Information Aggregation, Rationality, and the Condorcet Jury Theorem

DAVID AUSTEN-SMITH and JEFFREY S. BANKS *University of Rochester*

The Condorcet Jury Theorem states that majorities are more likely than any single individual to select the "better" of two alternatives when there exists uncertainty about which of the two alternatives is in fact preferred. Most extant proofs of this theorem implicitly make the behavioral assumption that individuals vote "sincerely" in the collective decision making, a seemingly innocuous assumption, given that individuals are taken to possess a common preference for selecting the better alternative. However, in the model analyzed here we find that sincere behavior by all individuals is not rational even when individuals have such a common preference. In particular, sincere voting does not constitute a Nash equilibrium. A satisfactory rational choice foundation for the claim that majorities invariably "do better" than individuals, therefore, has yet to be derived.

Although never achieving the notoriety of his discovery that majority voting can be cyclic, the Jury Theorem proposed by Condorcet ([1785] 1994) has received periodic attention and in the last decade been the subject of substantial interest and analysis. Loosely speaking, the theorem is as follows. Suppose there are two mutually exclusive alternatives, $\{A, B\}$, such that one of these alternatives is unequivocally better for all of n individuals in a group but the identity of the better alternative (i.e., A or B) is unknown. Suppose further that for all individuals $i \in N = \{1, \ldots, n\}$, the probability that i votes for the better alternative is $p > 1/2$ and is independent of the probability that any $j \neq i$ votes for the better alternative. Then the Jury Theorem states that the probability that a majority votes for the better alternative exceeds p and approaches 1 as n goes to infinity (Black 1958, 164–65; McLean and Hewitt 1994, 34–40). Given this, the Jury Theorem has been used as a positive argument for decision making by majority rule: Majorities are more likely to select the "correct" alternative than any single individual when there is uncertainty about which alternative is in fact the best (e.g., Berg 1993; Ladha 1992, 1993; Grofman and Feld 1988; Miller 1986; Nitzan and Paroush 1985; Paroush 1994; Young 1988).

The proofs of the Jury Theorem and subsequent extensions are entirely statistical in nature (see Berg 1993; Ladha 1992, 1993; Nitzan and Paroush 1985). Specifically, they take the individual probabilities of correct decisions parametrically and then aggregate them in a particular way. An important—but largely implicit—assumption of such proofs then is that an individual deciding which of the two outcomes to select when voting as a member of a collective (or jury) behaves in exactly the same manner as when that individual alone selects the outcome. Alternatively (to use the language of voting theory), individuals are taken to vote "sincerely." This implicit behavioral assumption is seen in the use of the same probability terms in the aggregation to generate both the likelihood that a majority chooses correctly and the likelihood that any one individual chooses correctly.

Our purpose here is to bring the "sincerity" assumption into the light and determine its relevance for the Jury Theorem. One intuition is that the assumption of sincere voting is innocuous. After all, every individual in the collective has the same fundamental preferences over the given binary agenda in that all want to choose the "correct" alternative. Consequently, there is neither incentive nor opportunity for any individual to manipulate the collective decision to their particular advantage at the expense of others. We show, however, that this intuition is mistaken. The argument is illustrated by the following example.

Consider a group of three individuals, all of whom share identical preferences over two alternatives, $\{A, B\}$, conditional on knowing the true "state of the world." There is, however, uncertainty about the true state of the world, which may (without ambiguity) be either state A or state B. In state A (B), individuals each receive a payoff of 1 if alternative A (B) is chosen and receive a payoff of 0 otherwise. There is a common prior probability that the true state is A. Individuals have private information about the true state of the world. Specifically, prior to any decision on which alternative to choose, each individual i privately observes a signal, $s_i = 0$ or $s_i = 1$, about the true state: if the true state is A, then it is more likely that the received signal is 0; and if the true state is B, then it is more likely that the received signal is 1. Once each individual has received his or her signal, the group chooses an alternative by majority vote (with no abstention). Three sorts of voting behavior are of particular interest: *sincere voting*, in which each individual selects the alternative yielding his or her highest expected payoff conditional on their signal; *informative voting*, in which each individual i votes for A if and only if receiving a signal $s_i = 0$; and *rational voting*, in which individuals' decision rules constitute a Nash equilibrium (i.e., given everyone else's rule, each individual votes to maximize the expected payoff). To complete the example, we make two assumptions on individuals' beliefs: (1) sincere voting is informative in that on receiving a signal of 0 (1) an individual thinks A (B) is the true state; and (2) the common prior belief that the true state is A is sufficiently strong that if any individual i were to observe *all three* individuals' signals, then i believes B is the true state only if *all* the available evidence supports

David Austen-Smith is Professor of Political Science and Jeffrey S. Banks is Professor of Political Science and Economics, University of Rochester, Rochester, NY 14627.

Austen-Smith thanks faculty and staff at the Research School in the Social Sciences, Australian National University, for their hospitality during the completion of this paper. Banks is similarly appreciative of members at CREED, University of Amsterdam, and the Tinbergen Institute. Both authors thank the referees for useful comments and the NSF for financial support.

the true state being B (i.e., $s_1 = s_2 = s_3 = 1$). Then despite it being common knowledge that individuals' preferences before receiving any private information are in fact identical, in this example *sincere voting by all individuals is not rational.*

To see why sincere voting is not rational here, consider any individual i and assume that the remaining individuals, say j and k, vote sincerely. By assumption (1), j and k are voting informatively, in which case individual i must be in one of three possible situations: Either (a) j and k have both observed a 0 and vote for A, (b) j and k have both observed a 1 and vote for B, or (c) j and k have observed different signal-values and vote for different alternatives. Under (a) or (b) the outcome is independent of i's vote, and if (c) obtains, i's vote is decisive. Consequently, the difference in i's expected payoffs from voting for A versus voting for B depends entirely on what i does conditional on situation (c). However, if the situation is (c), then i infers that exactly one of the remaining individuals (it does not matter which) has observed a signal of 0 in which case, by assumption (2), i has an unequivocally best decision: vote for A *irrespective* of the value of i's signal. Hence all individuals voting sincerely cannot be a Nash equilibrium.

Before going on to the general formal analysis, it is worth emphasizing two points. The first point is that when any one individual votes "insincerely," she is in fact acting in *everyone's* best interest (given the common ex post preferences) and not just her own. Thus the distinction between sincere and rational voting when there are only two alternatives derives entirely from the information-based heterogeneity of individuals' preferences consequent on receiving their private information (*interim* preferences), in that the existence of such heterogeneity allows for valuable information to be inferred in equilibrium.

The second point is that the result that, in general, sincere and informative voting by all individuals cannot be a Nash equilibrium does *not* say that sincere or informative voting by *any one* individual is inconsistent with Nash equilibrium. Further, there may be equilibria in which some vote sincerely and others do not and where the resulting majority-rule outcome improves on individual decision making. Our goal here, however, is less ambitious than that of identifying all Nash equilibria of certain games, determining which generate better outcomes compared with individual decision making and which do not. Rather, we simply demonstrate that while the canonical statement and proofs of the Condorcet Jury Theorems are correct as they stand *given* the additional assumption that all individuals vote sincerely and informatively, such an assumption is inconsistent with a game-theoretic view of collective behavior. A satisfactory rational choice foundation for the claim that majorities invariably "do better" than individuals, therefore, has yet to be derived.[1]

We shall develop the basic framework, essentially elaborating on the setup underlying this example. Subsequently, we consider two extensions of the basic model to explore the robustness of the main result (i.e., that sincere voting by all individuals cannot generally be both informative and rational) to variations in the structure of individuals' information. Finally, we shall consider some implications of the results.

