

# IGNORANCE AND NAIVETE IN LARGE ELECTIONS

CÉSAR MARTINELLI

**ABSTRACT.** We consider a two-alternative election with voluntary participation and nearly common interests in which voters may acquire information about which alternative is best. Voters may be *rational* or *naive* in the sense of being able, or not, to update their beliefs about the state of the world conditioning on the behavior of others. We show that there is full information equivalence if all voters are rational and there is arbitrarily cheap information. Electoral participation converges to zero if and only if information is costly for all voters. Per contra, if some voters are naive, participation remains bounded away from zero, and full information equivalence requires that information is free for some voters. Increasing the number of naive voters is bad for information aggregation if naive voters are few, but may be good if there are already many.

*Keywords:* Poisson games, rational ignorance, cursed equilibrium, Condorcet jury theorem. *JEL* D72, D83.

---

CIE-ITAM. Email: [martinel@itam.mx](mailto:martinel@itam.mx). This paper is partly motivated by experimental work carried jointly with Tom Palfrey, Alexander Elbittar and Andrei Gomberg; conversations with all of them have been extremely helpful. Audiences at Boston University, the University of Rochester (Wallis Institute), Northwestern University (MEDS), and the Applied Game Theory Workshop at CIDE provided useful comments and suggestions. Part of this research has been conducted while visiting the University of Chicago with a Tinker fellowship. Financial support from Asociación Mexicana de Cultura is gratefully acknowledged.

*Date:* October 11, 2011.

On trouve de plus, que si la probabilité de la voix de chaque Votant est plus grande que  $\frac{1}{2}$ , c'est-à-dire, s'il est plus probable qu'il jugera conformément à la vérité, plus le nombre des Votans augmentera, plus la probabilité de la vérité de la décision sera grande : la limite de cette probabilité sera la certitude [...]

Une assemblée très-nombreuse ne peut pas être composée d'hommes très-éclairés; il est même vraisemblable que ceux qui la forment joindront sur bien des objets beaucoup d'ignorance à beaucoup de préjugés.

Condorcet (1785)[1986, p. 29-30]

## 1. INTRODUCTION

It is not unusual for a political election to be dominated by a single complex issue, say a deep recession, or crime, to such an extent that disagreements between voters about which candidate to support stem more from different opinions about the effectiveness of the policies espoused by the candidates toward achieving largely shared goals (say, avoid an economic depression, or reduce a crime wave) than from purely distributive concerns. In some cases, the dominant issue may be the past performance of an elected official who is seeking reelection, with disagreements between voters reflecting their different opinions about the ability or the disposition displayed by the incumbent to pursue shared goals. To the extent that voters' opinions reflect to some degree different information held by voters, the ability of elections to successfully aggregate information dispersed among voters is crucial for elections both to lead to good policy choices and to discipline elected officials.

The idea that elections may serve to make good collective choices by aggregating the information dispersed among the voters was first given a statistical foundation by Condorcet and has been studied in the last fifteen years from a game-theoretic viewpoint, starting with the pioneer work of Austen-Smith and Banks (1996), Feddersen and Pesendorfer (1996, 1997), Myerson (1998b), and McLennan (1998). The conclusion of this literature has been, for the most, in the positive with respect to the ability of elections to lead to good choices under a variety of circumstances in which there is an element of common interest in voters' preferences.

The usual approach to large elections in the literature consists in looking at elections with increasing numbers of voters, keeping constant the information held by each voter. This approach leaves out two potential sources of difficulty for information aggregation. First, since each voter has little probability of influencing the outcome, voters may decline acquiring any

costly information if the electorate is large enough. That is, it may be a rational decision for many or most voters to remain ignorant if becoming informed requires any time or effort. Second, realizing that there is little at stake in casting a single vote, voters may not behave in a manner fully consistent with rational behavior when facing difficult inference problems related to the behavior of other voters at the voting booth. As a result, the connection between voters' opinions and information about the alternatives may weaken as the electorate increases.<sup>1</sup>

In this paper, we propose a model of large elections that recognizes both difficulties for successful information aggregation. In the model, voters may have different access to free or costly information about the alternatives. In addition, they may be rational or naive. Rational and naive voters are assumed to play best responses to their expectation about the behavior of other voters, but to form expectations differently. Unlike rational voters, naive voters are assumed to be unable to update their beliefs about the value of the alternatives, conditional on the behavior of other voters, and in particular conditional on the event of being decisive. That is, naive voters are "cursed" in the sense of Eyster and Rabin (2005).<sup>2</sup>

Naive behavior in the sense described is not equivalent to the notion of "sincere voting," conventionally used in the political economy literature to refer to a voter behaving as if the voter were the sole decision maker. Naive voters behave sincerely at the voting booth in the sense of conditioning only on their private information and not on the event of being decisive as rational voters would, but they do calculate correctly the (unconditional) distribution of voting profiles for other voters, and use this calculation to decide whether to acquire costly information. That is, unlike sincere voters, naive voters play best responses to beliefs about other voters' actions which are consistent with repeated observation.

Formally, we describe the election as a Poisson game of population uncertainty, using the setup introduced by Myerson (1998ab, 2000). As shown by Myerson (2000), this has the advantage of allowing to obtain rather directly properties of ratios of probabilities that are important for voters' decisions and that would hold even if there was no population uncertainty. We consider the limit of sequences of equilibria corresponding to the game described above as the expected number of voters goes to infinity. We are interested in the ability of large electorates to reach correct decisions with

---

<sup>1</sup>Indeed, as the initial quotes make it clear, Condorcet was aware of the possibility of both ignorance and biased judgment clouding the opinion of voters as electoral bodies grow large.

<sup>2</sup>Eyster and Rabin (2005) introduced the notion of cursed equilibrium to explain apparent overbidding in affiliated value auctions commonly referred to as the "winner's curse." See Sponda (2007) and Jehiel and Koessler (2008) for similar notions.

		Voters' behavior	
		<i>rational</i>	<i>rational and naive</i>
Information	<i>free and costly</i>	success; large turnout	success; large turnout
	<i>costly</i>	success; small turnout	failure; large turnout

TABLE 1. Information equivalence in large elections

large probability, and in the impact of introducing naive voters, both directly through their behavior regarding information acquisition and voting, and indirectly through the change in the behavior of rational voters.

Keeping with the literature, we say that the election satisfies full information equivalence if there is a sequence of equilibria of the game as the expected number of voters increases such that along that sequence the probability of reaching the correct decision converges to one. We say that a voter participates in the election if the voter votes for one of the alternatives (rather than abstaining), and define turnout as the probability that a voter participates. Table 1 summarizes our findings with respect to information equivalence and voter participation. In all cases described in Table 1, the incentive to acquire information decline in equilibrium with the expected number of voters, so the fraction of voters who acquire costly information declines in equilibrium to zero. This, however, has very different implications depending on whether there are or not naive voters.

If all voters are rational, the election satisfies full information equivalence under a very general condition that puts bounds on the preferences of voters in relation to the quality of the information possibly held by one voter, so that voters are (i) willing to abstain rather than vote for their favorite alternatives whenever they are uninformed, and (ii) willing to vote for an alternative different from their favorite whenever they are informed. In particular, if all information is costly, full information equivalence relies on a small fraction of voters acquiring cheap information and participating in the election, so that most voters abstain as the size of the electorate increases.

If there are both rational and naive voters, the condition for rational voters to be willing to abstain becomes more stringent. The reason is that naive voters do not abstain when they are uninformed. This clouds the relationship between votes and information, reducing the information contained in the event of a voter's favorite alternative being close but behind the other alternative in the vote tally. In particular, if all information is costly, there is (approximately) no information contained in that or any other event in which the voter may be decisive, since most voters who participate in the

election are actually uninformed. Thus, in large elections, rational voters do as naive voters and vote for their favorite alternatives rather than abstaining when they are uninformed. As a result, whichever alternative is the *ex ante* favorite wins the election with large probability, and information equivalence fails.

If there are both rational and naive voters, and the probability of receiving free information is positive, the election satisfies full information equivalence under two very different sets of circumstances. First, if there are relatively *few* naive voters, uninformed rational voters may offset any bias introduced by naive voters and abstain otherwise. This happens under a condition similar to that of the pure rational voter case, but recognizing that the quality of information held by the average voter who participates in the election is lower now, since uninformed naive voters and some (but not all) uninformed rational voters participate. Second, if there are relatively *many* naive voters, information equivalence may rely on informed naive voters. The reason is that even though naive voters are less willing than rational voters to abstain when they are uninformed, they are more willing than rational voters to vote for the alternative they favor the least conditional on being informed.

