Journal of Economic Perspectives—Volume 9, Number 1—Winter 1995—Pages 51–64

Optimal Voting Rules

Peyton Young

Democracy, E. B. White (1946) wrote, "is the recurrent suspicion that more than half the people are right more than half the time." The suspicion has been around for a long time. It is, in fact, the central thesis of a remarkable work published in 1785 by the French mathematician and political philosopher, Jean Antoine Nicholas Caritat, Marquis de Condorcet. Condorcet set out to prove that majority rule is not only a fair way to make political choices, it is actually the best way to do so—the way most likely to yield "optimal" results. Though this notion may at first sound strange to modern ears, it turns out to be a surprisingly fruitful way of thinking about the design of voting rules.

Condorcet begins with the premise that the object of government is to make decisions that are in the best interest of society. This leads naturally to the question: what voting rules are most likely to yield good outcomes? To analyze this problem he applied the then novel science of probability theory. Imagine that a group of voters must decide between two alternatives, one of which is objectively best. (Whether this is a meaningful assumption will be considered below.) Each individual makes a judgment about this question and registers an opinion. Sometimes, of course, people judge incorrectly. But let us assume —perhaps too optimistically—that each voter is more likely to make the right choice than the wrong choice in any given situation. Condorcet showed that, if the voters make their choices independently, then the laws of probability imply that the choice with the most votes is the one most likely to be correct. In other

■ *Peyton Young is Professor of Economics, Johns Hopkins University, Baltimore, Maryland.*

words, majority rule is a statistically optimal method for pooling individual judgments about a question of fact.

Consider the following example: 100 individuals are choosing between two alternatives, *a* and *b*. Let the outcome be 55 votes for *a* and 45 votes for *b*. Assume that each of these individuals is right 60 percent of the time. Now there are two possibilities: either *a* is really the best choice or *b* is the best choice. In the first case the observed voting pattern (55 for *a*, 45 for *b*) would occur with probability $(100!/55!45!)\ .6^{55}\ .4^{45}$, while in the second case it would occur with probability $(100!/45!55!)\ .6^{45}\ .4^{55}$. As the former is about 58 times more probable than the latter, we conclude that *a* is more likely to be correct than *b*. This style of reasoning is known in statistics as maximum likelihood estimation. The general idea is to estimate how likely an observed event would be in different (unobservable) states of the world. From this we infer that the most likely state is the one that would have produced the observation with highest probability.[1]

Why should we buy the idea, though, that there really is such a thing as an objectively "best" choice? Aren't values relative, and isn't the point of voting to strike a balance between conflicting opinions, not to determine a correct one? This relativistic point of view, which has reigned supreme in economic and social choice theory for most of this century, seems completely at odds with Condorcet's perspective. Yet there are many situations where Condorcet's premises make good sense. Consider a trial by jury. Let *a* be the proposition that the defendant is guilty, *b* that she is innocent. If all twelve members of the jury vote for conviction and if the probability that each is correct is 0.60, then the probability that the defendant is in fact innocent is less than one in 50,000 (assuming equal prior probability of guilt or innocence). This "unanimity rule" is called for if a false conviction is deemed much worse than letting a guilty party go free.

But if the objective is simply to reach the correct decision with highest probability, then clearly this is not the best one can do. The probability of making the correct choice under unanimity rule is only slightly greater than one-half (about .501). But under simple majority rule the probability of a correct choice is about .665. More generally, it can be shown that, *among all group decision rules on two alternatives (one of which is in fact correct), simple majority rule is most likely to identify the correct outcome* (Nitzan and Paroush, 1982; Shapley and Grofman, 1984). Furthermore, as the size of group becomes larger, the

[1]This approach is similar to Bayesian inference with uniform priors, though formally it does not rely on the notion of priors at all. A Bayesian analysis of the situation would proceed as follows. Suppose that *a* and *b* are equally likely to be best a priori. Let *p* be the conditional probability that the vote would occur given that *a* is best, and let *q* be the conditional probability that it would occur given that *b* is best. Then the posterior probability that *a* is best is $p/(p + q) = .983$, while the posterior probability that *b* is best is $q/(p + q) = .017$. Thus their likelihood ratios are $.983/.017 = 58$, so under either method we would choose *a* rather than *b*.

probability that the majority decision is correct approaches unity—a result first proved by Condorcet.