THE BASIC MODEL

One of the difficulties with previous work on the Jury Theorem is that from a behavioral perspective much of the analysis essentially "starts in the middle" of an information accumulation and aggregation process. The individual likelihoods of making correct decisions are often described as posterior probabilities, implying the existence of a prior belief, as well as some observed event statistically related to the true best alternative (e.g., Ladha 1992). In contrast, we describe in detail a model of prior beliefs and events that permits identifying optimal individual behavior. And because we wish the model to fall within the domain of problems to which previous work on the Jury Theorem speaks, it is constrained to generating a "middle" or interim stage consistent with the features of existing research on the Jury Theorem (e.g., individual probabilities of correct decisions, correlations between such probabilities).

There is a pair of alternatives $\{A, B\}$ and two possible states of the world; without ambiguity, we also label the states $\{A, B\}$. There is a set $N = \{1, \ldots, n\}$ of individuals; assume n is odd and $n \geq 5$. Individuals have identical preferences over alternatives and states of the world, represented by

$$\forall i \in N,\ u_i(A, A) = u_i(B, B) = 1 \textit{ and }$$
$$u_i(A, B) = u_i(B, A) = 0, \quad \textbf{(1)}$$

where the first argument of u_i describes the alternative selected and the second describes the state. Hence, all individuals prefer to select alternative A when the state is A, and alternative B when the state is B.[2]

The true state of the world is unknown to the individuals; let $\pi \in (0, 1)$ denote the common prior probability that the true state is A. Before any decision is made over $\{A, B\}$, individual $i \in N$ receives a private signal, $s_i \in \{0, 1\}$, about the true state of the world. Individuals' signals are independent draws from a state-dependent distribution satisfying

$$\Pr[s_i = 0 | A] = q_a \in (1/2, 1) \textit{ and }$$
$$\Pr[s_i = 1 | B] = q_b \in (1/2, 1). \quad \textbf{(2)}$$

[1] But see Feddersen and Pesendorfer 1995 and Myerson 1994 for large population results.

[2] The imposition of symmetry on the payoffs for making the "wrong" decision is an expository convenience only. If, as Condorcet assumed, $u_i(A, B) \neq u_i(B, A)$, then some of the definitions to follow need to be modified in obvious ways, and the algebra supporting the results becomes correspondingly messier (McLean and Hewitt 1994). The qualitative results themselves, however, are unaffected by the symmetry assumption. To see this, recall the first example presented. Here, the intuition behind (general) sincere voting not being rational is independent of the payoffs. The key thing is that sincere voting is informative. With asymmetric payoffs, the condition governing when sincere voting is also informative will be modified, but that is all.

That is, if the true state of the world is A, then it is more likely that a signal of $s_i = 0$ will be observed, whereas if the true state is B, it is more likely that $s_i = 1$ will be observed. We refer to this description of preferences and information collectively as model I.

After observing her signal, each individual votes for either alternative A or B. Thus a *voting strategy* for individual i is a map $v_i:\{0, 1\} \rightarrow \{A, B\}$ describing which of the two alternatives i votes for as a function of her information. Let $v:\{0, 1\}^n \rightarrow \{A, B\}^n$ defined by $v(s) = (v_1(s_1), \ldots, v_n(s_n))$ denote a *voting profile*.[3] Based on her signal, an individual can update her prior belief of π and determine which of the two alternatives would provide the higher expected utility if she alone were making the decision. In particular, from relationship 1, the expected utility of the alternative A being chosen, given the signal s_i, is simply the probability that the true state is A, and similarly for B. Employing Bayes' Rule and relationship 2, these probabilities are given by

$$\Pr(A|s_i = 0) = \frac{\pi q_a}{\pi q_a + (1 - \pi)(1 - q_b)},$$

$$\Pr(B|s_i = 0) = \frac{(1 - \pi)(1 - q_b)}{\pi q_a + (1 - \pi)(1 - q_b)},$$

$$\Pr(B|s_i = 1) = \frac{(1 - \pi)q_b}{\pi(1 - q_a) + (1 - \pi)q_b},$$

$$\Pr(A|s_i = 1) = \frac{\pi(1 - q_a)}{\pi(1 - q_a) + (1 - \pi)q_b}.$$

Therefore we have that

$$E[u_i(A, \cdot)|s_i = 0] > E[u_i(B, \cdot)|s_i = 0] \Leftrightarrow \pi q_a > (1 - \pi)(1 - q_b) \quad (3)$$

$$E[u_i(B, \cdot)|s_i = 1] > E[u_i(A, \cdot)|s_i = 1] \Leftrightarrow (1 - \pi)q_b > \pi(1 - q_a). \quad (4)$$

DEFINITION. *A voting strategy* v_i *is* sincere *if* $v_i(s_i) = A$ *(B) if and only if*

$$E[u_i(A, \cdot)|s_i] > (<)\ E[u_i(B, \cdot)|s_i].^4$$

Note that we can evaluate the individuals' preferences over the choices of alternative A and B at various hypothetical stages in the process, where these stages differ according to the information possessed by the individuals. For example, at the ex ante stage, before they have received their private signals, individuals have identical beliefs, as characterized by the prior π, and hence identical preferences over A and B. Similarly, at some ex post stage, where either the state is known with certainty or else remains unknown but where all of the private signals have been revealed, individuals again have identical preferences. On the other hand, at the interim stage, where individuals possess their private information but *only* their private information (and where this is the stage when actual decisions are made), preferences over the choice of A or B can diverge; that is, some individuals may now prefer alternative A, whereas others may prefer B (this difference of course emanating from their private signals). Thus we have heterogeneous policy preferences being generated endogenously within the model. Furthermore, the fact that this heterogeneity is *information-based* will play an important role in our results.

Second, note that if the prior π is sufficiently high relative to q_a and q_b, then sincere voting would prescribe choosing alternative A regardless of the signal observed; and similarly, a choice for B if π is sufficiently small. Under either of these circumstances it is clear that if individuals vote sincerely they will all vote for the same alternative, implying the probability of a correct majority decision will be exactly the same as the probability that any one individual chooses correctly. Hence, the Jury Theorem will not hold. In the present context, that is, the Condorcet Jury Theorem assumes not only that voting is sincere but also that sincere voting is *informative*.

DEFINITION. *A voting strategy* v_i *is* informative *if* $v_i(0) = A$ *and* $v_i(1) = B$.

Thus when all individuals adopt informative strategies all of their private information is revealed through their voting decisions.[5]

While recent extensions of the Condorcet Jury Theorem speak only to certain characteristics of majority rule, we wish to allow for other types of rules as well. Define an *aggregation rule* to be a map f: $\{A, B\}^n \rightarrow \{A, B\}$ describing the outcome of the process as a function of the individuals' votes. We restrict attention here to aggregation rules that are *anonymous* (i.e., that treat all individuals the same) and *monotonic* (if B is chosen when it receives k votes, then it is also chosen when it receives more than k votes). For any such aggregation rule f, we can define a nonnegative integer k_f such that B is the outcome if and only if B receives more than k_f votes.[6] For example, the *majority voting* aggregation rule is given by $k_f = (n - 1)/2$.

Given an aggregation rule f, we now have a well-defined *Bayesian game B(f)* in which N is the set of players, $\{A, B\}$ is the action set for each $i \in N$, $\{0, 1\}$ is the set of "types" each individual can be, the appropriate probabilities over "types" and utilities over vectors of actions, are induced from relationships 1 and 2 and f, and this structure is taken to be common knowledge among the participants (see appendix).

[3] Throughout, we will say that a voting profile satisfies a certain property if each of its individual components satisfies the property.

[4] We ignore instances where (π, q_a, q_b) are such that individuals are indifferent between choosing A and B, since these are (in a certain well-defined sense) not typical. Similarly, all inequalities below implicitly involving (π, q_a, q_b) will be taken to be strict.

[5] Of course, such information is also revealed when another strategy is played: $v_i(0) = B$ and $v_i(1) = A$. However, it is easily shown that this "contrary" strategy is weakly dominated.

