INFORMATION, ELECTORATES, AND DEMOCRACY:

SOME EXTENSIONS AND INTERPRETATIONS OF THE CONDORCET JURY THEOREM

Nicholas R. Miller

1. INTRODUCTION

Probably the most persistent descriptive theme in voting studies over the last three decades or so has been that individual voters are poorly informed and generally fail abysmally to meet, or even to approach, the requirements of citizenship postulated by the "classical doctrine" of democracy (in the sense of Schumpeter, 1942, pp. 250 ff.). But at the same time, sophisticated observers of politics might be hard put to name a single recent national election in which the vote division plausibly would have been substantially different, even if all voters had in fact made more complete use of the information potentially available to them. To quote V. O. Key, Jr. (1966, p. 7):

> To be sure, many individual voters act in odd ways indeed; yet in the large the electorate behaves about as rationally and responsibly as we should expect, given the clarity of the alternatives presented to it and the character of the information available to it.

This discrepancy between inferior "micro-level" performance and apparently superior "macro-level" performance has been asserted by others as well. Thus, in a noted quotation, Berelson et al. (1954, p. 312) say:

> *Individual voters* today seem to be unable to satisfy the requirements for a democratic system of government outlined by political theorists. But the *system of democracy* does meet certain requirements for a going political organization. The individual members may not meet all the standards, but the whole nevertheless survives and grows. This suggests that where the classic theory is defective is in its concentration on the *individual citizen.* What are undervalued are certain collective properties that reside in the electorate as a whole and in the political and social system in which it functions.

However, if the contrast between micro- and macro-level performance has fairly frequently been noted, the precise mechanisms that produce it seem to be less well understood. Certainly references to "certain collective properties" don't greatly advance understanding.

I argue here that there is nothing mystical or even surprising about the relative competence of the electorate as a whole, even though it may be composed largely of relatively incompetent individual voters. I argue further that the fundamental mechanism at work is of a "statistical" nature and is a consequence of a generalization of the "Jury Theorem" due to Condorcet (1785). This theorem and its relevance for a variety of problems in political theory and analysis have recently been rediscovered by a number of political scientists and public choice theorists—in some cases knowingly (see especially Grofman, 1975b, 1978; also Black, 1958, pp. 159 ff; Barry, 1964, pp. 9–14, 1965, pp. 292–293; Miller, 1977a; and Urken, 1980), in other cases apparently unknowingly (e.g., Swaby, 1939, especially Chap. 1; Kazmann, 1973; Allen, 1974; and Weissberg, 1978; also see Penrose, 1946, and Niemi and Weisberg, 1972, for related types of arguments).

2. THE CONDORCET JURY THEOREM

In its simplest form, the Jury Theorem says this: suppose that every person in a group has a given level of "competence," that is, a certain probability p of making a "correct" decision in a binary choice situation; then, assuming only that these individuals are at least minimally competent (i.e., that p is greater than .5—the "competence" of a flipped coin) and they choose independently of each other, the probability that the group, deciding on the basis of majority rule, makes the "correct" decision is greater than p, the level of individual competence, and furthermore "collective competence" increases as the size of the group increases and quite rapidly approaches perfection.

Formally, the matter may be stated in this way. There are n voters, each with a probability p of voting correctly on a given measure. Of course, $0 \le p \le 1$, and let $q = 1 - p$. Now let x be the number of individuals who vote correctly; then x is a binomially distributed random variable:

$$f(x) = \binom{n}{x} p^x q^{n-x}. \tag{1}$$

Let P_n be the probability that a group of size n, deciding on the basis of majority rule, decides correctly, that is, the probability that a majority of individuals vote correctly. If n is even and a tie vote occurs, we may suppose that the tie is broken by an even-chance lottery. Thus for an odd number n of voters, we have

$$P_n = \text{prob}\left(x \ge \frac{n+1}{2}\right) = \sum_{x=(n+1)/2}^{n} \binom{n}{x} p^x q^{n-x}, \tag{2a}$$

while for an even number n of voters we have

$$\begin{aligned} P_n &= \text{prob}\left(x \ge \frac{n}{2} + 1\right) + \frac{1}{2}\text{prob}\left(x = \frac{n}{2}\right) \\ &= \sum_{x=(n/2)+1}^{n} \binom{n}{x} p^x q^{n-x} + \frac{1}{2}\binom{n}{n/2} p^{n/2} q^{n/2}. \end{aligned} \tag{2b}$$

In fact, however, for any even value of n, the value of P_n is the same as the value of P_n for $n - 1$ (odd), so we can use (2a) in all cases.

For all except very small values of n (say $n < 15$) or extreme values of p in conjunction with smaller values of n, f(x) can be well approximated by a normal distribution with a mean of np and a variance of npq. Thus:

$$P_n \simeq \text{prob}\left(x > \frac{n}{2}\right) = 1 - \Phi\left(\frac{n/2 - np}{\sqrt{npq}}\right) = \Phi\left(\frac{p - .5}{\sqrt{pq/n}}\right), \tag{3}$$

where $\Phi(z)$ is the area under the normal curve from $-\infty$ to z standard deviation units.

Now we have the following:

CONDORCET JURY THEOREM.

1. If $.5 < p < 1$ and $n \ge 3$, then (i) $P_n > p$, (ii) P_n increases as n increases; and (iii) $P_n \to 1$ as $n \to \infty$.
2. If $0 < p < .5$ and $n \ge 3$, then (i) $P_n < p$; (ii) P_n decreases as n increases; and (iii) $P_n \to 0$ as $n \to \infty$.
3. If $p = 0$, $p = .5$, or $p = 1$, then $P_n = p$ for all n.

