Consequences of the Condorcet Jury Theorem for Beneficial Information Aggregation by Rational Agents — Source link

Andrew McLennan

Institutions: University of Minnesota

Published on: 01 Jun 1998 - American Political Science Review (Cambridge University Press)

Topics: Condorcet's jury theorem, Social choice theory, Condorcet method, Population and Nash equilibrium

Consequences of the Condorcet Jury Theorem for Beneficial Information Aggregation by Rational Agents

by

Andrew McLennan

Discussion Paper No. 289, October 1996

Center for Economic Research
Department of Economics
University of Minnesota
Minneapolis, MN 55455

Consequences of the Condorcet Jury Theorem for Beneficial Information Aggregation by Rational Agents †

by

Andrew McLennan

Department of Economics
University of Minnesota
271 19th Avenue South
Minneapolis, MN 55455

mclennan@atlas.socsci.umn.edu

October 22, 1996

Abstract

The Condorcet Jury Theorem pertains to elections in which the agents have common preferences but diverse information. We show that, whenever "sincere" voting leads to the conclusions of the Theorem—decisions superior to those that would be made by any individual based on private information, and asymptotically correct decisions as the population becomes large—there are also Nash equilibria with these properties, and in symmetric environments the equilibria may be taken to be symmetric. These conclusions follow from a simple property of common interest games: a mixed strategy profile of a (symmetric) common interest game that is optimal in the set of (symmetric) mixed strategy profiles is a Nash equilibrium.

† I would like to thank Richard McKelvey for bringing literature on the Condorcet Jury Theorem to my attention, and for very helpful and encouraging conversations at the inception of this project. I would also like to thank Timothy Fedderson for insightful comments.

Consequences of the Condorcet Jury Theorem for Beneficial Information Aggregation by Rational Agents

by

Andrew McLennan

1. Introduction

In addition to his contributions to the theory of elections in which the various agents have different preferences, Condorcet ([1785] 1994) established a result, known as the Condorcet Jury Theorem, giving certain conditions under which majority rule in two candidates elections with *sincere voting*, in the sense of voting for the candidate that seems best based on one's own information, yields better decisions, for a group of individuals with identical preferences but diverse information, than would be made by any one of the individuals acting on her own information. Recent interpretations, variations, and extensions of this result include Miller (1986), Grofman and Feld (1988), Young (1988), Ladha (1992), Ladha (1993), and Berg (1993). In these models it is generally the case that the probability of making the (full information) optimal choice converges to one as the number of voters goes to infinity.

Austen-Smith and Banks (1996) point out that the pertinence, for rational agents, of this style of result depends, or at least seems to depend, on an assumption (that had rarely been discussed explicitly) that sincere voting constitutes a Nash equilibrium of the game induced by majority rule. They demonstrate that, in fact, this can easily fail to be the case.

Thus there arises the question of the extent to which the conclusions of the Condorcet Jury Theorem continue to hold when the game induced by the voting procedure is played rationally in the sense that the profile of strategies is a Nash equilibrium. Myerson (1994) and Wit (1996) give models in which some of the symmetric equilibria have the desired properties: equilibrium decisions are superior to any individual's (benevolent but only privately informed) dictatorship, and the probability of an incorrect decision goes to zero as the number of agents becomes large. Ladha, Miller, and Oppenheimer (1996) give examples of asymmetric equilibria that outperform the symmetric equilibria, and therefore also yield the conclusions of interest. The papers by Fedderson and Pesendorfer (1996a, 1996b, 1996c) emphasize information aggregation by voters with common, or similar, preferences, but in settings in which there are other voters with different values. All three papers display equilibria in which (asymptotically, as the population becomes large) the outcomes are those that would be chosen by a fully informed electorate.

We demonstrate that this phenomenon is very general. By assumption all individuals have the same preferences, so the game induced by any voting procedure is a game of common interest: there is a single common utility function. Such games occur in the economic theory of teams, and the key observation applied here has been stated, in the language of that theory, as follows:

> We shall call a team decision function *person-by-person-satisfactory* (pbps) if it cannot be improved—i.e., its expected payoff cannot be increased—by changing any single member's decision function, δ_i. It is clear that an optimal team decision function is necessarily person-by-person-satisfactory. However, the converse is not in general true ... (Radner (1972, p. 195))

That is, while there may be Nash equilibria of a game of common interest that are not optimal, an optimal mixed strategy profile is always a Nash equilibrium, and gives decisions that are at least as good as "sincere" voting, or any other mode of behavior. Thus, whenever sincere voting is a better aggregator than any individual's dictatorship, an optimal strategy profile is both better still and an equilibrium. Optimal strategy profiles yield asymptotically perfect decisions when sincere voting has this property. Theorem 1 asserts that the set of such "optimal equilibria" contains stable sets of equilibria (Kohlberg and Mertens 1986) so that these equilibria cannot be excluded using the various refinements of Nash equilibrium (e.g., van Damme 1987).