[6] Note that this is exactly the class of rules considered by Condorcet, who, in the explicitly jury context, suggests that a defendant be convicted if and only if the number of votes for conviction exceeds some critical number depending, in general, on the parameters of the situation (McLean and Hewitt 1994).

DEFINITION. *A voting profile is* rational *in model I if it constitutes a Nash equilibrium of the Bayesian game* B(f).

It is worth emphasizing that, unlike the definitions of sincere and informative voting, rationality is a property of voting *profiles* and not of individual voting strategies.

The most important feature of Nash equilibrium voting in this environment is that there is, in principle, information about other individuals' private signals that can be incorporated into the decision to select alternative A or alternative B. This additional information comes about as follows: in computing whether A or B is the better response to other individuals' voting strategies, individual i only concerns herself with those situations where she is "pivotal"—that is, where her vote makes a difference in the collective choice, where here a situation is a particular list of the others' private information. Suppose, for example, that all other individuals were adopting the informative strategy described above and collective decision making is by majority rule. In those situations where at least $n/2$ of the others have observed 1s, i's vote is immaterial, because regardless of how she votes a majority will vote for B—and similarly for those situations where at least $n/2$ of the others have observed 0s. Therefore the only instance in which i's vote matters is when exactly $(n - 1)/2$ of the other individuals have observed 1s (and hence $(n - 1)/2$ have observed 0s). But then in making her voting decision, i can essentially *presume* that exactly $(n - 1)/2$ of the others have observed 1s, thereby generating this additional "equilibrium" information (see Austen-Smith 1990; Feddersen and Pesendorfer 1994). This argument is made precise in the appendix.

Finally, for any vector of signals

$$s = (s_1, \ldots, s_n),\ \text{let } k(s) = \sum_N s_i.$$

Just as in the determination of i's sincere voting strategy, we can compute via Bayes' Rule the probability that the true state is either A or B—and hence the expected utility from alternative A or B being chosen, conditional on the vector s:

$$E[u_i(A, \cdot)|s] > E[u_i(B, \cdot)|s] \text{ iff}$$

$$\frac{\pi}{1-\pi} > [q_b/(1 - q_a)]^{k(s)}[(1 - q_b)/q_a]^{n-k(s)} \quad \textbf{(5)}$$

and

$$E[u_i(A, \cdot)|s] < E[u_i(B, \cdot)|s] \text{ iff}$$

$$\frac{\pi}{1-\pi} < [q_b/(1 - q_a)]^{k(s)}[(1 - q_b)/q_a]^{n-k(s)}. \quad \textbf{(6)}$$

Of course, since all individuals have the same preferences, relationships 5 and 6 give the optimal decision for any individual given all of the available information about the true state—and hence for the collective as a whole.

Because

$$k(s) = \sum_N s_i,$$

relationship 2 implies that the greater is $k(s)$, the more likely it is that the true state is B and the less likely it is that the true state is A. Suppose equation 5 holds when $k(s) = 0$ and suppose relationship 6 holds when $k(s) = n$. Then there must exist some critical value of $k(s)$, say k^*, such that $k(s) \le k^*$ implies $E[u_i(A, \cdot)|s] > E[u_i(B, \cdot)|s]$, and $k(s) > k^*$ implies $E[u_i(A, \cdot)|s] < E[u_i(B, \cdot)|s]$. Then, k^* is defined implicitly by:

$$[q_b/(1 - q_a)]^{k^*+1}[(1 - q_b)/q_a]^{n-k^*-1} > \frac{\pi}{1-\pi} > [q_b/(1 - q_a)]^{k^*}[(1 - q_b)/q_a]^{n-k^*}, \quad \textbf{(7)}$$

where the dependence of k^* on (q_a, q_b, π) is understood. In general, then, k^* describes the optimal method of aggregating individuals' private information (if this information were known) in that the group unanimously prefers the collective decision to be A, rather than B at s if and only if $k(s) \le k^*$. Of course, k^* may not exist for all parameterizations. In particular, if relationship 5 holds at $k(s) = n$, then all individuals would prefer A over B irrespective of the vector of signals s—and conversely if relationship 6 holds at $k(s) = 0$. To avoid such trivialities, hereafter assume the parameters (q_a, q_b, π) are such that relationship 5 holds at $k(s) = 0$ and relationship 6 holds at $k(s) = n$.

We now turn to the question of when the explicit as well as implicit assumptions of the Condorcet Jury Theorem are satisfied in model I. Note first that the probability that i votes correctly, given that her strategy is informative, is simply

$$p_i \equiv \Pr[s_i = 0|A]\Pr[A] + \Pr[s_i = 1|B]\Pr[B] = \pi q_a + (1 - \pi)q_b > 1/2, \text{ all } i \in N, \quad \textbf{(8)}$$

while the probability that any pair of individuals (i, j) both vote correctly is given by

$$r_{ij} \equiv \Pr[s_i = 0 \text{ and } s_j = 0|A]\Pr[A] + \Pr[s_i = 1 \text{ and } s_j = 1|B]\Pr[B] = \pi q_a^2 + (1 - \pi)q_b^2. \quad \textbf{(9)}$$

(Because p_i and r_{ij} are identical across all i, j $(i \neq j)$, hereafter we write $p \equiv p_i$ and $r \equiv r_{ij}$.) Now suppose $q_a = q_b = q$; that is, the probability that $s_i = 0$ when the true state is A is the same as the probability $s_i = 1$ when the true state is B (this will be relaxed later). Under this assumption, if individuals' voting strategies are informative then the probability that individual i votes correctly is independent from that of voter j: $p = \pi q + (1 - \pi)q = q$ and, therefore, $r = \pi q^2 + (1 - \pi)q^2 = q^2 = p^2$.

LEMMA 1. *Assume* $q_a = q_b$. *Then sincere voting in model I is informative if and only if* $k^* = (n - 1)/2$.

Proof. Assume $q_a = q_b = q$. By relationships 3 and 4, if sincere voting is informative, then

$$q/(1 - q) > \frac{\pi}{1-\pi} > (1 - q)/q \quad \textbf{(10)}$$

(or, equivalently, $q > \pi > 1 - q$). Now set $q_a = q_b = q$ in relationship 7 and collect terms to yield k^* implicitly defined as the integer such that

$$[q/(1-q)]^{2(k^*+1)-n} > \frac{\pi}{1-\pi} > [(1-q)/q]^{n-2k^*}. \qquad \textbf{(11)}$$

Since $q/(1 - q) > \pi/(1 - \pi)$ by relationship 10,

$$[q/(1-q)]^{2(k^*+1)-n} > \frac{\pi}{1-\pi}$$

implies $k^* \geq (n - 1)/2$; and similarly, since relationship 10 requires

$$\pi/(1-\pi) > (1-q)/q, \frac{\pi}{1-\pi} > [(1-q)/q]^{n-2k^*}$$

implies $k^* \leq (n - 1)/2$. Therefore, if sincere voting is informative at (q_a, q_b, π) when $q_a = q_b$, $k^* = (n - 1)/2$ necessarily. To check sufficiency, substitute $(n - 1)/2$ for k^* in relationship 11 to yield relationship 10 and, by relationships 3 and 4, note that relationship 10 implies sincere voting is informative. *Q.E.D.*

Therefore if $q_a = q_b$ the only time sincere voting will be informative is when the optimal decision rule (i.e., the optimal way to aggregate individuals' information) is majority rule.

Lemma 1 states precisely when informative voting is sincere. The next result states precisely when informative voting is rational.

LEMMA 2. *For all* $(q_a, q_b) \in (1/2, 1)^2$, *informative voting in model I is rational if and only if the aggregation rule* f *is such that* $k_f = k^*$.