Table 1. (a) Values of P_n for Selected Values of n and p; (b) Values of p_i for Selected Values of k_i and Δ_i

	p (Δ_i)												
n (k_i)	*.5050*	*.5100*	*.5250*	*.5500*	*.6000*	*.6250*	*.7500*	*.8750*	*.9000*	*.9500*	*.9750*	*.9950*	*1.0*
0	.5000	.5000	.5000	.5000	.5000	.5000	.5000	.5000	.5000	.5000	.5000	.5000	.5000
1	.5050	.5100	.5250	.5500	.6000	.6250	.7500	.8750	.9000	.9500	.9750	.9950	1.0
3	.5075	.5150	.5375	.5748	.6480	.6836	.8438	.9570	.9720	.9928	.9982	.9999	1.0
5	.5094	.5187	.5468	.5931	.6826	.7248	.8965	.9839	.9914	.9988	.9998	.9999	1.0
7	.5109	.5219	.5546	.6083	.7102	.7570	.9294	.9938	.9973	.9998	.9999	.9999	1.0
9	.5123	.5246	.5613	.6214	.7334	.7834	.9510	.9975	.9991	.9999	.9999	.9999	1.0
15	.5154	.5309	.5769	.6514	.7854	.8419	.9873	.9999	.9999	.9999	.9999	.9999	1.0
25	.5199	.5398	.5988	.6924	.8462	.9017	.9981	.9999	.9999	.9999	.9999	.9999	1.0
75	.5345	.5688	.6677	.8079	.9614	.9873	.9999	.9999	.9999	.9999	.9999	.9999	1.0
250	.5628	.6241	.7857	.9440	.9994	.9999	.9999	.9999	.9999	.9999	.9999	.9999	1.0
1,000	.6241	.7365	.9433	.9993	.9999	.9999	.9999	.9999	.9999	.9999	.9999	.9999	1.0
10,000	.8413	.9772	.9999	.9999	.9999	.9999	.9999	.9999	.9999	.9999	.9999	.9999	1.0
1,000,000	.9999	.9999	.9999	.9999	.9999	.9999	.9999	.9999	.9999	.9999	.9999	.9999	1.0

Note: Here and elsewhere, most ".9999" entries actually round off to "1.0000" but this is not done to indicate residual uncertainty.

Given at least minimal competence, that is, $p > .5$, our interest is, of course, in statement (1) of the theorem. (For proof, see Grofman, 1975b and 1978, and references therein.)

Table 1 displays illustrative values of P_n—which we may characterize as "collective competence"—for various combinations of individual competence p and group size n. (The alternative symbols p_i, Δ_i, and k_i will be explained later.)

In a general and more interesting version of the theorem, different individuals may have different levels of competence, p_i, each greater than .5 or, in any case, with an average greater than .5. The generalized Jury Theorem then states that the group, deciding on the basis of majority rule, is more competent than the average individual and, quite possibly, more competent than the "best" (most competent) individual. (For proof see Owen, Grofman, and Feld, 1983.) Again collective competence increases with the size of the group (so that adding members may increase collective competence even if it reduces average individual competence) and quite rapidly approaches perfection. Thus—and hence the "Jury Theorem" designation—it may be entirely reasonable to entrust an important binary decision for which there is in principle a "correct" decision (e.g., convicting or acquitting a criminal defendant or finding for a plaintiff or defendant) to a group of individuals of lesser competence (e.g., a jury) rather than to a single individual of greater competence (e.g., a judge).

3. THE CONDORCET JURY THEOREM AND THE ELECTORAL PROCESS

The applicability of the Jury Theorem to political decision making in general appears to be highly limited, for it appears to demand acceptance of the "idealist" assumption that it is meaningful to speak of a "correct" political decision, for example, that there is an "objective" public interest independent of individual interests and/or that all individuals share the same "true" individual interests. Though "interesting," the Jury Theorem appears to be inapplicable to the analysis of electoral (or other political) decision making under the more "pluralistic" assumptions entailing conflicting individual interests that characterize American political science and most modern political theory. (On all this, see Held, 1970.) As Black (1958; p. 163) says of the Condorcet Jury Theorem:

> Now whether there is much or little to be said in favor of a theory of juries arrived at in this way, there seems to be nothing in favor of a theory of elections that adopts this approach. When a judge, say, declares an accused person to be either guilty or innocent, it would be possible to conceive of a test which, in principle at least, would be capable of telling us whether his judgment had been right or wrong. But in the case

of elections, no such test is conceivable; and the phrase "the probability of the correctness of a voter's opinion" seems to be without definite meaning.

But in fact the Jury Theorem, or a straightforward extension of it, can be applied to the "pluralistic" case in which we allow individual interests to conflict. In essence, we need only to admit that a decision that is "correct" for one individual may be "incorrect" for another (with conflicting interests). "The probability of the correctness of a voter's opinion" now refers to the probability that he has accurately perceived his own individual interest—not the public interest or "true" interests shared by all individuals. With individual probabilities so interpreted, we can use an extension of the Jury Theorem to explain, in a persuasive, nonmystical, and unsentimental way, the relatively superior performance of an electorate composed of relatively inferior individual voters.

The Jury Theorem can be extended from "juries" (in which the same decision is correct for all individuals or, equivalently, all individuals have the same interests) to "electorates" (in which the same decision may not be "correct" for all individuals or, equivalently, individuals may have conflicting interests) in the following fashion. In any binary political choice situation, such as a referendum or a two-party election, voters can be divided into two groups—those whose "true" interests lie in one direction and those whose "true" interests lie in the other direction. In this context, a *voter's "true" interest is to be thought of as the "subjective" preference he would have in the event that he were completely informed.* And the "competence" of an individual voter is now the probability that he correctly votes for the position or party that would best serve his "true" interests; for all the reasons identified in the empirical literature on voting behavior, this probability likely falls far below 1.