It may happen that optimal equilibria succeed in aggregating information when sincere voting would not. For example, suppose that there are two possible states of nature, L and R, two possible signals, also denoted by L and R, and two possible social alternatives, again denoted by L and R. Suppose that the two states have equal prior probability, the probability that an agent receives signal L when the true state is L is $2/3$, the probability that an agent receives signal R when the true state is R is $2/3$, and, conditional on the state, the various agents' signals are statistically independent. Suppose L is the preferred action when L is the state while R is the preferred action when R is the true state. If it is disastrous to choose R when L is the true state, but only mildly disadvantageous to choose L when the true state is R, then sincere voters will always vote for L. If the population is large, however, expected payoffs will be higher if all voters vote their signal, since the law of large numbers implies that the best alternative will be chosen with very high probability. Such behavior will not, in general, be an equilibrium (Austen-Smith and Banks (1996, Lemma 2, p. 38)) but optimal equilibria will be even more successful in aggregating information.

When the environment is symmetric, in that the agents are interchangeable, equilibria assigning different strategies to identical agents embody a degree of coordination that may seem implausible when the population is large. Myerson (1993) has argued that population uncertainty, in that the number of players of each type is uncertain, is best modelled by a framework in which asymmetric behavior is very difficult to describe. Theorem 2 establishes that a symmetric mixed strategy profile that is optimal in the set of such profiles is a Nash equilibrium. Since sincere voting is a symmetric profile of strategies, when it yields better decisions than any individual's dictatorship there is a symmetric equilibrium that does at least as well.

Thus the relevance of the Condorcet Jury to rational agents is revived: a "sincere" statistical version of the result implies a "rational" strategic version. However, the test that majority rule passes in this way is seen to be quite weak, in that it is only required that there be *some* mode of behavior that leads to decisions that are superior to any individual's dictatorship, or asymptotically perfect.

In the next section we introduce basic notation, which is essentially all that needs to be done to make our main point precise. Section 3 analyzes the optimal (mixed) profiles of voting strategies from the point of view of stability. The final section demonstrates the existence of suitable symmetric equilibria in symmetric environments.

2. Optimal Profiles of Voting Strategies

Fix a finite set of *alternatives* A, a set of *agents* $I = \{1, \ldots, n\}$, spaces of *types* $T_1, \ldots, T_n$, a probability distribution P on $\mathbf{T} := T_1 \times \ldots \times T_n$, and a function $u : \mathbf{T} \times A \to I\!R$ which is interpreted as the common von Neumann-Morgenstern utility function of all agents. We assume throughout that each type of each agent has positive probability. A *voting scheme* consists of nonempty finite sets $V_1, \ldots, V_n$ of allowed votes and an *aggregation rule* $f : \mathbf{V} \to A$ where $\mathbf{V} := V_1 \times \ldots \times V_n$. The tuple

$$((T_1, \ldots, T_n), P, A, u, (V_1, \ldots, V_n), f)$$

is called a *voting environment.*

A (pure) *voting strategy* for agent i is a function $s_i : T_i \to V_i$, and a *voting profile* is an n-tuple $s = (s_1, \ldots, s_n)$ of voting strategies, one for each agent. A map $c_i : T_i \to A$ is *sincere* if

$$(\forall t_i \in T_i) \quad c_i(t_i) \in \operatorname{argmax}_{a \in A} \mathbf{E}\Big(u((t_1, \ldots, t_n), a) | t_i\Big).$$

A voting profile s is an *acceptable aggregator*, in the sense that otherwise some individual's privately informed decisions would be unanimously (weakly) preferred, if

$$\mathbf{E}\Big(u((t_1, \ldots, t_n), f(s_1(t_1), \ldots, s_n(t_n)))\Big) > \mathbf{E}\Big(u((t_1, \ldots, t_n), c_i(t_i))\Big)$$

whenever, for some i, $c_i : T_i \to A$ is sincere. We extend this terminology to profiles of mixed voting strategies (viewed here as mixtures over pure voting strategies, though behavior strategies (Kuhn 1953) would be equivalent) in the obvious way, by requiring that the induced expected utility is greater than what any agent could attain with only private information.