Necessity Proof. Suppose instead that $k_f \neq k^*$, where without loss of generality $k^* < k_f$. Consider the decision by individual i when all other individuals are voting informatively; if we can show that i does *not* want to vote informatively, then we will be done. The only time i is pivotal (i.e., the only time i's vote makes a difference in the collective decision) is when B receives exactly k_f votes. Assuming all other individuals are voting informatively, this only occurs when exactly k_f of the individuals other than i have observed 1s. But then given $k^* < k_f$, the optimal decision is to select alternative B regardless of i's signal, and therefore i's best response is $v_i(s_i) = B$ for $s_i = 0$ *and* $s_i = 1$; that is, i's best response is not to vote informatively.

Sufficiency Proof. Employing the same logic, if $k_f = k^*$, all other individuals are voting informatively and i is pivotal, then exactly k^* of the other individuals must have observed 1s. Thus if $s_i = 1$ then the optimal decision is to select B, so that i should vote for B, whereas if $s_i = 0$, the optimal decision is A, and i should vote for A. Hence i's best response to informative voting by the others is also to vote informatively, and informative voting constitutes a Nash equilibrium. *Q.E.D.*

Combining these two results, we get the following.

THEOREM 1. *Assume* $q_a = q_b$*; then sincere voting in model I is informative and rational if and only if* $k_f = k^* = (n - 1)/2$.

That is, sincere voting is both informative and rational precisely when (a) majority rule is being used to aggregate individuals' votes and (b) majority rule is the optimal method of aggregating individuals' information.

As a corollary, we can identify when the parameters of the model (q, π) are such that the implicit, as well as explicit, assumptions of the Condorcet Jury Theorem hold: Under majority rule sincere voting is both informative and rational when $k^* = (n - 1)/2$, that is, when majority rule is the optimal method of aggregating individuals' information.[7] For example, if $\pi = 1/2$, then $k^* = (n - 1)/2$ and hence, given a uniform prior, sincere voting will be both informative and rational under majority rule.

Conversely, whenever it is the case that k^* does *not* equal $(n - 1)/2$, one of the explicit or implicit assumptions of the Jury Theorem must be violated. In particular, we know that sincere voting is not informative when $k^* \neq (n - 1)/2$ (by Lemma 1) and so, even if sincere voting is rational, the probability that a majority makes a correct decision will be exactly the same as that of any individual (since sincere voting in this case requires voting for either A or B regardless of the private information). And if sincere voting is not rational, then the implicit behavioral assumption in the proofs of the Jury Theorem is inconsistent with Nash equilibrium behavior. On the other hand, if we ignore this behavioral assumption, for any value of k^* (or, more properly, for any values of (q, π)), there will exist a voting rule for which the *conclusion* of the Jury Theorem (i.e., that the collective performs better than any individual) is true. The identity of this voting rule follows immediately from lemma 2: Whatever the value of k^*, set $k_f = k^*$. Then informative voting will be rational and—by definition of k^* and the fact that there is more than one individual (so the collective has strictly more draws than does an individual)—the probability that the collective makes a correct decision is strictly higher than that of any one individual. Thus an alternative view of the Condorcet Jury Theorem, from the perspective of model I, is this: For any value of k^* there exists a voting rule that the Jury Theorem conclusion obtains; and majority rule is that rule precisely when $k^* = (n - 1)/2$, that is, when majority rule is the optimal method of aggregating the individuals' private information.

The symmetry assumption on q_a and q_b found in Theorem 1 is very special. For the Jury Theorem to say anything more than that majority voting aggregates information effectively when majority voting is the optimal way to aggregate information, it must apply in situations in which the latter is not the case. And by the preceding argument, this necessarily involves $q_a \neq q_b$. Unfortunately, under such circumstances, the theorem cannot generally assume that individuals are rational. To see this, recall that the original statement of the Condorcet Jury Theorem presumes that the probability that any individual votes for the better alternative is statistically independent of the same probability for any other

[7] Of course, the converse of this statement also holds: If $k^* = (n - 1)/2$, then sincere voting is informative and rational if and only if f is majority rule.

individual. However, when q_a does not equal q_b, there is necessarily some amount of correlation between these probabilities (see relationship 9). But independence is only a sufficient condition. Ladha (1992) provides the following upper bound on the correlation r, denoted $\hat{r}$, for the Jury Theorem to apply:

$$\hat{r} \equiv p - \frac{n}{n-1} \frac{1-p}{p}(p - .25). \tag{12}$$

That is, if the average correlation between the likelihoods of any two individuals (sincerely) voting for the correct alternative is less than $\hat{r}$, then the Condorcet Jury Theorem goes through. Now consider the following example:

Example 1. Let $N = \{1, \ldots, 21\}$, $q_a = .7$, $q_b = .8$, and $\pi = .5$. Making the appropriate substitutions, we calculate

i. from relationships 3 and 4, that sincere voting is informative
ii. from relationships 8 and 9, that $p = .75$ and $r = .565$
iii. from relationship 12, that $r = .575$
iv. from relationship 7, that $k^* = (n + 1)/2$.

By Ladha 1992, proposition 1, properties i–iii imply that the Jury Theorem applies to this case. But by property iv and Lemma 2 (which does not require $q_a = q_b$), informative—and hence sincere—voting under majority rule is not rational here.

ADDING DIVERSITY

In the model of the previous section, each individual was one of only two possible "types" following their private observations: those who believed A was the more likely state (relative to the prior) and those who believed B was the more likely state. These types then correspond to the two different posterior beliefs the individuals might hold. We shall modify the basic setup to allow for a third type to exist as well, namely, those whose posterior belief turns out to be "in the middle" of the first two. We do this by assuming that individuals observe not one draw from the true distribution but rather *two*. Thus a private signal is now a pair $s_i = (s_{i1}, s_{i2}) \in \{0, 1\}^2$ describing the observations of two independent draws from the true distribution found in relationship 2, and so an individual's strategy is now a mapping v_i: $\{0, 1\}^2 \rightarrow \{A, B\}$. (As before, $v(\cdot)$ will denote a voting profile.) Further, we make the symmetry assumption that $q_a = q_b \equiv q$, where $q \in (1/2, 1)$. Finally, we let the prior π on state A take on any value *except* 1/2 (the importance of this last assumption will become apparent) and restrict attention only to the majority aggregation rule. Together we refer to these assumptions as model II.

The definition of sincere voting is just as before, namely, that an individual chooses alternative A or B depending on which is the more likely state based on her private information. Since an individual's two draws are assumed to be independent, sincere voting can, without loss of generality, be written as a function of the number of 1s observed, so let $S_i = s_{i1} + s_{i2}$ and redefine v_i: $\{0, 1, 2\} \rightarrow \{A, B\}$. Further, since $q_a = q_b$, it is immediate from Bayes' Rule that the posterior belief associated with $S_i = 1$ is simply equal to the prior belief π. Therefore, sincere voting upon observing $S_i = 1$ prescribes voting for A if $\pi > 1/2$ and voting for B if $\pi < 1/2$.[8] For $S_i \in \{0, 2\}$, sincere voting is determined by the following inequalities, which are analogous to those found in relationships 3 and 4:

$$E[u_i(A, \cdot)|S_i = 0] > E[u_i(B, \cdot)|S_i = 0]$$

$$\Leftrightarrow \frac{\pi}{1-\pi} > \left[\frac{1-q}{q}\right]^2 \tag{13}$$

$$E[u_i(B, \cdot)|S_i = 2] > E[u_i(A, \cdot)|S_i = 2]$$

$$\Leftrightarrow \frac{\pi}{1-\pi} < \left[\frac{q}{1-q}\right]^2. \tag{14}$$

Two comments on these equations are in order. First, since $q > 1/2$, one of the relationships 13 and 14 must always hold: if $\pi > 1/2$ then necessarily relationship 13 is true, whereas if $\pi < 1/2$, then necessarily relationship 14 is true. Second, if relationship 13 holds but relationship 14 does not, then it must be that sincere individuals vote for alternative A regardless of their private information—and similarly for B if relationship 14 holds but relationship 13 does not. As previously, such behavior immediately implies that the Jury Theorem will not hold if all vote sincerely, because a majority decision is no more likely to be correct than any individual's decision. We thus have the following generalization of the earlier definition of an informative strategy:

DEFINITION. *A voting strategy* v_i *is* informative *if* $v_i(0) = A$ *and* $v_i(2) = B$.