Among the game-theoretic work on information aggregation in elections, most relevant for this paper is Feddersen and Pesendorfer's (1996) analysis of the "swing voter's curse." This is, by analogy with the winner's curse in auctions with common values, the fact that rational voters will find it optimal to abstain rather than to vote sincerely when, conditional on being pivotal, they would regret changing the result of the election. Guarnaschelli, McKelvey and Palfrey (2000) and Battaglini, Morton and Palfrey (2010) study experimentally jury situations and provide some empirical support for the swing voter's curse. Eyster and Rabin (2005) introduce the notion of cursed equilibrium trying to capture precisely the experimental evidence in the context of auctions that bidders may fall prey of the winner's curse; they also discuss the possibility that the experimental evidence of Guarnaschelli et al. (2000) may be consistent with cursed behavior.

Rational ignorance in large elections have been considered in our previous work (Martinelli 2006, 2007); extensions are provided by Oliveros (2011ab). Previous results on full information equivalence with costly information have been limited to the case of mandatory voting with only rational voters.<sup>3</sup> As argued above, allowing for abstention is crucial to appreciate

---

<sup>3</sup>Krishna and Morgan (2010) consider a model of costly voting with free information whose equilibrium shares some features of the costly information/only rational voters case described in Table 1.

the impact of naive voters on the behavior of rational voters. Mukhopadhyaya (2005), Persico (2004) and Gerardi and Yariv (2008) have worked on the closely related subject of endogenous information in committees. Experimental evidence from Elbittar et al. (2011) suggests that rational and naive voters coexist in the context of voting in committees with costly information, and provides some of the motivation for this work.

The remainder of this paper is organized as follows. Section 2 describes the model. Section 3 looks into belief formation and best response behavior. Section 4 presents the results for the benchmark case in which all voters are rational. Section 5 presents the results for the case in which some voters are rational and some are naive. Section 6 gathers concluding remarks.

## 2. THE MODEL

We analyze an election with two alternatives,  $A$  and  $B$ . The number of voters is a random variable drawn from a Poisson distribution with mean  $n$ ; that is, the probability that the number of voters is  $z$  is given by

$$P(z|n) \equiv e^{-n} n^z / z!$$

for all nonnegative integers.

There are two possible states of the world,  $\omega_A$  and  $\omega_B$ . Both states of the world are equally likely. Let

$$V(d, \omega_{d'}) = \begin{cases} 1 & \text{if } d = d' \\ 0 & \text{if } d \neq d' \end{cases}$$

for  $d, d' \in \{A, B\}$ , where the first argument is the winning alternative and the second argument is the state of the world.  $V$  represents the *common value* component of voters' payoffs.

Each voter has a bias  $\varepsilon$  in favor of  $A$  that is independently drawn from  $\{-\underline{\varepsilon}, \bar{\varepsilon}\}$ , where  $-\underline{\varepsilon} < 0 < \bar{\varepsilon}$ . We denote  $\underline{p} = \Pr(\varepsilon = -\underline{\varepsilon})$  and  $\bar{p} = \Pr(\varepsilon = \bar{\varepsilon})$ . Let

$$W(d, \varepsilon) = \begin{cases} \varepsilon & \text{if } d = A \\ 0 & \text{if } d = B \end{cases}$$

for  $d \in \{A, B\}$  and  $\varepsilon \in \{\varepsilon_k\}$ .  $W$  represents the *partisan* component of voters' payoffs.

Voters do not know the realization of the state of the world. Each voter, however, may acquire some costly information. The cost of information is independently and identically distributed across voters according to a distribution function  $F$ .  $F$  is strictly increasing and continuously differentiable over the interval  $(0, \bar{c})$  for some  $\bar{c} \in \mathfrak{R}_+$ , with  $F(0) \geq 0$  and  $F(\bar{c}) = 1$ . Note that, if  $F(0) > 0$ , then voters receive free information with positive probability. After learning their idiosyncratic cost of information, each voter must

decide whether to acquire information or not. Each voter who acquires information receives a signal  $s \in \{s_A, s_B\}$ . The probability of receiving signal  $s_d$  in state  $\omega_d$  is equal to  $1/2 + q$  for  $d \in \{A, B\}$ , where  $q \in (0, 1/2)$ .

The election takes place after voters receive their signals. A voter may vote for  $A$ , vote for  $B$ , or abstain. The alternative with most votes is chosen, with ties broken by a fair coin toss.

Given the winning alternative  $d$ , the state of the world  $\omega$ , a voter cost of information  $c$ , and a voter bias  $\varepsilon$ , the utility of the voter is

$$V(d, \omega) + W(d, \varepsilon) - c$$

if the voter acquires information, and

$$V(d, \omega) + W(d, \varepsilon)$$

otherwise.

Since prior beliefs give equal probability to each state of the world, the posterior probability of state  $\omega_d$  conditional on having observed a signal  $s_d$  is equal to  $1/2 + q$ . Thus, a voter with bias  $\varepsilon$  acting like a single decision-maker would choose  $A$  if  $\varepsilon > 0$  and would choose  $B$  if  $\varepsilon < 0$  if the voter has not acquired information about the state of the world, and would choose whichever alternative is favored by the voter's private signal if the voter has acquired information and  $|\varepsilon| < 2q$ . Of course, a voter is not a single decision-maker. Whether or not casting a vote influences the winning alternative, and hence the utility of the voter, depends on other voters' decisions. Therefore, in order to choose optimally whether to acquire information and which alternative to support, the voter must form some beliefs about other voters' actions, and how these actions depend on the state of the world.

We assume that each voter is either *rational* or *naive*. Both rational and naive voters play best responses to their beliefs about other voters' actions, but form these beliefs differently. Rational voters understand that the probability distribution of other voters' decisions is influenced by the state of the world, and calculate correctly this probability distribution for the two states of the world. Naive voters mistakenly believe that other voters' actions are independent of the state of the world, and calculate the probability distribution of other voters' decisions in each state of the world to be equal to the unconditional distribution. Each voter's behavioral type  $\beta$  is independently drawn from  $\{\rho, \nu\}$  (for rational and naive, respectively), where  $\Pr(\beta = \rho) = r$ .

A voter's type is a vector  $t = (\beta, \varepsilon, c, s)$  specifying the voter's behavioral type, partisan bias, cost of information and private signal (which the voter only gets to observe if the voter acquires information). An action is a pair  $a = (i, v)$ ,  $i \in \{1, 0\}$ ,  $v \in \{A, B, 0\}$ , indicating whether the voter acquires or

not information and whether the voter votes for alternative  $A$ , for alternative  $B$ , or abstains.

A *strategy function* is a measurable mapping  $\sigma$  assigning to each voter type a probability distribution over the set of actions, where  $\sigma(a|t)$  is interpreted as the probability that a voter will choose action  $a$  given type  $t$ , with the constraint

$$\sigma((0, v)|(\beta, \varepsilon, c, s_A)) = \sigma((0, v)|(\beta, \varepsilon, c, s_B))$$

for all  $v$ ,  $\beta$ ,  $\varepsilon$  and  $c$ . That is, an uninformed voter's behavior is independent of the voter's signal (and hence, independent of the state of the world).

Let  $\tau(v|\sigma, \omega)$  be the probability that a voter chooses  $v \in \{A, B, 0\}$  given strategy  $\sigma$  conditional on the state of the world  $\omega$ . From the independent action property of Poisson games (Myerson 1998a), the number of votes for  $A$ , votes for  $B$  and abstentions in state  $\omega$  are independent Poisson random variables with mean  $n\tau(A|\sigma, \omega)$ ,  $n\tau(B|\sigma, \omega)$ , and  $n\tau(0|\sigma, \omega)$ . Let

$$\tau(\sigma, \omega) = (\tau(v|\sigma, \omega))_{v \in \{A, B, 0\}}.$$

Then the probability of the voting profile  $x = (x(A), x(B), x(0))$ , describing how many voters vote for  $A$ , vote for  $B$ , or abstain, is equal to

$$P(x|n\tau(\sigma, \omega)) = \prod_{v \in \{A, B, 0\}} P(x(v)|n\tau(v|\sigma, \omega)).$$

From the environmental equivalence property of Poisson games (Myerson 1998a),  $P(x|\lambda(\sigma, \omega))$  is also the probability that the voting profile of other voters is given by  $x = (x(A), x(B), x(0))$  from the perspective of any given voter.

