau, (1, \emptyset, ., f^*)) \geq EU_i(\tau, (0, ., ., f_i))$ to hold. This condition is that i^* prefers L to R conditional on having an l signal, $s_l^* = s_r^* + 1$ and $n_l - n_r \in \{n_i - 1, n_i\}$. Using Lemma 5 we obtain that i^* strictly prefers L to R for $n_l - n_r < n_i - 1$ keeping all else (namely $\boldsymbol{s}^*$ and i^*'s signal) equal.

We need to show next that if i^* prefers L to R under the condition named above, then he will prefer sending his signal l to keeping it secret. To show this we need to first identify the range of cases in which a switch from silence to sending changes the outcome. As above for all $\boldsymbol{s}^*$ with either $s_l^* > s_r^* + 1$ or $s_l^* < s_r^*$ the outcome remains the same. So let us investigate how the outcome changes in the remaining cases that $s_l^* = s_r^* + 1$ with $X = R$ and $s_l^* = s_r^*$ with $X = L$. As before the case that $s_l^* = s_r^*$ is easier to deal with, and we attack this case first.

If $s_l^* = s_r^*$ and i^* remains silent then all independents will vote for R, by our arguments above we know that this includes the i^*. To the contrary if i^* publishes his signal the public signal L decides directs all independents will to vote for L. Under the assumption that $s_l^* = s_r^*$ the signals do not reveal anything about the state of the world. The symmetry of the problem implies that the decision whether to reveal the signal or not does not entail a utility change for i^*.

So let us now have a look at the alternative case $s_l^* = s_r^* + 1$ and $X = R$. In this case i^* will sway the vote of the independents by sending his l signal. To evaluate whether he should do so i^* has to come up with a list of cases in which his sending of the signal changes the outcome of the election

[9] A plausible case in which i^* would have an incentive to vote according to the hidden signal is that n is very small and $s_l = s_r = 0$. In this case it is very unlikely there there are any other voters out there, so pivotality concerns do not matter much, i^* should follow his own signal

(switch from R to L). If i^* does not send his signal the right wing candidate receives $n_i + n_r - 1$ votes whereas the left wing candidate receives $n_l + 1$ votes. On the other hand if i^* sends the signal the right wing candidate receives n_r votes whereas the left wing candidate receives $n_l + n_i$ votes. The signal of the independent is pivotal if $n_l - n_r \leq n_i - 2$ and $n_l - n_r > -n_i$. In other words the voter's signal is pivotal if $n_l - n_r \in \{-n_i + 1, ..., n_i - 2\}$. By our above arguments we know that i^* strictly prefers L to R for any single one of these cases. It is therefore true that $EU_i(\tau, (1, l, ., f_i)) > EU_i(\tau, (1, \emptyset, ., f^*))$ for $EU_i(\tau, (1, \emptyset, ., f^*)) = \max_f EU_i(\tau, (1, \emptyset, ., f)) \geq EU_i(\tau, (0, ., ., f_i))$. An analogous argument holds for the case that the independent observed an r signal.

$\square$

9 Existence of Equilibrium

We have now reduced the problem of showing that an equilibrium exists to a problem of showing that there exist probabilities of information acquisition, λ and π, such that neither the extremists nor the independents would change their information acquisition behavior given every one else's information acquisition behavior and that all voters follow τ after the information has been acquired. For the proofs in this section it is convenient to define Δ^+ and Δ^- as functions of the information acquisition probabilities. For two given information acquisition probabilities of the independents and extremists we denote the probability increase of the correct and incorrect choice when one more signal is sent by $\Delta^+(\pi, \lambda)$ and $\Delta^-(\pi, \lambda)$ respectively.

Theorem 3 *An equilibrium τ exists.*

Proof Define a correspondence $f : [0, 1] \times [0, 1] \rightarrow [0, 1] \times [0, 1]$, by $f = h \circ g$ with g being a function and $g : [0, 1] \times [0, 1] \rightarrow \mathbb{R} \times \mathbb{R}$ such that

$$g(\pi, \lambda) = \begin{bmatrix} \frac{p}{2}\Delta^+(\pi, \lambda) \\ \frac{p}{2}\Delta^+(\pi, \lambda) - \frac{(1-p)}{2}\Delta^-(\pi, \lambda) \end{bmatrix}$$

and h being a correspondence $h : \mathbb{R} \times \mathbb{R} \rightarrow [0, 1] \times [0, 1]$ with

$$h_i(x, y) = \begin{cases} 1 \text{ if } g(x, y) > c \\ [0, 1] \text{ if } g(x, y) = c \\ 0 \text{ if } g(x, y) < c \end{cases}$$

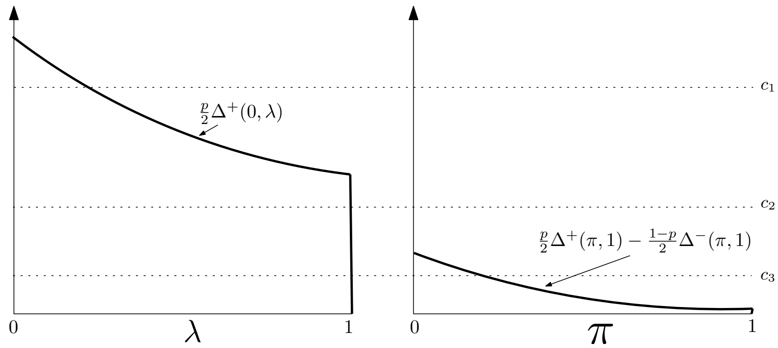


Figure 1. MB and MC

Observe that g is continuous and h is upperhemicontinuous. So f is upperhemicontinuous. We conclude by Kakutani's fixed point theorem that a fixed point exists. $\square$

The reason why we could not apply the existence result from Myerson (2000) is that strategy spaces in our game are not finite since the set of voting rules is infinite. However, we reduce the existence problem early on to a problem that would fit Myerson's existence result. As soon as we know that players will play f_{t_1} in equilibrium, we could apply Myerson's result.

We have a strong intuition that the equilibrium is unique. Indeed if the marginal benefit of information to extremists $\frac{p}{2}\Delta^+(0,\lambda)$ is strictly decreasing in λ for the case that only extremists acquire information ($\pi = 0$) and the marginal benefit to independents, $(\frac{p}{2}\Delta^+(\pi,1) - \frac{(1-p)}{2}\Delta^-(\pi,1))$ is strictly decreasing in π for the case that all extremists obtain information ($\lambda = 1$), there exists a unique equilibrium. The intuition is most clear using a diagram. Figure (1) shows the values of $\frac{p}{2}\Delta^+(0,\lambda)$ and $\frac{p}{2}\Delta^+(\pi,1) - \frac{(1-p)}{2}\Delta^-(\pi,1)$ for all equilibrium of the form $(0,\lambda)$ and of the form $(1,\pi)$. The constant marginal cost of information acquisition can cross these curves at most once if they are decreasing. If c does not cross, either information is so cheap that everyone acquires the signal, it is so expensive that no one acquires a signal, or it is between the extremist and independent curves so that extremists always acquire and independents never do.