Thus sincere voting will be informative if and only if both relationships 13 and 14 hold.

Finally, as before, we will say that a voting profile is *rational* if it constitutes a Nash equilibrium of the Bayesian game associated with model II. Again, the crucial distinction between sincere and rational voting will be due to the additional "equilibrium" information associated with the latter.

In model I we saw how sincere voting under majority rule can be both informative and rational in certain situations, namely, when majority rule was the optimal method of aggregating information. The following result states that no such conditions exist in model II.

THEOREM 2. *Sincere voting in model II cannot be both informative and rational.*

Proof. Suppose all $i \neq j$ vote sincerely and informatively, $S_j = 1$, and j is pivotal. Without loss of generality, let $\pi < 1/2$; thus a sincere vote by j would be to choose B. Since f is majority rule and j is presumed pivotal, she can infer that exactly $(n - 1)/2$ individuals have $S_i = 0$. Let $d \equiv (n - 1)/2$ and let $\mathscr{E}(s)$ denote the event $\{S_j = 1,$

[8] Hence, assuming $\pi \neq 1/2$ ensures that individuals with $S_i = 1$ are not indifferent.

$|\{i \in N \backslash \{j\}: S_i = 0\}| = d\}$. The expected utility, conditional on $\mathscr{E}(s)$, for j of alternative A being selected is equal to the probability A is the true state conditional on $\mathscr{E}(s)$—and similarly for B. Using Bayes' Rule and the Binomial Theorem, we have

$$E[u_j(A, \cdot)|\mathscr{E}(s)] = \frac{X}{X+Y}, E[u_j(B, \cdot)|\mathscr{E}(s)] = \frac{Y}{X+Y},$$

where

$$X = \pi(1-q)q\binom{2d}{d}q^{2d} \times \left\{\sum_{j=0}^{j=d}\binom{d}{j}(2q(1-q))^{d-j}(1-q)^{2j}\right\}$$

$$Y = (1-\pi)q(1-q)\binom{2d}{d}(1-q)^{2d} \times \left\{\sum_{j=0}^{j=d}\binom{d}{j}(2q(1-q))^{d-j}q^{2j}\right\}.$$

Therefore $E[u_j(A, \cdot)|\mathscr{E}(s)] > E[u_j(B, \cdot)|\mathscr{E}(s)]$ if and only if

$$\pi(1-q)q\binom{2d}{d}q^{2d}\left\{\sum_{j=0}^{j=d}\binom{d}{j}(2q(1-q))^{d-j}(1-q)^{2j}\right\} > (1-\pi)q(1-q)\binom{2d}{d}(1-q)^{2d}\left\{\sum_{j=0}^{j=d}\binom{d}{j} \times (2q(1-q))^{d-j}q^{2j}\right\} \Leftrightarrow \sum_{j=0}^{j=d}\binom{d}{j}(2q(1-q))^{d-j} \times [\pi(1-q)^{2j}q^{2d} - (1-\pi)q^{2j}(1-q)^{2d}] > 0.$$

Since $q > 1/2$ and relationship 13 holds,

$$\forall j = 0, \ldots, d-1, \quad [\pi(1-q)^{2j}q^{2d} - (1-\pi)q^{2j}(1-q)^{2d}] > 0.$$

Therefore all but the last bracketed term in this summation is positive, whereas the last term is negative. Moreover, considering only the first and last terms in the summation, we have

$$\begin{aligned}
&\binom{d}{0}(2q(1-q))^d[\pi q^{2d} - (1-\pi)(1-q)^{2d}] \\
&+ \binom{d}{d}[\pi(q(1-q))^{2d} - (1-\pi)(q(1-q))^{2d}] \\
&> \{[\pi q^{2d} - (1-\pi)(1-q)^{2d}] + [\pi(q(1-q))^d \\
&- (1-\pi)(q(1-q))^d]\}[q(1-q)]^d \\
&= \{\pi q^d[q^d + (1-q)^d] - (1-\pi)(1-q)^d[q^d \\
&+ (1-q)^d]\}[q(1-q)]^d > 0 \\
&\Leftrightarrow \pi q^d > (1-\pi)(1-q)^d,
\end{aligned}$$

where the last inequality holds by relationship 13, $q > 1/2$, and the fact that $n \geq 5$ implies $d \geq 2$. Hence, $E[u_j(A, \cdot)|\mathscr{E}(s)] > E[u_j(B, \cdot)|\mathscr{E}(s)]$, and therefore, conditional on being pivotal, j's best response is A; that is, if everyone else is voting sincerely, then j should not vote sincerely. *Q.E.D.*

Thus in model II sincere voting by all individuals is either uninformative (i.e., everyone votes for A or B independent of their private information), so that the *conclusion* of the Condorcet Jury Theorem that majority voting does strictly better than any one individual does not hold, or is not rational, in which case an implicit *assumption* of the theorem (i.e., that all individuals vote sincerely in collective decision making) prescribes irrational behavior.

Intuitively, it is apparent where the "problem" with the Jury Theorem lies in model II, in contrast to model I. In the former, there are now three possible "types" of voters (i.e., three possible posterior beliefs they could have), yet there are only two possible decisions, "vote for A" or "vote for B." Thus assuming sincere and informative voting, and $\pi < 1/2$, only one type ($S_i = 0$) votes for A, so that the private information possessed by those voting for A can be precisely inferred, whereas the remaining two types ($S_i = 1$ or 2) vote for B, thereby not allowing such a precise inference. This is in contrast to the informative voting found in model I, where such precise inferences always obtain. The problem, therefore, reduces to one of the "size of the message space," in that under a more complicated aggregation rule, individuals might simply announce their private information or (equivalently) announce their posterior beliefs.

This distinction between sincere and rational voting is to a certain degree reminiscent of that found in the complete-information voting literature. If, for example, individuals first vote on alternative x versus y, and then the winner of this contest is pitted against alternative z, then depending on the individual preferences one can have instances where sincere voting under majority rule is distinct from (and generates different outcomes from) rational or "sophisticated" voting. Such examples require at least three alternatives as well as a certain amount of heterogeneity in individual preferences. By contrast, in models I and II sincere and rational voting can differ when there are only two alternatives and where the ex ante and ex post preferences of the individuals are the same. As remarked in the introduction, the wedge between sincere and rational voting here rests on private information leading to heterogeneity in individuals' *interim* policy preferences, and such heterogeneity permits individuals to make inferences about the general distribution of private information in equilibrium.

Theorem 2 states that when relationships 13 and 14 hold, sincere voting does not constitute a Nash equilibrium of the Bayesian game associated with model II. This leaves open the question, What *are* the equilibria? Two candidates immediately come to mind, namely, everyone voting for A regardless of their private information and, similarly, everyone voting for B. As with most majority-rule voting games, either of these voting profiles is evidently a Nash equilibrium, because if everyone is (say) voting for A, then any one individual's vote is immaterial, so that voting for A regardless of private information is a best response. Now in complete-information voting games, these "trivial" equilibria are eliminated by requiring individual strategies to be undominated. Thus, if individual j prefers A to B, then voting for B is dominated. On the other hand, in the current incomplete-information voting game, such a refinement argument does not work. To see this, let $d = (n - 1)/2$, and have d individuals play the strategy "always vote for B" and d individuals play the strategy $v_i(S_i) = A$ if and only if $S_i = 0$. What is individual j's best response? Just as in the Nash calculations found in the proof of Theorem 2, j only need consider when she is pivotal; further, if one can show that a best response for j when $S_j = 2$ is to choose A, then A will be the best response for $S_j \leq 1$ as well. So suppose $S_j = 2$, and j is pivotal; then according to the others' strategies, it must be that the latter group of d individuals must all have private signals of $S_i = 0$, whereas j cannot make *any* inference about the former group's private information. Therefore, employing Bayes' Rule, we get that voting for A will be the better response as long as

$$\pi(1-q)^2q^{2d} > (1-\pi)q^2(1-q)^{2d}$$

$$\Leftrightarrow \frac{\pi}{1-\pi} > \left[\frac{1-q}{q}\right]^{2d-2},$$

where the latter inequality follows from $q > 1/2$, $d \geq 2$, and relationship 13. Therefore everyone voting for A regardless of their private information constitutes an undominated Nash equilibrium, and a similar logic shows the same to be true with respect to B.