We say that a strategy  $\sigma'$  is a *best response* to the strategy  $\sigma$  if

- (i) for almost every triple  $\rho, \varepsilon, c$  (that is, for each possible rational voter) the strategy  $\sigma'$  maximizes the expected utility of the voter given the distributions  $P(x|n\tau(\sigma, \omega_A))$  and  $P(x|n\tau(\sigma, \omega_B))$  of voting profiles of other voters in state  $\omega_A$  and  $\omega_B$ , respectively, and
- (ii) for almost every triple  $v, \varepsilon, c$  (that is, for each possible naive voter) the strategy  $\sigma'$  maximizes the expected utility of the voter given the distribution  $\frac{1}{2}P(x|n\tau(\sigma, \omega_A)) + \frac{1}{2}P(x|n\tau(v|\sigma, \omega_B))$  of voting profiles of other voters in both states of the world.

An *equilibrium* is a strategy that is a best response to itself. Note that, in equilibrium, both rational and naive voters play best responses, but naive voters ignore the dependence of the distribution of other voters' profiles on the state of the world.

If  $r = 1$ , that is, if all voters are rational, then the equilibrium definition is similar to Myerson's (1998b) definition of equilibrium for extended Poisson games. If  $r = 0$ , that is, if all voters are naive, then the equilibrium definition



is similar to Eyster and Rabin's (2005) definition of fully cursed equilibrium, adapted to a context of private information acquisition and population uncertainty. If  $0 < r < 1$ , then the equilibrium definition is analogous to the generalized cursed equilibrium of Eyster and Rabin (2005).

We say that the election satisfies *full information equivalence* if there is a sequence of equilibria of the game as the expected number of voters increases such that along that sequence the probability of choosing alternative  $A$  in state  $\omega_A$  and the probability of choosing alternative  $B$  in state  $\omega_B$  converge to one. We say that a voter participates in the election if the voter votes for  $A$  or  $B$  rather than abstaining, and define turnout as the probability that a voter participates.

### 3. BELIEFS AND BEST RESPONSES

Given any strategy  $\gamma$ , the number of votes for  $A$  minus the number of votes for  $B$  in state  $\omega$  is the difference between two independent Poisson random variables with mean  $n\tau(A|\sigma, \omega)$  and  $n\tau(B|\sigma, \omega)$ , respectively. This random variable is sometimes referred to as a Skellam random variable<sup>4</sup> (after Skellam 1946) with parameters equal to the means of the two independent Poisson variables.

In particular, the probability that the difference between the number of votes for  $A$  and the number of votes for  $B$  is equal to some integer  $z$  is  $S(z|n\tau(A|\sigma, \omega), n\tau(B|\sigma, \omega))$ , where

$$S(z|m_1, m_2) = e^{-(m_1+m_2)} \left(\frac{m_1}{m_2}\right)^{z/2} I_{|z|}(2\sqrt{m_1 m_2})$$

is the *Skellam distribution function* with parameters  $m_1$  and  $m_2$ , and

$$I_z(y) = \sum_{k=0}^{\infty} \frac{1}{k!(k+z+1)!} \left(\frac{y}{2}\right)^{2k+z}$$

is a *modified Bessel function of the first kind* (Bronshtein et al 2003). Note that  $I_z(y)$  is increasing in  $y$ , and  $I_0(0) = 1$  and  $I_z(0) = 0$  for  $z > 0$ .

We will make use of the following asymptotic approximation

$$I_z(y) = \frac{e^y}{\sqrt{2\pi y}} (1 + O(1/y)),$$

where  $g(y) = O(1/y)$  stands for  $\lim_{y \rightarrow \infty} yg(y) = C$  for some constant  $C \neq 0$  (Bronshtein et al 2003).

We will also make use of the following known result. Consider a sequence  $(\lambda_{1n}, \lambda_{2n})$  of positive numbers for increasing values of  $n$ , and let  $z_n$  be a Skellam random variable with parameters  $n\lambda_{1n}$  and  $n\lambda_{2n}$  for each  $n$ . If

<sup>4</sup>E.g. by Karlis and Ntzoufras (2003, 2006).

$n\lambda_{1n}$  and  $n\lambda_{2n}$  go to infinity and  $\lambda_{1n}/\lambda_{2n}$  converges to a number larger than one as  $n$  increases, then the probability that  $z_n$  is strictly larger than zero converges to one.<sup>5</sup>

Let

$$D(z|\rho, \sigma, \omega) = S(z|n\tau(A|\sigma, \omega), n\tau(B|\sigma, \omega))$$

and

$$D(z|v, \sigma, \omega) = \frac{1}{2} \sum_{\omega'} S(z|n\tau(A|\sigma, \omega'), n\tau(B|\sigma, \omega'))$$

represent, respectively, the beliefs of a rational and a naive voter about the distribution of the net difference of votes in favor of  $A$  of other voters in state  $\omega$  if other voters are playing according to the strategy  $\sigma$ . That is, rational voters assess correctly the distribution of the net difference in favor of  $A$  while naive voters believe incorrectly this distribution is independent of the state of the world, but assess correctly the unconditional distribution.

The difference in expected utility between voting for alternative  $A$  after receiving signal  $s_A$  and abstaining after receiving signal  $B$  versus abstaining regardless of the signal for a voter of behavioral type  $\beta$  and bias  $\varepsilon$  is

$$\begin{aligned} G(s_A|\beta, \varepsilon, \sigma) &\equiv \frac{1+\varepsilon}{4} \left( \frac{1}{2} + q \right) (D(0|\beta, \sigma, \omega_A) + D(-1|\beta, \sigma, \omega_A)) \\ &\quad - \frac{1-\varepsilon}{4} \left( \frac{1}{2} - q \right) (D(0|\beta, \sigma, \omega_B) + D(-1|\beta, \sigma, \omega_B)). \end{aligned}$$

Similarly, the difference in expected utility between abstaining after signal  $A$  and voting for alternative  $B$  after receiving signal  $s_B$  versus abstaining regardless of the signal for a voter of behavioral type  $\beta$  and bias  $\varepsilon$  is

$$\begin{aligned} G(s_B|\beta, \varepsilon, \sigma) &\equiv -\frac{1+\varepsilon}{4} \left( \frac{1}{2} - q \right) (D(0|\beta, \sigma, \omega_A) + D(1|\beta, \sigma, \omega_A)) \\ &\quad + \frac{1-\varepsilon}{4} \left( \frac{1}{2} + q \right) (D(0|\beta, \sigma, \omega_B) + D(1|\beta, \sigma, \omega_B)). \end{aligned}$$

The difference in expected utility between voting for alternative  $A$  and abstaining, regardless if the signal in both cases, is

$$\begin{aligned} G(A|\beta, \varepsilon, \sigma) &\equiv \frac{1+\varepsilon}{4} (D(0|\beta, \sigma, \omega_A) + D(-1|\beta, \sigma, \omega_A)) \\ &\quad - \frac{1-\varepsilon}{4} (D(0|\beta, \sigma, \omega_B) + D(-1|\beta, \sigma, \omega_B)). \end{aligned}$$

Finally, the difference in expected utility between voting for alternative  $B$  and abstaining, regardless if the signal in both cases, is

$$\begin{aligned} G(B|\beta, \varepsilon, \sigma) &\equiv -\frac{1+\varepsilon}{4} (D(0|\beta, \sigma, \omega_A) + D(1|\beta, \sigma, \omega_A)) \\ &\quad + \frac{1-\varepsilon}{4} (D(0|\beta, \sigma, \omega_B) + D(1|\beta, \sigma, \omega_B)). \end{aligned}$$

---

<sup>5</sup>This follows from the fact that the Skellam distribution is asymptotically normal. Bounds on the normal approximation can be obtained treating the Skellam random variable  $z_n$  as a summation of  $n$  iid random variables as in Fisz (1955ab) and using the Berry-Esseen theorem (Feller 1971).

We claim that, if a best-responding voter acquires information at some positive cost, then the voter either votes for  $A$  if the signal is  $s_A$  and abstains if the signal is  $s_B$ , or votes for  $B$  if the signal is  $s_B$  and abstains if the signal is  $s_A$ , or votes according to the signal received regardless of the signal. To see this, note that after receiving signal  $d$  a voter must be strictly better off voting for  $d$  or abstaining rather than voting for the other alternative. Otherwise, the voter is strictly better off voting for the other alternative after receiving the other signal, which decreases the expected utility of voting for alternative  $d$  or abstaining for any strategy of other players. But then the voter is better off voting for the other alternative and not acquiring costly information. Moreover, the voter cannot prefer abstaining or be indifferent between abstaining and voting for alternative  $d$  if the signal is  $s_d$  for both alternatives, since then the voter would be better off abstaining and not acquiring costly information.

Let

$$c(\beta, \varepsilon, \sigma) \equiv \max\{G(s_A|\beta, \varepsilon, \sigma), G(s_B|\beta, \varepsilon, \sigma), G(s_A|\beta, \varepsilon, \sigma) + G(s_B|\beta, \varepsilon, \sigma)\} \\ - \max\{0, G(A|\beta, \varepsilon, \sigma), G(B|\beta, \varepsilon, \sigma)\}.$$

From the preceding argument it follows that this is the utility gain of acquiring information, net of the cost of information acquisition.