If it were the case that the extremists had no vote, but were just information gatherers we can show that this intuition is is indeed fact as the Δ functions are decreasing (and look like the above figure).[10] Unfortunately once we attempt to take into account the small probability that an extremist group

[10] A proof of this claim follows along the same lines as our proof of Lemma 11 and is therefore omitted.

outnumbers the rest of the electorate, manipulations of the Δ function become intractable.To see this consider that just $\Delta^{+}(\pi,\lambda)$ can be written as:

$$\frac{1}{2}\sum_{n_i,n_r,n_l=0}^{\infty} e^{-n}\frac{(n\eta)^{n_r+n_l}(n(1-2\eta)^{n_i})}{n_r!n_l!n_i!}I(n_r\leq n_i+n_l,n_l\leq n_i+n_r)$$
$$\cdot[\Pr(s_l=s_r|\tau,n_i,n_r,n_l,r)+\Pr(s_l=s_r+1|\tau,n_i,n_r,n_l,r)]$$

As such uniqueness is left uncertain.

10 Comparative Statics

In this section we show that the total amount of public information converges to a positive number x as n grows large. This number does not depend on η as long as η is positive. For large n only extremists acquire information. If there are no extremists, $\eta = 0$, then the total amount of public information converges to a positive number $y < x$. It is useful to define the following events for this discussion:

Definition 2 *The events $P_1, \overline{P_1}$ and P_2 are defined as follows:*

$$\begin{aligned}
P_1 &:= \{\boldsymbol{n}|n_r \leq n_i + n_l \text{ and } n_l \leq n_r + n_i\}\\
\overline{P_1} &:= \{\boldsymbol{n}|n_r > n_i + n_l \text{ or } n_l > n_r + n_i\}\\
P_2 &:= \{\boldsymbol{s}, X|f_i(\boldsymbol{s}, X) = R, f_i((s_l, s_r - 1), X) = L\}
\end{aligned}$$

P_1 denotes the event that the independents decide the vote. $\overline{P_1}$ denotes the complementary event in which a group of extremists holds an absolute majority. P_2 denotes the event that an r signal is pivotal in the sense the independents would vote R if the signal was sent but would vote L if it was withheld. Denote the probability that event E occurs for the parameters n, η, p and the strategy profile τ by $Prob(E|n, \eta, p, \tau)$. Fixing η and p we denote a sequence of equilibrium profiles for every $n \in \mathbb{N}$ by τ_n.

The broadcast of an r signal changes the outcome of the vote in the event $P_1 \cap P_2$. If τ_n is such that the extremists mix between acquiring and not acquiring information, then we have that $\frac{p}{2}Prob(P_1 \cap P_2|n, \eta, p, \tau_n) = c$. We show next that we only need to be concerned with this case in our analysis of information acquisition for large n. For large n all independents will rely on the information acquired by the extremists they will not acquire any information themselves.

Lemma 7 *There exists an $\overline{n} \in \mathbb{N}$ such that $\pi_n = 0$ for all $n > \overline{n}$.*

Proof Suppose not. Then we can find an infinite subsequence of τ'_n of τ_n such that $\pi'_n > 0$ for all elements of this sequence. This implies that $\frac{p}{2} Prob(P_1 \cap P_2 | n, \eta, p, tau'_n) > c$ for all elements of the sequence. Observe that $Prob(P_2|n, \eta, p, \tau'_n) > Prob(P_1 \cap P_2|n, \eta, p, \tau'_n)$. As a consequence of Lemma 10 in the appendix $Prob(P_2|n, \eta, p, \tau'_n)$ converges to 0 as n grows large, a contradiction with $Prob(P_2|n, \eta, p, \tau'_n) > c$ for all $n \in \mathbb{N}$. $\square$

We infer $\frac{p}{2} Prob(P_1 \cap P_2|n, \eta, p, \tau_n) = c$ is true in equilibrium for large n. This is (as shown in the end of the preceding section) an unwieldy expression. Luckily the event that a group of extremists decides the election ($\overline{P_1}$) becomes more and more unlikely as n grows large. As n becomes large the expression $Prob(P_1 \cap P_2|n, \eta, p, \tau_n)$ can be approximated by $Prob(P_2|n, \eta, p, \tau_n)$, the probability that an additional r signal sways the vote of the independents. We define $x(p) = n\lambda\eta$ as the expected amount of information acquisition for which $\frac{p}{2} Prob(P_2|n, \eta, p, \tau) = c$ holds. Formally we define $x(p)$ as the solution to

$$c = \frac{p}{4} \sum_{k=0}^{\infty} e^{-x} \Big(\frac{(px)^k ((1-p)x)^k}{k!k!} + \frac{(px)^k ((1-p)x)^{k+1})}{(k+1)!k!} \Big) \tag{7}$$

It is to be expected that $n\lambda_n\eta$ converges to $x(p)$ for all sequences of equilibria τ_n. We claim exactly that in our next lemma.

Lemma 8 *$n\lambda_n\eta$ converges to $x(p)$ as n goes to infinity.*

Proof As a first observation we show in the appendix that $x(p)$ is well defined and unique (see appendix A). Observe then that for any sequence of equilibria τ_n we have that $|Prob(P_1 \cap P_3|n, \eta, p, \tau_n) - Prob(P_3|n, \eta, p, \tau_n)|$ converges to 0. To see this observe that the difference between these two can be rewritten as

$$\begin{aligned} Prob(\overline{P_1} \cap P_2|n, \eta, p, \tau_n) = \\ \sum_{\boldsymbol{n}} Prob(\boldsymbol{n}) I(n_r > n_i + n_l \text{ or } n_l > n_r + n_i) Prob(P_2|\boldsymbol{n}, p, \tau_n) \leq \\ \sum_{\boldsymbol{n}} Prob(\boldsymbol{n}) I(n_r > n_i + n_l \text{ or } n_l > n_r + n_i) \end{aligned}$$

we show in the appendix that this latter term converges to zero (see appendix B).

By Lemma 7 we also know that there exists an $\overline{n} \in \mathbb{N}$ such that $Prob(P_1 \cap P_2|n, \eta, p, \tau_n) = c$ for all $n > \overline{n}$ as from some n on only the extremists will acquire information. They are willing to mix if and only if the above equation holds. Taken together with the above observation on the limit behavior we know that $Prob(P_2|n, \eta, p, \tau_n)$ converges to c as n grows large. Finally observe that $Prob(P_2|n, \eta, p, \tau_n)$ is a continuous function in $n\eta\lambda_n$. So $n\eta\lambda_n$ must converge to a unique $x(p)$. $\square$

When $\eta = 0$ there are only independents in the electorate. For any equilibrium in which independents don't always acquire information (that is $\pi_n < 1$) the expected amount of public signals is $y(p)$, the solution to the following equation, which equalizes the cost and the benefit of information acquisition for the independents.

$$
\begin{aligned}
c = \frac{p}{4}\sum_{k=0}^{\infty} e^{-x}\Big(\frac{(px)^k((1-p)x)^k}{k!k!} + \frac{(px)^k((1-p)x)^{k+1})}{(k+1)!k!}\Big) - \\
\frac{1-p}{4}\sum_{k=0}^{\infty} e^{-x}\Big(\frac{(px)^k((1-p)x)^k}{k!k!} + \frac{(px)^{k+1}((1-p)x)^k)}{(k+1)!k!}\Big) = \\
\frac{2p-1}{4}\sum_{k=0}^{\infty} e^{-x}\frac{(px)^k((1-p)x)^k)}{k!k!} + \\
\frac{1}{4}\sum_{k=0}^{\infty} e^{-x}\Big(\frac{p(px)^k((1-p)x)^{k+1})}{k!(k+1)!} - \frac{(px)^{k+1}((1-p)x)^k(1-p))}{k!(k+1)!}\Big) = \\
\frac{2p-1}{4}\sum_{k=0}^{\infty} e^{-x}\frac{(px)^k((1-p)x)^k)}{k!k!}
\end{aligned}
$$