What is most striking about these Nash equilibria is that they actually *reverse* the conclusion of the Condorcet Jury Theorem—that is, any one individual, acting alone (i.e. voting sincerely) will have a higher probability of making a correct decision than will a majority acting in accordance with one of these Nash equilibria. This follows from the fact that a sincere strategy maximizes the probability that a single individual makes a correct choice, and by relationships 13 and 14, such a maximizing strategy will depend nontrivially on the private information, in contrast to the "always vote A (or B)" strategy.[9]

Finally, Theorem 2 leaves open the possibility that whenever sincere voting is not rational, the Condorcet Jury Theorem fails to apply because the relevant correlations between individuals' likelihoods of voting correctly, when they vote sincerely, are too high. Were this invariably the case, Theorem 2 would not bear on the Jury Theorem directly. However this is not the case, as Example 2 below aptly demonstrates.

[9] Of course, other equilibria involving mixed or asymmetric strategies might leave the conclusion of the theorem intact. But then a complete explanation for why majorities are more likely to choose the "better" alternative than any individual requires some sort of equilibrium selection argument.

Example 2. Let $N = \{1, \ldots, 21\}$, $q = .9$, and $\pi = .6$. Then since $p = \pi[q^2 + 2q(1-q)] + (1-\pi)q^2$ and $r = \pi[q^4 + 4(1-q)^3q + 4q^2(1-q)^2] + (1-\pi)q^4$, we have

i. Informative voting is sincere
ii. $p = .918$, and $r = .8505$
iii. $\hat{r} = .8554$.

Thus, just as with Example 1, there exist environments where all of the explicit assumptions of the Jury Theorem are met, yet by Theorem 2 the implicit assumption of sincere voting prescribes irrational behavior.

A MODEL WITH PUBLIC AND PRIVATE SIGNALS

We have found existing proofs of the Condorcet Jury Theorem to be applicable and consistent with rational voting only in those circumstances in which majority voting is ex ante the optimal method for aggregating information. However, the results are derived in a model without public signals, whereas, typically there are public signals. For instance, "opinion leaders," the media, or acts of nature can be commonly observed and lead people to update their beliefs about the true state of the world. Hence, we shall extend the model to include a public signal and argue that in this setting, the consistency of the Jury Theorem with rational behavior breaks down in a similar fashion to that seen in the previous section.

In model III, the source and structure of an individual's private information is as in model II: two independent draws, where $q_a = q_b = q \in (1/2, 1)$. Therefore, there are again three types of individuals, depending on their private information (or, equivalently, on their posterior beliefs). In addition, we assume for analytic simplicity that the prior π is equal to 1/2. Now, however, subsequent to observing their private draws but before voting on the two alternatives, all individuals observe one public draw $s_p \in \{0, 1\}$. A voting strategy is now a mapping v_i: $\{0, 1\}^3 \rightarrow \{A, B\}$, with again $v(\cdot)$ denoting a voting profile.

We allow this public signal to differ in its "informational content" from that of individuals' private signals, by assuming

$$\Pr[s_p = 0|A] = Q = \Pr[s_p = 1|B],\ Q \in (1/2, 1).$$

That is, private and public signals can be drawn from different (state-dependent) distributions.

As before, sincere voting describes an individual's optimal voting decision based solely on her own private information, as well as (here) on the public information associated with the value of s_p. As in model II, for such behavior we can characterize an individual's private information by $S_i = s_{i1} + s_{i2}$, and then define a sincere strategy as a map from $\{0, 1, 2\} \times \{0, 1\}$ into $\{A, B\}$. Employing Bayes' Rule (and recalling that π is set equal to 1/2 and $Q > 1/2$), we know that if $S_i = 1$, then sincere voting prescribes voting in accordance with the public signal, that is, for A if $s_p = 0$ and for B if $s_p = 1$. Likewise, if $S_i = 0$ and $s_p = 0$, sincere voting would select A, whereas if $S_i = 2$ and $s_p = 1$, sincere voting would select B. The remaining cases are when $S_i = 0$ and $s_p = 1$, and when $S_i = 2$ and $s_p = 0$, that is, where there is a certain amount of conflict in an individual's private and public signals. In these two cases, by the various symmetry assumptions (i.e., $q_a = q_b$, $\pi = 1/2$), we have

$$E[u_i(A,\cdot)|S_i = 0, s_p = 1] > E[u_i(B,\cdot)|S_i = 0, s_p = 1] \Leftrightarrow q^2(1-Q) > (1-q)^2Q \Leftrightarrow E[u_i(B,\cdot)|S_i = 2, s_p = 0] > E[u_i(A,\cdot)|S_i = 2, s_p = 0] \quad \textbf{(15)}$$

If the inequalities do not hold, then sincere voting would have the individuals ignoring their private information and simply choosing between A and B based on the realization of the public draw. As previously, such a situation would render the probability of a correct majority decision under sincere voting equal to that of a correct individual decision. Therefore we again require the voting strategies to be *informative*:

DEFINITION. *A voting strategy in model III is* informative *if* $v_i(0, 1) = A$ *and* $v_i(2, 0) = B$.

From relationship 15, we have that sincere voting is informative if and only if $q^2(1-Q) > (1-q)^2Q$; that is, the public signal cannot be "too" informative relative to the private signals.

As before, we say that a voting profile is rational if it constitutes a Nash equilibrium of the Bayesian game associated with model III.

THEOREM 3. *Sincere voting in model III cannot be both informative and rational.*

Proof. The proof is almost identical to that for Theorem 2. Suppose all $i \neq j$ vote sincerely and informatively, $S_j = 1$, and j is pivotal; without loss of generality, let $s_p = 1$. Since f is majority rule and j is presumed pivotal, j can infer that exactly $(n-1)/2$ individuals other than j have $S_i = 0$. Let $d \equiv (n-1)/2$ and let $\mathscr{E}(s)$ denote the event, $\{S_j = 1, s_p = 1, |\{i \in \mathcal{N}\backslash\{j\}: S_i = 0\}| = d\}$. The expected utility, conditional on $\mathscr{E}(s)$, for j, of alternative A being selected is equal to the probability A is the true state conditional on $\mathscr{E}(s)$—and similarly for B. Using Bayes' Rule and the Binomial Theorem,

$$E[u_j(A,\cdot)|\mathscr{E}(s)] > E[u_j(B,\cdot)|\mathscr{E}(s)]$$
$$\Leftrightarrow q(1-q)(1-Q)\binom{2d}{d}q^{2d} \times \left\{\sum_{j=0}^{j=d}\binom{d}{j}(2q(1-q))^{d-j}(1-q)^{2j}\right\}$$
$$> q(1-q)Q\binom{2d}{d}(1-q)^{2d} \times \left\{\sum_{j=0}^{j=d}\binom{d}{j}(2q(1-q))^{d-j}q^{2j}\right\}$$
$$\Leftrightarrow \sum_{j=0}^{j=d}\binom{d}{j}(2q(1-q))^{d-j}[(1-Q)(1-q)^{2j}q^{2d} - Qq^{2j}(1-q)^{2d}] > 0.$$