This result reaches far beyond jury trials. It applies to any choice problem in which people agree about the objective, but disagree about the best means to achieve that objective. For example, the members of the Federal Reserve Board may all want to maximize the long-run rate of economic growth, but at any given moment they may be uncertain whether lowering or raising interest rates is the best way to achieve this. Similarly, the directors of a corporation may agree that their task is to maximize the firm's long-run profitability, but they may have different views about which candidate for chief executive is most likely to realize this goal. In these situations Condorcet's premises make perfect sense, and simple majority rule is the best way to estimate the optimal choice. Moreover, this is how such decisions are usually made in practice.[2]

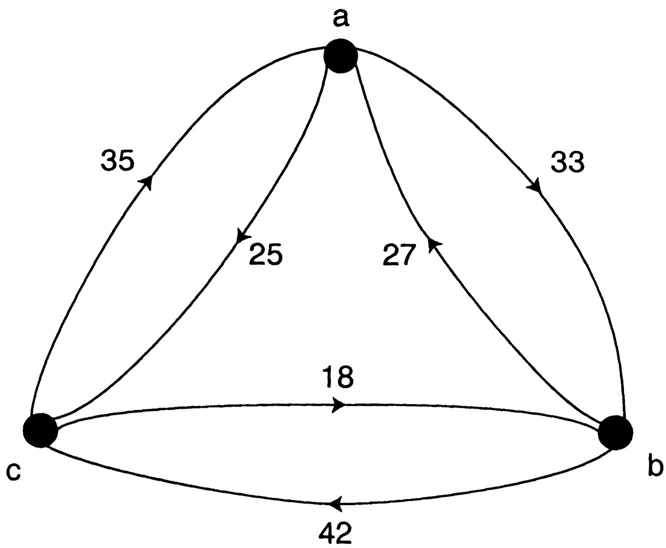
Extension to Three or More Alternatives

When we try to extend this reasoning to three or more alternatives, however, matters become more complicated. Ideally we would like to choose the alternative that has a simple majority over every other. (Such an alternative is called a "majority" or "Condorcet" alternative.) As Condorcet was the first to show, such an alternative may not exist. Consider the following example from Condorcet's *Essay*. There are three policy alternatives and 60 voters: 23 voters rank the policies *abc*, 17 choose *bca*, 2 *bac*, 10 *cab*, and 8 *cba*. Here majority rule comes up empty-handed because it leads to a voting cycle: *a* beats *b* by 33 to 27, *b* beats *c* by 42 to 18, and *c* beats *a* by 35 to 25.

Condorcet set himself the problem of determining the optimal decision rule under these circumstances, that is, the procedure most likely to identify the correct ranking of the policies—assuming there is such a correct ranking and that the expressed differences in opinion are just differences in the voters' information or judgment, not in their objectives. Suppose, for example, that the common objective is to reduce crime, and the three proposed policies are a) hire more police, b) increase prison sentences, and c) offer training programs for ex-convicts. Voters may agree on the objective, yet differ in their judgment about which of these policies is in fact more likely to reduce crime per dollar spent.

[2] If some voters are a priori more likely than others to make correct judgments, then it makes sense to give their votes greater weight. But it is not necessarily a good idea to give the most competent voters all the weight. The reason is that there is statistical value in having a large number of independent opinions on a question, even though some of these opinions are more likely to be correct than others. If p_i is the probability that voter i is correct in any given judgment, and if the judgments are independent, then the maximum likelihood rule for two alternatives is to use weighted majority rule, where the weight on individual i's vote is $\log p_i/(1 - p_i)$. See Nitzan and Paroush (1982), and Shapley and Grofman (1984). Related work may be found in Poisson (1837), Urken and Traflet (1974), Grofman, Owen, and Feld (1983), and Grofman and Feld (1988).

Figure 1
A Vote Graph



A convenient way to visualize the problem is to draw a vote graph like that shown in Figure 1. There is one vertex for each alternative, and between every pair of vertices there are two edges, one in each direction. The edge going in a particular direction tells the number of votes that the alternative at its base gets over the alternative at its tip. Thus, the two edges between a and b show that a has 33 votes over b, and b has 27 votes over a.

