Would Rational Voters Acquire Costly Information?

César Martinelli*

September 8, 2003

Abstract

In order to study the aggregate implications of Downs's rational ignorance hypothesis, we analyze an election in which voters are uncertain about which of two alternatives is better for them. Voters can acquire some costly information about the alternatives. In agreement with Downs's contention, as the number of voters increases, individual investment in political information declines to zero. However, the election outcome is likely to correspond to the interest of the majority if the marginal cost of information acquisition approaches zero as the information acquired becomes nearly irrelevant. Under certain conditions, the election outcome corresponds to the interests of the majority with probability approaching one. Thus, "rationally ignorant" voters are consistent with a well-informed electorate. Moreover, a result like Condorcet's jury theorem may hold even if information is costly. *JEL* D72, D82.

*Centro de Investigación Económica, Instituto Tecnológico Autónomo de México. *Mailing address:* CIE- ITAM, Camino a Sta Teresa 930, 10700 México D.F., MEXICO. *Phone:* +52 55 56284197. *Fax:* +52 55 56284058. *E-mail:* `martinel@itam.mx`. I would like to thank seminar audiences at Caltech, ITAM, SMU and University of Oklahoma for useful comments.

1 Introduction

One of the most influential contributions of Anthony Downs's *An Economic Theory of Democracy* to the economic modelling of politics is the concept of "rational ignorance." Given that each individual voter has a negligible probability of affecting the outcome in a large election, voters will not have an incentive to acquire political information before voting. In a situation in which discovering their interests or "true views" takes time and effort from individual citizens, the result may be a failure of democracy to produce a result consistent with the interests of the majority. In Downs's words,

> If all others express their true views, he [the voter] gets the benefit of a well-informed electorate no matter how well-informed he is; if they are badly informed, he cannot produce those benefits himself. Therefore, as in all cases of individual benefits, the individual is motivated to shirk his share of the costs: he refuses to get enough information to discover his true views. Since all men do this, the election does not reflect the true consent of the governed. (Downs 1957, p. 246)

We can actually draw a distinction between two versions of the rational ignorance hypothesis. The "weak version" is that individual voters, realizing that each vote has a negligible probability of affecting the outcome of the election, invest very little or no effort in acquiring political information. The "strong version" is that the election outcome itself will not be more likely to reflect the interests of the majority than, say, a fair coin toss. In this paper, we develop a formal model that is consistent with the weak version of the rational ignorance hypothesis, but contradicts the strong version.

A good deal of the literature on the influence activities of interest groups assumes that a decisive fraction of the electorate is uninformed because individual voters have little incentive to get political information (see e.g. Becker 1983 for an explicit discussion). Becker (1985) argues that efficiency may be restored in the voting market because of the activity of influence groups. Coate and Morris (1995) point out that the reelection motive may induce incumbent politicians to behave efficiently unless voters are uncertain about politicians' types. (In their view, and Becker's, efficiency does not mean that transfers from the majority to interest groups do not occur; it only means that those transfers are carried out with minimum dead weight costs.) Closer to our point, Wittman (1989) calls into question the idea that the costs of information fall on the voter instead of on political entrepreneurs.

We provide a different rationale for elections to reflect the interests of the majority. In our model, there are no interest groups or active politicians. Voters do not have access to free information. Instead, they may acquire some information, at a cost. Crucially, acquiring poor information is cheap. We show that, as the number of voters increases, voters acquire less and less information. However, under some conditions detailed below, the outcome of the election is very likely to correspond to the interests of a majority of voters. Thus, the electorate may be quite well-informed even if individual voters are (at least asymptotically) rationally ignorant.

We study an election in which "moderate" or "swing" voters do not know which of two alternatives is better for them. Voters may acquire a costly signal about the alternatives. The signal is correct with probability $1/2 + x$, where x is chosen by the voter. We refer to x as the quality of the signal. The cost of acquiring the signal is given by some convex function $C(x)$. Our first three results describe information acquisition and information aggregation in the context of this model.

Theorem 1 shows that the quality of information acquired by individual voters goes to zero as the size of the electorate increases. However, if $C'(0) = 0$, then the quality of information is positive for an arbitrarily large electorate. The reason is simple: the probability of being pivotal is not exactly zero. (If the probability of being pivotal were zero, instrumentally rational voting behavior would be unconstrained.)

Theorem 2 provides an estimate of the limit probability of choosing the best alternative. If $C'(0) = 0$ and $C''(0) < \infty$, this probability is strictly larger than 1/2. Moreover, this probability goes to one as $C''(0)$ approaches zero, or as the importance attached by moderate voters to the election grows unboundedly, and it increases with the fraction of moderate voters in the society. If $C'(0) = C''(0) = 0$, the limit probability of choosing the best alternative is actually one. Successful information aggregation is possible because the information acquired by each moderate voter goes to zero but it does so slowly enough to allow the effect of large numbers to kick in.

However, voters acquire too little information (with respect to a symmetric optimal strategy) even in the limit unless $C''(0) = 0$.

It is reasonable to believe that voters are involuntarily exposed to a flow of political information in the course of everyday activities – a point already acknowledged by Downs (1957, p. 245), who relies on the unwillingness of voters to assimilate even freely available information in order to support the rational ignorance hypothesis. If the function C simply reflects the cost of "paying a little attention," the conditions for at least partially successful information aggregation, that is $C'(0) = 0$ and $C''(0) < \infty$, do not appear unduly restrictive.

Theorem 3 shows that elections with information acquisition will be almost always very close. On one hand, elections must be close to keep individual voters acquiring some information. On the other hand, the fact that voters acquire vanishingly little information keeps elections close even as the number of voters increases. Theorems 2 and 3 are illustrated by looking at the distribution of the fraction of votes received by the best alternative for moderate voters. As the number of voters increases, the mean of this distribution converges (from above) to 0.5, reflecting the fact that individual moderate voters invest less and less in discovering their "true views." However, the variance of the distribution shrinks very fast, ensuring that there is partially or even completely successful information aggregation in the limit.

Taken together, our results support the idea that elections serve the interests of the majority better than what the rational ignorance hypothesis would seem to indicate at first glance. They suggest that models of public opinion that take into account the production of information by the media, interest groups, and the like, can be enriched by considering the aggregate implications of voters investing some small (but positive) effort in costly information processing.