For the moment, let us say that *the electoral process "succeeds" when the interests of the majority prevail*—put otherwise, *when the victorious position or party is the one that would win in the event that all voters were completely informed.*

A straightforward extension of the Condorcet Jury Theorem then states that if all voters are equally competent, whatever that level of competence (greater than .5), or more generally if the two groups of voters have the same average competence, majority interests will probably prevail and—once the electorate achieves some minimal size—this probability is greater than the average competence of all voters, increases further as the size of the electorate further increases, and in due course (though not as rapidly as in the case of the original Jury Theorem) approaches perfection. Moreover, the same conclusions may be reached if the two groups are of unequal average competence, provided only that the size of the majority group exceeds the size of the minority group by a ratio greater than the ratio of average minority competence minus .5 to average majority competence minus .5. (If the second ratio exceeds the first, the electoral process will probably "fail," by the criterion of realizing majority interests, and is more certain to fail as the size of the electorate increases.)

Formally, we may develop the argument as follows. For the moment, let us suppose that all voters have the same competence p, as in the basic Jury Theorem, but voting "correctly" now means in light of the voter's own interests. Again let $q = 1 - p$. We suppose that, in any given case, all voters are divided into two blocs—those whose true interests are served by passage of the measure or victory by party A and those whose true interests are served by defeat of the measure or victory by party B. Let n_A and n_B be the numbers of voters in these two blocs; thus $n_A + n_B = n$. By convention, but without loss of generality, we assume $n_A > n_B$; that is, majority interests are served by passage of the measure or victory by party A.

Let x be the number of votes for the measure or for party A. The expected vote E(x) in favor of the majority position is

$$E(x) = n_A p + n_B q. \tag{4}$$

Let $n_A = n/2 + e_1$ and let $p = .5 + e_2$, where of course $e_1, e_2 > 0$. Then

$$E(x) = \left(\frac{n}{2} + e_1\right)(.5 + e_2) + \left(\frac{n}{2} - e_1\right)(.5 - e_2) = \frac{n}{2} + e_1 e_2 > \frac{n}{2}.$$

That is, it is expected that majority interests will prevail, whatever the value of p (>.5).

Majority interests are expected to prevail, but with what actual probability? What is the probability that the electorate will make a correct decision, in the sense that majority interests will prevail?

Let $p^* = E(x)/n$, that is, the expected proportion of the vote in favor of the majority position. Note that the value of p^* depends only on p and the ratio n_A/n, not on the absolute size of the electorate n.

Let x_A be the number of voters in the majority bloc who vote ("correctly") for the measure or for party A, and let x_B be the number of voters in the minority bloc who vote ("incorrectly") for the measure or for party A. Of course, $x_A + x_B = x$, $x_A \leq n_A$, and $x_B \leq n_B$.

Each of x_A and x_B is distributed binomially in the manner of x in the original Jury Theorem, except that p and q are reversed in the case of x_B.

$$f(x_A) = \binom{n_A}{x_A} p^{x_A} q^{n_A - x_A}; \tag{5}$$

$$f(x_B) = \binom{n_B}{x_B} q^{x_B} p^{n_B - x_B}. \tag{6}$$

Thus,

$$f(x = x_A + x_B) = \sum_{x_B=k}^{n_B} \binom{n_A}{x_A}\binom{n_B}{x_B} p^{n_B+x_A-x_B} q^{n_A-x_A+x_B}$$

where $k = x - n_A$ if $x - n_A > 0$ and $k = 0$ otherwise.

Let P'_n be the probability that the measure passes or that party A wins in an electorate of size n, that is, the probability that majority interests prevail, which for the moment we count as "success" of the electoral process. Then

$$P'_n = \text{prob}\left(x \geq \frac{n+1}{2}\right)$$
$$= \sum_{x=(n+1)/2}^{n} \sum_{x_B=k}^{n_B} \binom{n_A}{x_A}\binom{n_B}{x_B} p^{n_B+x_A-x_B} q^{n_A-x_A+x_B}. \tag{7}$$

As before, for all but small n, $f(x_A)$ and $f(x_B)$ can be approximated by normal distributions with means of $n_A p$ and $n_B q$, respectively, and variances of $n_A pq$ and $n_B pq$, respectively.

Now, given two independent normally distributed random variables with means of m_1 and m_2 and variances of s_1^2 and s_2^2, then their sum is a normally distributed random variable with a mean of $m_1 + m_2$ and a variance of $s_1^2 + s_2^2$. Therefore, $f(x = x_A + x_B)$ can be approximated by a normal distribution with a mean of $n_A p + n_B q = E(x) = np^*$ and a variance of $n_A pq + n_B pq = npq$. Thus

$$P'_n \simeq \text{prob}\left(x > \frac{n}{2}\right) = 1 - \Phi\left(\frac{n/2 - np^*}{\sqrt{npq}}\right) = \Phi\left(\frac{p^* - .5}{\sqrt{pq/n}}\right). \tag{8}$$

Now we have the following:

CONDORCET JURY THEOREM (EXTENDED). If $.5 < p < 1$ and $n \geq 3$, then for any given ratio n_A/n, (i) $P'_n > p^*$; (ii) P'_n increases as n increases; and (iii) $P_n \to 1$ as $n \to \infty$.

Essentially the same considerations apply here as in the original theorem. For given values of n_A and n, P'_n increases as p increases; and for given values of p and n, P'_n increases as n_A increases. Slightly less obviously, for given values of p and a given ratio n_A/n, P'_n increases as n increases, in the manner of the original Jury Theorem (where, in effect, $n_A/n = 1$).