We now introduce notation for (normal form) common interest games for the given set of agents I. For each i let S_i be a finite set of *pure strategies*, let $\mathbf{S} := S_1 \times \ldots \times S_n$ be the set of *pure strategy profiles*, and let

$U : \mathbf{S} \to I\!R$ be a function that is interpreted as the common von Neumann-Morgenstern utility function of all agents in I. Such a framework is derived from the general voting environment above by letting $S_i = A^{T_i}$ be the set of voting rules for each agent i, and letting

$$U(s_1, \ldots, s_n) = \mathbf{E}\Big(u\big((t_1, \ldots, t_n), f(s_1(t_1), \ldots, s_n(t_n))\big)\Big).$$

The set of mixed strategies for i is denoted by $\Delta(S_i)$, with typical elements $\mu_i, \mu_i', \ldots$. The set of *mixed strategy profiles* is $\Sigma := \Delta(S_1) \times \ldots \times \Delta(S_n)$. Abusing notation, we let U denote also the mixed extension of U, i.e., the unique multilinear function $U : \Sigma \to I\!R$ that agrees with the original U when we identify the elements of $\mathbf{S}$ with the corresponding vertices of $\Delta(S_i)$.

We reiterate our main finding. Any optimal mixed strategy vector $(\mu_1^*, \ldots, \mu_n^*)$ of a common interest game is a Nash equililibrium, and if there is any acceptable aggregator, then $(\mu_1^*, \ldots, \mu_n^*)$ must also be acceptable. If, along some sequence of voting environments, the probability of a correct decision under any sequence of aggregators goes to one, then the probability of a correct decision under any optimal mixed strategy will also converge to unity. Note that, in searching for voting models for which the existence of acceptable or asymptotically correct aggregators is guaranteed, one is not restricted to models in which each agent's set of allowed votes is A.

If so inclined, the reader may proceed directly to §4 without loss of understanding.

3. Stability

The optimal strategy profiles may prescribe behavior that is complex and highly coordinated, or implausible on other grounds. However, these profiles do survive the criteria posed by the literature on refinements of Nash equilibrium, as we now explain.

If $\hat{\sigma} \in \Sigma$ and $\epsilon = (\epsilon_1, \ldots, \epsilon_n) \in (0,1)^n$, the *perturbed game* induced by $(\hat{\sigma}, \epsilon)$ is the common interest game with the original pure strategy sets and the perturbed utility function

$$U_{(\hat{\sigma},\epsilon)}(s) := U((1-\epsilon_1)s_1 + \epsilon_1\hat{\sigma}_1, \ldots, (1-\epsilon_n)s_n + \epsilon_n\hat{\sigma}_n).$$

Recall (Selten 1975) that a mixed strategy profile σ^* is a *perfect equilibrium* if there are sequences $\{\hat{\sigma}^r\}$, $\{\epsilon^r\}$, $\{\sigma^r\}$ such that: (a) for each r, σ^r is a Nash equilibrium of the perturbed game induced by $(\hat{\sigma}^r, \epsilon^r)$; (b) $\epsilon_i^r \to 0$ for each i; (c) $\sigma^r \to \sigma^*$. A compact set E of Nash equilibria is *robust with respect to trembles* if, for any neighborhood W or E, there is $\delta > 0$ such that for all $\hat{\sigma}$ and ϵ with $\epsilon_i < \delta$ for all i, W contains an equilibrium of the perturbed game induced by $(\hat{\sigma}, \epsilon)$. The arguments below use the (obvious) fact that if E is robust with respect to trembles, and F is a compact set of Nash equilibria that contains E, then F is robust with respect to trembles. If E is robust with respect to trembles and has no compact proper subset that is robust with respect to trembles, then E is *stable* (Kohlberg and Mertens 1986). Ideally one would like to find stable singletons (when this happens the unique element is said to be a *strictly perfect equilibrium*) but it is possible that there are none (e.g., §1.5 of van Damme 1987).

Let $M := \max_{(s_1,\dots,s_n)\in S} U(s_1,\dots,s_n)$ and let $K := U^{-1}(M)$ be the set of optimal mixed strategy profiles. It is interesting to note that K can be characterized as the union of the various sets $\Delta(R_1) \times \dots \times \Delta(R_n)$ where $R_1,\dots,R_n$ is a collection of subsets of $S_1,\dots,S_n$, respectively, such that $U(r_1,\dots,r_n) = M$ whenever $r_1 \in R_1,\dots,r_n \in R_n$.

Theorem 1: K is robust with respect to trembles.