Given a strategy  $\sigma$ , we can characterize completely best responses using the using the definition of  $G(\cdot|\beta, \varepsilon, \sigma)$  as follows.

**Lemma 1.** *A strategy  $\sigma'$  is a best response to  $\sigma$  if for almost every triple  $\beta, \varepsilon, c$ , under the strategy  $\sigma'$ ,*

- (i) *if  $c < c(\beta, \varepsilon, \sigma)$  then the voter acquires information, and after signal  $s_d$  the voter votes for  $d$  if  $G(s_d|\beta, \varepsilon, \sigma) > 0$ , abstains if  $G(s_d|\beta, \varepsilon, \sigma) < 0$ , and either votes for  $d$  or abstains if  $G(s_d|\beta, \varepsilon, \sigma) = 0$*
- (ii) *if  $c > c(\beta, \varepsilon, \sigma)$ , then the voter does not acquire information, and votes for  $d$  only if  $G(d|\beta, \varepsilon, \sigma) = \max\{0, G(A|\beta, \varepsilon, \sigma), G(B|\beta, \varepsilon, \sigma)\}$  and abstains only if  $0 = \max\{0, G(A|\beta, \varepsilon, \sigma), G(B|\beta, \varepsilon, \sigma)\}$ .*

Since  $D(z|v, \sigma, \omega_A) = D(z|v, \sigma, \omega_B)$  for all  $z$ , we have that, for any given  $\sigma$ ,

$$\text{sgn}(G(A|v, \varepsilon, \sigma)) = -\text{sgn}(G(B|v, \varepsilon, \sigma)) = \text{sgn}(\varepsilon).$$

Thus, (best-responding) *uninformed naive voters always vote for the alternative favored by their bias regardless of the strategy followed by other voters*. In particular, uninformed naive voters always participate in the election. In contrast, uninformed rational voters may decide to abstain if the strategy followed by other voters is such that the voter's expected utility

worsens by casting an uninformed vote, in fact delegating the collective decision to other voters as in the swing voter's curse model of Feddersen and Pesendorfer (1996).

#### 4. RATIONAL VOTERS

In this section we consider as a benchmark the case in which all voters are rational. We use the following condition in the statements of theorems below:

$$(A1) \quad \max\{\underline{\varepsilon}, \bar{\varepsilon}\} < \frac{2q}{1 + 2\sqrt{\frac{1}{4} - q^2}}.$$

Note that condition (A1) implies  $|\varepsilon| < 2q$ ; i.e., biases in favor of either alternative are small enough compared to the quality of a signal so that a single signal would sway any voter if the voter were a single decision-maker. As we will show, condition (A1) is sufficient for (i) an uninformed rational voter to be willing to abstain rather than vote for the voter's favorite alternative, and (ii) for an informed rational voter to vote according to the signal received rather than abstaining, if all other uninformed voters abstain in a large election. Moreover, if the opposite inequality holds, neither (i) nor (ii) are possible for voters with the largest (in absolute terms) partisan bias.

To provide an intuition for condition (A1), suppose  $\bar{\varepsilon} > \underline{\varepsilon}$  and consider the problem of an uninformed rational voter who is biased in favor of  $A$  and is considering voting for  $A$  rather than abstaining. There are two possibilities for a vote for  $A$  to be decisive: either  $A$  and  $B$  are tied, or  $A$  is behind by one vote. In either case, a vote for  $A$  would change the result from  $A$  winning the election to  $B$  winning the election with probability  $1/2$ . If the voter knew that  $A$  and  $B$  are tied, the voter would think both states are equally likely, and the voter would gain

$$(1/2)[(1 + \bar{\varepsilon}) - 0] + (1/2)[\bar{\varepsilon} - 1] = \bar{\varepsilon}$$

in expected terms by changing the result in favor of  $A$ . If the voter knew that  $A$  is behind by one vote, the voter would think that the probability of state  $\omega_A$  is  $1/2 - q$  and the probability of state  $\omega_B$  is  $1/2 + q$ , and the voter would lose

$$(1/2 - q)[(1 + \bar{\varepsilon}) - 0] + (1/2 + q)[\bar{\varepsilon} - 1] = 2q - \bar{\varepsilon}$$

in expected terms by changing the result in favor of  $A$ . The voter will be willing to abstain, then, if

$$\frac{2q - \bar{\varepsilon}}{\bar{\varepsilon}} > \frac{\Pr(A \text{ and } B \text{ are tied})}{\Pr(A \text{ is behind by one vote})},$$

where the probabilities in the right-hand side are unconditional. In the proof of the theorem we show that the ratio in the right-hand side converges to  $2\sqrt{1/4 - q^2}$  if all voters who participate are informed, yielding condition (A1) as necessary for rational abstention. The quality of information play a double role of the quality of information  $q$  in facilitating rational abstention when preferences among voters are not perfectly aligned. Increasing the value of  $q$  both makes abstention more attractive when a voter's favorite alternative is behind by one vote, since the information conveyed by one vote is increasing in  $q$ , and makes the event of one's favorite alternative being behind by one vote more likely in relation to the event of both alternatives being tied.

Now suppose  $\bar{\varepsilon} > \underline{\varepsilon}$  and consider the problem of a rational voter who is biased in favor of  $A$ , who has received signal  $s_B$  and is considering voting for  $B$  rather than abstaining. If the voter knew that  $A$  and  $B$  are tied, the voter would think that the probability of state  $\omega_A$  is  $1/2 - q$  and the probability of state  $\omega_B$  is  $1/2 + q$ , and the voter would gain  $2q - \bar{\varepsilon}$  in expected terms by changing the result in favor of  $B$ . If the voter knew that  $B$  is behind by one vote, the voter would think both states are equally likely, and the voter would lose  $\bar{\varepsilon}$  in expected terms by changing the result in favor of  $B$ . Thus, for a rational voter, the decision of abstaining when uninformed is entirely analogous to the decision of voting for the voter's least favorite alternative when the voter has received a signal favoring that alternative. As discussed in the next section, these two situations are very different for a naive voter.

**Theorem 1.** *If  $r = 1$  and (A1) holds, there is some  $\underline{n}$  such that, for every  $n > \underline{n}$ , there is an equilibrium such that voters acquire information if their cost is below some cutoff  $c_n$ , and vote for  $A$  if they observe  $s_A$  and for  $B$  if they observe  $s_B$ , and do not acquire information and abstain if their cost is above  $c_n$ . As  $n$  goes to infinity, along any sequence of such equilibria the probability that a voter acquires information converges to  $F(0)$ , the probability of abstention converges to  $1 - F(0)$ , and the probability of choosing alternative  $A$  in state  $\omega_A$  and alternative  $B$  in state  $\omega_B$  converges to 1.*

This is a special case of Theorem 3 below, but we include a separate proof as it makes the arguments more transparent and describes in detail the equilibrium strategy.

*Proof.* Suppose  $r = 1$  and (A1) holds. Let  $\sigma(\hat{c})$  denote the strategy of acquiring information if the cost is below  $\hat{c}$  independently of  $\varepsilon$ , and voting for  $A$  if the signal is  $s_A$  and for  $B$  if the signal is  $s_B$ , and do not acquire

information and abstain if their cost is above  $\hat{c}$ , for any  $\hat{c} \geq 0$ . We have

$$G(s_A|\rho, \varepsilon, \sigma(\hat{c})) = e^{-nF(\hat{c})} \left[ \left( \frac{q}{2} + \frac{\varepsilon}{4} \right) I_0 \left( 2nF(\hat{c}) \sqrt{\frac{1}{4} - q^2} \right) + \frac{\varepsilon}{2} \sqrt{\frac{1}{4} - q^2} I_1 \left( 2nF(\hat{c}) \sqrt{\frac{1}{4} - q^2} \right) \right]$$

and

$$G(s_B|\rho, \varepsilon, \sigma(\hat{c})) = e^{-nF(\hat{c})} \left[ \left( \frac{q}{2} - \frac{\varepsilon}{4} \right) I_0 \left( 2nF(\hat{c}) \sqrt{\frac{1}{4} - q^2} \right) - \frac{\varepsilon}{2} \sqrt{\frac{1}{4} - q^2} I_1 \left( 2nF(\hat{c}) \sqrt{\frac{1}{4} - q^2} \right) \right].$$