Lemma 9 *$n\pi_n$ converges to $y(p)$ as n goes to infinity.*

Proof The proof of Lemma 7 implies that there exists an $\overline{n} \in \mathbb{N}$ such that $\pi_n < 1$ in equilibrium for all $n > \overline{n}$. So for $n > \overline{n}$

$$
c = \frac{2p-1}{4}\sum_{k=0}^{\infty} e^{-y(p))}\frac{(py(p))^k((1-p)y(p))^k)}{k!k!}
$$

has to hold in equilibrium, for $y(p) = n\pi_n$. In the appendix we show that $y(p)$ is well defined and unique (see appendix A). $\square$

The main result of this section is that for any positive share of extremists $\eta > 0$ there exists a number $\overline{n} \in \mathbb{N}$ such that the expected information acquisition in equilibrium is larger than the expected equilibrium information acquisition if there where only independents.

Theorem 4 *$x(p) > y(p)$ for all $p \in (\frac{1}{2}, 1)$.*

Proof See appendix C $\square$

So extremists do not only have a higher incentive to acquire information in any equilibrium, Theorem 4 shows that large electorates with extremists acquire strictly more information than large electorates without extremists. This difference can be significant for some parameter values. For instance if $p = 0.51$ and $c = 0.01$, when there are no extremists the expected number of broadcast signals is 2.66. When there are some extremists in the environment the expected amount of signals broadcast jumps to 580.75. As expected, as p approaches 1, representing fully revealing signals, the difference between the two cases converges to zero.

11 Less extreme extremists

The assumption that extremists always weakly prefer their candidate is only a simplifying assumption in our model. In this section we argue that it does not drive the main result. Even if all extremists *strictly* prefer candidate L in state l and candidate R in r; their incentive to acquire information should be stronger than the independent's. Matters become much more complicated; so we do not provide formal proofs of all our claims.

Consider the model described in section 2 and replace the assumption that the parameter δ equals 0 for the left wing extremists and 1 for right wing extremists by the assumption that $\delta = \epsilon$ for left wing extremists and $\delta = 1 - \epsilon$ for right wing extremists for some small positive ϵ. So left wing extremists suffer a small loss ϵ if L is elected in r, right wing extremists suffer the same small loss in case of the alternative mistake in which R is elected in state l.

Just as for the model in the main body of the text define $\Delta^z_{Y,\omega}$ as the change in probability that candidate Y is being chosen in state ω when another z signal is being sent. Formally

$$\Delta^z_{Y,\omega} := Pr(Y + z|\tau, \omega) - Pr(Y|\tau, \omega).$$

Assuming that all voters behave symmetrically[11] we are still able to define $\Delta^+ := Pr(R+r|\tau, r) - Pr(R|\tau, r) = Pr(L+l|\tau, l) - Pr(L|\tau, l)$ and $\Delta^- := Pr(L+r|\tau, l) - Pr(L|\tau, l)$, the increase in probability of the correct and the wrong choice when another signal is being sent. Assuming furthermore that more left wing signals make a victory of L more likely we have that $\Delta^+, \Delta^- > 0$.

Using the same arguments as in the main model we focus only on the set of pure strategies consisting of $X = \{(0, \emptyset, \emptyset, f), (1, r, \emptyset, f), (1, \emptyset, l, f), (1, r, l, f)\}$. Since inaction is the outside option; we compare all strategies to $(0, \emptyset, \emptyset, f)$.

$$\begin{aligned}
EU_\delta(\tau, (1, r, \emptyset, f)) - EU_\delta(\tau, (0, \emptyset, \emptyset, f)) &= \frac{p}{2}\delta\Delta^+ - \frac{1-p}{2}(1-\delta)\Delta^- - c \\
EU_\delta(\tau, (1, \emptyset, l, f)) - EU_\delta(\tau, (0, \emptyset, \emptyset, f)) &= \frac{p}{2}(1-\delta)\Delta^+ - \frac{1-p}{2}\delta\Delta^- - c \\
EU_\delta(\tau, (1, r, l, f)) - EU_\delta(\tau, (0, \emptyset, \emptyset, f)) &= \frac{p}{2}\Delta^+ - \frac{(1-p)}{2}\Delta^- - c
\end{aligned}$$

Observe that a left winger sending only left wing signals receives the same utility of a right winger sending only right wing signals, as the parameters of the left and right wingers are $\delta = \epsilon$ and $\delta = 1-\epsilon$. For an extremist, the utility difference between only sending a signal in their favor and not sending any signal is $\frac{p}{2}(1-\epsilon)\Delta^+ - \frac{1-p}{2}\epsilon\Delta^- - c$. Furthermore, left wing extremists would never only send right wing signals and vice versa, as $EU_\epsilon(\tau, (1, \emptyset, l, f)) - EU_\delta(\tau, (1, r, \emptyset, f)) = \frac{p}{2}\Delta^+ + \frac{1-p}{2}\Delta^- > 0$.

Finally and most importantly observe that for small ϵ, the case of an extremist, $EU_\epsilon(\tau, (1, \emptyset, l, f)) - EU_\epsilon(\tau, (0, \emptyset, \emptyset, f)) = \frac{p}{2}(1-\epsilon)\Delta^+ - \frac{1-p}{2}\epsilon\Delta^- - c$ is close to $\frac{p}{2}\Delta^+ - c$ which is of course larger than $EU_\delta(\tau, (1, r, l, f)) - EU_\delta(\tau, (0, \emptyset, \emptyset, f)) = \frac{p}{2}\Delta^+ - \frac{(1-p)}{2}\Delta^- - c$ the utility differential between sending both signals and not acquiring information. Notice that $EU_\delta(\tau, (1, r, l, f)) - EU_\delta(\tau, (0, \emptyset, \emptyset, f))$ has the same value for voters of all types. Since $(1, l, r, f)$ is the independent strategy, we can now conclude that for small ϵ and symmetric strategies the value of information acquisition is strictly higher for extremists.

Note that the arguments given here stayed somewhat informal. This is due to our difficulty to explicitly characterize τ. In the main text we imposed sincere voting to derive that Δ^+ and Δ^- are well-defined and positive, which are the

[11] We call a strategy profile τ symmetric if $\tau = \overline{\tau}$, where $\overline{\tau}$ is generated using τ be replacing l with r and R with L. For a relaxation of the model that allows for asymmetries see section 12

only assumptions necessary to derive the above result. We also showed that sincere voting was consistent with strategic voting. We cannot pursue this route here as sincere and strategic voting need not be consistent in the model with less extreme extremists. To see this consider a left wing extremist with $\delta = \epsilon \in (0, 1/2)$. Observe that this extremist would prefer that R wins the election when the probability that the state of the world is r is sufficiently high.