P'_n (the probability that majority interests will prevail) always exceeds p^* (the expected proportion of voters voting for the majority position), but P'_n does not always exceed p (the competence of the individual voter or, equivalently, the expected proportion of voters voting "correctly" in support of their interests) for smaller values of n. That is, in the case of an electorate with conflicting interests, a small electorate is less "competent" (in the sense of realizing majority interests) than individual voters are "competent" (in the sense of voting to support their true interests). This, of course, stands in contrast to the case of a jury without conflicting interests, for which $P_n > p$ even for small values of n. It is also clear that, for any given ratio n_A/n, P'_n approaches 1 rather slowly compared with a jury with the same p—and this is especially true as n_A/n approaches .5. Roughly, we may look at these matters in this way: in the generalized theorem applied to electorates, p^* plays the mathematical role that p does in the original theorem applied to juries [compare (3) and (8)]. But in the electoral case, if $p < 1$ always, $p^* < p$. Thus while we do always have $P'_n > p^*$, we may not have $P'_n > p$; that is, we may have $p^* < P'_n < p$. And, if we compare an electorate and a jury with the same p, collective competence P'_n is less and approaches 1 less rapidly as group size increases in the former case than P_n in the latter case.

As with the original theorem, if different voters have different levels of competence, that is, different probabilities of voting correctly in support of their interests, nothing changes substantially, *provided that voters in the two blocs have the same average competence $\bar{p}$*, in which case we can simply substitute $\bar{p}$ for p in the previous formulas.

If, on the other hand, the two blocs differ in average competence, voters in the first bloc voting correctly with an average probability of $\bar{p}_A$ and those in the second with an average probability $\bar{p}_B$, majority interests can be expected to prevail if and only if

$$E(x) = n_A\bar{p}_A + n_B\bar{q}_B > \frac{n}{2} > n_B\bar{p}_B + n_A\bar{q}_A$$

or, by simple algebraic manipulation, if and only if

$$\frac{n_A}{n_B} > \frac{\bar{p}_B - .5}{\bar{p}_A - .5}. \tag{9}$$

In any case, the probability that majority interests prevail is

$$P'_n = \sum_{x=(n+1)/2}^{n} \sum_{x_B=k}^{n_B} \binom{n_A}{x_A}\binom{n_B}{x_B} \bar{p}_A^{x_A}\bar{q}_B^{n_A-x_A}\bar{p}_A^{x_b}\bar{p}_B^{n_B-x_B}, \tag{10}$$

where $k = x - n_A$ if $x - n_A > 0$ and $k = 0$ otherwise.

As before, P'_n is well approximated, for larger n, by

$$P'_n \simeq \Phi\left(\frac{n_A\bar{p}_A + n_B\bar{q}_B - n/2}{\sqrt{n_A\bar{p}_A\bar{q}_A + n_B\bar{p}_B\bar{q}_B}}\right). \tag{11}$$

If condition (9) is met, then P'_n is always greater than $p^* = E(x)/n$ and increases as n increases (as before); but if the inequality in (9) is reversed, then P'_n is always less than p^* and decreases as n increases.

To this point, we have supposed that the outcome in which majority interests prevail should always be counted as electoral "success" and conversely that the outcome in which minority interests prevail should always be counted as electoral "failure." But it is not at all clear that this supposition is always appropriate. This is so because the notion of individual "competence" becomes more complicated when we move from the case of a jury to the "pluralistic" electoral case.

Let us first return to the original jury example. The probability that an individual juror makes a correct decision in a given case, what we have called his "competence," depends of course on certain factors particular to that individual—his state of information in particular, as well as such personal qualities as cognitive capacity, judgment, and the like. But the magnitude of this probability that we have dubbed "competence" also depends on something quite apart from the individual, namely the nature of the case to be decided. In an open-and-shut case, all jurors may be highly "competent" (say with probabilities of deciding correctly of .9 or better), whatever their individual circumstances or qualities; whereas in a very close and difficult case, the same jurors with the same amount of information and the same qualities of judgment may be only marginally competent (say with probabilities of deciding correctly of close to .5). Accordingly, it is not really appropriate to refer to these probabilities as "competences," since their magnitudes depend on the difficulty of the task as well as on the circumstances and qualities of the individual. However, in the case of jury decision making (or electoral decision making under "idealist" assumptions), this inappropriate language is not especially misleading, because all individuals are faced with the *same* (easy or difficult) task, and individual *differences* in probabilities must be attributed to *individual differences in circumstances and qualities* and may thus be taken as an indication of differences in "competence."

But in the electoral case under "pluralistic" assumptions, things are not so simple. Different individuals are faced with different choice making tasks, as their differing interests are differentially at stake in the electoral contest. In the sense of Downs (1957, pp. 39 ff.), different voters have different *party* (or generally *interest*) *differentials*, and plausibly these differ not only in direction but also in magnitude. Thus even if all voters are equally well (or badly) informed, have equal cognitive capacities, and so forth, their probabilities of voting correctly in support of their interests will plausibly vary with the magnitude of their party differentials—those with large differentials have an easy task and are more likely to vote correctly, and those with small differentials have a more difficult task and are less likely to vote correctly.

This means that if voters in the two (majority and minority) blocs differ, on the average, in their probabilities of voting correctly, this may reflect only (or primarily or in part) a difference in their average party interest differentials. And in that case, what we have previously counted as "failure" of the electoral process might, in some cases, be otherwise evaluated, namely as a successful resolution of the "intensity problem" in democratic theory (see Dahl, 1956, pp. 48 ff.). Of course, this need not be the case; minority interests may prevail not because the minority voters have more at stake but only because they are better informed, or whatever, in which case it is no doubt appropriate to continue to speak of electoral "failure."

This complexity in interpreting "competence" in the electoral case can be clarified if we move from the postulated individual probabilities on which the several versions of the Jury Theorem are based to an information sampling model that produces the probabilities as a consequence of varying party-interest differentials and varying levels of information. The development of such a model, and a preliminary investigation of its properties, constitute the focus of the remaining section of this paper. This model also allows us to examine, from the point of view of the collective competence of the electorate as a whole, the consequences of various patterns of distribution of information among voters.