Note that political information in our model is a public good, at least from the point of view of moderate voters. As in other instances of privately provided public goods, there is an incentive to free ride on other voters, and in fact moderate voters underinvest in political information in relation to a symmetric optimal profile. In the traditional problem of private provision of public goods in large economies (as described e.g. in Andreoni 1988), the marginal cost of contributing is constant, and the contributions of others reduce the marginal utility of additional units of the public good up to the point where it does not compensate most agents to contribute. In our model, the marginal cost is small for small contributions, and the marginal benefit of contributing is bounded below because the probability of being pivotal is nonzero even if all moderate voters acquire perfect information due to uncertainty about voters' preferences. Opposite to what happens in the traditional problem, aggregate investment in political information grows unboundedly and in some cases approximate efficiency can be obtained in the limit.

Our model is related to the literature on information aggregation in elections inspired by Condorcet's jury theorem (e.g. Miller 1986, Austen-Smith and Banks 1996, Feddersen and Pesendorfer 1997, McLennan 1998, Duggan and Martinelli 2001). This literature typically assumes that there is some information dispersed among the voters, while in our paper the distribution of information arises endogenously through the actions of voters. As a consequence, we obtain that larger electorates do better than a single decision maker in some circumstances, but not in others.

Recently, Persico (1999) has proposed another model of endogenous information in collective decision making. In Persico's model, the quality of the signal is given; voters can either acquire or not acquire information. As a consequence, in his model it is not possible to have arbitrarily large numbers of voters acquiring arbitrarily poor information. Persico is concerned with the optimal design of committees, i.e. the optimal selection of committee size and voting rule, while we consider an environment where majority rule is optimal and concern ourselves with the positive issue of endogenous production and aggregation of information in large elections.

2 The Model

We analyze an election with two alternatives, A and B. There are $2n+1$ voters ($i = 1, \ldots, 2n+1$). A voter's utility depends on the chosen alternative $d \in \{A, B\}$, a preference parameter $t \in \{t_A, t_M, t_B\}$, the state $z \in \{z_A, z_B\}$, and the quality of information acquired by the voter before the election $x \in [0, 1/2]$. Acquiring information of quality x has a utility cost given by $C(x)$, so the utility of a voter can be written as

$$U(d, t, z) - C(x).$$

At the beginning of time, nature selects the state and the type of each voter. Both states are equally likely ex ante. Each voter's type is equal to t_A with probability ϵ, to t_B with probability ϵ, and to t_M with probability $1-2\epsilon$, where $0<\epsilon<1/2$. Voters' types are independent from each other and from the realization of the state. (The symmetry assumptions on the ex ante probability of states and the fraction of partisan voters simplify the presentation, but, as shown in Section 6, are not necessary for our asymptotic results on information aggregation.)

Each voter knows her preference type but is uncertain about the type of other voters. Voters are also uncertain about the realization of the state. After learning her type, a voter decides the quality of her information. After deciding on x, the voter receives a signal $s \in \{s_A, s_B\}$. The probability of receiving signal s_A in state A is equal to the probability of receiving signal s_B in state B and is given by $1/2+x$. That is, the likelihood of receiving the "right" signal is increasing in the quality of information acquired by the voter; if the voter acquires no information the signal is uninformative. Signals are private information.

The election takes place after voters receive their signals. A voter can either vote for A or vote for B. The alternative with most votes is chosen. (Again, as discussed in Section 6, the assumption that there are no abstentions can be relaxed without consequences.)

Let

$$v(t,z) = U(A,t,z) - U(B,t,z).$$

We assume that

$$\begin{array}{rcccr} v(t_M, z_A) & = & -v(t_M, z_B) & = & r, \\ v(t_A, z_A) & = & v(t_A, z_B) & = & q, \\ v(t_B, z_A) & = & v(t_B, z_B) & = & -q \end{array}$$

where r and q are two positive real numbers. Voters of type t_A and t_B are "extremists," who favor alternative A or alternative B regardless of the possible circumstances or states. Voters of type t_M are "moderates," willing to support alternative A or alternative B depending on the circumstances.

The cost function C is strictly increasing, strictly convex, and twice continuously differentiable on $(0, 1/2)$. We assume that $C(0) = 0$, so that acquiring no information is costless. Note that $C'(0) \in [0, \infty)$. If $C''(x)$ grows unboundedly as x goes to zero, we use the notation $C''(0) = \infty$. Thus, $C''(0) \in [0, \infty]$.

After describing the environment, we turn now to the description of strategies and the definition of equilibrium in the model. A pure strategy is a pair a_x, a_v, where

$$a_x : \{t_A, t_M, t_B\} \to [0, 1/2]$$

is a mapping from a voter's type to a quality of information x, and

$$a_v : \{t_A, t_M, t_B\} \times \{s_A, s_B\} \to \{A, B\}$$

is a mapping from a voter's type and the signal received to a decision to vote for A or for B. A mixed strategy for voter i is a probability distribution α_i over the set of pure strategies.

A *voting equilibrium* $\overline{\alpha}$ ($\alpha_i = \alpha$ for all i) is a symmetric Nash equilibrium in which no voter uses a weakly dominated strategy.

Clearly, an equilibrium strategy will only assign positive probability to pure strategies such that $a_x(t_A) = a_x(t_B) = 0$ and, from elimination of weakly dominated strategies, $a_v(t_A, s) = A$, $a_v(t_B, s) = B$ for $s \in \{s_A, s_B\}$, so we can restrict our attention to pure strategies satisfying those constraints. It remains to determine the equilibrium behavior of moderate voters.

Let $P_{\alpha_1,\ldots,\alpha_{2n+1}}(A|z_A)$ and $P_{\alpha_1,\ldots,\alpha_{2n+1}}(B|z_B)$ be the probability of alternative A winning the election if the state is z_A and the probability of alternative B winning the election if the state is z_B, for a given strategy profile. Let $E_{\alpha_i}(C(x_i))$ be the expected cost of information acquisition for voter i given her own strategy, conditional on her type being t_M. Then, the ex ante utility for voter i is given by

$$(1-2\epsilon)\left[\frac{1}{2}P_{\alpha_1,\ldots,\alpha_{2n+1}}(A|z_A) + \frac{1}{2}P_{\alpha_1,\ldots,\alpha_{2n+1}}(B|z_B)\right] r - (1-2\epsilon)E_{\alpha_i}(C(x_i))$$

plus some constant term which we ignore hereafter. We refer to the term in brackets as the *probability of choosing the right alternative*. We are particularly interested in the limit value of this probability as the size of the electorate increases.

3 Rational Ignorance

In this section we describe the equilibrium behavior of moderate voters. We show that, according to the weak version of the rational ignorance hypothesis, in large elections voters acquire vanishingly little information or no information at all.