Proof: For a given neighborhood W of K choose $\gamma > 0$ small enough that $K_\gamma := \{\,\sigma \in \Sigma : U(\sigma) > M - \gamma\,\}$ is contained in W. Choose $\delta > 0$ small enough that

$$U(\sigma) - \gamma/2 < U_{(\hat{\sigma},\epsilon)}(\sigma) < U(\sigma) + \gamma/2$$

for all σ, $\hat{\sigma}$, and ϵ with $\epsilon_1,\dots,\epsilon_n < \delta$. Fixing $(\hat{\sigma},\epsilon)$ with $\epsilon_1,\dots,\epsilon_n < \delta$, let σ^* be a maximizer of $U_{(\hat{\sigma},\epsilon)}(\cdot)$. Just as earlier, σ^* is clearly a Nash equilibrium of the perturbed game induced by $\hat{\sigma}$ and ϵ. We have $U_{(\hat{\sigma},\epsilon)}(\sigma) > M - \gamma/2$ whenever $\sigma \in K$, so $U_{(\hat{\sigma},\epsilon)}(\sigma^*) > M - \gamma/2$ and thus $U(\sigma^*) > M - \gamma$, whence $\sigma^* \in K_\gamma \subset W$, as desired. ∎

Corollary: K contains a stable set of equilibria.

Proof: Let $\{O_1, O_2, \dots\}$ be a countable base[1] of the topology of Σ. Set $K_0 := K$, and define $K_1, K_2, \dots$ inductively by setting $K_{j+1} := K_j \backslash O_j$ if $K_j \backslash O_j$ is robust with respect to trembles, and otherwise setting $K_{j+1} = K_j$. Then each K_j is compact and robust with respect to trembles, and simple topological arguments show that $K_\infty := \bigcap_j K_j$ is also compact and robust with respect to trembles. If K_∞ is not minimal with respect to these properties, then, for some j, $K_\infty \backslash O_j \neq K_\infty$ is robust with respect to trembles. But $K_\infty \subset K_j$, so $K_j \backslash O_j$ is robust with respect to trembles and $K_\infty \subset K_{j+1} = K_j \backslash O_j$, so that $K_\infty = K_\infty \backslash O_j$. This contradiction completes the proof. ∎

4. Symmetry

We now consider symmetric voting environments. The aggregation rule f is *anonymous* if $V_i =: V$ is the same for all i and the chosen alternative depends only on the numbers of agents casting the various possible votes. More formally, $f(v_{\sigma(1)},\dots,v_{\sigma(n)}) = (v_1,\dots,v_n)$ for all $(v_1,\dots,v_n) \in V^n$ and all permutations σ :

[1] That is, any open set is a union of members of $\{O_1, O_2, \dots\}$. For example we may take the open balls of rational radii whose centers have rational coordinates.

$\{1, \ldots, n\} \to \{1, \ldots, n\}$. The notion of symmetry relevant to our analysis requires also that the information structure and u are symmetric in the sense that $T_1 = \ldots = T_n =: T$ and

$$P(t_{\sigma(1)}, \ldots, t_{\sigma(n)}) = P(t_1, \ldots, t_n) \quad \text{and} \quad u\big((t_{\sigma(1)}, \ldots, t_{\sigma(n)}), a\big) = u\big((t_1, \ldots, t_n), a\big)$$

for all $(t_1, \ldots, t_n) \in T^n$, all permutations σ, and all $a \in A$. We say that the voting environment is *symmetric* when all these symmetry conditions are satisfied.

On the other hand the common interest game given by $S_1, \ldots, S_n$ and U is *symmetric* if $S_1 = \ldots = S_n =: S$ and

$$U(s_{\sigma(1)}, \ldots, s_{\sigma(n)}) = U(s_1, \ldots, s_n)$$

for all $(s_1, \ldots, s_n) \in S^n$ and all permutations σ. We leave to the reader the conceptually simple, but tedious, algebraic exercise of showing that a symmetric voting environment induces a symmetric common interest game. When the game given by $S_1, \ldots, S_n$ and U is symmetric we may say that a mixed strategy profile $(\mu_1, \ldots, \mu_n)$ is *symmetric* if $\mu_1 = \ldots = \mu_n$.

Theorem: *Suppose that the game given by $S_1, \ldots, S_n$ and U is symmetric. If a symmetric mixed strategy profile $(\mu, \ldots, \mu)$ is maximal for U in the set of such profiles, then it is a Nash equilibrium*[2].

The corollary, that symmetric equilibria always exist, is a (very) special case of a result of Nash (1951).