We can calculate a precise cutoff on the signals such that a sincere left wing voter would vote for R if this threshold was met. However, a strategic left wing voters should use a lower threshold. To see this observe that a left wing voter would cast the pivotal vote only if there are (approximately) equally many left and right wing votes. What can this voter infer from the event of being pivotal? Given that the right wing signals outnumber the left wing signals it can be inferred that all right wing extremists and all independents voted for the right wing candidate. For there to be an electoral tie half of the electorate must consist of left wing extremists; implying the number of left wing extremists is at least as large as the number of right wing extremists. On the other hand, more right wing signals have been sent. The pivotal left winger deduces that many more right wing signals remained hidden compared to left wing signals. In short, if a sincere left wing voter is indifferent between voting for either candidate and if all voters vote sincerely; a strategic left wing extremist should vote for the right wing candidate. We can therefore no longer use sincere voting as a refinement criterium to solve for equilibria of the voting game.

Given that we cannot use sincere voting as a refinement criterium we cannot solve for τ, however we showed above that the result that extremists have higher incentives to acquire information holds as long as τ is symmetric. The arguments in section 12 extend our analysis to asymmetric environments. A combination of the present section with that section could extend our argument here to a case in which τ is not symmetric. Finally observe that there should be equilibrium strategies τ that do not differ "too much" from sincere voting. To see this observe that the case in which left wing extremists prefer R should happen only very rarely for the case of small enough ϵ. First of all for a small ϵ the right state would have to be a lot more probable than the left state for a left wing extremist to prefer R. The super-majority of right wing signals should be large to convince a left wing extremist to vote for the right. However given that information acquisition is costly the expected total number of signals acquired will not grow infinitely large even as the electorate grows beyond all bounds. Thus if ϵ is small enough such a super-majority is very unlikely to arise. We conclude that while the assumption of "extreme extremists" allows us to solve the model, the main results on the incentives to acquire information should also hold true in a model with less extreme extremists.

12 Asymmetries

The model contains three symmetry assumptions. The two states are assumed to happen with equal probability, both signals are equally strong, and the expected share of right and left wingers is equal. In the solution concept we impose two further symmetry conditions: in equilibrium both types of extremists are equally likely not to acquire information and the independents will pass on both signals. What are these assumptions needed for? How would the results change if we were to drop some (or all) of these assumptions?

As an answer to the first question observe that our proof that strategic voters do not have an incentive to deviate from sincere voting depend on our symmetry assumptions. To see this, consider an electorate in which there are an average 10 leftists and 1000 rightists. Consider the case of a prior that slightly favors state l and assume that both signals are equally strong and in equilibrium all extremists acquire information. If $s_r = s_l = 0$, sincere independent should vote for R. He should expect that there are about 10 leftists and 1000 rightists in the electorate. Since all extremists stayed silent and since all extremists acquire information according to the strategy profile, the independent should infer that there are about 10 people hiding a right wing signal and 1000 people hiding a left wing signal. Since the prior favors l only slightly, the information contained in $s_l = s_r = 0$ should outweigh the prior. Now let us reconsider this question from the point of view of a strategic independent. Assume that the electorate is small and assume that independents are "nearly never" born. So the pivotality event contains the information that there are approximately equally many rightists and leftists. This implies that they hide approximately equally many right and left signals. So a strategic voter should vote according to the prior for L since $s_r = s_l = 0$ together with the pivotality event do not reveal any information about the state of the world.

In sum, some of our results do not extend to the general case, since equilibrium requirements 1, 3, and 4 need not be consistent without the symmetry conditions. We therefore follow a different route in this section. We complement the relaxation of the symmetry assumptions with the imposition of two behavioral assumptions: We assume that voters vote sincerely, and that independents share all their information. We do not require that these two assumption be compatible with strategic voting and strategic information transmission, these assumptions need not constitute a refinement as they do in the main body of the text. With these behavioral assumptions we are able to establish a version of our main result for the case of asymmetric equilibria in asymmetric models. To deal with the proposed asymmetries we need more notation, which we will introduce as part of the setup of our the modified model. For brevity we refer the reader to the prior sections for a discussion of our setup.

12.1 Asymmetric Model

Take the Poisson game $\{\Omega, \alpha, T, n, p, C, U\}$ game defined in Section 2 and modify it such that the common prior that the state is l is $\alpha \in (0, 1)$. Assume that $p_1(t_1 = l|\omega) = \eta_l,\ p_1(t_1 = r|\omega) = \eta_r$ and $\eta_l+\eta_r < 1$ for $\omega \in \{l, r\}$ implying that a voter is a leftist with probability η_l and a rightist with probability η_r. Finally let $p_2(t_2 = l|\omega = l) = p_l;\ p_2(t_2 = r|\omega = r) = p_r$, this implies that the signals l and r might have different strength. It might for example be easier to obtain evidence that a dictator owns weapons of mass destruction, in case that he indeed owns some, than it is obtain evidence that he has no such weapons in the complementary case.

The prior model arises as a special case for $\alpha = \frac{1}{2}$, $p_l = p_r$, $\eta_r = \eta_l$ and the assumption that the electorate is "large" $((1-2\eta)n > 1)$. The last assumption was only used to establish that independents have no strategic incentive to deviate from sincere voting. We now have the behavioral assumption of sincere voting. Consequently the assumption that the electorate is large is no longer needed.

12.2 Strategies and Equilibrium

With the behavioral assumptions strategies can be described by shorter vectors in the modified game. In fact the strategy of independents can be described by their probability to acquire information π alone (their voting and information sharing behavior is fixed according to our modelling assumption.) Strictly speaking the strategy vectors of extremists should have two entries as we did not make any behavioral assumption on their information sharing behavior. However, just as in the above model it is straightforward to show that extremists will always broadcast signals in favor of their candidate and never broadcast signals in favor of the other candidate. We therefore suppress this strategy variable and define the strategy of leftists by λ_l the probability that leftists acquire information and λ_r respectively for rightists.

We solve this game for an equilibrium following Myerson's definition of an equilibrium in an extended Poisson game (Myerson 1998). The vector $(\lambda_l, \lambda_r, \pi)$ is an equilibrium if no voter has an incentive to deviate from their (information acquisition) strategy given everyone else's (information acquisition) strategy. It is important to note that this definition of equilibrium does allow for asymmetric strategies, we do not impose that $\lambda_r = \lambda_l$. Also note that Myerson's equilibrium existence result applies directly to the games in this family.

12.3 Solution

Voters are assumed to vote sincerely. Just as in the original model this implies that all extremists will vote for their preferred candidate. We calculate an asymmetric analog of the independents cutoff rule. Just as in the prior setup s_r and s_l are Poisson distributed random variables with parameters x_l^l, x_l^r, x_r^r and x_l^r, where the superscript denotes the state and the subscript denotes the signal:

$$\begin{aligned} x_l^l &= n\eta_l\lambda_l p_l + n(1-\eta_l-\eta_r)\pi p_l \\ x_r^l &= n\eta_r\lambda_r(1-p_l) + n(1-\eta_l-\eta_r)\pi(1-p_l) \\ x_l^r &= n\eta_l\lambda_l(1-p_r) + n(1-\eta_l-\eta_r)\pi(1-p_r) \\ x_r^r &= n\eta_r\lambda_r p_r + n(1-\eta_l-\eta_r)\pi p_r \end{aligned}$$

So state l is at least as likely as state r if and only if

$$\alpha\frac{e^{-x_l^l}[x_l^l]^{s_l}}{s_l!}\cdot\frac{e^{-x_r^l}[x_r^l]^{s_r}}{s_r!} \geq (1-\alpha)\frac{e^{-x_r^l}[x_r^l]^{s_l}}{s_l!}\cdot\frac{e^{-x_r^r}[x_r^r]^{s_r}}{s_r!}$$