To evaluate the probability of a ranking such as *abc*, consider the three pairwise propositions; a is better than b, b is better than c, and a is better than c. These correspond to the three directed edges $a \rightarrow b$, $b \rightarrow c$, and $a \rightarrow c$, which have weights 33, 42, and 25 respectively. Thus the total *pairwise support* for the ranking *abc* is $33 + 42 + 25 = 100$. The support for the other rankings is found in like fashion, and the results are shown below:

abc	100	*bca*	104
acb	76	*cab*	86
bac	94	*cba*	80

Condorcet proved that, when there are three alternatives, the ranking that is most likely to be correct is the one that has maximum pairwise support. Thus in the example the answer is *bca*. This is known as *Condorcet's rule of three*.

To see why this rule identifies the most likely ranking, let us compute the likelihood ratios explicitly. We begin by assuming that, given any two of the policies to compare, each voter has a fixed probability $p > 1/2$ of choosing the best one. Assume further that each voter's judgment about any pair is independent of the other voters' judgments and that his judgment about a given pair is independent of his judgment about the other pairs. (In particular, his judgments may not form a transitive order.) While one can quibble with the realism of these assumptions, they serve to simplify the calculations. The method of analysis is quite general.

For each possible ranking of the alternatives, we want to compute the conditional probability that the above voting pattern would occur given that the ranking is correct. Suppose, for instance, that the correct ranking is *abc*. The probability of the above vote is the product of three terms: the probability that *a* gets 33 votes over *b*, the probability that *b* gets 42 votes over *c*, and the probability that *a* gets 25 votes over *c*. Hence the probability of the vote in this example is proportional to $p^{33}(1-p)^{27} \times p^{42}(1-p)^{18} \times p^{25}(1-p)^{35} = p^{100}(1-p)^{80}$. This gives the relative likelihood that *abc* is correct. The likelihoods of the other five orderings are calculated similarly. In each case, the exponent of p is the total pairwise support for the ranking. Since $p > 1/2$, the maximum likelihood ranking is the one with the greatest total support.

When there are more than three alternatives, Condorcet seems to have become confused; at any rate he did not get the correct answer.[3] Nevertheless the solution can be found by a straightforward extension of the previous argument. Suppose that there are n voters and m alternatives. Given a voting outcome and a ranking R of the alternatives, the conditional probability of observing the vote, given that the true ranking is R, is proportional to $p^{s(R)}(1-p)^{M-s(R)}$, where $M = nm(m-1)/2$ and $s(R)$ is the total pairwise support for R. Hence R has maximum likelihood if and only if it has maximum support. To compute a ranking with maximum support, it suffices to find the maximum-weight set of edges in the vote graph that does not contain a cycle. This is the solution to Condorcet's problem (Young, 1986, 1988).[4]

Borda's Rule and Condorcet's Response

At this point we need to pick up a second strand in our story that actually begins somewhat earlier. In 1770, 15 years before Condorcet published his work on voting theory, his colleague Jean-Charles de Borda read a paper on the design of voting procedures to the French Academy of Sciences. Like Condorcet, Borda was a prominent figure in scientific circles, with interests that

[3]For a discussion of Condorcet's somewhat obscure argument in this case, see Young (1988) and Crepel (1990).

[4]The maximum likelihood rule can be formulated as an integer programming problem. Define a variable x_{ij} for each directed edge $i \rightarrow j$ in the vote graph, that is, one variable for each ordered pair of alternatives. Let w_{ij} be the weight on edge $i \rightarrow j$, that is, the number of votes for alternative i over alternative j. A maximum likelihood ranking corresponds to a solution x that maximizes $\Sigma w_{ij} x_{ij}$ subject to $x_{ij} + x_{ji} = 1$, $x_{ij} + x_{jk} + x_{ki}^{2} \leq 2$, and all $x_{ij} = 0$ or 1. If the voters have different competencies, individual i's vote is weighted by $\log(p_i/(1-p_i))$, where p_i is the probability that i is correct and $1/2 < p_i < 1$ (Young, 1986).

spanned a wide variety of subjects. Unlike Condorcet, he had a strong practical bent; along with voting, he did important research in mechanics, hydraulics, and optics, and was a leader in developing the metric system. This put him in the applied faction of the Academy, which was often at loggerheads with purists like Condorcet. Borda began by observing that, when there are three or more alternatives, the one that achieves the most first-place votes is not necessarily the one that has the highest standing overall. As an example, consider the following situation of 21 voters ranking three alternative candidates: 7 choose *bca*, 7 *acb*, 6 *cba*, and 1 *abc*. Under the conventional plurality method, *a* receives 8 votes, *b* 7, and *c* 6. But this fails to take into account that all but one of the people who prefer *a* like *c* better than *b*, and everyone who prefers *b* likes *c* better than *a*. One could therefore argue that *c* is the most natural compromise candidate *even though it receives the fewest first-place votes*.

