Discussion Paper No. 637

A CASE AGAINST BULLET, APPROVAL AND PLURALITY VOTING

by

Donald G. Saari
Jill Van Newenhizen

January, 1985

A CASE AGAINST BULLET, APPROVAL AND PLURALITY VOTING

Donald G. Saari
Jill Van Newenhizen

Department of Mathematics
Northwestern University
Evanston, Illinois 60201

January, 1985

It is well known that plurality voting need not reflect the true sentiments of the voters. When there are more than two candidates running for the same office, it is probable that the candidate approved by most of the voters won't win. Examples abound; perhaps the best known one is the Senatorial election in the state of New York during 1970. Conservative James Buckley benefited from a split vote for his two liberal opponents; he was elected with 39% of the vote even though 61% of the electorate preferred a liberal. In the 1983 Democratic Party primary campaign for the mayor of Chicago, the black candidate, Harold Washington, was elected because of a split vote for his two white opponents, Jane Bryne and Richard Daley. Indeed, examples can be found in many closely contested elections among three or more candidates.

To illustrate the problem, consider the following example of fifteen voters and three candidates A, B, and C. Suppose the voters' preferences are split in the following way: Six of the voters have the ranking A>C>B; five have the ranking B>C>A, and four have the ranking C>B>A. The result of a plurality election is A>B>C with a tally 6:5:4. But, although A is elected, a majority (60%) of these voters prefer B to A. More seriously, a majority (60%) of these voters prefer C, the last place candidate, to A, and 2/3 of them prefer C to B! So, C is the preferred candidate, but this fact isn't reflected in the election results. This is because with a plurality vote the voters can vote for only their top ranked candidate.

Voters are aware of this phenomena and react by using strategic voting behavior. This is manifested by the common refrain near election time of "Don't waste your vote, mark your ballot for ---." To remain viable, candidates must devote valuable campaign time to counter this effect. For instance, during the 1984 presidential primaries, Jesse Jackson urged his followers to vote for him rather than strategically voting for Walter Mondale.

There is little question that this instrument of democracy is in need of reform. But, what should be its replacement? One reform measure, which is intended to capture how strongly a voter prefers his favorite candidate, is bullet voting. Here a voter has two votes; he can vote for his top ranked candidate, his top two candidates, or he can cast both votes for his top ranked candidate. For instance, in the above example, if all of the voters had sufficient regard for candidate C, then she would be elected. Bullet voting was used in Illinois for certain legislative offices.

A second proposed reform method is approval voting. This is where a voter votes for all of the candidates he approves of. As such, when there are N candidates, the voter has N choices; he can vote approval for his top i candidates, $i=1,..,N$. Again for the above example, depending upon the degree which the voters favor candidate C, she may, or may not, emerge victorious.

Approval voting enjoys the support of several experts in this field. It was employed for a straw ballot during the Pennsylvania Democratic party conference in December, 1983, [1], and it was used to select faculty members to the Northwestern University Presidential search committee in November, 1983. In addition, it "..is now used in academic societies such as the Econometric Society, in the selection of members of the National Academy of Science during final balloting, and by the United Nations Security Council in the election of a Secretary General. Bills to enact this reform are now before the state legislatures of New York and Vermont." [2].

Much of this support is a consequence of the analysis of its properties by two of its foremost advocates, Steven Brams and Peter Fishburn. Most of their conclusions, which highlight several of the desirable properties of this system, are summarized in their book "Approval Voting"[3]. To demonstrate the strength of approval voting, often they compared it with those commonly used systems which distinguish between two sets of candidates - the top k and the rest. Plurality voting is the special case where $k=1$; it distinguishes between the top ranked candidate and all others. Because of the technical difficulties involved, not all of the properties of "approval voting" were found, and in [3] it wasn't compared with all other voting systems.

At a conference in July, 1984, one of us (DGS) presented several negative results concerning the behavior of approval voting (see [4,5]) as part of a lecture discribing what can occur with voting methods. Later, S. Brams privately asked whether the techniques developed to obtain these results could expose other properties of approval voting. We started this project with the expectation that approval voting is, in some sense, better than most systems. But, we found that it has several disturbing features which makes it worse than even the plurality voting system. Indeed, these properties appear to be sufficiently bad to disqualify approval voting as a viable reform alternative.