Define

$$G(x) = \frac{(2n)!}{n!n!}\left(\frac{1}{4} - (1-2\epsilon)^2 x^2\right)^n r - C'(x).$$

Intuitively, this expression gives us the marginal benefit of acquiring quality of information x for a given voter when every other voter is acquiring x. The first term in the definition of G is the probability that a given voter is pivotal multiplied by the gain in reaching the right decision. The second term is the marginal cost of quality of information x. Note that G is strictly decreasing.

Let

$$x_M = \begin{cases} 0 & \text{if } G(0) \leq 0 \\ 1/2 & \text{if } G(1/2) \geq 0 \\ G^{-1}(0) & \text{otherwise.} \end{cases}$$

The first term in the definition of G is strictly positive and converges to zero as n goes to infinity for any sequence of $x \in [0, 1/2]$. Thus, if $C'(0) = 0$, we get $G(x) > 0$ for every x and then $x_M > 0$. However, if we let n go to infinity while keeping ϵ, r and the function C constant, x_M should converge to 0.

If $C'(0) > 0$, let $n(r, C)$ be the minimum n such that

$$\frac{(2n)!}{n!n!}\left(\frac{1}{4}\right)^n r \leq C'(0).$$

Note that for any $n \geq n(r, C)$, we get $x_M = 0$.

We have

Theorem 1
(i) If $C'(0) = 0$, there is a unique voting equilibrium. In this equilibrium, the pure strategy given by $a_x(t_M) = x_M$, $a_v(t_M, s_A) = A$ and $a_v(t_M, s_B) = B$ is played with probability one.
(ii) If $C'(0) > 0$ and $n \geq n(r, C)$, every equilibrium assigns probability one to the set of pure strategies such that $a_x(t_M) = 0$.

PROOF. Suppose that every voter other than i adopts the strategy α, and let $P_\alpha(\text{piv}|z_A)$ and $P_\alpha(\text{piv}|z_B)$ be the probabilities that n voters other than i vote for A and n voters other than i vote for B in state z_A and in state z_B, respectively. The expected utility for voter i of adopting the pure strategy $a_x(t_M) = x$, $a_v(t_M, s_A) = A$ and $a_v(t_M, s_B) = B$ for any $x \in [0, 1/2]$ is given by $(1 - 2\epsilon)$ times

$$\left[\frac{1}{2}P_\alpha(\text{piv}|z_A)\left(\frac{1}{2} + x\right) + \frac{1}{2}P_\alpha(\text{piv}|z_B)\left(\frac{1}{2} + x\right)\right] r - C(x) \tag{1}$$

plus a term that does not depend on the action chosen by i. Note that the expected utility is a strictly concave function of x.

We can show that, in equilibrium, it has to be the case that the pure strategy with $a_v(t_M, s_A) = A$ and $a_v(t_M, s_B) = B$ is strictly superior to any other pure strategy for a given choice $x > 0$ of information quality. For suppose that it is not superior to the pure strategy with $a_v(t_M, s_A) = A$ and $a_v(t_M, s_B) = A$ (other cases are treated similarly). Then

$$P_\alpha(\text{piv}|z_A)\left(\frac{1}{2} + x\right) + P_\alpha(\text{piv}|z_B)\left(\frac{1}{2} + x\right) \leq P_\alpha(\text{piv}|z_A),$$

that is

$$\frac{P_\alpha(\text{piv}|z_A)}{P_\alpha(\text{piv}|z_B)} \geq \frac{1/2 + x}{1/2 - x} > 1.$$

Then, for every choice of information quality, every pure strategy such that $a_v(t_M, s_A) = B$ is strictly dominated by the pure strategy with $a_v(t_M, s_A) = A$ and $a_v(t_M, s_B) = A$. That is, voter i assigns probability one to the set of pure strategies with $a_v(t_M, s_A) = A$. Now let $\beta(x')$ be the probability that i plays a pure strategy with $a_v(t_M, s_B) = B$ and with quality of information smaller or equal than x', for any $x' \in (0, 1/2]$, as induced by voter i's strategy. Let p_A and p_B be the probabilities with which voter i votes for A in state z_A and for B in state z_B, as induced by voter i strategy. Then $p_A = (1-2\epsilon)(1-\int_0^{1/2}(1/2-x')d\beta(x'))+\epsilon$ and $p_B = (1-2\epsilon)(\int_0^{1/2}(1/2+x')d\beta(x'))+\epsilon$. It follows that $|p_A - 1/2| \geq |p_B - 1/2|$. But then, $p_A^n(1 - p_A)^n \leq p_B^n(1 - p_B)^n$. Since equilibrium is symmetric, we get $P_\alpha(\text{piv}|z_A) \leq P_\alpha(\text{piv}|z_B)$, a contradiction.

From the previous paragraph, we can restrict our attention to pure strategies with $a_x(t_M) = x$, $a_v(t_M, s_A) = A$ and $a_v(t_M, s_B) = B$ for any $x > 0$.

From equation (1), if any such pure strategy is optimal for voter i, it is the unique optimal pure strategy among strategies with $a_x(t_M) = x$, $a_v(t_M, s_A) = A$ and $a_v(t_M, s_B) = B$ for any $x \geq 0$. Moreover, the argument in the previous paragraph shows that $P_\alpha(\text{piv}|z_A) = P_\alpha(\text{piv}|z_B)$. But this implies that all pure strategies with no information acquisition have the same expected payoff. Thus, if there is some information acquisition, it has to be the case that the voting equilibrium is a pure strategy equilibrium with $a_x(t_M) = x^*$, $a_v(t_M, s_A) = A$ and $a_v(t_M, s_B) = B$ for some $x^* > 0$.