Observe this holds expression if and only if

$$ln(\frac{\alpha}{1-\alpha}) + x_r^r + x_l^r - x_l^l - x_r^l \geq s_l ln\left(\frac{x_l^r}{x_l^l}\right) - s_r ln\left(\frac{x_r^l}{x_r^r}\right)$$

Substituting in the values for x_l^l, x_l^r, x_r^r and x_l^r we obtain that

$$s_l ln\left(\frac{p_l}{1-p_r}\right) \geq ln(\frac{1-\alpha}{\alpha}) + n(p_r+p_l-1)(\eta_l\lambda_l - \eta_r\lambda_r) + s_r ln\left(\frac{p_r}{1-p_l}\right)$$

This expression reduces to the previous cutoff rule if $\alpha = \frac{1}{2},\ p_l = p_r,\ \eta_r = \eta_l$, and $\lambda_r = \lambda_l$. This condition has some nice intuitive properties: to make their decision the independents will multiply the raw data s_l, s_r with factors $ln\left(\frac{p_l}{1-p_r}\right)$ and $ln\left(\frac{p_r}{1-p_l}\right)$ that reflect the relative ease or difficulty to obtain l or r signals in either state. The larger is $\frac{p_r}{1-p_l}$ the more meaningful are the r signals. In the expression the total amount of r signals is multiplied by this indicator of the informativeness. In the extremely noninformative case that $p_r = p_l = \frac{1}{2}$ the signals are meaningless, and the decision will be based on the prior alone. The above inequality is either always true or never true, depending on the value of α.

If r is more likely following the prior more l-signals are needed to sway the independents to vote L. This is reflected by the term $ln\left(\frac{\alpha}{1-\alpha}\right)$. Finally the term $n(p_r + p_l - 1)(\eta_r\lambda_r - \eta_l\lambda_l)$ is equal to zero if $\eta_r\lambda_r = \eta_l\lambda_l$. In this case any voter is equally likely to be either a right winger that acquires information or a left winger that acquires information; the only asymmetries in the decision rule should be attributable to the prior (α) and to the different signal strength (the factors $ln\left(\frac{p_r}{1-p_l}\right)$ and $ln\left(\frac{p_l}{1-p_r}\right)$ as discussed above). If this does not hold, say if $\eta_r\lambda_r > \eta_l\lambda_l$, then independents need more r signals to be convinced in R's favor, as there are on average more right wingers that send such signals than there are left wingers. This intuition holds true as $\eta_r\lambda_r - \eta_l\lambda_l$ is multiplied by a positive factor $n(p_l + p_r - 1)$ in this equation. As expected this factor depends on the average size of the electorate (n) and a measure of the value of signals.

Recall that $\Delta^z_{Y,\omega}$ was defined as the change in probability that candidate Y is being chosen in state ω when another z signal is being sent. Just as in the asymmetric case it is true that the probability that a candidate wins is increasing in the number of signals sent in his favor: $\Delta^l_{L,\omega}, \Delta^r_{R,\omega}$ for $\omega = l, r$ Without the various symmetry assumptions it is however not necessarily true that $\Delta^r_{R,l} = \Delta^l_{L,r}$ and $\Delta^l_{L,l} = \Delta^r_{R,r}$.

Following the arguments given in section 6 we can establish a weaker version of the result that in equilibrium extremists have stronger incentives to acquire information. We state this result as a separate theorem.

Theorem 5 *In an equilibrium the utility gain from acquiring information for an independent is never greater than the utility gain of either type of extremist.*

The proof is similar to the proof of Theorem 1. It is therefore kept very short.

Proof The utility of acquiring information for the right and left wing extremists are respectively:

$$X = \frac{p}{2}\Delta^r_{R,r} - c \text{ and } Y = \frac{p}{2}\Delta^l_{L,l} - c.$$

We calculate the utility of information acquisition for an independent as

$$Z = \frac{p}{4}(\Delta^l_{L,l} + \Delta^r_{R,r}) - \frac{1-p}{4}(\Delta^r_{R,l} + \Delta^l_{L,r}) - c = \frac{X+Y}{2} - \frac{1-p}{4}(\Delta^r_{R,l} + \Delta^l_{L,r}).$$

Finally observe that $\Delta^r_{R,l}, \Delta^l_{L,r} > 0$ so we have that $Z < \frac{X+Y}{2}$ and it cannot be true that both extremists receive a lower utility from information acquisition than the independent does. □

Corollary 2 *If the equilibrium probability that independents acquire information is positive then at least one set of extremists acquires information in equilibrium. If the left and the right wing extremists abstain form information acquisition with positive probability, then no independent acquires any information in equilibrium.*

Proof In analogy to the proof of corollary 1 for the symmetric case. □

13 Conclusion

We believe the main contribution of this paper is to show that using standard environmental assumptions it is possible to theoretically explain the information acquisition by extreme voters that has been noted in the empirical literature. We also view this as an early attempt to examine the effects of communication in an electoral model. This paper shows that a simple communication mechanism can significantly alter the predictions of a model without communication; leading us to believe that models that integrate communication and voting deserve more attention in future work.

The communication structure in this model is about as simple as we could think of and could be enriched. One way to envision this is to add some form of search or matching mechanism so that any voter's message is restricted to a voter-specific subset of the electorate. We believe that this would be interesting, however it is not clear if it would change the results of the model. An extremist would have less opportunity to convey their signal which would decrease the incentive to obtain a signal. On the other hand the incentive to acquire information would be increased by the fact each voter would now be less informed, so for an extremist each signal sent would carry more weight for its recipient.

An application of this model to the media industry could provide a theoretical foundation for media bias. Baron (2006) and Bernhardt, Krasa and Polborn (2006) offer two different explanations for media bias. Bernhardt et al. (2006)'s explanation for media bias is demand driven; they use the assumption that more extreme voters derive a higher consumption value from being informed to explain media bias. In contrast, Baron (2006)'s explanation for media bias is supply driven; Baron (2006) assumes that journalists prefer to offer biased stories. Baron justifies this assumption with the argument that more biased and possibly scandalous stories increase the career perspectives of journalists. Our model could be used as a micro foundation of a very similar assumption.

Say that news agencies sell signals. They acquire signals from independent journalists and pay a competitive market rate per signal. In such a model the sale of a signal to the news agency can be equated with the the broadcast of a signal in our model. Any person who intends to become a journalist has to compare the utility of not being a journalist with the utility of the wage received, the learning cost incurred, and the benefit from broadcasting a signal. The latter is highest for the most extreme members of the polity. Consequently only the most extreme members of a polity will take up journalism in a competitive market. Our results also indicate that it may be optimal for a large independent majority to pay a few opposing extremists to gather information rather than pay some independents do this. The extremists are willing to acquire the same amount of information at a cheaper price.

Finally we believe that it would be useful to consider ways to join this type of model to the results of Feddersen and Pesendorfer (1996). We know that extremists tend to be better informed and the swing voters curse predicts that voters that are more informed are more likely to vote. This would lead one to say that by the swing voters curse, extremists should end up voting more often. This claim would be verified by the data as this is another major observation of Palfrey and Poole (1987). In our model a swing voters boon prevails, no voters will abstain, and can therefore not yield this result. It would be a useful endeavor to set up a model of pre-electoral communication which generates both the correlation between extremism and information and voting and information.