To assess the true strength of the various candidates, said Borda, one must look at their overall standing in the individual rankings. This led him to propose the following rule. Let each voter rank the candidates. (For simplicity of exposition, assume there are no ties.) In each voter's list, assign a score of zero to the alternative ranked last, a score of 1 to the alternative ranked next to last, a score of 2 to the alternative next above that, and so forth. The "Borda score" of an alternative is its total score summed over all voter lists, and "Borda's rule" is to rank the alternatives from highest to lowest Borda score. In the above example the scores are 26 for *c*, 21 for *b*, and 16 for *a*. Thus, according to Borda's rule, the proper ordering of the alternatives is *cba*, which is exactly the opposite of the one implied by the number of first-place votes.[5] (Condorcet's rule yields the same result in this case, though in other examples they differ, as we shall soon see.)

Borda's paper was not published until 1784, one year before Condorcet's treatise on voting appeared. Condorcet took strong exception to Borda's method on the merits, and then added a certain amount of personal venom to the attack.[6] To illustrate his objections to Borda's method, Condorcet introduced the following example. Table 1 shows a situation with three candidates named Peter, Paul, and Jack, and 81 voters with the given preferences.

[5]Although it may not be obvious at first, Borda's rule (like Condorcet's) actually depends only on the pairwise votes between the various alternatives. This point is explained in the overview to this symposium by Levin and Nalebuff.

[6]Throughout the *Essay*, Condorcet refers sarcastically to the "method of a famous mathematician" but fails to mention him by name. The implication is that the method is not worthy of a mathematician. Condorcet's contemptuous view of Borda is borne out in his private correspondence. For example, in a letter to Turgot he wrote (Henry, 1883): "[M. Malesherbes] makes a great case for Borda, not because of his memoirs, some of which suggest talent (although nothing will ever come of them, and no one has ever spoken of them or ever will) but because he is what one calls a good academician, that is to say, because he speaks in meetings of the Academy and asks for nothing better than to waste his time doing prospectuses, examining machines, etc., and above all because, feeling eclipsed by other mathematicians, he, like d'Arcy, has abandoned mathematics for petty physics."

Table 1
Condorcet's Counterexample to Borda

30	*1*	*29*	*10*	*10*	*1*
Peter	Peter	Paul	Paul	Jack	Jack
Paul	Jack	Peter	Jack	Peter	Paul
Jack	Paul	Jack	Peter	Paul	Peter

According to Borda's rule (as the reader should be able to verify) the proper ordering is Paul, Peter, Jack. But this is absurd, said Condorcet, because Peter obtains a simple majority over both Paul and Jack. Surely this means that Peter should be ranked first. More generally, Condorcet formulated the "majority principle," which states that if there exists a majority alternative —one that obtains a simple majority over every other alternative in pairwise competitions—then it should be ranked first.

Condorcet (1788) explained why Borda's rule gives the "wrong" result in this case.[7]

> [H]ow is it that Paul is not the clear winner when the only difference between himself and Peter is that Peter got 31 first places and 39 second, while Paul got 39 first and 31 second? Well, out of the 39 voters who put Peter second, 10 preferred him to Paul, whereas only one of the 31 voters who put Paul second preferred him to Peter. The points method [of Borda] confuses votes comparing Peter and Paul with those comparing either Peter or Paul to Jack and uses them to judge the relative merits of Peter and Paul. As long as it relies on irrelevant factors to form its judgments, it is bound to lead to error, and that is the real reason why this method is defective for a great many voting patterns, regardless of the particular values assigned to each place. The conventional method [plurality] is flawed because it ignores elements which should be taken into account and the new one [Borda's] because it takes into account elements which should be ignored.

In other words, Condorcet is saying that the comparison between Peter and Paul should depend only on the relative ordering of these two candidates in the voters' lists, not on their relation to other candidates. In modern theory, starting with Arrow's work, this principle has been called "independence of irrelevant alternatives."