Now, suppose that every voter other than i adopts the pure strategy with $a_x(t_M) = \tilde{x}$, $a_v(t_M, s_A) = A$ and $a_v(t_M, s_B) = B$ for some $\tilde{x} \geq 0$. Then the probabilities of voter i being pivotal in states z_A and z_B are

$$\begin{aligned} P_\alpha(\text{piv}|z_A) &= \binom{2n}{n}(\epsilon + (1-2\epsilon)(1/2+\tilde{x}))^n(\epsilon + (1-2\epsilon)(1-1/2-\tilde{x}))^n \\ &= \frac{(2n)!}{n!n!}\left(\frac{1}{4} - (1-2\epsilon)^2\tilde{x}^2\right)^n \\ &= P_\alpha(\text{piv}|z_B). \end{aligned}$$

Replacing these probabilities in equation (1), we get that the expected utility for voter i of adopting the pure strategy with $a_x(t_M) = x$, $a_v(t_M, s_A) = A$ and $a_v(t_M, s_B) = B$ is a positive affine function of

$$\frac{(2n)!}{n!n!}\left(\frac{1}{4} - (1-2\epsilon)^2\tilde{x}^2\right)^n\left(\frac{1}{2} + x\right)r - C(x).$$

The first derivative of this expression with respect to x is

$$H(x, \tilde{x}) = \frac{(2n)!}{n!n!}\left(\frac{1}{4} - (1-2\epsilon)^2\tilde{x}^2\right)^n r - C'(x).$$

The second derivative is negative for $x > 0$. Note that $H(x, x) = G(x)$. Thus, the distribution that gives probability one to the pure strategy $a_x(t_M) = x_M$, $a_v(t_M, s_A) = A$ and $a_v(t_M, s_B) = B$ constitutes a voting equilibrium. Moreover, this is the only equilibrium in which information acquisition is possible.

To check that there is no equilibrium without information acquisition if $C'(0) = 0$, note that the probability of being pivotal is positive for any choice of strategy by other voters. Thus, from equation (1), the marginal benefit of acquiring information is larger than the marginal cost for x sufficiently close to 0. Finally, if $C'(0) > 0$, we know that $x_M = 0$ for $n \geq n(r, C)$. It follows that there is no equilibrium with information acquisition for $n \geq n(r, C)$.

□

4 Information Aggregation

In this section, we let n go to infinity while keeping the other parameters of the model (ϵ, q, r) and the function C constant. We study the limit behavior of the probability of choosing the right alternative along the sequence of voting equilibria thus obtained.

From the previous section, we know that if $C'(0) > 0$, there is no information acquisition for n large enough. Thus, the probability of choosing the right alternative converges to $1/2$ – the only possibility for an uninformed electorate since the two states are equally likely. However, if $C'(0) = 0$, moderate voters acquire some information for every n. In this section we show that, if $C'(0) = C''(0) = 0$, the quality of information acquired by moderate voters declines slowly enough to allow the probability of choosing the right alternative to converge to one. In other words, even though in the limit voters are rationally ignorant, the electorate is perfectly well-informed. If $C'(0) = 0$ and $C''(0) = c \in (0, \infty)$, information aggregation is still possible; in this case the limit value of the probability of choosing the right alternative increases with r/c (the gain of choosing the right alternative divided by the limit of the second derivative of the cost function). Finally, if $C'(0) = 0$ but $C''(0) = \infty$, the quality of information acquired by voters declines very fast so the probability of choosing the right alternative converges to $1/2$.

As an example, consider the cost function $C(x) = x^\gamma$, with $\gamma > 1$. Theorem 2 below establishes that for $\gamma < 2$, the probability of choosing the right alternative converges to $1/2$. For $\gamma = 2$, the probability of choosing the right alternative converges to some value between $1/2$ and one. (This value is about 0.74123 for $r = 1$ and ϵ close to zero.) For $\gamma > 2$, the probability of choosing the right alternative converges to one.

If $C'(0) = 0$ and $C''(0) = c \in (0, \infty)$, let $k(\epsilon, r, c)$ be the solution to

$$k/\phi(k) = 4(r/c)(1 - 2\epsilon),$$

where ϕ is the standard normal density. Note that $k(\epsilon, r, c) \in (0, \infty)$, and moreover, $k(\epsilon, r, c)$ is increasing in r/c and grows unboundedly as r/c goes to infinity. As we will see below, $k(\epsilon, r, c)$ is an indicator of the information held by the electorate in large elections. It is equal to the limit of the product of the information acquired by each individual and the square root of the size of the electorate, multiplied by a constant term.

We have

Theorem 2
(i) If $C'(0) = C''(0) = 0$, the probability of choosing the right alternative converges to one as the size of the electorate increases.
(ii) If $C'(0) = 0$ and $C''(0) = c \in (0, \infty)$, the probability of choosing the right alternative converges to $\Phi(k(\epsilon, r, c))$, where Φ is the standard normal distribution.
(iii) If $C'(0) > 0$ or $C''(0) = \infty$, the probability of choosing the right alternative converges to $1/2$.

Note that in case (ii) the probability of choosing the right alternative goes to one as c goes to zero or as r (the importance of the election for moderate voters) goes to infinity. Moreover, it is increasing in the expected fraction of moderate voters in the society, $1 - 2\epsilon$, but remains bounded away from one even if ϵ approaches zero.

We prove the theorem via two lemmas. In the two lemmas we write x_n to represent the value of x_M (as defined in the previous section) for a given n. We know from the previous section that if $C'(0) = 0$, then x_n is positive but converges to zero as n grows to infinity. The first lemma tells us how fast x_n converges to zero in each of the three cases of the theorem. The second lemma uses a version of the central limit theorem to establish the desired results. A direct application of the central limit theorem is not possible because the distribution representing the decision of a given voter changes with n. Instead, we use a normal approximation result for finite samples, the Berry-Esseen theorem.

Lemma 1
(i) If $C'(0) = C''(0) = 0$, then $n^{1/2}x_n$ goes to $+\infty$ as n grows arbitrarily large.
(ii) If $C'(0) = 0$ and $C''(0) = c < \infty$, then
$\lim_{n\to\infty} n^{1/2}x_n k(\epsilon, r, C)/(2\sqrt{2}(1-2\epsilon))$.
(iii) If $C'(0) = 0$ and $C''(0) = \infty$, then $\lim_{n\to\infty} n^{1/2}x_n = 0$.

PROOF. For large n, if $C'(0) = 0$ then x_n is given by the solution to $G(x_n) = 0$ or

$$\frac{(2n)!}{n!n!}\left(\frac{1}{4} - (1-2\epsilon)^2 {x_n}^2\right)^n r = C'(x_n).$$

Letting $y_n = n^{1/2}x_n$ we get

$$\frac{(2n)!}{n!n!}\left(\frac{1}{4} - (1-2\epsilon)^2 \frac{{y_n}^2}{n}\right)^n r = C'(n^{-1/2}y_n).$$

Using the mean value theorem for C' and rearranging slightly we have

$$\frac{(2n)!}{n!n!}\frac{n^{1/2}}{2^{2n}}\left(1 - 4(1-2\epsilon)^2 \frac{{y_n}^2}{n}\right)^n r = y_n\, C''(\xi_n) \tag{2}$$

for some ξ_n between zero and $n^{-1/2}y_n$.