Proof: If $(\mu, \ldots, \mu)$ is not Nash, then there is a mixed strategy ν such that $U(\nu, \mu, \ldots, \mu) > U(\mu, \ldots, \mu)$. Applying multilinearity and symmetry, for all $\epsilon \in [0, 1]$ we have

$$\begin{aligned} &U((1-\epsilon)\mu + \epsilon\nu, \ldots, (1-\epsilon)\mu + \epsilon\nu) \\ &\qquad = U(\mu, \ldots, \mu) + n\epsilon\big(U(\nu, \mu, \ldots, \mu) - U(\mu, \ldots, \mu)\big) + \epsilon^2 P(\epsilon), \end{aligned}$$

where P is a polynomial function of ϵ. Therefore

$$U((1-\epsilon)\mu + \epsilon\nu, \ldots, (1-\epsilon)\mu + \epsilon\nu) > U(\mu, \ldots, \mu)$$

for sufficiently small $\epsilon > 0$, contrary to hypothesis. ∎

Therefore, to produce an equilibrium symmetric (mixed) acceptable aggregator, it suffices to prove the existence of a (possibly nonequilibrium) symmetric acceptable aggregator. For this condition to make sense the model must give all agents the same set of allowed votes, but again this set need not be A.

[2] Although I know of no precedent, at this date at least (October 22, 1996) I would be very unsurprised if this fact had already been pointed out. Anyone who knows of a predecessor is urged to contact me.

5. Concluding Remarks

Certainly our framework is very general in comparison with most models considered in related literature. Additional generality can be attained easily by allowing voting schemes that yield nontrivial lotteries over the set of alternatives; this has no effect on our arguments. In particular, by identifying a probability that an agent does not vote with a probability that her vote does not count, one may represent situations in which the population of voters is random. From a mathematical point of view the main losses of generality in our framework are those associated with the finiteness of the sets of agents, types, and alternatives. We hope that mathematically sophisticated readers will agree that our argument is not critically dependent on these assumptions, at least in the sense that only minor and mundane technical details arise in the extensions to several sufficiently well behaved infinite dimensional settings.

References

Austen-Smith, David. 1990. "Information Transmission in Debate," *American Journal of Political Science* 34 (1) February: 124-52.

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

Berg, Sven. 1993. "Condorcet's Jury Theorem, Dependency among Jurors." *Social Choice and Welfare* 10 (1) January: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 à la probabilité des decisions rendues à la pluralité des voix.* Trans. Iain McLean and Fiona Hewitt. Paris.

Feddersen, Tim, and Wolfgang Pesendorfer. 1996a. "The Swing Voter's Curse." *American Economic Review* 86 (3) June: 408-24.

Feddersen, Tim, and Wolfgang Pesendorfer. 1996b. "Voting Behavior and Information Aggregation in Elections with Private Information." Northwestern University. September: Typescript.

Feddersen, Tim, and Wolfgang Pesendorfer. 1996c. "Abstention and Common Values." Northwestern University. September: Typescript.

Grofman, Bernard, and Scott Feld. 1988. "Rousseau's General Will: A Condorcetian Perspective." *American Political Science Review* 82 (2) June: 567-76.

Kohlberg, Elon, and Jean-Francois Mertens. 1986. "On the Strategic Stability of Equilibria." *Econometrica* 54 (5) September: 1003-1038.

Kuhn, Harold. 1953. "Extensive Games and the Problem of Information." In *Contributions To the Theory of Games*, Vol. 2, ed. by H. Kuhn and A. Tucker. Princeton: Princeton University Press.

Ladha, Krish. 1992. "The Condorcet Jury Theorem, Free Speech, and Correlated Votes." *American Journal of Political Science* 36 (3) August: 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 (1) January: 69-86.

Ladha, Krish, Gary Miller, and Joe Oppenheimer. 1996. "Democracy" Turbo-charged or Shackled? Information Aggregation by Majority Rule," *mimeo*, Washington University, St. Louis.

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 Theory." In *Information Pooling and Group Decision Making*, ed. B. Grofman and G. Owen. Greenwich, CT: JAI Press.

Nash, John. 1951. "Non-cooperative games." *Annals of Mathematics* 54 (2) September: 286-295.

Radner, Roy. 1972. "Teams." In *Decision and Organization: A Volume in Honor of Jacob Marschak*, ed. C. B. McGuire and R. Radner. Amsterdam: North Holland.

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

Selten, Reinhard. 1975. "Reexamination of the Perfectness Concept for Equilibrium Points in Extensive Games," *International Journal of Game Theory* 4 Issue 1: 25-55.

van Damme, Eric. 1987. *Stability and Perfection of Nash Equilibria.* Springer-Verlag.

Young, Peyton, 1988. "Condorcet's Theory of Voting." *American Political Science Review* 82 (4) December: 1231-44.