APPENDIX

A Proof that $x(p)$ and $y(p)$ are Unique

$x(p)$ is defined as the solution to:

$$c = \frac{p}{4}\sum_{k=0}^{\infty} e^{-x}\Big(\frac{(px)^k((1-p)x)^k}{k!k!} + \frac{(px)^k((1-p)x)^{k+1})}{(k+1)!k!}\Big) \tag{A.1}$$

and $y(p)$ as the solution to

$$c = \frac{2p-1}{4}\sum_{k=0}^{\infty} e^{-y}\frac{(py)^k((1-p)y)^k)}{k!k!} \tag{A.2}$$

It is clear that the above equations are continuous in x and y respectively, however to show that this equation has a solution we need to do some more work.

A.1 Mathematical Preliminary

Before going into the proof of this we review a few established mathematical results. The first is a result due to Skellam (1946):

$$\sum_{k=0}^{\infty} e^{-x}\frac{(px)^k((1-p)x)^{k-n}}{k!(k-n)!} = s(n, px, (1-p)x)$$
$$= e^{-(px+(1-p)x)}\left(\frac{px}{(1-p)x}\right)^{\frac{n}{2}} I_n\left(2x\sqrt{p(1-p)}\right) \;\forall\; n = 0, 1, 2, \ldots \tag{A.3}$$

Where $I_n(z)$ is a Modified Bessel function of the first kind. Here we list some interesting properties found in Abramowitz and Stegun, eds (1972). Where possible for the following expressions the equation numbers will be the same as in Abramowitz and Stegun, eds (1972).

$$I_0(0) = 1 \tag{A.4}$$
$$I_n(0) = 0 \ \forall \ n = 1, 2, 3, \ldots \tag{A.5}$$
$$I_{-n}(z) = I_n(z) \tag{9.6.6}$$
$$I_{n+2}(z) = I_n(z) - \frac{2(n+1)}{z} I_{n+1}(z) \tag{9.6.26}$$
$$I_n'(z) = \frac{I_{n-1}(z) + I_{n+1}(z)}{2} \tag{9.6.26}$$
$$e^z = I_0(z) + 2\sum_{n=1}^{\infty} I_n(z) \tag{9.6.37}$$

A.2 Characteristics of x(p)

We will show that when c is not too high, $x(p)$ is defined and unique. To do this we will show that the right hand side of equation (A.1) is strictly decreasing and converges to zero.

First we can use equation (A.3) to describe equation (A.1) as:

$$c = \frac{p}{4}[s(0, px, (1-p)x) + s(-1, px, (1-p)x)] \tag{A.6}$$

Let $\theta_x = 2x\sqrt{p(1-p)}$. Then using the definition of $s(n, px, (1-p)x)$:

$$c = \frac{p}{4}e^{-x}\left[I_0(\theta_x) + \left(\frac{p}{1-p}\right)^{-\frac{1}{2}} I_{-1}(\theta_x)\right]$$

Remark 1 With the above expression, we can use equations (A.4) and (A.5) to see that when $x = 0$ the right hand side evaluates to $\frac{p}{4}$.

Lemma 10 $\lim_{x\to\infty}\left\{\frac{p}{4}e^{-x}\left[I_0(\theta_x) + \left(\frac{p}{1-p}\right)^{-\frac{1}{2}} I_{-1}(\theta_x)\right]\right\} = 0$

Proof Using equation (9.6.6) and distributing we can rewrite the expression as:

$$\lim_{x\to\infty}\left\{\frac{p}{4}\left[\frac{I_0(\theta_x)}{e^x} + \left(\frac{p}{1-p}\right)^{-\frac{1}{2}}\frac{I_1(\theta_x)}{e^x}\right]\right\}$$

Using equation (9.6.37) we can expand e^x and end up with:

$$\lim_{x\to\infty}\left\{\frac{p}{4}\left[\frac{I_0(\theta_x)}{I_0(x) + 2\sum_{n=1}^{\infty} I_n(x)} + \left(\frac{p}{1-p}\right)^{-\frac{1}{2}}\frac{I_1(\theta_x)}{I_0(x) + 2I_1(x) + 2\sum_{n=2}^{\infty} I_n(x)}\right]\right\}$$

Note that $I_n(2x\sqrt{p(1-p)}) \le I_n(x)\ \forall\ x \ge 0,\ n = 0, 1, 2, \ldots$ since $I_n(\cdot)$ is monotonically increasing in its argument and $2x\sqrt{p(1-p)} \le x$. Since this is so we have:

$$0 \le \lim_{x\to\infty} \frac{I_0(\theta_x)}{I_0(x) + 2\sum_{n=1}^{\infty} I_n(x)} \le \lim_{x\to\infty} \frac{I_0(x)}{I_0(x) + 2\sum_{n=1}^{\infty} I_n(x)} = 0$$

$$0 \le \lim_{x\to\infty} \frac{I_1(\theta_x)}{I_0(x) + 2I_1(x) + 2\sum_{n=2}^{\infty} I_n(x)} \le \lim_{x\to\infty} \frac{I_1(x)}{I_0(x) + 2I_1(x) + 2\sum_{n=2}^{\infty} I_n(x)} = 0$$

Which is true since $\forall\ x \ge 0,\ n = 0, 1, 2, \ldots\ I_n(x) \ge 0$ and $I_n'(x) > 0$. This proves the claim. $\square$

Now we know that the right hand side of equation (A.1) is equal to $\frac{p}{4}$ when $x = 0$ and approaches 0 as $x \to \infty$. If the expression is strictly decreasing in x we have shown that $x(p)$ is defined and unique.

Lemma 11 $\frac{d}{dx}\left\{\frac{p}{4}e^{-x}\left[I_0(\theta_x) + \left(\frac{1-p}{p}\right)^{\frac{1}{2}} I_{-1}(\theta_x)\right]\right\} < 0$

Proof

$$\begin{aligned}
&\frac{d}{dx}\left\{\frac{p}{4}e^{-x}\left[I_0(\theta_x) + \left(\frac{1-p}{p}\right)^{\frac{1}{2}} I_{-1}(\theta_x)\right]\right\} \\
&= \frac{p}{4}e^{-x}\left\{2\sqrt{p(1-p)}I_1(\theta_x) + (1-p)[I_0(\theta_x) + I_2(\theta_x)]\right\} \\
&\quad -\frac{p}{4}e^{-x}\left\{I_0(\theta_x) + \left(\frac{1-p}{p}\right)^{\frac{1}{2}} I_1(\theta_x)\right\} \\
&= \frac{p}{4}e^{-x}\left[(1-p)I_2(\theta_x) - pI_0(\theta_x) + (2p-1)\left(\frac{1-p}{p}\right)^{\frac{1}{2}} I_1(\theta_x)\right]
\end{aligned}$$

Using equation (9.6.26) we can replace $I_2(\theta_x)$ with $I_0(\theta_x) - \frac{2}{\theta_x}I_1(\theta_x)$ to get:

$$\begin{aligned}
&\tfrac{p}{4}e^{-x}\left[(1-p)I_0(\theta_x) - (1-p)\tfrac{2}{\theta_x}I_1(\theta_x) - pI_0(\theta_x) + (2p-1)\left(\tfrac{1-p}{p}\right)^{\frac{1}{2}} I_1(\theta_x)\right] \\
&= \tfrac{p}{4}e^{-x}\left[(1-2p)I_0(\theta_x) + [(2p-1)\left(\tfrac{1-p}{p}\right)^{\frac{1}{2}} - (1-p)\tfrac{2}{\theta_x}]I_1(\theta_x)\right]
\end{aligned}$$