Note that

$$\frac{(2n)!}{n!n!}\frac{n^{1/2}}{2^{2n}} \to \pi^{-1/2}$$

(from Stirling's formula) and

$$0 < \left(1 - 4(1-2\epsilon)^2 \frac{{y_n}^2}{n}\right)^n < 1$$

(because $0 < y_n < n^{1/2}$).

Now consider the case $C'(0) = C''(0) = 0$. Suppose that along some subsequence y_n converges to a finite $L \geq 0$. Then, along the subsequence the right hand side of equation (2) converges to zero. However, the left hand side converges to a positive number, as can be seen from the fact that

$$\left(1 - 4(1-2\epsilon)^2 \frac{{y_n}^2}{n}\right)^n \to \exp\{-4(1-2\epsilon)^2 L^2\}$$

(see e.g. Durrett 1991, p. 94). Thus, y_n diverges to $+\infty$, which establishes case (i).

Consider the case $C'(0) = 0$ and $C''(0) = c < \infty$. Suppose that along some subsequence y_n converges to a finite $K \geq 0$. Following the steps of the previous case, we get that K must satisfy $\pi^{-1/2} \exp\{-4(1-2\epsilon)^2 K^2\} r = Kc$ or, equivalently, $K = k(\epsilon, r, C)/(2\sqrt{2}(1-2\epsilon))$. It remains to show that along no subsequence y_n diverges to $+\infty$. To see this, note that the right hand side of equation (2) grows without bound if y_n goes to infinity, while for any positive δ, the left hand side is smaller than $(\pi^{-1/2} + \delta)r$ for n large enough. This establishes case (ii).

Finally, consider the case $C'(0) = 0$ and $C''(0) = \infty$. If along some subsequence y_n converges to a finite $L > 0$ or diverges to $+\infty$, the right hand side of equation (2) grows without bound, while the left hand side is bounded by the argument above. □

Lemma 2 *Suppose that $C'(0) = 0$. If $\lim_{n\to\infty} n^{1/2} x_n = K < \infty$, the probability of choosing the right alternative converges to $\Phi(2\sqrt{2}(1-2\epsilon)K)$, where Φ is the standard normal distribution. If $n^{1/2} x_n$ diverges to $+\infty$, the probability of choosing the right alternative converges to one.*

Proof. Suppose the state is z_A (similar calculations hold if the state is z_B). Given the equilibrium strategy described in Theorem 1(i), the event of a given voter voting for A in state z_A corresponds to a Bernoulli trial with probability of success

$$(1-2\epsilon)(1/2 + x_n) + \epsilon = 1/2 + (1-2\epsilon)x_n.$$

For $n = 1, 2, \ldots$ and $i = 1, \cdots, 2n+1$ define the random variables

$$V_i^n = \begin{cases} 1/2 - (1-2\epsilon)x_n & \text{if voter } i \text{ votes for } A, \\ -1/2 - (1-2\epsilon)x_n & \text{if voter } i \text{ votes for } B. \end{cases}$$

For each n, the random variables V_i^n are iid. Moreover,

$$\begin{aligned} &E(V_i^n) = 0, \\ &E((V_i^n)^2) = 1/4 - (1-2\epsilon)^2 x_n^2, \quad \text{and} \\ &E(|V_i^n|^3) = 2\left(1/16 - (1-2\epsilon)^4 x_n^4\right). \end{aligned}$$

Let F_n stand for the distribution of the normalized sum

$$(V_1^n + \cdots + V_{2n+1}^n)/\sqrt{E((V_i^n)^2)(2n+1)}.$$

Note that A loses the election if it obtains n or fewer votes, that is, if

$$V_1^n + \cdots + V_{2n+1}^n + (2n+1)(1/2 + (1-2\epsilon)x_n) \leq n$$

or equivalently

$$V_1^n + \cdots + V_{2n+1}^n \leq -1/2 - (2n+1)(1-2\epsilon)x_n.$$

Then, the probability of A winning the election is $1 - F_n(J_n)$, where

$$J_n = \frac{-1/2 - (2n+1)(1-2\epsilon)x_n}{\sqrt{E((V_i^n)^2)(2n+1)}}.$$

Now, from the Berry-Esseen theorem (see Feller 1971, p. 542 or Durrett 1991, p. 106), for all w,

$$|F_n(w) - \Phi(w)| \leq \frac{3E(|V_i^n|^3)}{E((V_i^n)^2)^{3/2}\sqrt{2n+1}}.$$

The right hand side of the equation above converges to zero as n goes to infinity, so we obtain an increasingly good approximation using the normal distribution even though the distribution of V_i^n changes with n. Thus,

$$\lim_{n\to\infty} |F_n(J_n) - \Phi(J_n)| = 0.$$

If $\lim_{n\to\infty} n^{1/2}x_n = K < \infty$, then J_n converges to $-2\sqrt{2}(1-2\epsilon)K$. Since Φ is continuous,

$$\lim_{n\to\infty} |\Phi(J_n) - \Phi(-2\sqrt{2}(1-2\epsilon)K)| = 0.$$

Thus, the probability of A winning converges to $1 - \Phi(-2\sqrt{2}(1-2\epsilon)K)$. The desired result follows from symmetry.

If $n^{1/2}x_n$ goes to infinity with n, then J_n goes to $-\infty$. Thus, for arbitrarily large L, the probability of A winning the election is larger than $1 - F_n(-L)$ for n large enough. Using the normal approximation above we can see that the probability of A winning must go to one. $\square$

Note that with full information aggregation, as in case (i) in Theorem 2, the voting equilibrium is at least approximately efficient for large electorates, since the expected payoff of voters converges to its supremum (corresponding to a probability one of choosing the right alternative and a per capita expenditure of zero in information acquisition). By the same token, with partial information aggregation, as in case (ii), the voting equilibrium is inefficient even among symmetric strategy profiles, since the expected payoff remains bounded away from the supremum. That is, individual contributions to the provision of political information are too small, even if the total expected contribution, $(2n+1)(1-2\epsilon)x_n$, grows unboundedly.

5 The Winning Margin

Define the *winning margin* to be a random variable representing the difference between the number of votes for the winner and the number of votes for the loser, divided by $2n+1$. In this section we show that, if $C'(0) = 0$, the winning margin is likely to be close to zero for large electorates. In other words, elections with information acquisition tend to be very close. Intuitively, information acquisition requires that the probability of a voter being pivotal should not decline too fast. Otherwise, voters would lose the incentive to acquire costly information.