In this paper, we report on some of these negative features which are shared by voting systems, such as approval voting, bullet voting, cardinal voting, etc., where there is more than one way to tally each voter's ranking of the candidates. The feature emphasized here is that the election outcome can be "random" in nature rather than being decisive. More precisely, if there are N candidates, then there are N! possible ways to rank them without ties. The purpose of an election is to determine which one of them is the group's choice. If a voting system is decisive, a given set of voters' profiles uniquely determines one of these rankings. However, for approval and bullet voting, there are a large number of examples where all N! outcomes occur for the same profile! That is, each voter votes honestly according to his fixed ranking of the candidates. But, as the voters vary their choice of how their ballots will be tallied, each of the possible N! rankings emerge!

This phenomenon can be illustrated with the above example. Let w,y,z denote, respectively, the number of voters from the three types of voters who vote approval for their top two candidates rather than just their top ranked candidate. Then, $0 \le w \le 6$, $0 \le y \le 5$, $0 \le z \le 4$, and the tally for A:B:C is $6:5+z:4+w+y$. It follows immediately that any election outcome is attainable from these voters. For instance the result

B>A>C occurs when $z \geq 2$ (at least two from the last set of voters vote for their top two ranked candidates) and $w+y \leq 1$. Even ties are possible. A deadlocked election of A=B=C results from z=1 and w+y=2, while the result B=C>A results from $w+y-1=z \geq 2$.

This example and the general result indicates that systems such as approval and bullet voting possess features which are more undesirable than even those of the plurality voting system! Just this stochastic, random nature of the election outcome raises serious questions whether approval voting, and related methods, truly offer any reform. The proposed cure seems to be much worse than the disease.

This result doesn't mean we are doomed to accept and to live with the failings of the plurality voting. There are other ways to tally a ballot which would reflect a voter's first, second, .., last ranked candidates. For instance, a Borda Count is where when there are N candidates, N points are tallied for a voter's top ranked candidate, (N-1) for his second ranked candidate, ..., (N-k) for his k^{TH} ranked candidate, .., and 1 point for his last ranked candidate. (The Borda ranking for the above example is the desired one of C>B>A with a tally of 34:29:27.) So, out of all possible ways there are to tally ballots, the problem is to isolate those ways which best capture the wishes of the electorate. It turns out that the unique solution for this problem is the Borda Count.[6]

2. The main results.

Assume there are $N \geq 3$ candidates denoted by $\{a_1,..,a_N\}$. Let $\underline{W}$ = $(w_1,...,w_N)$ be a <u>voting vector</u> where its components satisfy the inequalities $w_j \geq w_k$ if and only if j<k, and $w_1 > w_N$. Furthermore, we require all of the weights to be rational numbers. (The only purpose of the last requirement is to simplify the proofs. Clearly it doesn't impose any practical limitations because for all commonly used methods the weights are fractions and/or integers.) Such a vector defines the tallying process for an election -- w_j points are tallied for a voter's j^{TH} ranked candidate. Then, the sum of the points tallied for a candidate determines her final ranking. For example, the vector (1,0,..,0) corresponds to the plurality vote. $\underline{B}^N=(N,N-1,...,1)$ defines the usual Borda Count procedure. (More generally, a Borda Vector is whenever the differences w_j-w_{j+1} are the same nonzero constant for j=1,..,N-1. An election using such a voting vector is a Borda Count.) Let $\underline{E}_N$ be the vector $N^{-1}(1,1,..,1)$. This isn't a voting vector because all of the components are equal, so it can't distinguish how candidates are ranked.

A simple voting system is where a specified voting vector is used to tally the voters' rankings of the candidates. A general voting system is where there is a specified set of at least two voting vectors, $\{\underline{W}_j\}$, where the difference between any two of these vectors isn't a scalar multiple of $\underline{E}_N$. (Hence, no two are the same.) Then, each voter selects a voting vector to tally his ballot.

Examples. For the bullet method, the set of three voting vectors are ((1,0,..,0), (1,1,0,..,0), (2,0,..,0)). For approval voting, the set of N vectors is ((1,0,..,0), (1,1,0,..,0),..., (1,1,..,1,0), (1,1,...,1)). For cardinal voting, the voter is free to select the values of the weights w_j subject to certain constraints. For instance, to standarize the choices, the weights might be required to sum to unity, or to be bounded above and below by specified constants.

The extreme indecisiveness for elections which was described above is characterized in the following definition.

Definition 1. A general voting system for N candidates is said to be stochastic if there exist profiles of voters where all possible rankings of the candidates (without ties) can result from the same profile as the voters vary their choice how their ballots are to be tallied.