4. PARTY DIFFERENTIALS AND INFORMATION SAMPLING

The model we work with has the following three basic elements.

1. There are two *alternatives*, A and B, in some dichotomous social choice situation. These might be alternative verdicts in a jury trial, the passage and defeat of a measure in a referendum, two parties or candidates in an electoral contest, and so on. The nature of these alternatives is fixed (e.g., if parties/candidates, they cannot adjust their platforms, etc.).
2. There is a set N of n *voters*, who are partitioned (on basis described below) into two subsets: N_A, those n_A voters whose "true" interests would be better served by selection of alternative A; and N_B, those n_B voters whose "true" interests would be better served by selection of alternative B.[1] (One of these two subsets may be empty, giving us the circumstances of the original jury theorem.)
3. There is a universe X of "bits" of political *information* bearing on the contest between A and B. We suppose for simplicity that these bits are equally weighted. For each voter $i \in N$, X is partitioned into two subsets: X_i^A, those bits that are "reasons" for voter i to favor alternative A; and X_i^B, those bits that are "reasons" for voter i to favor alternative B.

Let $x = |X|$, $x_i^A = |X_i^A|$, and so on. We now define the *party*, or more generally *interest*, *differential* Δ_i of voter i as the proportion of bits in the universe of political information that are "reasons" for i to favor A; that is, $\Delta_i = x_i^A/x$. Thus $\Delta_i = 1$ if every bit of potential information is a reason for i to favor A; $\Delta_i = 0$ if every bit is a reason to favor B; and $\Delta_i = .5$ if, for every reason to favor A, there is a countervailing reason to favor B.

We now define a voter's "true" interests in terms of this differential; that is, voter i belongs to N_A if and only if $\Delta_i > .5$ and to N_B if and only if $\Delta_i < .5$.[2]

For each voter $i \in N$, we define $K_i \subseteq X$ as the sample (subset of the universe) of bits of information that voter i actually has at the time he votes. We suppose that all voters take independent random samples out of the universe of political information. Let $k_i = |K_i|$, which then indicates how well informed voter i is—that is, $0 \leq k_i \leq x$, ranging from total ignorance to complete information.

For each voter $i \in N$, we define $K_i^A = K_i \cap X_i^A$ and $K_i^B = K_i \cap X_i^B$—in words, the bits of information that voter i actually has that are reasons to favor A and B, respectively. Perhaps a bit misleadingly, but for ease of exposition, we refer to a bit of information as *correct for i* if it belongs to X_i^A and $i \in N_A$ or if it belongs to X_i^B and $i \in N_B$.

Finally we suppose that each voter $i \in N$ uses the following *voter decision rule* or *VDR*[3]: (1) voter i votes for A if $k_i^A > k_i^B$; (2) voter i votes for B if $k_i^B > k_i^A$; and (3) voter i votes for each alternative with a probability of one half if $k_i^A = k_i^B$.[4]

We now consider the conjunction of various circumstances pertaining to interest differentials and the distribution of information. With respect to the former, we distinguish among these circumstances: (A) differentials uniform in both *direction*, for example, $\Delta_i > .5$ for all $i \in N$, and *magnitude*, that is, $\Delta_i = \Delta_j$ for all $i, j \in N$; (B) differentials uniform in direction but diverse in magnitude, that is, in general $\Delta_i \neq \Delta_j$; and (C) differentials diverse in both direction, for example, $\Delta_i > .5$ and $\Delta_j < .5$, and magnitude. Circumstance (A), which implies either $N_A = \emptyset$ or $N_B = \emptyset$, clearly puts us in the context of jury decision making, and circumstance (C) puts us in the context of electoral decision making. Circumstance (B), which also implies either $N_A = \emptyset$ or $N_B = \emptyset$ and in that respect is consistent with the jury case, may never arise in practice but is included for theoretical completeness.

With respect to the distribution of information, we distinguish among three circumstances: (I) *equal* and *minimal* information, that is, $k_i = k_j = 1$ for all $i, j \in N$ (as is noted below, $k_i = 2$ also is effectively minimal); (II) equal and more than minimal information, that is, $k_i = k_j \geq 3$ for all $i, j \in N$; and (III) *unequal* information, that is, in general $k_i \neq k_j$.

Let us now look at the simplest possible case of a single voter i who is minimally informed; that is, $k_i = 1$. What is the probability p_i that he will vote correctly? Given that i follows the VDR, this is the same as the probability that a random bit of information drawn from X is correct for i, which in turn is simply the proportion of all bits that are correct for i, which is equivalent to his differential Δ_i. Thus, $p_i = \Delta_i$ if $\Delta_i > .5$ and $p_i = 1 - \Delta_i$ if $\Delta_i < .5$ (see row I of Table 2, where, however, the possibility that $p_i = 1 - \Delta_i$ is not noted to save space).

Now suppose that the completeness of voter i's information increases. Clearly his probability p_i of voting correctly also increases. What exactly is the nature of this relationship?

Suppose $k_i = 2$. Then, according to the VDR, voter i will vote correctly in the event that both bits are correct, and he also has a probability of $\frac{1}{2}$ of voting correctly in the event that exactly one bit is correct. Thus

$$\begin{aligned} p_i &= (\Delta_i)^2 + \tfrac{1}{2}(\Delta_i)(1 - \Delta_i) + \tfrac{1}{2}(1 - \Delta_i)(\Delta_i) \\ &= (\Delta_i)^2 + (\Delta_i)(1 - \Delta_i) = (\Delta_i)^2 + \Delta_i - (\Delta_i)^2 = \Delta_i. \end{aligned}$$

That is, p_i for $k_i = 2$ is the same as p_i for $k_i = 1$.