From the previous section, we know that, if $C'(0) = C''(0) = 0$, the probability that the right alternative wins the election goes to one as the size of the electorate increases. Theorem 2 and Theorem 3 below imply that, if $C'(0) = C''(0) = 0$, the percentage of votes for the right alternative will be very likely to be barely above 1/2. The reason is that the distribution of the percentage of votes for the right alternative concentrates very fast around its central terms, near $1/2 + (1-2\epsilon)x_n$, with x_n going to zero as n increases.

We have

Theorem 3 *If $C'(0) = 0$, then for any $\kappa > 0$ the probability that the winning margin is larger than κ converges to zero as the size of the electorate increases.*

PROOF. Suppose the state is z_A (similar calculations hold if the state is z_B). Using the notation of the proof of lemma 2, the number of votes for A is given by

$$V_1^n + \cdots + V_{2n+1}^n + (2n+1)(1/2 + (1-2\epsilon)x_n).$$

Then, the winning margin is

$$\left| \frac{2\left(\sum_{i=1}^{2n+1} V_i^n + (2n+1)(1/2 + (1-2\epsilon)x_n)\right) - (2n+1)}{2n+1} \right|$$

or equivalently,

$$2\left|\tfrac{1}{2n+1}\sum_{i=1}^{2n+1} V_i^n + (1-2\epsilon)x_n\right|.$$

Therefore, the probability that the winning margin is smaller or equal to κ is equal to $F_n(D_n) - F_n(I_n)$, where

$$D_n = \frac{(2n+1)(\kappa/2 - (1-2\epsilon)x_n)}{\sqrt{E((V_i^n)^2)(2n+1)}}$$

and

$$I_n = \frac{(2n+1)(-\kappa/2 - (1-2\epsilon)x_n)}{\sqrt{E((V_i^n)^2)(2n+1)}}.$$

Note that D_n goes to $+\infty$ and I_n goes to $-\infty$ with n. Following the last steps of the proof of lemma 2, we have that the probability that the winning margin is smaller or equal to κ must go to one. $\square$

Theorem 3 is reminiscent of a similar result by Feddersen and Pesendorfer (1997). In their work, the tightness of the electoral race is brought about by the fact that a vanishing fraction of voters takes into account their private information when casting a vote. That is, only a small fraction of voters takes informative actions in large elections. In our model, the fraction of swing voters is constant because of our assumption that moderate voters have common preferences. However, in large elections, the actions of individual voters carry very little information.

To illustrate Theorems 2 and 3, consider the following *cubic example*: $C(x) = x^3$, $r = 1$ and $\epsilon = 0$, corresponding to case (i) in Theorem 2. Figure 1 represents the probability distribution of the fraction of votes for the right alternative for $n = 50$ (101 voters) and $n = 500$ (1001 voters). (We use the same normal approximation for the distribution of $\sum V_i$ as in the proof of Lemma 2.) The mean of the distribution is equal to the probability of a moderate voter having the right opinion. For 101 voters, this probability is 0.582161; for 1001 voters, it is 0.532279. As the electorate increases, the mean converges to 0.5, but the variance shrinks fast enough so that the probability of choosing the wrong outcome (that is, the area below each curve to the right of 0.5) goes from 0.047016 for 101 voters to 0.020305 for 1001 voters to 0 in the limit. Figure 2 represents the *quadratic example*: $C(x) = x^2$, $r = 1$ and $\epsilon = 0$, corresponding to case (ii) in Theorem 2. In this case, the probability of choosing the wrong outcome goes from 0.25785 for 101 voters to 0.25849 for 1001 voters to 0.25877 in the limit.

In both examples, a single voter would acquire perfect information and the probability of choosing the right alternative decreases when going from one to three voters. In the cubic example the probability of choosing the right alternative increases with the number of voter afterwards, and the expected utility of a voter is always increasing in the number of voters, while in the quadratic example the probability of choosing the right alternative keeps declining and the expected utility is *decreasing* in the number of voters for small numbers.

Note that when ϵ is actually equal to zero, as in the examples above, there are other, trivial equilibria without information acquisition, which we are ignoring. In those equilibria, all voters vote for A, or all of them vote for B, so the probability of being pivotal is zero.

6 Extensions

In this section, we discuss several possible extensions of the model presented in Section 2 and their implications with regards to information aggregation.

General cost functions. Suppose that C is any nondecreasing real-valued function on $[0, 1/2]$ that is twice continuously differentiable in a neighborhood of zero. Then, if $C'(0) > 0$, or if $C'(0) = 0$ and the marginal cost is increasing near zero, the equilibrium described in Section 3 holds for n large enough. The reason is that the incentive to acquire information depends linearly on the probability of being pivotal. This probability, in turn, is strictly positive but converges to zero under any sequence of symmetric strategy profiles. If the marginal cost is equal to 0 in a neighborhood of zero, that is, if voters receive some free information, whether or not there is information acquisition becomes asymptotically irrelevant. Since the probability of a voter casting a vote for the right alternative is bounded below by a number larger than 1/2, the probability of choosing the right alternative must converge to one. Thus, the limit probabilities described in Theorem 2 hold under very general assumptions on the cost of acquiring information.

Asymmetries in the probabilities of extremist voters. The symmetric setup described in previous sections seems to favor information aggregation under majority rule. Conversely, if we deviate from the symmetry assumptions we could expect majority rule to do worse, especially if we keep the assumption that only symmetric Nash equilibria are played. This intuition turns out to be correct, though in a limited sense. To see this, suppose that the probability that each voter's type is t_A is ϵ and the probability that it is t_B is ϵ', for $0 < \epsilon' < \epsilon < 1/2$. That is, the expected fraction of extremists favoring A is larger than the expected fraction of extremists favoring B. Suppose also that $C'(0) = 0$.