[7]The translation is by Sommerlad and McLean (1989), who were the first to call attention to the importance of this passage and its connection with independence of irrelevant alternatives.

Condorcet noted the remarkable fact that *any* scoring system leads to the same outcome as Borda's rule in this example, and is therefore subject to the same criticism. In general, a *scoring method* is defined by a sequence of real numbers $s_1 > s_2 > \cdots > s_m$, one for each alternative. Given an individual's ranking of the alternatives, assign a score of s_1 to the alternative that occupies first position, a score of s_2 to the alternative in second position, and so forth. The total score of each alternative is the sum of its scores over all voter lists, and the alternatives are ordered according to their total scores. Borda's rule corresponds to the scoring system $s_i = m - i$; in fact, it is equivalent to any scoring system in which the successive differences $s_i - s_{i+1}$ are equal and positive.

Consider now any scoring system for three alternatives with descending scores $s_1 > s_2 > s_3$. In the Condorcet counterexample, the score for Peter is $31s_1 + 39s_2 + 11s_3$ whereas the score for Paul is $39s_1 + 31s_2 + 11s_3$. Therefore, Paul obtains a higher score than Peter, even though Peter is the majority candidate. From this we conclude that any scoring system will sometimes violate the majority principle. Moreover, it shows that any scoring system yields outcomes that are based on "irrelevant factors."

Local Independence of Irrelevant Alternatives

Condorcet's broadside against Borda is fine as far as it goes. But how far does it go? Arrow's (1951) theorem shows that, when there are more than two alternatives, *every* reasonable decision rule sometimes violates independence of irrelevant alternatives. Why then should we believe that Condorcet's approach is any better than Borda's? In this section we shall argue that it is, in fact, considerably better.

Let us first consider why independence of irrelevant alternatives is worth bothering about at all. Essentially, it says that the way a given group of alternatives is ordered should depend only on opinions about those alternatives. There are at least two reasons why this is desirable from a practical standpoint. First, if it does not hold, then it is possible to manipulate the outcome by introducing extraneous alternatives (Gibbard, 1973; Satterthwaite, 1975). Second, independence allows the electorate to make sensible decisions within a restricted range of choices without worrying about the universe of all possible choices. It is desirable to know, for example, that the relative ranking of candidates for political office would not be changed if purely hypothetical candidates were included on the ballot.

Again, Arrow's theorem shows that independence cannot be fully realized by any democratic rule. However, it can be realized to a significant extent. Consider the vote between Peter, Paul, and Jack. The real contest here is between Peter and Paul; Jack has many fewer first place votes and many more last place votes than either alternative. We could argue that Jack ought to be

Table 2
Illustrating the Local Independence of Irrelevant Alternatives

	a	*b*	*c*	*d*	*e*	*f*	*Sum*
a	—	51	54	58	60	62	285
b	49	—	68	56	52	58	283
c	46	32	—	70	66	75	289
d	42	44	30	—	41	64	221
e	40	48	34	59	—	34	215
f	38	42	25	36	66	—	207

"irrelevant" to the choice between Peter and Paul because Jack is inferior to both. Moreover, under Condorcet's rule, this is actually the case.

The key point here is that Peter and Paul occur "together" in the consensus ranking; they are not separated by other alternatives. More generally, an *interval* of an ordering is any subset of alternatives that occurs in succession in that ordering. This suggests a weaker form of the independence condition. Suppose we insist only that independence hold within every interval of the proposed ordering. In other words, the ordering within each interval should remain fixed when the alternatives outside the interval are ignored. For example, the ordering of alternatives toward the top of the list is unaffected by the removal of those at the bottom, and so forth. We shall say that such a ranking rule satisfies *local independence of irrelevant alternatives* (LIIA).

Remarkably enough, the maximum likelihood method satisfies LIIA.[8] Moreover, as a later section will argue, it is the only reasonable ranking rule that does so.

To illustrate this idea, consider the pairwise comparison matrix in Table 2 involving six alternatives and 100 voters.[9] Here we may think of a, b, and c as being the real choices under discussion, while d, e, and f are red herrings that have been dragged in by political strategists to attempt to manipulate the outcome. (Note by looking at the upper right-hand quarter of the matrix that each of a, b, c has a majority over each of d, e, f, so the latter three are weaker than the former.) This attempt to muddy the agenda will succeed if Borda's rule is used. The row sums determine the Borda ordering *cabdef*. Now suppose that the three red herrings had not been introduced into the debate. Then the vote matrix would be the one enclosed by the dashed lines. The Borda scores

[8]See Young (1988), where the LIIA condition was called local stability.