For $k_i = 3$:

$$p_i = (\Delta_i)^3 + 3(\Delta_i)^2(1 - \Delta_i),$$

and so forth. By inspection then, it turns out that what we have in the case of a single voter is the original Jury Theorem "writ small," where Δ_i replaces p, k_i replaces n, and p_i replaces P_n (see Grofman and Mackelprang, 1974, pp. 4 ff., who use an identical model of individual choice). Put otherwise, there is a precise correspondence between (i) a single individual drawing n bits of information out of a universe in which p is the proportion of bits that are correct and making a decision based on the VDR (which prescribes "majority rule" on information bits) and (ii) a group of n individuals, each drawing a single bit of information out of the same universe, voting

Table 2. Information Distribution and Interest Differentials

	Interest Differentials		
Information Distribution	*A. Uniform in Direction and Magnitude*	*B. Uniform in Direction; Diverse in Magnitude*	*C. Diverse in Both Direction and Magnitude*
I. Equal and minimal	IA. Basic Jury Theorem ($p = \Delta$)	IB. Generalized Jury Theorem ($p_i = \Delta_i$)	IC. Extended Jury Theorem ($p_i = \Delta_i$)
II. Equal and more than minimal	IIA. Basic Jury Theorem ($p > \Delta$)	IIB. Generalized Jury Theorem ($p_i > \Delta_i$)	IIC. Extended Jury Theorem ($p_i > \Delta_i$)
III. Unequal	IIIA. Generalized Jury Theorem ($p_i \neq p_j$)	IIIB. Generalized Jury Theorem	IIIC. Extended Jury Theorem

accordingly, with a decision taken by majority rule. Thus Table 1, showing illustrative values of P_n (the probability that a jury makes a correct decision) for selected values of n (jury size) and p (individual competence), also shows illustrative values of p_i (the probability that a voter, using the VDR, votes correctly) for selected values of k_i (the number of bits of information he has) and Δ_i (his interest differential).

Referring to Table 2 then, box IA gives us the circumstances of the basic Jury Theorem, when all voters are minimally informed. By increasing the information of all voters equally, we move down to box IIA; we are still in the circumstances of the basic Jury Theorem, but now $p > \Delta$.

We reach the circumstances of the generalized Jury Theorem (allowing differing levels of individual competence p_i) by moving in either (or both) of two directions—down to row III, while remaining in column A (that is, by allowing unequal information while keeping differentials uniform), or over to column B, while remaining in rows I or II (that is, by allowing varying magnitudes of individual differentials, while keeping information levels equal). (Or we can move in both directions, to box IIIB.) Presumably, however, box IIIA most plausibly interprets the generalized Jury Theorem (cf. the discussion toward the end of Section 3).

We have seen that—by virtue of the two different interpretations of Table 1 noted above—a group of n voters, each with the same differential, each with a single bit of information, and each choosing on the basis of the VDR and collectively deciding on the basis of majority rule, does as well as a single well-informed voter (with the same differential) with n bits of information making an individual decision on the basis of the VDR. Insofar as information leads to correct decisions, the two situations are theoretically equivalent.[5]

But it is important to note that the second case is one of strictly individual decision making. If we keep the total amount of information constant but add in an additional $n - 1$ totally uninformed individuals who nevertheless vote (randomly according to the VDR), collective competence is of course considerably undermined. What this illustrates is that, *for a voting body of given size n and with a uniform differential in the manner of a jury and with a fixed total amount of information, collective competence P increases the more equally that information is distributed among the voters.* Table 3 is a convenient small-scale illustration (where $\Delta_1 = \Delta_2 = \Delta_3 = .75$).

The four rows (1)-(4) are distinguished in terms of four distinct patterns of distribution of 9 bits of information among 3 voters, as shown in the column labeled k_1, k_2, and k_3. The mean amount of information per voter is, of course, constant, that is, $\bar{k} = 3$; but the distributions may be ranked from least to most equal, say in terms of their standard deviations SD_k. The next three columns show the individual competences of the voters; these entries come from Table 1. In the remaining columns, we see that, as

Table 3. Illustration of Effect of Information Distribution on Collective Competence

	k_1	k_2	k_3	$\bar{k}$	SD_k	p_1	p_2	p_3	$\bar{p}$	SD_p	P_n
(1)	9	0	0	3	4.24	.9510	.5000	.5000	.6503	.2126	.7255
(2)	7	1	1	3	2.83	.9294	.7500	.7500	.8098	.0846	.9111
(3)	5	3	1	3	1.63	.8965	.8438	.7500	.8301	.0606	.9269
(4)	3	3	3	3	0.00	.8438	.8438	.8438	.8438	.0000	.9344

information is distributed more equally, mean individual competence ($\bar{p}$) and collective competence (P_n) both increase (and, of course, variability in individual competence, SD_p, decreases).

The explanation for this is, fundamentally, that information is subject to diminishing marginal returns, in terms of its effect on individual competence. (This pattern of diminishing returns is a familiar result in sampling theory.) Thus redistributing information away from the information-rich to the information-poor reduces the competence of the former less than it increases the competence of the latter and increases mean competence overall and accordingly collective competence as well. (The argument is reminiscent of—indeed, formally identical to—the conventional utilitarian argument, based on postulated diminishing marginal utility of income, according to which an equal distribution of income maximizes total social utility.) In this sense, then, effective political equality fosters collective competence.

Thus far we have assumed that the interest differential Δ is uniform across voters; that is, we have considered only cases in column I of Table 2. Now we allow the differential to vary from voter to voter. If the differential varies in magnitude but not in direction, that is, individual interests vary in intensity but not in direction, we remain within the circumstances of the Jury Theorem (in column II of the diagram). But, on the whole, if differentials vary at all, it seems likely that they will vary in direction as well as magnitude; that is, if interests vary in intensity they may also conflict. Thus it is reasonable to shift our attention directly to column III in the diagram—that is, to the circumstances of the extended Jury Theorem and to electoral and other political decision making. For terminological convenience, we will conduct the remaining discussion in the context of a two-party electoral contest.