Thus,

$$G(s_A|\rho, \varepsilon, \sigma(\hat{c})) + G(s_B|\rho, \varepsilon, \sigma(\hat{c})) = qe^{-nF(\hat{c})} I_0 \left( 2nF(\hat{c}) \sqrt{\frac{1}{4} - q^2} \right) > 0$$

for all  $\hat{c} \geq 0$ . Moreover, using the asymptotic approximation for  $I_0(\cdot)$ , there is some  $\tilde{n}$  such that

$$G(s_A|\rho, \varepsilon, \sigma(\bar{c})) + G(s_B|\rho, \varepsilon, \sigma(\bar{c})) = qe^{-nF(\bar{c})} I_0 \left( 2nF(\bar{c}) \sqrt{\frac{1}{4} - q^2} \right) < \bar{c}$$

for all  $n > \tilde{n}$ . Thus, for large enough  $n$  there is some  $c_n \in (0, \bar{c})$  such that

$$(1) \quad e^{-nF(c_n)} q I_0 \left( 2nF(c_n) \sqrt{\frac{1}{4} - q^2} \right) = c_n.$$

Now let  $n$  go to infinity. We claim that  $c_n$  converges to zero. For suppose there is a subsequence such that along that subsequence  $c_n$  converges to any number in  $(0, \bar{c}]$ . Then, using the asymptotic approximation for  $I_0(\cdot)$ , along that subsequence the expression in the left-hand side of equation 1 converges to 0, but the expression in the right-hand side converges to a positive number. Similarly, we claim that  $nF(c_n)$  increases without bound. For suppose there is a subsequence such that along that subsequence  $nF(c_n)$  converges to either zero or a positive number. Then, since  $I_0(y) \geq 1$  for  $y \geq 0$ , along that subsequence the expression in the left-hand side of equation 1 converges to a positive number, but the expression in the right-hand side converges to 0. (Note that  $nF(c_n)$  increases without bound regardless of whether  $F(0) = 0$  or  $F(0) > 0$ . In the former case,  $F(c_n)$  decreases at a rate slower than  $1/n$ .)

We have

$$G(s_A|\rho, \varepsilon, \sigma(c_n)) = e^{-nF(c_n)} \left[ \left( \frac{q}{2} + \frac{\varepsilon}{4} \right) I_0 \left( 2nF(c_n) \sqrt{\frac{1}{4} - q^2} \right) + \frac{\varepsilon}{2} \sqrt{\frac{1}{4} - q^2} I_1 \left( 2nF(c_n) \sqrt{\frac{1}{4} - q^2} \right) \right]$$

and

$$G(A|\rho, \varepsilon, \sigma(c_n)) = e^{-nF(c_n)} \times \left[ \frac{\varepsilon}{2} I_0 \left( 2nF(c_n) \sqrt{\frac{1}{4} - q^2} \right) + \left( \frac{1 + \varepsilon}{4} \left( \frac{\frac{1}{2} - q}{\frac{1}{2} + q} \right)^{1/2} - \frac{1 - \varepsilon}{4} \left( \frac{\frac{1}{2} + q}{\frac{1}{2} - q} \right)^{1/2} \right) I_1 \left( 2nF(c_n) \sqrt{\frac{1}{4} - q^2} \right) \right].$$

Using the fact that  $nF(c_n)$  increases without bound and the asymptotic approximation for  $I_z(\cdot)$ , we have that

$$I_0 \left( 2nF(c_n) \sqrt{\frac{1}{4} - q^2} \right) / I_1 \left( 2nF(c_n) \sqrt{\frac{1}{4} - q^2} \right)$$

converges to one as  $n$  goes to infinity. But then under condition (A1) the terms in brackets in the expression for  $G(s_A|\rho, \varepsilon, \sigma(c_n))$  becomes positive for  $n$  large enough, and the term in brackets in the expression for  $G(A|\rho, \varepsilon, \sigma(c_n))$  becomes negative for  $n$  large enough. Therefore, for  $n$  large enough,

$$G(s_A|\rho, \varepsilon, \sigma(c_n)) > 0 \quad \text{and} \quad G(A|\rho, \varepsilon, \sigma(c_n)) < 0.$$

A similar argument shows that, for  $n$  large enough,

$$G(s_B|\rho, \varepsilon, \sigma(c_n)) > 0 \quad \text{and} \quad G(B|\rho, \varepsilon, \sigma(c_n)) < 0.$$

Thus, there is some  $\underline{n}$  such that for  $n > \underline{n}$ ,  $\sigma(c_n)$  is an equilibrium.

Now, pick  $n > \underline{n}$ . Given the equilibrium  $\sigma(c_n)$ , the probability that a voter acquires information and votes according to the signal received is  $F(c_n)$ , and the probability that a voter abstains is  $1 - F(c_n)$ . Since  $c_n$  converges to zero, the probability that a voter is informed converges to  $F(0)$ , and the probability that a voter abstains converges to  $1 - F(0)$ . Moreover, the probability that  $A$  wins the election in state  $\omega_A$  is equal to the probability that  $B$  wins the election in state  $\omega_B$  and is equal to the probability that a Skellam random variable with parameters  $nF(c_n)(1/2 + q)$  and  $nF(c_n)(1/2 - q)$  is positive. Since

$$\frac{nF(c_n)(1/2 + q)}{nF(c_n)(1/2 - q)} = \frac{1/2 + q}{1/2 - q} > 1,$$

the probability that this random variable is positive converges to one as  $n$  goes to infinity.  $\square$

If (A1) fails, we may still obtain full information equivalence because, say, the group of voters with the largest bias in absolute terms is small enough so that voters in the other group can offset any bias introduced by the former. If a condition like (A1) fails for both possible realizations of the partisan bias, however, full information equivalence becomes unattainable.

**Theorem 2.** *If  $r = 1$ ,*

$$\min\{\underline{\varepsilon}, \bar{\varepsilon}\} > \frac{2q}{1 + 2\sqrt{\frac{1}{4} - q^2}}$$

*and  $\bar{p} \neq p$ , then as  $n$  goes to infinity, along any sequence of equilibria the probability of choosing the alternative favored by most voters' partisan biases converges to one regardless of the state.*

This is just a special case of theorem 4 below.

## 5. RATIONAL AND NAIVE VOTERS

In this section we consider the case in which some voters are naive. From lemma 1, we know that naive voters, whenever ignorant, will vote for the alternative they are biased in favor. The fact that uninformed voters participate in the election reduces the information that rational voters can deduce conditional on being pivotal, and makes voting for the alternative they are biased in favor more attractive. We use the following condition, expressing this additional constraint for rational abstention, in the statement of theorems below:

$$(A2) \quad \max\{\underline{\varepsilon}, \bar{\varepsilon}\} < \frac{2H(F, r, \underline{p}, \bar{p})q}{1 + 2\sqrt{\frac{1}{4} - H(F, r, \underline{p}, \bar{p})^2 q^2}},$$

where  $H(F, r, \underline{p}, \bar{p})$  stands for

$$\frac{F(0)}{2(1-r)(\max\{\underline{p}, \bar{p}\})(1-F(0)) + F(0)}.$$

Note that  $H(F, r, \underline{p}, \bar{p}) \leq 1$ , with strict inequality unless  $r = 1$  or  $F(0) = 1$ . Thus, (A2) is equivalent to (A1) if  $r = 1$  or  $F(0) = 1$ , and is more restrictive otherwise. In particular, (A2) fails if  $r < 1$  and  $F(0) = 0$ .

To provide an intuition for condition (A2), note that if uninformed rational voters are (close to fully) offsetting the biased voting of naive voters, the probability that an uninformed voter participates in the election must be nearly equal to  $2(1-r)\max\{\underline{p}, \bar{p}\}$ . Since the probability that a voter is uninformed converges to  $1 - F(0)$ ,  $H$  represents, in the limit, the probability that a voter who participates in the election is in fact informed.



In particular, if a voter knew that  $A$  is behind by one vote, the voter would think that the probabilities of state  $\omega_A$  and  $\omega_B$  are, respectively,

$$(1 - H)(1/2) + H(1/2 - q) = 1/2 - Hq \quad \text{and} \quad 1/2 + Hq.$$

Thus, condition (A2) is similar to condition (A1), taking into account that the informativeness of a single vote is now  $Hq$  rather than  $q$ .

Condition (A2) guarantees that, if rational voters were to fully offset the biased voting of naive voters, uninformed rational voters would prefer to abstain, and informed voters would prefer to vote according to the signal received. In the equilibrium we describe below, uninformed rational voters do not fully offset the biased voting of naive voters in order to keep some uninformed voters participating in the election in the first place. This makes equilibrium behavior more involved than in the previous section.