It is easy to check that the pure strategy with $a_v(t_M, s_A) = A$ and $a_v(t_M, s_B) = B$ is strictly superior to the pure strategy with $a_v(t_M, s_A) = B$ and $a_v(t_M, s_B) = A$ for a given choice $x > 0$ of information quality for any given symmetric strategy profile α played by other voters. Moreover, among pure strategies with $a_v(t_M, s_A) = A$ and $a_v(t_M, s_B) = B$ for any $x \geq 0$, the strategy with $x = \hat{x}$ solving

$$\left[\frac{1}{2}P_\alpha(\text{piv}|z_A) + \frac{1}{2}P_\alpha(\text{piv}|z_B)\right] r = C'(\hat{x}) \tag{3}$$

is strictly superior to the others. (For n large enough, the previous equation has a unique, positive solution regardless of the strategy played by moderate voters, as the probability of being pivotal can be shown to be positive and to converge uniformly to zero.) Also, the pure strategy with $a_v(t_M, s_A) = A$, $a_v(t_M, s_B) = A$ and $x = 0$ is strictly superior to any other pure strategy where moderate voters only vote for A, and similarly for $a_v(t_M, s_A) = B$, $a_v(t_M, s_B) = B$ and $x = 0$. Finally, it is not possible that the pure strategies of playing for A no matter the signal, and playing for B no matter the signal, are both played with positive probability in equilibrium. If that were the case, we would have $P_\alpha(\text{piv}|z_A) = P_\alpha(\text{piv}|z_B)$. But then, the strategy of playing A

no matter the signal would have the same return than the strategy of playing $a_v(t_M, s_A) = A$, $a_v(t_M, s_B) = B$ and $x = 0$, which is strictly dominated by $a_v(t_M, s_A) = A$, $a_v(t_M, s_B) = B$ and $x = \hat{x}$. Similarly, neither the pure strategy of playing A no matter the signal nor the pure strategy of playing B no matter the signal can be played with probability one in equilibrium, because then we would have $P_\alpha(\text{piv}|z_A) = P_\alpha(\text{piv}|z_B)$.

Now consider any symmetric strategy such that moderate voters play the pure strategy $a_v(t_M, s_A) = A$, $a_v(t_M, s_B) = B$ and $x = \hat{x}$ with positive probability, the pure strategy $a_v(t_M, s_A) = A$, $a_v(t_M, s_B) = A$ and $x = 0$ with nonnegative probability, and every other pure strategy with probability zero. It is easy to check that then $P_\alpha(\text{piv}|z_A) < P_\alpha(\text{piv}|z_B)$, so playing A regardless of the signal would be strictly dominated by playing B regardless of the signal, and the proposed strategy cannot be an equilibrium.

From the previous two paragraphs, it follows that in equilibrium moderate voters play $a_v(t_M, s_A) = A$, $a_v(t_M, s_B) = B$ and $x = \hat{x}$ with probability $1-\pi$, and $a_v(t_M, s_A) = B$, $a_v(t_M, s_B) = B$ and $x = 0$ with probability π, for some $\pi \in (0, 1)$. The equilibrium conditions are given by equation (3) and:

$$(4) \quad \left(\frac{1}{2} + \hat{x}\right)\left[\frac{1}{2}P_\alpha(\text{piv}|z_A) + \frac{1}{2}P_\alpha(\text{piv}|z_B)\right] r - C(\hat{x}) = \frac{1}{2}P_\alpha(\text{piv}|z_B) r.$$

This last condition equalize the payoffs of both pure strategies played with positive probability. Note that now $P_\alpha(\text{piv}|z_A)$ is equal to

$$\binom{2n}{n}(\epsilon + (1-\epsilon-\epsilon')(1-\pi)(\tfrac{1}{2}+\hat{x}))^n(\epsilon' + (1-\epsilon-\epsilon')(\pi + (1-\pi)(\tfrac{1}{2}-\hat{x})))^n,$$

and $P_\alpha(\text{piv}|z_B)$ is equal to

$$\binom{2n}{n}(\epsilon + (1-\epsilon-\epsilon')(1-\pi)(\tfrac{1}{2}-\hat{x}))^n(\epsilon' + (1-\epsilon-\epsilon')(\pi + (1-\pi)(\tfrac{1}{2}+\hat{x})))^n.$$

Since both probabilities of being pivotal converge to zero as n goes to infinity, from equation (3), $\hat{x}$ converges to zero in any equilibrium sequence. But then, from equation (4), π converges to $(\epsilon-\epsilon')/(1-\epsilon-\epsilon')$. That is, the probability of a random voter taking informative action converges to $1-2\epsilon$.

It is simple to show that, for any $\kappa > 0$ and for n large enough, both probabilities of being pivotal are larger than

$$\binom{2n}{n}(\epsilon + (1-2\epsilon-\kappa)(\tfrac{1}{2}-\hat{x}))^n(\epsilon + (1-2\epsilon-\kappa)(\tfrac{1}{2}+\hat{x}))^n.$$

But then, using equation (3),

$$\frac{(2n)!}{n!n!}\left(\frac{1}{4}-(1-2\epsilon-\kappa)^2\hat{x}^2\right)^n < C'(\hat{x}).$$

Using the definition of x_M in Section 3, we can see that the information acquired by moderate voters in the asymmetric model is eventually larger than that acquired by voters in a symmetric situation with a fraction $1-2\epsilon-\kappa$ of moderate voters. Similarly, the information acquired by voters is smaller than that acquired by voters in a symmetric situation with a fraction $1-2\epsilon+\kappa$ of moderate voters. Since κ can be made arbitrarily small, the limit probabilities identified in Theorem 2 must hold.

Note, however, that in the model with asymmetric partisan voters, $1-2\epsilon$ is strictly less than the fraction of moderate voters, $1-\epsilon-\epsilon'$. So, even in the limit, information aggregation is less successful in the asymmetric model whenever $C'(0)=0$ and $C''(0)\in(0,\infty)$. As in the symmetric model, no information is "wasted" in the sense that informed moderate voters vote according to the signals received. But, unlike the symmetric model, a positive fraction of moderate voters do not acquire information.

Other extensions. In the symmetric situation described in the previous sections abstentions can be allowed without consequence as no voter would find it optimal to abstain. In asymmetric situations in which, for instance, the signal is biased and favors one of the two states, abstention may happen with positive probability in equilibrium. We do not explore formally this possibility, but note that the analysis of Feddersen and Pesendorfer (1996) suggests that abstention helps rather than hinders information aggregation in models of elections with private information.

7 Final Remarks

In a setting in which acquiring political information is costly, we have shown that the electorate as a whole may be much better informed than individual voters. In some circumstances, a result analogous to Condorcet's jury theorem is upheld: majorities are more likely to select the better of two alternatives when there is uncertainty about which of the alternatives is preferred.

(In fact, they get asymptotically closer to selecting the better alternative with probability one.) In some other circumstances, though, Condorcet's contention about the superiority of elections fails: majorities are *less* likely to select the better alternative than a single individual voter would be *if* the social decision were to rest on that voter. More generally, under the conditions described by case (i) in Theorem 2, majorities will beat any decision maker that would not acquire perfect information. Under those described by case (ii), majorities may be beaten by an imperfectly informed decision maker.