As before, let P'_n be the probability that the "right" party wins the election, that is, that the electoral process "succeeds," *if we accept the criterion that majority interests ought always to prevail*; that is, A should win if $n_A > n_B$ and B should win if $n_B > n_A$. Equivalently, P'_n is the probability that the party that would win the election in the event that all voters had complete information will in fact win the election.

Also let P''_n be the probability that the "right" party wins the election, that is, that the electoral process "succeeds," *if we reject the majority interest criterion and instead weigh interests according to their "intensity"* as indicated by the absolute magnitude of party differentials. By the intensity criterion, A should win if $\sum_{i \in N} x_i^A > \sum_{i \in N} x_i^B$ and conversely for B.

We may note several points concerning P'_n and P''_n and the corresponding definitions of electoral success.

First, if intensities are uniform across voters, that is, either $\Delta_i = \Delta_j$ or $\Delta_i = 1 - \Delta_j$ for all $i, j \in N$, the two criteria make the same prescription and $P'_n = P''_n$.

Second, the same is also true if average intensities are the same in the majority bloc, say N_A, and the minority bloc, say N_B.

Thus, the two criteria make contrary prescriptions only when these average intensities differ and when they are sufficiently greater (on the average) on the minority side (which is, of course, precisely the sense of the "intensity problem" discussed by Dahl, 1956, pp. 48 ff., and referred to previously).

Finally, we may note that, whenever the two criteria do conflict, *the electoral process*—based as it is on "one-man, one-vote" majority voting—*can succeed by the intensity criterion only if voters are incompletely informed,* since by definition, when all voters are completely informed, the electoral process always succeeds by the former criterion. Thus, insofar as we find the second criterion ever appealing, we can take some comfort from the fact that voters' information is in practice always incomplete. (Incomplete information is an example of what Dahl, 1961, p. 305, calls "slack in the system," which allows highly motivated actors to achieve political goals beyond what their "power" assures.)

Consider the following two examples. In both cases, n = 100, members of the majority bloc N_A have an (average) party differential of .6, and members of the minority bloc N_B have an (average) party differential of .25. In the first case, N_A constitutes three-fourths of the electorate; in the second, two-thirds of the electorate. By the majority interest criterion, party A should win in either case. By the intensity criterion, party A should win in the first case (the average of all differentials is .5125) and party B should win in the second case (the average of all differentials is .4833). As Table 4 shows, A is expected to win in the first case even when voters are only minimally informed; and as all voters become equally better informed, party A becomes more and more likely to win ($P'_n = P''_n$ increases) and its expected vote (p^*) more and more closely approaches its "deserved" margin of 75 percent. In the second case, party A is expected to lose when all voters are minimally informed, which violates the majority interest criterion and complies with the intensity criterion. But as all voters become equally better informed (beyond k = 3), the expected vote for party A and its probability of winning both increase, crossing the $p^* = 50$ percent and $p'_n = .5000$

Table 4. Constant Differentials, Varying Bloc Sizes
($\bar{\Delta}_A = .6$; $\bar{\Delta}_B = .25$; $n = 100$)

			$n_A = .75n$		$n_A = 2n/3$		
k	p_A	p_B	p^*	$P'_n = P''_n$	p^*	P'_n	P''_n
1	.6000	.7500	51.25%	.6035	48.33%	.3619	.6381
3	.6480	.8438	52.51%	.7104	48.41%	.3597	.6403
5	.6826	.8965	53.78%	.8010	48.96%	.4016	.5984
7	.7102	.9294	55.03%	.8883	49.70%	.5220	.4970
9	.7334	.9510	56.23%	.9413	50.53%	.5549	.4451
15	.7854	.9873	59.22%	.9948	52.78%	.7925	.2075
25	.8462	.9981	63.51%	.9999	56.48%	.9857	.0143
75	.9614	.9999	72.11%	.9999	64.10%	.9999	.0001
250	.9994	.9999	74.96%	.9999	66.63%	.9999	.0000
1000	.9999	.9999	75.00%	.9999	66.67%	.9999	.0000

thresholds at $k = 9$, and party A ultimately approaches its "deserved" 66.67 percent margin.[6]

Table 5 displays a basically similar pair of examples, except that here the two blocs are of constant size, $n_A = 60$ and $n_B = 40$, while their (average) differentials vary. The first case is symmetric: $\bar{\Delta}_A = .75$ and $\bar{\Delta}_B = .25$; in the second case, the minority is more intense: $\bar{\Delta}_A = .60$ and $\bar{\Delta}_B = .125$. Again, of course, by the majority interest criterion, party A should win in both cases, while by the intensity criterion party A should win in the first case and party B in the second. The pattern of expected votes and probabilities is the same in Table 5 as in Table 4.