Since $q > 1/2$ and relationship 15 holds,

$$\forall j = 0, \ldots, d-1,$$
$$[(1-Q)(1-q)^{2j}q^{2d} - Qq^{2j}(1-q)^{2d}] > 0.$$

Thus all but the last terms in the above summation are positive, while the last term is negative. Moreover, considering only the first and last terms in the summation, we have

$$\binom{d}{0}(2q(1-q))^d[(1-Q)q^{2d} - Q(1-q)^{2d}] + \binom{d}{d}[(1-Q)(q(1-q))^{2d} - Q(q(1-q))^{2d}]$$
$$> \{[(1-Q)q^{2d} - Q(1-q)^{2d}] + [(1-Q)(q(1-q))^d - Q(q(1-q))^d]\}[q(1-q)]^d$$
$$= \{(1-Q)q^d[q^d + (1-q)^d] - Q(1-q)^d[q^d + (1-q)^d]\}[q(1-q)]^d > 0$$
$$\Leftrightarrow (1-Q)q^d > Q(1-q)^d,$$

where the last inequality holds by relationship 15, $q > 1/2$, and the fact that $n \geq 5$ implies $d \geq 2$. Hence, $E[u_j(A,\cdot)|\mathscr{E}(s)] > E[u_j(B,\cdot)|\mathscr{E}(s)]$, and therefore, conditional on being pivotal, j's best response is A; that is, if everyone else is voting sincerely, then j should not vote sincerely. *Q.E.D.*

A number of remarks about this result are in order. The proof demonstrated that those individuals whose private signals were split (i.e., one 0 and one 1) vote against the public signal if all others are voting sincerely. If we assume that $n \geq 9$, then the same logic yields the same result when an individual's private signal is perfectly consistent with the public signal: if j privately observes two 1s, the public signal is 1, and all other individuals vote sincerely, then j's best response is to vote for A, rather than the sincere choice of B.

Second, the result does not rely critically on the assumption that $\pi = 1/2$, that is, each state is equally likely ex ante. Setting π equal to 1/2 merely allowed us to cancel terms that would be close as long as π was close to 1/2. Hence, as long as the ex ante likelihood of the states are sufficiently close, the result remains the same.

Finally, the assumption that there is only a single public signal is not important either. Suppose instead that there were m public draws, were m is greater than 1 and odd and suppose these consisted of $(m+1)/2$ 0s and $(m-1)/2$ 1s. Then, given our symmetry assumptions, Bayesian updating reveals that this generates the same posterior belief as having observed one 0 and no 1s (recall the equivalence of relationships 7 and 10 when $q_a = q_b$). So, upon observing such public draws, sincere voting will not be rational. A single public signal, therefore, is used only for emphasis, because in such an environment sincere voting is not rational regardless of the public draw.

As before, Theorem 3 raises the question, What *are* the Nash equilibria? One possibility is that the individuals ignore the public signal and vote solely on the information contained in their private signal. However, it is apparent that for any value of π other than 1/2, such behavior is equivalent to that discussed previously, and therefore sincere voting based only on private information is not a Nash equilibrium. Conversely, ignoring all private information and voting exclusively on the basis of the public signal is clearly a Nash equilibrium. Moreover, such behavior is undominated. To see this, suppose that if the public signal is 0, the first $d \equiv (n - 1)/2$ individuals vote for A regardless of their private information, and the next d vote for B if and only if their private information is $S_i = 2$; and if the public signal is 1, the first d individuals vote for B regardless of their private information and the next d vote for A if and only if $S_i = 0$. What is individual n's best response when the public signal is, say, 1? Just as in the Nash calculations found in the proof of Theorem 3, individual n only needs consider when she is pivotal. Further, if we can show that n's best response when $S_n = 0$ is to choose B, then B will be the best response for $S_n \geq 1$ as well, thereby demonstrating the claim. Making the relevant calculations, we find

$$E[u_n(B,\cdot)|\cdot] > E[u_n(A,\cdot)|\cdot]$$

$$\Leftrightarrow (1-q)^2 Q q^{n-1} > q^2(1-Q)(1-q)^{n-1}$$

$$\Leftrightarrow \frac{Q}{1-Q} > \left(\frac{1-q}{q}\right)^{n-3}.$$

Because $Q > 1/2$, $q > 1/2$, and $n \geq 5$, this inequality necessarily holds.

Theorem 3 leaves open the possibility that whenever rational voting is not sincere, the Condorcet Jury Theorem fails to apply because the relevant correlations between individuals' likelihoods of voting correctly, when they vote sincerely, are too high. Just as with Theorem 2, however, we provide an example to show that this is not invariably the case:

Example 3. Let $N = \{1, \ldots, 21\}$, $q = Q = .9$ and $\pi = .5$. Then the sincere voting rule is informative, and we calculate:

i. $p = q^3 + 3q^2(1 - q) = .972$
ii. $r = (1 - q)q^4 + q[4q^2(1 - q)^2 + q^4 + 4q^3(1 - q)] = .9472$
iii. $\hat{r} = .9502$ (from relationship 12).

Once more, i–iii imply the Jury Theorem goes through assuming that all individuals vote sincerely (Ladha 1992). But by Theorem 3, sincere voting here is not rational.

SUMMARY AND IMPLICATION

The Condorcet Jury Theorem and its extensions assert that under certain conditions, the probability that a collective chooses the correct alternative by majority vote exceeds the probability that any constituent member of the collective would unilaterally choose that alternative. Implicit in these conditions is the assumption that individuals behave in the collective decision exactly as they would if choosing alone (and that such voting is informative). The intuition that an assumption of "sincere" voting is innocuous here turns out to be faulty. Although there is certainly no incentive or opportunity for individual gain at the expense of others, it does not follow that rational individuals behave identically in collective and in autarkic decision-making environments.

We have looked at the role of the "sincerity" assumption in the Jury Theorems with three variations of an extremely simple model, each of which differs from the others only in the specification of individuals' information. And in each case, the model is set up so that the features of the Jury Theorems typically taken as primitive (e.g., individuals' probability assessments on which alternative is best, the correlation between such assessments) are generated both endogenously within the model and consistent with the parametric restrictions imposed by the Jury Theorems per se. In only one circumstance is it the case that sincere voting is informative and rational. In model I, where individuals' information consists of a single independent and private draw from a state-dependent distribution, the Condorcet Jury Theorem obtains when and only when majority voting is the optimal way to aggregate individuals' private information. Moreover, if majority voting is not the optimal way to aggregate information, then sincere voting under majority rule cannot be rational. In particular, we provide an example in which the explicit assumptions of the Jury Theorem hold yet sincere voting is not rational. Indeed, when all other individuals are voting sincerely, any one individual has an incentive to vote *against* the advice of her private information. Within model I, therefore, the Condorcet Jury Theorem as usually formulated is either trivial or necessarily precludes Nash equilibrium behavior.

The situation is more stark in model II, where individuals' information comprises two independent draws from a state-dependent distribution, and model III, where individuals' model II information is augmented by a public signal from some (possibly) distinct state-dependent distribution. In both these environments we prove that sincere voting cannot be both informative and rational. Moreover, in both models II and III, all individuals voting for the same alternative irrespective of their information is an undominated Nash equilibrium. With respect to the probability of selecting the correct alternative, that is, majority voting can easily do worse than any individual acting alone (and this is true even when the parametric conditions for the Jury Theorem to hold are satisfied).

Two immediate implications follow from these results. The first is that ignoring the sources of information that support individuals' beliefs precludes analysis of individually rational behavior. And if the objective of the analysis is to understand how people behave under specific institutional constraints, this is clearly undesirable. Second, the appropriateness of majority rule (or, for that matter, any voting rule) in generating "good" collective outcomes will depend on the details of the situation of concern. For example, as was illustrated in model I, the identity of the optimal voting rule hinged critically on parameters governing individuals' information.