As described below, if condition (A2) holds, for successful information aggregation to occur in equilibrium it is sufficient that there are enough rational voters to offset uninformed naive voters. The following condition, which we use below, expresses this constraint:

$$(A3) \quad r \geq \frac{|\bar{p} - \underline{p}|}{1 + |\bar{p} - \underline{p}|}.$$

**Theorem 3.** *If (A2) and (A3) hold, there is some  $\bar{n}$  such that, for every  $n > \bar{n}$ , there is an equilibrium such that type  $(\beta, \varepsilon, c, s)$  voters acquire information and vote according to the signal received if and only if  $c \leq c_n(\beta, \varepsilon)$  for some vector*

$$(c_n(\beta, \varepsilon)) \text{ such that } 0 < c(\beta, \varepsilon) < \bar{c} \text{ for each pair } \beta, \varepsilon,$$

*naive voters vote for their favorite alternative if they are uninformed, and rational voters abstain with positive probability if they are uninformed. As  $n$  goes to infinity, along any sequence of such equilibria the probability that a voter acquires information converges to  $F(0)$ , and the probability of choosing alternative  $A$  in state  $\omega_A$  and alternative  $B$  in state  $\omega_B$  converges to 1.*

*Proof.* The case  $r = 1$  is covered in Theorem 2, so for the remainder of the proof suppose (A2) hold and  $r < 1$ . Note that (A2) and  $r < 1$  imply  $F(0) > 0$ .

Consider the case  $\bar{p} > \underline{p}$  and  $r \geq 1 - \underline{p}/\bar{p}$ . For any given vector

$$(c(\beta, \varepsilon)) \text{ such that } 0 \leq c(\beta, \varepsilon) \leq \bar{c} \text{ for each pair } \beta, \varepsilon,$$

and any real number  $w \in [0, 1]$ , let  $\sigma((c(\beta, \varepsilon)), w)$  denote the strategy such that voters acquire information and vote according to the signal received if  $c \leq c(\beta, \varepsilon)$ , naive voters vote for their favorite alternative if they are uninformed, rational voters with bias  $\bar{\varepsilon}$  abstain if they are uninformed, and

rational voters with bias  $\underline{\varepsilon}$  abstain with probability  $w$  if they are uninformed and vote for  $B$  with probability  $1 - w$  if they are uninformed.

Let

$$\Lambda : [0, \bar{c}]^4 \times [0, 1] \rightarrow [0, \bar{c}]^4 \times [0, 1]$$

with

$$\Lambda((c(\beta, \varepsilon)), w) = ((\hat{c}(\beta, \varepsilon)), \hat{w})$$

be a mapping such that

$$\hat{c}(\rho, \varepsilon) = \min\{\bar{c}, \max\{0, G(s_A|\rho, \varepsilon, \sigma((c), w)) + G(s_B|\rho, \varepsilon, \sigma((c), w))\}\}$$

and

$$\begin{aligned} \hat{c}(v, \varepsilon) = \min\{\bar{c}, \max\{0, G(s_A|\rho, \bar{\varepsilon}, \sigma((c), w)) + G(s_B|\rho, \varepsilon, \sigma((c), w)) \\ - \max\{G(A|v, \varepsilon, \sigma((c), w)), G(B|v, \varepsilon, \sigma((c), w))\}\}\} \end{aligned}$$

for  $\varepsilon \in \{-\underline{\varepsilon}, \bar{\varepsilon}\}$ , and  $\hat{w}$  solves the equation

$$(2) \quad G(B|v, -\underline{\varepsilon}, \sigma((\hat{c}), \hat{w})) = 0$$

if a solution in the interval  $[0, 1]$  exists, and  $\hat{w} = 1$  otherwise.

We claim that for any vector  $(c(\beta, \varepsilon))$ , each cutoff  $\hat{c}(\beta, \varepsilon)$  is arbitrarily close to zero for large  $n$ , where the bound on the approximation depends on  $n$ ,  $q$  and  $F(0)$ . This is straightforward and follows from the fact that  $G(s_A|\rho, \bar{\varepsilon}, \sigma((c), w))$  and  $G(s_B|\rho, \varepsilon, \sigma((c), w))$  converge to zero.

Now let  $\tau(v|(\hat{c}), w, \omega)$  represent the probability that a voter votes for alternative  $v \in \{A, B\}$  in state  $\omega$  given a strategy  $\sigma((\hat{c}), w)$ . From the previous claim, letting  $n$  go to infinity,  $\tau(A|(\hat{c}), w, \omega_A)$ ,  $\tau(B|(\hat{c}), w, \omega_A)$ ,  $\tau(A|(\hat{c}), w, \omega_B)$  and  $\tau(B|(\hat{c}), w, \omega_B)$  converge respectively to

$$\begin{aligned} \bar{\tau}(A|(\hat{c}), w, \omega_A) &= (1 - r)(1 - F(0))\bar{p} + F(0)(1/2 + q) \\ \bar{\tau}(B|(\hat{c}), w, \omega_A) &= (1 - r)(1 - F(0))\underline{p} + r(1 - F(0))\underline{p}w + F(0)(1/2 - q), \\ \bar{\tau}(A|(\hat{c}), w, \omega_B) &= (1 - r)(1 - F(0))\bar{p} + F(0)(1/2 - q), \end{aligned}$$

and

$$\bar{\tau}(B|(\hat{c}), w, \omega_B) = (1 - r)(1 - F(0))\underline{p} + r(1 - F(0))\underline{p}w + F(0)(1/2 + q).$$

Next, we claim that there is some  $\underline{n}$ , depending on the parameters  $F(0)$ ,  $\underline{\varepsilon}$  and  $q$ , such that for any  $n \geq \underline{n}$  a solution to equation 2 in the interval  $[0, 1]$  exists for any vector  $(c(\beta, \varepsilon)) \in [0, \bar{c}]^4$ . To see this, note that equation 2 is equivalent to

$$\frac{D(0|\rho, \sigma((\hat{c}), \hat{w}), \omega_A) + D(1|\rho, \sigma((\hat{c}), \hat{w}), \omega_A)}{D(0|\rho, \sigma((\hat{c}), \hat{w}), \omega_B) + D(1|\rho, \sigma((\hat{c}), \hat{w}), \omega_B)} = \frac{1 + \underline{\varepsilon}}{1 - \underline{\varepsilon}}.$$

Let

$$\begin{aligned}
 L(w) &= \exp(-n(\tau(A|(\hat{c}), w, \omega_A) + \tau(B|(\hat{c}), w, \omega_A)) \\
 &\quad + n(\tau(A|(\hat{c}), w, \omega_B) + \tau(B|(\hat{c}), w, \omega_B))) \\
 &= \frac{I_0(2n\sqrt{T((\hat{c}), w, \omega_A)}) + \sqrt{\frac{\tau(A|(\hat{c}), w, \omega_A)}{\tau(B|(\hat{c}), w, \omega_A)}} I_1(2n\sqrt{T((\hat{c}), w, \omega_A)})}{I_0(2n\sqrt{T((\hat{c}), w, \omega_A)}) + \sqrt{\frac{\tau(A|(\hat{c}), w, \omega_B)}{\tau(B|(\hat{c}), w, \omega_B)}} I_1(2n\sqrt{T((\hat{c}), w, \omega_A)})}
 \end{aligned}$$

where

$$T((\hat{c}), w, \omega_A) = \tau(A|(\hat{c}), w, \omega_A)\tau(B|(\hat{c}), w, \omega_A).$$

Equation 2 has a solution if and only if there is some  $w \in [0, 1]$  such that  $L(w) = (1 + \underline{\varepsilon})/(1 - \underline{\varepsilon})$ . From the previous paragraph,  $L(w)$  is arbitrarily close to

$$\begin{aligned}
 &\exp\left(-n\left(\left(\sqrt{\bar{\tau}(A|(\hat{c}), w, \omega_A)} - \sqrt{\bar{\tau}(B|(\hat{c}), w, \omega_A)}\right)^2\right.\right. \\
 &\quad \left.\left. - \left(\sqrt{\bar{\tau}(A|(\hat{c}), w, \omega_B)} - \sqrt{\bar{\tau}(B|(\hat{c}), w, \omega_B)}\right)^2\right)\right) \\
 &\times \frac{\left[\sqrt{\bar{\tau}(A|(\hat{c}), w, \omega_B)}\bar{\tau}(B|(\hat{c}), w, \omega_B)\left(1 + \sqrt{\frac{\bar{\tau}(A|(\hat{c}), w, \omega_A)}{\bar{\tau}(B|(\hat{c}), w, \omega_A)}}\right)\right]}{\left[\sqrt{\bar{\tau}(A|(\hat{c}), w, \omega_A)}\bar{\tau}(B|(\hat{c}), w, \omega_A)\left(1 + \sqrt{\frac{\bar{\tau}(A|(\hat{c}), w, \omega_B)}{\bar{\tau}(B|(\hat{c}), w, \omega_B)}}\right)\right]},
 \end{aligned}$$

where the bound on the approximation depends on  $n$ ,  $q$  and  $F(0)$ .