In the environment we study, a small deviation from rationality by voters – ignoring completely the effects of a single opinion – would have important negative effects on the responsiveness of collective decision making to the interests of the majority. Akerlof (1989) has approached the issue of rational ignorance from that perspective. However, deviations from strictly rational beliefs may be as likely to occur in the direction of overestimating the importance of a single opinion as in the direction of underestimating it. Voters may derive some satisfaction from the belief that their vote counts for more than it actually does, and overinvest in political information for that reason.

We have represented information acquisition by voters as a strictly individual endeavor. If voters can communicate their information to others before the election, if different voters have access to the same sources of information, or if sources of information compete for subscribers, strategic considerations will differ from those in the current framework in nontrivial ways. There is clearly a need for more formal research on the issue of endogenous production and aggregation of information in large elections, perhaps in connection with the recent interest in pre-election communication by privately informed voters, and our individualistic framework is meant as a first step.

References

[1] Akerlof, G. (1989) The Economics of Illusion, *Economics and Politics* 1: 1-15.

[2] Andreoni, J. (1988) Privately Provided Public Goods in a Large Economy: The Limits of Altruism, *Journal of Public Economics* 35: 57-73.

[3] Austen-Smith, D. and J. Banks (1996) Information Aggregation, Rationality, and the Condorcet Jury Theorem, *American Political Science Review* 90: 34-45.

[4] Becker, G. (1983) A Theory of Competition among Pressure Groups for Political Influence, *Quarterly Journal of Economics* 98: 371-400.

[5] Becker, G. (1985) Public Policies, Pressure Groups, and Dead Weight Costs, *Journal of Public Economics* 28: 329-347.

[6] Coate, S. and S. Morris (1995) On the Form of Transfers to Special Interests, *Journal of Political Economy* 103: 1210-1235.

[7] Downs, A. (1957) *An Economic Theory of Democracy*, HarperCollins Publishers, New York.

[8] Duggan, J. and C. Martinelli (2001) A Bayesian Model of Voting in Juries, *Games and Economic Behavior* 37: 259-294.

[9] Durrett, R. (1991) *Probability: Theory and Examples*, BrooksCole Publishing Company, Pacific Grove, California.

[10] Feddersen, T. and W. Pesendorfer (1996) The Swing Voter's Curse, *American Economic Review* 86: 408-424.

[11] Feddersen, T. and W. Pesendorfer (1997) Voting Behavior and Information Aggregation in Elections with Private Information, *Econometrica* 65: 1029-1058.

[12] Feller, W. (1971) *An Introduction to Probability Theory and Its Applications*, Vol. II, second edition, John Wiley&Sons, New York.

[13] McLennan, A. (1998) Consequences of the Condorcet Jury Theorem for Beneficial Information Aggregation by Rational Agents, *American Political Science Review* 92: 413-418.

[14] Miller, N. (1986) Information, Electorates, and Democracy: Some Extensions and Interpretations of the Condorcet Jury Theorem, in B. Grofman and G. Owen, editors, *Information Pooling and Group Decision Making*, JAI Press, Greenwich, Connecticut.

[15] Persico, N. (1999) Committee Design with Endogenous Information, *Review of Economic Studies*, forthcoming.

[16] Wittman, D. (1989) Why Democracies Produce Efficient Results, *Journal of Political Economy* 97: 1395-1424.

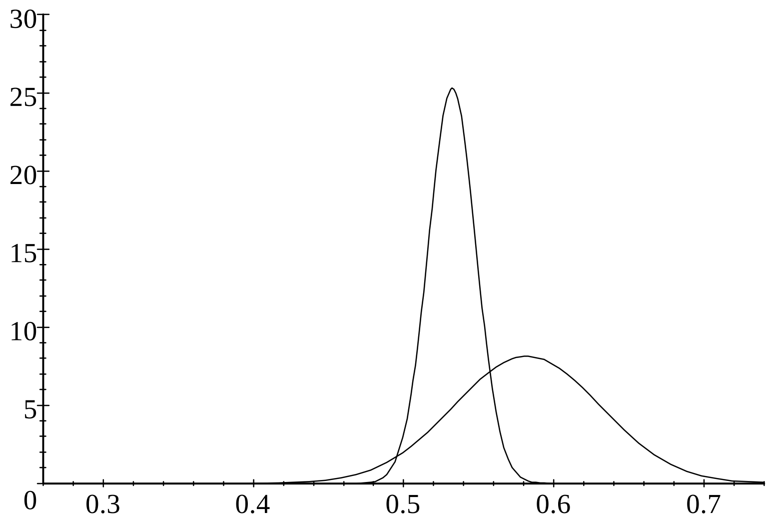

FIGURE 1: PDF OF % VOTE FOR RIGHT ALTERNATIVE
FOR 101 AND 1001 VOTERS (CUBIC COST FUNCTION)

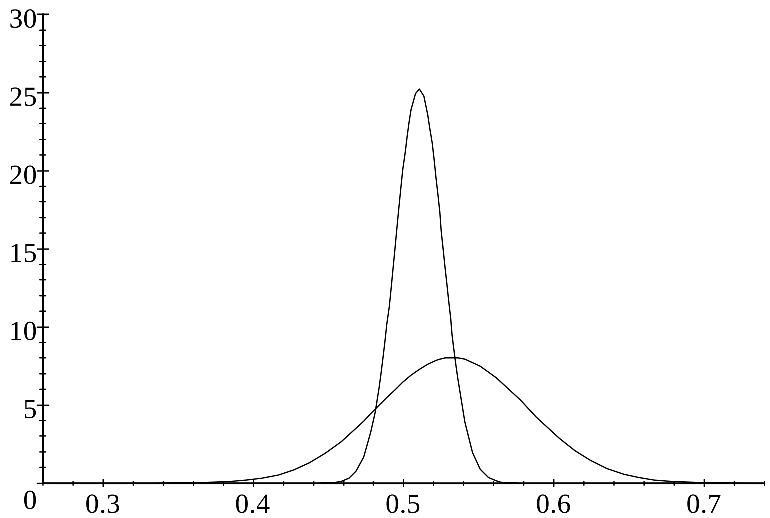

FIGURE 2: PDF OF % VOTE FOR RIGHT ALTERNATIVE FOR 101 AND 1001 VOTERS (QUADRATIC COST FUNCTION)