We want to show that this is negative. Since $I_n(x) \geq \; 0 \forall x \geq 0$ we can say:

$$\begin{aligned}
&(1-2p)I_0(\theta_x) + [(2p-1)\left(\tfrac{1-p}{p}\right)^{\frac{1}{2}} - (1-p)\tfrac{2}{\theta_x}]I_1(\theta_x) \\
&\leq (1-2p)I_0(\theta_x) + (2p-1)\left(\tfrac{1-p}{p}\right)^{\frac{1}{2}} I_1(\theta_x) \\
&\leq \quad (1-2p)I_0(\theta_x) + (2p-1)I_1(\theta_x) \\
&= \quad (2p-1)I_1(\theta_x) - (2p-1)I_0(\theta_x) \\
&= \quad (2p-1)[I_1(\theta_x) - I_0(\theta_x)] < 0
\end{aligned}$$

Which shows that equation (A.1) is strictly decreasing in x. □

As long as $c \leq \frac{p}{4}$, $x(p)$ is well defined and unique.

A.3 Characteristics of y(p)

Following the techniques used above we can rewrite equation (A.2) as:

$$c = \left(\frac{2p-1}{4}\right)s(0, px, (1-p)x)$$

If we then let $\theta_y = 2y\sqrt{p(1-p)}$ we can say:

$$c = \left(\frac{2p-1}{4}\right)e^{-y}I_0(\theta_y)$$

Note that when $y = 0$ the right hand side evaluates to $\frac{2p-1}{4}$. We can use the same argument as above to show that

$$\lim_{y\to\infty}\left(\frac{2p-1}{4}\right)e^{-y}I_0(\theta_y) = 0$$

which only leaves us with showing that the right hand expression is decreasing in y.

Lemma 12 $\frac{d}{dy}\left\{\left(\frac{2p-1}{4}\right)e^{-y}I_0(\theta_y)\right\} < 0$

Proof

$$\begin{aligned}
&\frac{d}{dy}\left\{\left(\frac{2p-1}{4}\right)e^{-y}I_0(\theta_y)\right\} \\
=&\left(\frac{2p-1}{4}\right)e^{-y}[2\sqrt{p(1-p)}I_1(\theta_y)-I_0(\theta_y)] \\
&\leq \left(\frac{2p-1}{4}\right)e^{-y}[I_1(\theta_y)-I_0(\theta_y)] < 0
\end{aligned}$$

Where the last inequality comes from the fact that $\forall\ x \geq 0,\ I_0(x) > I_1(x)$. □

So $y(p)$ is also defined and unique when $c \leq \frac{2p-1}{4}$.

B Limit Properties of P_2

Lemma 13 *As $n \to \infty$, the probability that the extremists are pivotal,* $\Pr[P_2]$*, is zero.*

Proof First we consider the conditional probability, $\Pr[P_2|t]$. Without loss of generality we look at the number of right extremists. Once the number of agents is drawn, a simple trinomial process draws the ideology type. By the weak law of large numbers we know that the proportion of right extremists converges to η as $t \to \infty$. Since $\eta < \frac{1}{2}$, $\lim_{t\to\infty} \Pr[P_2|t] = 0$.

Now we must show that as the poisson parameter, n, approaches infinity the probability of an extremist majority approaches zero. Formally we need to show that:

$$\lim_{n\to\infty} \sum_{t=1}^{\infty} \frac{e^{-n}n^t}{t!} \Pr[P_2|t] = 0 \tag{B.1}$$

First consider the poisson distribution itself:

$$\sum_{t=1}^{\infty} \frac{e^{-n}n^t}{t!} = 1$$

This equality holds for all n so we can write:

$$\lim_{n\to\infty} \sum_{t=1}^{\infty} \frac{e^{-n}n^t}{t!} = 1$$

for any $N > 0$ we can rewrite as:

$$\lim_{n\to\infty}\left[\sum_{t=1}^{N-1}\frac{e^{-n}n^t}{t!}+\sum_{t=N}^{\infty}\frac{e^{-n}n^t}{t!}\right]=1$$

since for all t, $\frac{e^{-n}n^t}{t!}$ is positive, both terms must be bounded so we can take the limit on the inside:

$$\lim_{n\to\infty}\sum_{t=1}^{N-1}\frac{e^{-n}n^t}{t!}+\lim_{n\to\infty}\sum_{t=N}^{\infty}\frac{e^{-n}n^t}{t!}=1$$

Notice that for any $N > 0$ the the left part of this sum evaluates to 0, so it must be the case that $\lim_{n\to\infty}\sum_{t=N}^{\infty}\frac{e^{-n}n^t}{t!}=1$. Now for any $N > 0$ we can rewrite equation (B.1) as:

$$\lim_{n\to\infty}\left[\sum_{t=1}^{N-1}\frac{e^{-n}n^t}{t!}\Pr[P_2|t]+\sum_{t=N}^{\infty}\frac{e^{-n}n^t}{t!}\Pr[P_2|t]\right]$$

if we take the maximal values of $\Pr[P_2|t]$ we can see that:

$$\lim_{n\to\infty}\left[\sum_{t=1}^{N-1}\frac{e^{-n}n^t}{t!}\Pr[P_2|t]+\sum_{t=N}^{\infty}\frac{e^{-n}n^t}{t!}\Pr[P_2|t]\right]$$
$$\leq\lim_{n\to\infty}\left[\max_{1\leq t<N}\Pr[P_2|t]\sum_{t=1}^{N-1}\frac{e^{-n}n^t}{t!}+\max_{N\leq t<\infty}\Pr[P_2|t]\sum_{t=N}^{\infty}\frac{e^{-n}n^t}{t!}\right]$$

by the same argument as before we can move the limit on the inside such that:

$$\lim_{n\to\infty}\left[\max_{1\leq t<N}\Pr[P_2|t]\sum_{t=1}^{N-1}\frac{e^{-n}n^t}{t!}+\max_{N\leq t<\infty}\Pr[P_2|t]\sum_{t=N}^{\infty}\frac{e^{-n}n^t}{t!}\right]$$
$$=\max_{1\leq t<N}\Pr[P_2|t]\lim_{n\to\infty}\sum_{t=1}^{N-1}\frac{e^{-n}n^t}{t!}+\max_{N\leq t<\infty}\Pr[P_2|t]\lim_{n\to\infty}\sum_{t=N}^{\infty}\frac{e^{-n}n^t}{t!}$$

Taking limits, by the arguments above and by the fact that $\Pr[P_2|t]$ does not depend on n we have that for all $N > 0$:

$$\lim_{n\to\infty}\sum_{t=1}^{\infty}\frac{e^{-n}n^t}{t!}\Pr[P_2|t]\leq\max_{N\leq t<\infty}\Pr[P_2|t]$$

Since this is for all $N > 0$ we can take the limit as $N\to\infty$. By the weak law of large numbers argument above $\max_{N\leq t<\infty}\Pr[P_2|t]\to 0$ as N goes to infinity and thus $\Pr[P_2|t]$ is squeezed to zero. As such $\Pr[P_2|t]$ is zero in the limit.