Let  $\bar{w} \equiv (1 - r)(\bar{p} - \underline{p})/(r\underline{p})$ . If  $w = \bar{w}$ , the argument of the exponential function in the expression above is zero, and, under condition (A3), the term in brackets is larger than or equal to  $(1 + \underline{\varepsilon})/(1 - \underline{\varepsilon})$ , so for large enough  $n$ ,  $L(\bar{w}) \geq (1 + \underline{\varepsilon})/(1 - \underline{\varepsilon})$ . Let instead  $w$  take any value smaller than  $\bar{w}$  and larger than or equal to

$$\max\left\{0, \frac{(1 - r)(\bar{p} - \underline{p})}{r\underline{p}} - \frac{2q}{r\underline{p}} \frac{F(0)}{1 - F(0)}\right\},$$

so that  $\bar{\tau}(A|(\hat{c}), w, \omega_A) > \bar{\tau}(B|(\hat{c}), w, \omega_B) \geq \bar{\tau}(A|(\hat{c}), w, \omega_B) > \bar{\tau}(B|(\hat{c}), w, \omega_A)$ . Then the argument in the exponential function in the expression above goes to  $-\infty$  as  $n$  goes to infinity, so for large enough  $n$ ,  $L(w) < 1$ . Thus, for large enough  $n$ , for any strategy  $\sigma((c(\beta, \varepsilon)), w)$  there is some  $\hat{w}$  that solves equation 2. Moreover,  $\hat{w}$  is arbitrarily close to  $\bar{w}$  from below.

Since  $\Lambda$  is a mapping from a compact set into itself and is continuous for large  $n$ , it has a fixed point  $((c_n(\beta, \varepsilon)), w_n)$ . Moreover, for large  $n$ ,  $c_n(\beta, \varepsilon)$  is

arbitrarily close to 0 from above for any pair  $\beta, \varepsilon$ , and  $w_n$  is arbitrarily close to  $\bar{w}$  from below.

We claim that for large enough  $n$ ,  $\sigma((c_n(\beta, \varepsilon)), w_n)$  is an equilibrium strategy. To see this, it is straightforward to check using condition (A2) that

$$G(s_A|\beta, \varepsilon, \sigma((c_n(\beta, \varepsilon)), w_n)) \geq 0, \quad G(A|\beta, \varepsilon, \sigma((c_n(\beta, \varepsilon)), w_n)) \leq 0,$$

$$G(s_B|\beta, \varepsilon, \sigma((c_n(\beta, \varepsilon)), w_n)) \geq 0, \quad G(B|\beta, \varepsilon, \sigma((c_n(\beta, \varepsilon)), w_n)) \leq 0.$$

Along the sequence of equilibria  $\sigma((c_n(\beta, \varepsilon)), w_n)$ , the probability of voting for alternative  $A$  in state  $\omega_A$  and the probability of voting for alternative  $B$  in state  $\omega_A$  converge respectively to  $\bar{\tau}(A|\bar{w}, \omega_A)$  and  $\bar{\tau}(B|\bar{w}, \omega_A)$ . Since

$$\frac{\bar{\tau}(A|\bar{w}, \omega_A)}{\bar{\tau}(B|\bar{w}, \omega_A)} > 1,$$

the probability that  $A$  wins the election in state  $\omega_A$  converges to one as  $n$  goes to infinity. A similar argument shows that the probability that  $B$  wins the election in state  $\omega_B$  converges to one as  $n$  goes to infinity.

The proof for the case  $\bar{p} > \underline{p}$  and  $1 - \underline{p}/\bar{p} > r \geq (1/2)(1 - \underline{p}/\bar{p})$  is similar to the one above, but the equilibrium involves rational voters biased in favor of  $B$  voting for  $B$  with probability one when uninformed, and rational voters biased in favor of  $A$  randomizing between abstention and voting for  $B$  when uninformed. The proof for the case  $\bar{p} = \underline{p}$  is also similar to the one above, but the equilibrium involves rational voters abstaining with probability one. Finally, the case  $\bar{p} < \underline{p}$  is analogous to the one above.  $\square$

Recall that uninformed naive voters consider both states equally likely regardless of whether  $A$  and  $B$  are tied or one of the alternatives is ahead of the other by one vote, so that they rather vote for their favorite alternative than abstaining. Informed naive voters, per contra, consider that state  $w_d$  has probability  $1/2 + q$  if they receive signal  $d$  regardless of whether  $A$  and  $B$  are tied or one of the alternatives is ahead of the other by one vote, so that they are willing to vote for their least favorite alternative if  $2q \geq \max\{\underline{\varepsilon}, \bar{\varepsilon}\}$ . Note that this condition is weaker than (A2). That is, *while naive voters are not willing to abstain when ignorant, they are more willing to vote with their signals than rational voters when informed*. If the opposite inequality to (A2) holds for each partisan bias, this means that full information equivalence is possible only there are enough naive voters.

**Theorem 4.** *If*

$$\min\{\underline{\varepsilon}, \bar{\varepsilon}\} > \frac{2H(F, r, \underline{p}, \bar{p})q}{1 + 2\sqrt{\frac{1}{4} - H(F, r, \underline{p}, \bar{p})^2}q^2} \quad \text{and} \quad 2q \geq \max\{\underline{\varepsilon}, \bar{\varepsilon}\},$$

there is a sequence of equilibria for large  $n$  such that along that sequence the probability of choosing alternative  $A$  in state  $\omega_A$  and alternative  $B$  in state  $\omega_B$  converges to 1 if and only if

$$2q > |\bar{p} - \underline{p}| \left( \frac{1}{(1-r)F(0)} - 1 \right).$$

*Proof.* We first prove sufficiency. For any pair

$$(c(\varepsilon)) \text{ such that } 0 \leq c(\varepsilon) \leq \bar{c} \text{ for each } \varepsilon,$$

let  $\sigma((c(\varepsilon)))$  denote the strategy such that naive voters with bias  $\varepsilon$  acquire information and vote according to the signal received if  $c \leq c(\varepsilon)$ , naive voters vote for their favorite alternative if they are uninformed, and rational voters do not acquire costly information and vote for their favorite alternative regardless of whether they are informed or uninformed. Let

$$N : [0, \bar{c}]^2 \rightarrow [0, \bar{c}]^2$$

with

$$N((c(\beta))) = ((\bar{c}(\beta)))$$

be a mapping such that

$$\begin{aligned} \bar{c}(\varepsilon) = \min\{ & \bar{c}, \max\{0, G(s_A|\rho, \bar{\varepsilon}, \sigma(c)) + G(s_B|\rho, \varepsilon, \sigma(c)) \\ & - \max\{G(A|v, \varepsilon, \sigma(c)), G(B|v, \varepsilon, \sigma(c))\}\} \} \end{aligned}$$

for  $\varepsilon \in \{-\underline{\varepsilon}, \bar{\varepsilon}\}$ . A familiar argument establishes that  $N$  has a fixed point  $((c_n(\beta, \varepsilon)))$ . Moreover, for large  $n$ ,  $c_n(\varepsilon)$  is arbitrarily close to 0 from above for any  $\varepsilon$ . Using

$$\min\{\underline{\varepsilon}, \bar{\varepsilon}\} > \frac{2H(F, r, \underline{p}, \bar{p})q}{1 + 2\sqrt{\frac{1}{4} - H(F, r, \underline{p}, \bar{p})^2}q^2}$$

and

$$2q \geq \max\{\underline{\varepsilon}, \bar{\varepsilon}\}$$

it is simple to check that for large enough  $n$ , rational and naive voters are playing best responses, so that  $\sigma(c_n(\varepsilon))$  is an equilibrium. Along any sequence of such equilibria, the probability that a voter votes for  $A$  in state  $\omega_A$  converges to  $(1 - F(0))\bar{p} + F(0)r\bar{p} + F(0)(1 - r)(1/2 + q)$ . Thus, as long as

$$2q > (\underline{p} - \bar{p}) \left( \frac{1}{(1-r)F(0)} - 1 \right),$$

the ratio of the probability of a voter voting for  $A$  to the probability of a voter voting for  $B$  in state  $\omega_A$  remains bounded above 1 as the number of

voters increases, so the probability of  $A$  winning the election in state  $\omega_A$  converges to one. Similarly, as long as

$$2q > (\bar{p} - \underline{p}) \left( \frac{1}{(1-r)F(0)} - 1 \right),$$

the probability of  $B$  winning the election in state  $\omega_B$  converges to one.