[9]Each cell of the vote matrix tells the number of votes received by the candidate in that row over the candidate in that column. Summing across a row therefore adds up all the votes for a given candidate in such pairwise comparisons. This sum is the Borda score, since being ranked ahead of, say, three other candidates will give one three Borda points, and also be tallied as three pairwise votes in the appropriate columns of the matrix. See the essay by Levin and Nalebuff (in this issue) for a more detailed discussion of this point.

for this three-alternative situation are 105 for a, 117 for b, and 78 for c, so the ranking would be *bac*. But this is exactly the reverse of how these three choices are ordered when all six alternatives are considered together. This example shows why Borda's rule is highly susceptible to manipulative practices, as Condorcet alleged.

Now consider the maximum likelihood solution to this problem. We begin by observing an important property of this rule: if some alternative has a simple majority over every other, then the maximum likelihood rule must rank it first.[10] This simple fact can be used to deduce the maximum likelihood solution to the above problem almost immediately. Since a is the majority alternative in every pairwise comparison, it must be ranked first. Among the remaining alternatives, b obtains a simple majority over c, d, e, and f. Hence a similar argument shows that b comes next, and then c. As for d, e, and f, they will be ordered relative to one another as if they were the only three alternatives. (Since they form an interval of the consensus ordering, LIIA applies). It is easy to see that the maximum likelihood solution for these three alternatives is *dfe*. Putting all of this together, we conclude that the maximum likelihood solution to the whole problem is *abcdfe*. This example also shows that the maximum likelihood solution is often quite easy to calculate even when there is more than a handful of alternatives.

At this point we can also see more clearly why the method satisfies LIIA, that is, why any interval of alternatives is ranked as it would be in the absence of the others. The reason is this: if it were not, then one could shuffle the alternatives *within the interval* and obtain an ordering that is supported by a larger number of pairwise votes, hence an ordering that has greater likelihood. This contradiction shows that the maximum likelihood rule cannot be manipulated from below by introducing inferior alternatives, nor can it be manipulated from above by introducing utopian (but infeasible) alternatives.

The Maximum Likelihood Method as a Form of Compromise

So far we have proceeded on the premise that there really is a best ordering to be estimated, or that assuming the existence of such an ordering provides a productive way to think about group choice problems. But in many situations, differences of opinion arise from differences in values, not erroneous judgments. In this case it seems better to adopt the view that group choice is an exercise in finding a compromise between conflicting opinions. Arrow's ax-

[10] The reason is straightforward. Suppose that x (the majority alternative) were not ranked first. Then it must be ranked immediately below some other alternative y. By assumption, x defeats y by a simple majority. Therefore, if we switch the positions of x and y, we obtain a new ranking that is supported by more pairwise votes. But then the new ranking is more likely than the original ranking, because it has more pairwise votes, which is a contradiction.

iomatic approach is one way of analyzing this issue, and we shall pick up this scent again in the next section. First, however, I want to draw attention to another interesting approach along these lines that was pioneered by John Kemeny (1959).

Kemeny viewed the voters' opinions as data and asked for the ordering that best represents or averages the data. For the notion of average to make sense, of course, we must have some way of measuring how far apart one ranking is from another, that is, we need a metric defined on the set of rankings. Kemeny proposed the following natural metric: the distance between two rankings R and R' is the number of pairs of alternatives on which they differ. For example, if $R = abcd$ and $R' = dabc$, then the distance between R and R' is 3, because they differ on exactly the three pairs (a, d), (b, d), and (c, d). If $R'' = cdab$, then the distance from R'' to R is 4, while the distance from R'' to R' is 3, and so forth.[11]

Suppose now that each member of a group of n voters submits a ranking of the alternatives. Given these data points, what is the best definition of a compromise ordering? A statistician would say there are two obvious answers: the mean and the median. When the data are real numbers, the mean minimizes the sum of squares from the observations, while the median minimizes the sum of absolute distances from the observations. By analogy, we may say that a *mean ranking* is one that minimizes the sum of squares of distances from a given set of n rankings. A median ranking is one that minimizes the sum of distances from the n given rankings.