Table 5. Constant Bloc Sizes, Varying Differentials
($n_A = 60$; $n_B = 40$)

	$\bar{\Delta}_A = .75$; $\bar{\Delta}_B = .25$			$\bar{\Delta}_A = .60$; $\bar{\Delta}_B = .125$				
k	p	p^*	$P'_n = P''_n$	p_A	p_B	p^*	P'_n	P''_n
1	.7500	55.00%	.8758	.6000	.8750	41.00%	.0189	.9811
3	.8438	56.88%	.9709	.6480	.9570	40.60%	.0082	.9918
5	.8965	57.93%	.9953	.6826	.9839	41.60%	.0114	.9886
7	.9294	58.59%	.9996	.7102	.9938	42.86%	.0221	.9779
9	.9510	59.02%	.9999	.7334	.9975	44.01%	.0432	.9568
15	.9873	59.75%	.9999	.7854	.9999	47.13%	.1833	.8167
25	.9981	59.96%	.9999	.8462	.9999	50.77%	.6088	.3912
75	.9999	60.00%	.9999	.9614	.9999	57.68%	.9999	.0001
250	.9999	60.00%	.9999	.9999	.9999	59.96%	.9999	.0000
1000	.9999	60.00%	.9999	.9999	.9999	60.00%	.9999	.0000

Table 6. An Increasingly Well-Informed Minority Bloc ($\bar{\Delta}_A = .75$; $\bar{\Delta}_B = .25$; $n_A = 60$; $n_B = 40$)

k_A	k_B	p_A	p_B	p^*	$P'_n = P''_n$
1	1	.7500	.7500	55.00%	.8759
1	3	.7500	.8438	51.25%	.6206
1	5	.7500	.8965	49.14%	.4120
1	7	.7500	.9294	47.82%	.2796
1	9	.7500	.9510	46.96%	.2006
1	15	.7500	.9873	45.51%	.0950
1	25	.7500	.9981	45.08%	.0717
1	75	.7500	.9999	45.00%	.0682

Cases can also arise in which both the majority interest criterion and the intensity criterion are violated. This can occur when information is unequally distributed, specifically when the minority bloc is, on the average, better informed than the majority. Table 6 provides an example. In this case, party differentials are symmetric: $\bar{\Delta}_A = .75$ and $\bar{\Delta}_B = .25$. Further $n_A = 60$ and $n_B = 40$. Thus, by either criterion, party A should win. With equal information, including equal minimal information, it will probably do so. (As information becomes more complete, while remaining equal, the probability will increase further and likewise the expected vote, as illustrated in the first part of Table 5.) But, if the information level of the majority bloc remains minimal while that of the minority bloc increases, party A's expected vote and probability of winning decrease, falling below 50 percent and .5000, respectively, at $k_B = 9$ and ultimately approaching 45 percent and .0682, respectively.

Note that, given $\Delta_A = .75$, party B can never be expected to win (regardless of the information advantage of the minority) if $n_A \geq 67$. Likewise, given $n_A = 60$, party B can never be expected to win if $\bar{\Delta}_A > .8333$. And in general party B can never be expected to win if $(n_a)(\bar{\Delta}_A) > n/2$—all this assuming that all voters are at least minimally informed.

5. FURTHER DEVELOPMENTS

Clearly the model presented here can and should be developed further and its properties investigated further and in a more analytical fashion. Moreover, the model itself is open to significant extensions in several directions, if factors that here are taken as given are allowed to vary. For example, in the present discussion, if A and B are parties or candidates, they are inert. But they might be allowed to "campaign," not necessarily by adjusting their platform positions (in the manner of the well-developed spatial models of electoral competition) but by changing the information levels of selected classes of voters.

Conversely, voters themselves might choose actively to seek additional information as a rational strategy (see Downs, 1957, Pt. III; Powell, 1979; and Calvert, 1980). Of course, if we allow individual voters to "purchase" (additional) information at some cost, it is clear that in an electorate of any size voters will not be instrumentally motivated to do so. This conclusion is consistent with much public choice literature going back to Downs (1957, pp. 214–218 especially). But what has often been overlooked—and what we have emphasized here—is that the apparent bad consequences for democracy and the electoral process of "rational ignorance" in the electoral context are at least mitigated and perhaps reversed by the "statistical" mechanism identified by the Condorcet Jury Theorem and its extensions presented here. Moreover, the same factor—the large size of electorates—that discourages voters from acquiring political information also reduces the need (from the point of view of the chances of "success" of the electoral process) for individual voters to be well informed.

Of course, this optimistic conclusion cannot be sustained if there are substantial inequalities (of a particular sort) in the distribution of information in the electorate. But the fundamental lesson of this paper is worth restating: it is inequalities or biases (of a particular sort) in the information levels in the electorate, and *not generally low levels of information*, that threaten the "success" of the electoral process.

This is an abridged version of a paper presented at the Annual Meeting of the Midwest Political Science Association, Chicago, April 24–26, 1980. A earlier version of portions of this paper was presented at the Annual Meeting of the Public Choice Society, New Orleans, March 1977. Additional tables which illustrate the various points made in the paper with numerical examples are available upon request from the author.

NOTES

1. In general, we might want to add a third category N_0 of voters whose "true" interests would be equally well—or badly—served by either alternative. Correspondingly, there might be a third category of information bits X_i^0 that have no bearing, for voter i, on the choice between A and B.
2. Since X may be taken to be infinite in size, or virtually so, we may disregard the possibility that $\Delta_i = .5$ exactly.
3. See Kelley and Mirer (1974, p. 574.) Kelley and Mirer, however, use party identification as a tie-breaking criterion.
4. Alternatively, voter i might abstain if $k_i^A = k_i^B$.

5. That is, P_n in the case $n = 1$ and $k_1 = t$ and in the case $n = t$ and $k_1 = \cdots = k_t = 1$ are the same. With an intermediate number n of voters and an intermediate level k of (equally distributed) information with $nk = t$, P_n is slightly less than in the two extreme cases, especially when either or both of n and k is an even number. Two simple examples are shown below. (In both $\Delta_i = .6$ for all i.)

n	k	p	P_n
1	15	.7869	.7869
3	5	.6826	.7617
5	3	.6480	.7617
15	1	.6000	.7869
1	12	.7535	.7535
2	6	.6826	.6826
3	4	.6480	.7155
4	3	.6480	.7155
6	2	.6000	.6826
12	1	.6000	.7535

6. If the electorate were larger, all probabilities P'_n and P''_n greater than .5 would become still greater, and those smaller than .5 would become still smaller.