We now turn to necessity of the condition in the statement of the theorem. From lemma 1, in any equilibrium there must be a cutoff for each bias such that naive voters acquire information if their cost is below the cutoff and vote according to the signal received, and do not acquire information and vote for their favorite alternative if their cost is above the cutoff. It is simple to check that, along any sequence of equilibria, cutoffs for both possible realizations of the bias go to zero. If

$$\max\{\underline{\varepsilon}, \bar{\varepsilon}\} \leq 2q \leq |\bar{p} - \underline{p}| \left( \frac{1}{(1-r)F(0)} - 1 \right),$$

the bias in favor of one of the alternatives introduced by uninformed naive voters is not offset by informed naive voters. It is also simple to check that along any sequence of equilibria the probability that rational voters acquire information converges to either  $F(0)$  or a number below. Using

$$\min\{\underline{\varepsilon}, \bar{\varepsilon}\} > \frac{2H(F, r, \underline{p}, \bar{p})q}{1 + 2\sqrt{\frac{1}{4} - H(F, r, \underline{p}, \bar{p})^2}q^2},$$

it is simple to show that regardless of the behavior of other rational voters, as long as the probability of information acquisition converges to either  $F(0)$  or a number below, it is a (unique) best response for every rational voter to vote for the voter's favorite alternative, which reinforces the bias in favor of one of the alternatives introduced by uninformed naive voters. Thus, along any sequence of equilibria, it must be the case that  $A$  wins the election regardless of the state if  $\bar{p} > \underline{p}$ , and  $B$  wins the election in the regardless of the state if  $\bar{p} < \underline{p}$ .  $\square$

## 6. FINAL REMARKS

We have shown that, when discovering their private information is costly for all voters, there is full information equivalence only if all voters are rational, and their preferences are close enough to being aligned for them to be willing to abstain when uninformed, which most of them do. In a way, the model resembles collective choice by means of committees that self-select on the basis of the opportunity cost of expertise and are small in relation to the potential population, perhaps like some faculty meetings or voluntary consumer reviews in the internet. If discovering private information is costly for all voters and some voters are naive, the incentive for

abstention is destroyed since uninformed naive voters rather vote for their favorite alternative than abstain, so that most voters who participate in the election are uninformed. As a result, if information is costly for all voters and some voters are naive, full information equivalence is not possible. That is to say, voters' opinions do not contain information and elections fail to aggregate the information potentially dispersed among the electorate.

When discovering their private information is costless for voters with some probability, there may be full information equivalence even if some voters are naive. One possibility is that rational voters offset any net bias given by naive voters participating in the election, and a condition for rational abstention is satisfied. An alternative set of circumstances relies on informed naive voters voting according to the information they have, even if rational voters are unwilling to abstain or indeed, unwilling to use their private information even if it is free. In both cases, full information equivalence is possible only if enough free information is available to voters to begin with. The role of naive voters is different in the two cases, though. Increasing the fraction of the population who is naive may be bad for information aggregation when there are few naive voters to begin with, but may be good for information aggregation when there are already many naive voters. Naive behavior may indeed be a curse or a blessing, depending on the circumstances.

Though we conduct the analysis in a sparse model, it should be clear that the main arguments as outlined in the discussion at the top of sections 4 and 5 survive the extensions that come naturally to mind, that is several partisan biases, multiple levels of cursedness, and multiple levels of information quality. In particular, if naive voters are not fully cursed, but cursed enough so that a condition similar to (A1) fails for them, they will be unwilling to abstain when uninformed but willing to use their private information when rational voters would not. Thus, naive voters would still have an ambivalent role for information aggregation.

Going back to the sort of environment mentioned in the introduction, that is, a large election dominated by a single issue about which voters have common preferences, the main thrust of this paper is that information aggregation may be successful even under "realistic" circumstances, including mass participation, widespread ignorance, and naive behavior. This opens up the question of what "realistic" institutions improve the performance of elections in the sense of increasing the probability of making a good choice, and the related question of how to best design polls with a large set of potential participants and not fully rational behavior.

## REFERENCES

- [1] David Austen-Smith and Jeffrey Banks (1996), Information aggregation, rationality, and the Condorcet jury theorem, *American Political Science Review* 90, 34-45.
- [2] Marco Battaglini, Rebecca Morton and Thomas R. Palfrey (2010), The swing voter's curse in the laboratory, *Review of Economic Studies*, 77: 61-89.
- [3] Ilja Bronshtein, Konstantin Semendyayev, Gerhard Musiol and Heiner Muehlig (2003), *Handbook of Mathematics*, 4th ed., Springer-Verlag: Berlin.
- [4] Jean Antoine Nicolas de Caritat, Marquis de Condorcet (1785)[1986], *Discours préliminaire de l'Essai sur l'application de l'analyse à la probabilité des décisions rendues à la pluralité des voix*, reprinted in Condorcet, *Sur les Élections et autres textes*, Librairie Artheme Fayard: Paris.
- [5] Alexander Elbittar, Andrei Gomborg, César Martinelli and Thomas R. Palfrey (2011), An experiment on collective decision with costly information, CalTech and ITAM, manuscript.
- [6] Ignacio Esponda (2007), Behavioral equilibrium in economies with adverse selection, *American Economic Review* 98, 1269-1291.
- [7] Erik Eyster and Matthew Rabin (2005), Cursed equilibrium, *Econometrica* 73: 1623-1672.
- [8] Timothy Feddersen and Wolfgang Pesendorfer (1996), The swing voter's curse, *American Economic Review* 86, 408-424.
- [9] Timothy Feddersen and Wolfgang Pesendorfer (1997), Voting behavior and information aggregation in elections with private information, *Econometrica* 65, 1029-1058.
- [10] William Feller (1971), *An Introduction to Probability Theory and its Applications*, vol. II, second ed., Wiley: New York.
- [11] Marek Fisz (1955a), The limiting distribution of a function of two independent random variables and its statistical application, *Colloquium Mathematicum* 3, 138-146.
- [12] Marek Fisz (1955b), Refinement of a probability limit theorem and its application to Bessel functions, *Acta Mathematica Hungarica* 6, 199-202.
- [13] Dino Gerardi and Leeat Yariv (2008), Information acquisition in committees, *Games and Economic Behavior* 62, 436-459.
- [14] Serena Guarnaschelli, Richard D. McKelvey and Thomas R. Palfrey (2000), An experimental study of jury decision rules, *American Political Science Review* 94, 407-423.
- [15] Philippe Jehiel and Frederic Koessler (2008), Revisiting games of incomplete information with analogy-based expectations, *Games and Economic Behavior* 62, 533-557.
- [16] Dimitris Karlis and Ioannis Ntzoufras (2003), Analysis of sports data using bivariate Poisson models, *Journal of the Royal Statistical Society: Series D (The Statistician)* 52, 381-393.
- [17] Dimitris Karlis and Ioannis Ntzoufras (2006), Bayesian analysis of the differences of count data, *Statistics in Medicine* 25, 1885-1905.
- [18] César Martinelli (2006), Would rational voters acquire costly information?, *Journal of Economic Theory* 129, 225-251.
- [19] César Martinelli (2007), Rational ignorance and voting behavior, *International Journal of Game Theory* 35, 315-335.
- [20] Andrew McLennan (1998), Consequences of the Condorcet jury theorem for beneficial information aggregation by rational agents, *American Political Science Review* 92, 413-418.



- [21] Kaushik Mukhopadhyaya (2005), Jury size and the free rider problem, *Journal of Law, Economics and Organization* 19, 24-44.
- [22] Roger B. Myerson (1998a), Population uncertainty and Poisson games, *International Journal of Game Theory* 27, 375-392.
- [23] Roger B. Myerson (1998b), Extended Poisson games and the Condorcet jury theorem, *Games and Economic Behavior* 25, 111-131.
- [24] Roger B. Myerson (2000), Large Poisson games, *Journal of Economic Theory* 94, 7-45.
- [25] Roger B. Myerson (2002), Comparison of scoring rules in Poisson voting games, *Journal of Economic Theory* 103, 219-251.
- [26] Santiago Oliveros (2011a), Abstention, ideology and information acquisition, UC Berkeley, mimeo.
- [27] Santiago Oliveros (2011b), Aggregation of endogenous information in large elections, UC Berkeley, mimeo.
- [28] Peter C. Ordeshook and Thomas R. Palfrey (1988), Agendas, strategic voting, and signalling with incomplete information, *American Journal of Political Science*, 32: 441-466.
- [29] Nicola Persico (2004), Committee design with endogenous information, *Review of Economic Studies* 71, 165-191.
- [30] John Skellam (1946), The frequency distribution of the difference between two Poisson variates belonging to different populations, *Journal of the Royal Statistical Society: Series A*, 109: 296.