By the above argument in the limit, $\sum_{\boldsymbol{n}} Prob(\boldsymbol{n})I(n_r > n_i + n_l \text{ or } n_l > n_r + n_i) = 0$ since for all $\boldsymbol{n}$ such that $I(\boldsymbol{n}) = 1$, $\Pr(\boldsymbol{n}) = 0$, and for all $\boldsymbol{n}$ such that $\Pr(\boldsymbol{n}) > 0$, $I(\boldsymbol{n}) = 0$. $\square$

C Comparison of $x(p)$ and $y(p)$

In section A.2 we showed that $x(p)$ is the solution to

$$c = \frac{p}{4}e^{-x}\Big[I_0(\theta_x) + \Big(\frac{p}{1-p}\Big)^{-\frac{1}{2}} I_{-1}(\theta_x)\Big]$$

and in section A.3 we showed that $y(p)$ is the solution to

$$c = \Big(\frac{2p-1}{4}\Big)e^{-y}I_0(\theta_y)$$

Using these expressions we can use

$$\frac{p}{4}e^{-x}\Big[I_0(\theta_x) + \Big(\frac{p}{1-p}\Big)^{-\frac{1}{2}} I_{-1}(\theta_x)\Big] = \Big(\frac{2p-1}{4}\Big)e^{-y}I_0(\theta_y) \tag{C.1}$$

to compare the magnitudes of $x(p)$ and $y(p)$.

If we divide equation (C.1) by p and recognize that

$$\forall\ p \in (\frac{1}{2}, 1),\ x \geq 0,\ e^{-x}\Big(\frac{1-p}{p}\Big)^{\frac{1}{2}} I_1(\theta_x) \geq 0$$

we can say that (C.1) implies that:

$$\Big(2 - \frac{1}{p}\Big)e^{-y}I_0(\theta_y) \geq e^{-x}I_0(\theta_x)$$

and

$$e^{-y}I_0(\theta_y) > e^{-x}I_0(\theta_x) \tag{C.2}$$

since $\forall\ p \in (\frac{1}{2}, 1),\ (2 - \frac{1}{p}) < 1$.

If $\forall z \geq 0,\ p \in (\frac{1}{2}, 1)\ \frac{d}{dz}[e^{-z}I_0\Big(2z\sqrt{p(1-p)}\Big)] < 0$ then the above inequality implies $x(p) > y(p)$.

$$\begin{aligned}
&\frac{d}{dz}[e^{-z}I_0\Big(2z\sqrt{p(1-p)}\Big)] \\
&= -e^{-z}I_0\Big(2z\sqrt{p(1-p)}\Big) + 2\sqrt{p(1-p)}e^{-z}I_1\Big(2z\sqrt{p(1-p)}\Big) \\
&= e^{-z}[2\sqrt{p(1-p)}I_1\Big(2z\sqrt{p(1-p)}\Big) - I_0\Big(2z\sqrt{p(1-p)}\Big)] \\
&< e^{-z}[I_1(2z\sqrt{p(1-p)}\Big) - I_0\Big(2z\sqrt{p(1-p)}\Big)] < 0 \\
&\forall z \geq 0,\ p \in (\frac{1}{2}, 1)
\end{aligned}$$

And so $x(p) > y(p)\ \forall\ p \in (\frac{1}{2}, 1)$

References

Abramowitz, Milton and Irene A. Stegun, eds, *Handbook of Mathematical Functions with Formulas, Graphs, and Mathematical Tables* Applied Mathematical Series 55, 10 ed., National Bureau of Standards, December 1972.

Austen-Smith, David and Jeffrey S. Banks, "Information Aggregation, Rationality, and the Condorcet Jury Theorem," *American Political Science Review*, March 1996, *90* (1), 34–45.

___ **and Timothy Feddersen**, *Social Choice and Strategic Decisions: Essays in Honor of Jeffrey S. Banks*, Berlin: Springer, 2005.

___ **and** ___ , "Deliberation, Preference Uncertainty, and Voting Rules," *American Political Science Review*, 2006, *100*, 209–217.

Baron, David, "Persistent Media Bias," *Journal of Public Economics*, 2006, *90*, 1–36.

Bernhardt, Dan, Stefan Krasa, and Mattias K. Polborn, "Political Polarization and the Electoral Effects of Media Bias," March 2006. CESifo Working Paper Series No. 1798, Mimeo.

Coughlan, Peter J., "In Defense of Unanimous Jury Verdicts: Mistrials, Communication, and Strategic Voting," *The American Political Science Review*, jun 2000, *94* (2), 375–393.

Crawford, Vincent P and Joel Sobel, "Strategic Information Transmission," *Econometrica*, November 1982, *50* (6), 1431–51. available at http://ideas.repec.org/a/ecm/emetrp/v50y1982i6p1431-51.html.

Degan, Arianna, "Policy Position, Information Acquisition, and Turnout," *Scandinavian Journal of Economics*, December 2006, *108* (4), 669–682.

Dewatripont, Mathias and Jean Tirole, "Advocates," *The Journal of Political Economy*, February 1999, *107* (1), 1–39.

Feddersen, Timothy J. and Wolfgang Pesendorfer, "The Swing Voter's Curse," *The American Economic Review*, June 1996, *86* (3), 408–424.

Gerardi, Dino and Leeat Yariv, "Information Acquisition in Committees," May 2006. Mimeo.

Glaeser, Jacob and Ariel Rubinstein, "Debates and Decisions: On a Rationale of Argumentation Rules," *Games and Economic Behavior*, 2001, *36*, 158–173.

Hafer, Catherine and Dimitri Landa, "Deliberation and Social Polarization," 2006. New York University, Mimeo.

Martinelli, Cesar, "Would Rational Voters Acquire Costly Information?," *Journal of Economic Theory*, 2006, *129*, 225–251.

——, "Rational Ignorance and Voting Behavior," *International Journal of Game Theory*, 2007, *35*, 315–335.

Mathis, Jerome, "Deliberation with Partially Verifiable Information," Technical Report 2006.

Milgrom, Paul and John Roberts, "Relying on the Information of Interested Parties," *Rand Journal of Economics*, 1986, *17*, 18–32.

Myerson, Roger B., "Extended Poisson Games and the Condorcet Jury Theorem," *Games and Economic Behavior*, 1998, pp. 111–131.

——, "Large Poisson Games," *Journal of Economic Theory*, September 2000, *94* (1), 7–45. available at http://ideas.repec.org/a/eee/jetheo/v94y2000i1p7-45.html.

Oliveros, Santiago, "Aggregation of Endogenous Information in Mandatory Elections," 2006. Haas School of Business, UC: Berkeley, Mimeo.

Palfrey, Thomas R. and Keith T. Poole, "The Relationship Between Information, Ideology, and Voting Behavior," *American Journal of Political Science*, August 1987, *31* (3), 511–530.

Shin, Hyun-Song, "The Burden of Proof in a Game of Persuasion," *Journal of Economic Theory*, 1994, *64*, 253–264.

Sidanius, Jim, "Political Sophistication and Political Deviance: A Structural Equation Examination of Context Theory," *Journal of Personality and Social Psychology*, 1988, *55* (1), 37–51.

Skellam, J. G., "The Frequency Distribution of the Difference Between Two Poisson Variates Belonging to Different Populations," *Journal of the Royal Statistical Society*, 1946, *109* (3), 296.

The National Election Studies, "2004 National Election Study," Ann Arbor, MI: University of Michigan, Center for Political Studies [producer and distributor] 2004.