Kemeny left open the question of whether the mean or the median was to be preferred. There can be little doubt, however, that the median is the better choice. To see why, consider the following example. There are three alternatives and 41 voters, where 21 choose *abc*, 5 *bca*, 4 *cab*, and 11 *cba*.

Alternative a has an absolute majority of first-place votes, so *a fortiori* it is the majority alternative. Indeed, the ranking *abc* is supported by a majority, so it has maximum likelihood; that is, it is the median ranking. (The reader can check that Borda's rule yields the same result.) A simple calculation shows, however, that the mean ranking is *bac*.[12] Thus b wins with only 5 of the 41 first-place votes, which hardly seems credible. The problem with the mean (with squaring the differences) is that it places a lot of weight on extreme

[11]Actually, Kemeny (1959) defined the distance between two rankings to be twice the number of pairs on which they differ. He also defined the distance between orderings when voters were indifferent between certain choices. For simplicity of exposition, I will ignore this case.

[12]The distance from *abc* (the median ranking) to the voters' expressed opinions are as follows: 0 to *abc*, 2 to *bca*, 2 to *cab*, and 3 to *cba*. Thus the sum of absolute distances is $21 \cdot 0 + 5 \cdot 2 + 4 \cdot 2 + 11 \cdot 3 = 51$, and it may be shown that this is a minimum. The sum of squared distances from *abc* to the voters opinions is $21 \cdot 0^2 + 5 \cdot 2^2 + 4 \cdot 2^2 + 11 \cdot 3^2 = 151$. This is not a minimum, because the sum of squares of distances from *bac* is smaller: $21 \cdot 1^2 + 5 \cdot 1^2 + 4 \cdot 3^2 + 11 \cdot 2^2 = 106$. It may be checked that this minimizes the sum of squares over all rankings, so *bac* is the mean.

observations: notice that b has only 4 last-place votes, compared with 21 for c and 16 for a. In the present case, the voters who announce the ordering cba shift the outcome in favor of b, not because they are especially attached to b, but because their top candidate is c, which is at odds with the views of most of the other voters. In other words, their opinion about b versus a is heavily weighted because their opinion about something else (namely c) differs from the opinion of the majority. We conclude that, if the object is to find a compromise between the various rankings reported by the voters, then the median is, in a statistical sense, the most appropriate solution. This reinforces the argument for the maximum likelihood rule, but from a different (and more modern) point of view.

An Axiomatic Justification of Maximum Likelihood Voting

The maximum likelihood rule can also be justified from an axiomatic standpoint. Indeed, it is the unique ranking rule that satisfies three standard axioms in the social choice literature plus local independence of irrelevant alternatives.

Define a ranking rule to be a function that associates one or more consensus rankings with every set of rankings reported by a group of individuals on a finite set of alternatives. The rule is anonymous if it treats all voters alike. It is neutral if it treats all alternatives alike. It is Pareto if, whenever everyone ranks one alternative above another, then so does the consensus ranking. Finally, a rule satisfies reinforcement if, whenever two distinct groups of voters each reach the same concensus ordering under separate votes, this ordering is also the consensus for the two groups merged together. For example, if the House of Representatives orders three choices abc, and the Senate also orders these choices abc (using the same voting rule), then abc should be the outcome when the rule is applied to both houses together and the votes remain as before.[13] (In practice, almost all rules have this property.) Young and Levenglick (1978, thm. 3) show that the maximum likelihood rule is the unique ranking rule that is anonymous, neutral, Pareto, and satisfies reinforcement and local independence of irrelevant alternatives.

Conclusion

The maximum likelihood method for ranking alternatives can be justified from several different points of view. It is arguably the best method if we think

[13] This idea was introduced by Young and Levenglick (1978). In the case of ties, reinforcement states that the rankings that are chosen by both groups separately (if any such exist) are precisely the rankings that result when the votes of the two groups are pooled. A variation of the concept characterizes scoring methods (Smith, 1973; Young, 1974, 1975).

of voting as a collective quest for truth—that is, as a way of estimating what decisions are most likely to be factually correct, or most likely to meet a common objective. This especially applies when the decision is being taken by a group of experts. But it also applies to many forms of political decision making—what bill is most likely to reduce crime, what foreign policy will minimize the prospect of war, and so forth. On the other hand, in situations where voting appears to be a way of compromising between conflicting values, maximum likelihood rule still makes sense because it represents the median opinion. Moreover it is reasonably resistant to strategic manipulation—the outcome cannot be changed by introducing inferior alternatives or superior but infeasible alternatives.