More generally, our results reveal the importance of addressing issues in collective decision making from a game-theoretic perspective. In particular, the appropriate approach to problems of information aggregation is through game theory and mechanism design, not statistics.

APPENDIX

Here we describe in detail the Bayesian game played by the individuals in models I and II, as well as the Nash equilibrium conditions; the description for model III is similar. Each individual has a set of *types* T_i, where in model I this set is equal to $\{0, 1\}$ whereas in model II it is $\{0, 1\}^2$; let

$$T = \times_{i \in N} T_i \text{ and } T_{-i} = \times_{j \neq i} T_j.$$

A *type profile* $t = (t_1, \ldots, t_n) \in T$ is drawn according to the prior $p(t)$ over T, where this prior is given by $p(t) = \pi \Pr[t|A] + (1 - \pi)\Pr[t|B]$, and where $\Pr[t|A]$ and $\Pr[t|B]$ are determined by relationship 2 and the presumed independence of the individual draws.

Let $\{A, B\} = D_i$ be the set of *decisions* available to individual i and let

$$D = \times_{i \in N} D_i.$$

Prior to making her decision, i observes her component t_i of the type profile t. Thus, a *strategy* for i is a function v_i: $T_i \to D_i$, where we let v: $T \to D$ denote a *strategy profile*; throughout, $v(t_{-i})$ will denote the vector of decisions by all individuals except i according to the strategy profile $v(\cdot)$.

From relationship 1 we can define preferences over vectors of decisions, d, and type profiles, t, in the following manner (for simplicity, we assume majority rule): Let $M(d) = A$ if $|\{i \in N: d_i = A\}| > n/2$ and $M(d) = B$ otherwise (recall n is assumed odd). Then we have

$$u(d, t) = \Pr[M(d)|t] = \frac{\Pr[M(d) \,\&\, t]}{p(t)}. \quad \textbf{(A-1)}$$

That is, the expected utility from (say) a majority selecting outcome A given the vector of types t is equal to the probability that A is the true state conditional on t, which by Bayes' Rule is given by the probability the state is A *and* the type profile t is observed divided by the probability that t is observed.

Upon observing t_i, individual i updates her belief about others' types in a Bayesian fashion:

$$p(t_{-i}|t_i) = \sum_{T_{-i}} \frac{p(t)}{p(t_i, t_{-i})}. \quad \textbf{(A-2)}$$

The expected utility from voting for $d_i \in D_i$, given the updated belief $p(t_{-i}|t_i)$ and the strategies of the others $v(t_{-i})$ is given by

$$EU(d_i; t_i, v) = \sum_{T_{-i}} p(t_{-i}|t_i)u(d_i, v(t_{-i}), t_i, t_{-i}).$$

DEFINITION. *A* (Bayesian) Nash equilibrium *of the above Bayesian game is a strategy profile* $v^*(\cdot)$ *such that for all* $i \in N$ *and all* $t_i \in T_i$, $v_i^*(t_i) = A$ *only if* $EU(A; t_i, v^*) - EU(B; t_i, v^*) \geq 0$.

It turns out that we can simplify significantly the inequality found in this definition. By relationships A-1 and A-2, we know that

$$EU(d_i; t_i, v) = \frac{1}{\sum_{T_{-i}} p(t_{-i}, t_i)} \left\{ \sum_{T_{-i}} \Pr[M(d_i, v(t_{-i})) \text{ and } t] \right\}. \quad \textbf{(A-3)}$$

More important, because only the *sign of the difference* in the expected utility in voting for A and B is relevant, we can ignore the denominator in relationship A-3 and write

$$EU(A; t_i, v) - EU(B; t_i, v) \propto \sum_{T_{-i}} \{\Pr[M(A, v(t_{-i})) \text{ and } t] - \Pr[M(B, v(t_{-i})) \text{ and } t]\} \quad \textbf{(A-4)}$$

(where $\propto$ means "is proportional to"). Now define

$$T_{-i}^{P}(v) = \{t_{-i} \in T_{-i} | \, |\{ j \in N \backslash \{i\}: v_j(t_j) = A\}| = (n - 1)/2\}.$$

$T_{-i}^{P}(v)$ thus gives those type subprofiles where, according to the strategy v, exactly half of the remaining individuals (i.e., all but i) vote for A. Thus, when $t_{-i} \in T_{-i}^{P}(v)$, voter i will be pivotal, so that for these subprofiles $M(d_i, v(t_{-i})) = d_i$, whereas for $t_{-i} \notin T_{-i}^{P}(v)$ i is not pivotal, so that the bracketed term in the summation on the right-hand side of relationship A-4 will be zero. Therefore, we can replace relationship A-4 with

$$\begin{aligned} EU(A; t_i, v) - EU(B; t_i, v) &\propto \sum_{T_{-i}^{P}(v)} \{\Pr[A \text{ and } t] - \Pr[B \text{ and } t]\} \\ &= \sum_{T_{-i}^{P}(v)} \{\pi \Pr[t|A] - (1 - \pi)\Pr[t|B]\}. \quad \textbf{(A-5)} \end{aligned}$$

It is relationship A-5, then, which is employed to determine when sincere voting does or does not constitute a Nash equilibrium. Further, the "equilibrium" restriction to only those type subprofiles in $T_{-i}^{P}(v)$ is the source of the additional information alluded to in the text.

REFERENCES

Austen-Smith, David. 1990. "Information Transmission in Debate." *American Journal of Political Science* 34:124–52.

Berg, Sven. 1993. "Condorcet's Jury Theorem, Dependency among Jurors." *Social Choice and Welfare* 10:87–96.

Black, Duncan. 1958. *The Theory of Committees and Elections*. Cambridge: Cambridge University Press.

Condorcet, Marquis de. [1785] 1994. *Essai sur l'application de l'analyse a la probabilité des decisions rendues a la pluralité des voix*. Trans. Iain McLean and Fiona Hewitt. Paris.

Feddersen, Tim, and Wolfgang Pesendorfer. N.d. "The Swing Voter's Curse." *American Economic Review*. Forthcoming.

Feddersen, Tim, and Wolfgang Pesendorfer. 1995. "Voting Behavior and Information Aggregation in Elections with Private Information." Northwestern University. Typescript.

Grofman, Bernard, and Scott Feld. 1988. "Rousseau's General Will: A Condorcetian Perspective." *American Political Science Review* 82: 567–76.

Ladha, Krish. 1992. "The Condorcet Jury Theorem, Free Speech, and Correlated Votes." *American Journal of Political Science* 36:617–34.

Ladha, Krish. 1993. "Condorcet's Jury Theorem in the Light of de Finetti's Theorem: Majority Rule with Correlated Votes." *Social Choice and Welfare* 10:69–86.

McLean, Iain, and Fiona Hewitt. 1994. *Condorcet: Foundations of Social Choice and Political Theory*. Brookfield, VT: Elgar.

Miller, Nicholas. 1986. "Information, Electorates, and Democracy: Some Extensions and Interpretations of the Condorcet Jury Theorem." In *Information Pooling and Group Decision Making*, ed. B. Grofman and G. Owen. Greenwich, CT: JAI Press.

Myerson, Roger. 1994. *Extended Poisson Games and the Condorcet Jury Theorem*. Discussion Paper 1103, Center for Mathematical Studies in Economics and Management Science, Northwestern University.

Nitzan, Shmuel, and Jacob Paroush. 1985. *Collective Decision Making*. Cambridge: Cambridge University Press.

Paroush, Jacob. 1994. "Stay Away from Fair Coins; or, The Correct Condorcet Jury Theorem." Bar-Ilan University. Typescript.

Young, Peyton. 1988. "Condorcet's Theory of Voting." *American Political Science Review* 82:1231–44.