The one remaining question is whether the maximum likelihood method is really practical. It is more complicated to calculate than traditional methods like plurality voting, single transferable vote, or Borda's rule. Given modern computing capabilities, however, this issue is largely moot. To find the maximum likelihood solution for six or fewer alternatives, for example, is a near triviality. Even for much larger numbers of alternatives, the method can be implemented in real time as people cast their votes. The more important issue is whether the method is intuitively easy to grasp, and whether it improves on methods currently in use. On both of these counts I think that the answer is affirmative, and I predict that the time will come when it is considered a standard tool for political and group decision making.

References

Arrow, Kenneth, J., *Social Choice and Individual Values*. New York: Wiley, 1951.

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

Borda, Jean-Charles de, "Memoire sur les Elections au Scrutin." *Histoire de l'Academie Royale des Sciences*, Paris, 1784.

Condorcet, Marquis de, "Essai sur l'Application de l'Analyse a la Probabilite des Decisions rendues a la Probabilite des Voix." Paris: L'Imprimerie Royale, 1785. In O'Connor, A. Condorcet, and M. F. Arago, eds., *Oeuvres de Condorcet*. Paris, 1847–9.

Condorcet, Marquis de, "On the Constitution and the Functions of Provincial Assemblies." 1788. In O'Connor, A. Condorcet, and M. F. Arago, eds., *Oeuvres de Condorcet*. Paris. 1847–9.

Crepel, P., *Le Dernier mot de Condorcet sur les Elections*. Paris: Centre d'Analyse et de Mathematique Sociales, 1990.

Gibbard, Allan, "Manipulation of Voting Schemes: A General Result," *Econometrica*, July 1973, 41, 587–601.

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

Grofman, Bernard, Guillermo Owen, and Scott Feld, "Thirteen Theorems in Search of the Truth," *Theory and Decision*, September 1983, *15*, 261–78.

Henry, Charles, ed., *Correspondance Inedite de Condorcet et de Turgot 1770–1779*. Paris: Charavay Freres, 1883.

Kemeny, John, "Mathematics Without Numbers," *Daedalus*, Fall 1959, *88*, 571–91.

Nitzan, Shmuel, and Jacob Paroush, "Optimal Decision Rules in Uncertain Dichotomous Situations," *International Economic Review*, June 1982, *23*, 289–97.

Poisson, Simeon-Denis, *Recherches sur la Probabilite des Jugements en Matiere Criminale at en Matiere Civile, Precedees des Regles Generale du Calcul des Probabilites*. Paris: Bachelier, 1837.

Satterthwaite, Mark, "Strategy-Proofness and Arrow's Conditions: Existence and Correspondence Theorems for Voting Procedures and Social Welfare Functions," *Journal of Economic Theory*, April 1975, *10*, 187–217.

Shapley, Lloyd S., and Bernard Grofman, "Optimizing Group Judgmental Accuracy in the Presence of Interdependencies," *Public Choice*, 1984, *43*:3, 329–43.

Smith, John, "Aggregation of Preferences with Variable Electorate," *Econometrica*, November 1973, *41*, 1027–41.

Sommerlad, F., and Iain McLean, "The Political Theory of Condorcet." Social Studies Faculty Centre Working Paper 1/89, Oxford University, 1989.

Urken, A. B., and Steven Traflet, "Optimal Jury Design," *Jurimetrics*, 1974, *24*, 218–35.

White, E. B., "The Wild Flag," *The New Yorker*, 1946.

Young, H. Peyton, "An Axiomatization of Borda's Rule," *Journal of Economic Theory*, September 1974, *9*, 43–52.

Young, H. Peyton, "Social Choice Scoring Functions," *SIAM Journal on Applied Mathematics*, June 1975, *28*, 824–38.

Young, H. Peyton, "Optimal Ranking and Choice from Pairwise Comparisons." In Grofman, B., and G. Owen, eds., *Information Pooling and Group Decision Making*. Greenwich, Conn.: JAI Press, 1986, pp. 113–22.

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

Young, H. Peyton, and Arthur Levenglick, "A Consistent Extension of Condorcet's Election Principle," *SIAM Journal on Applied Mathematics*, September 1978, *35*, 285–300.