In other words, for these examples of voters' profiles, the outcome is random and reflects the voters' fluctuations in their choice of tallying procedure rather than their rankings of the candidates. Thus, the election outcome could be

$a_1 > a_2 > \ldots > a_N$, or $a_N > a_{N-1} > \ldots > a_1$, or $a_1 > a_N > \ldots$, etc. where the determining factor is the voters' choice of how their ballots are to be tallied, <u>not</u> their rankings of the candidates. As in the introductory example, all rankings, even those with ties can occur. This random feature is a property we want to avoid, so it is important to characterize those general voting systems which have it.

Theorem 1. Assume there are $N \geq 3$ candidates. All general voting systems are stochastic.

As illustrated by the introductory example, a major criticism of the plurality vote is that it need not reflect the voters' true wishes. To quantify this, we need a measure of the true sentiment of the voters. The one most commonly used is the Condorcet winner.

Definition 2. Assume that the $N \geq 3$ candidates are $(a_1, a_2, .., a_N)$. Candidate a_K is called a Condorcet winner if, in all possible pairwise comparisions, a_K always wins by a majority vote. A Condorcet loser is an candidate which always loses by a majority vote in all possible pairwise comparisions with the other candidates.

A Condorcet winner appears to capture the true choice of the voters; after all, she was the choice of a majority of the electorate when she was compared with any other candidate. But, the introductory example shows that the plurality vote can rank a Condorcet winner in last place and a Condorcet loser in first place! Actually, as it is shown in [6], this type of behavior is characteristic of all simple voting systems with the sole exception of the Borda Count.

Theorem 2.[6]. Suppose there are $N \geq 3$ candidates. Then, for any simple voting system other than a Borda Count, there exist examples of voters' prefences where the Condorcet winner is ranked in last place and the Condorcet loser is ranked in first place. The Borda Count is the unique method which never ranks a Condorcet winner in last place, and never ranks a Condorcet loser in first place.

How does a general voting system fare? It turns out that it can be even worse.

Definition 3. A general voting system $\{\underline{W}_J\}$ is "plurality like" if there are non-negative scalars $\{b_J\}$ such that when the differences between successive components of $\sum b_J\underline{W}_J$ are computed, all but one are zero.

The summation defines a voting vector, and the condition is that this vector distinguishes between only two sets of candidates. This condition is satisfied automatically when the general voting method includes a voting vector of this type, e.g., a plurality voting vector. Thus, both approval and bullet voting are plurality like. Also, the condition is trivially satisfied should the general method include N-1 voting vectors which, along with $\underline{E}_N$, form a linearly independent set. This is because the vectors span R^N.

Theorem 3. Suppose there are $N \geq 3$ candidates. Choose a ranking for each of the $N(N-1)/2$ pairs of candidates in any manner you wish. (This may be done in a random fashion; the rankings need not be transitive.) Assume that all subsets of more than two candidates are to be ranked with a plurality like, general voting method. Then, there exist examples of voters' profiles so that

1) for each of the pairs of candidates, a majority of the voters have the indicated preference, and

2) for each subset of three or more candidates, the outcome is stochastic.

(As it will become clear in the proof, the constraint that the general voting method is "plurality like" isn't necessary; the conclusion holds for almost all general voting systems. We impose this assumption because it significantly simplifies the proof without incurring much of a sacrifice to generality -- the general methods which have been seriously considered or used, such as approval or bullet voting, satisfy this condition.)

This theorem means that there exist examples of voters' profiles where for each pair (a_J, a_K), a majority of the voters prefer the candidate with the smaller subscript. Thus, the rankings of the pairs are highly transitive and [illegible] group ranking of $a_1 > a_2 > .. > a_N$. In spite of this, for the same voters, the approval voting rankings of <u>all</u> subsets of three or more candidates are stochastic. Thus the approval voting outcome over a set of more than two candidates could depend more on the quirks and fortunes of how the voters select to have their ballots tallied than on how they rank the candidates. The introductory example illustrates this for N=3.

Our goal was to compare approval voting with the Condorcet winner. This is a corollary of the theorem when the rankings of the pairs define a Condorcet winner. For instance, when N=4, there exist examples of voters so that whenever a_1 is compared with any other candidate, she always wins with a majority vote. Yet, when these same voters use approval voting to rank the candidates, the outcome is stochastic whether the set of candidates is $\{a_1,a_2,a_3,a_4\}$, $\{a_1,a_2,a_3\}$, $\{a_1,a_2,a_4\}$, $\{a_1,a_3,a_4\}$, or $\{a_2,a_3,a_4\}$. In other words, it is probable that a Condorcet winner could win an approval election only by accident. Moreover, the approval voting outcome over any of these subsets of candidates is random, so it isn't clear who is the approval winner. This type of an example is important because it illustrates that a general voting method can be random even when there is a Condorcet winner; a general voting method need not reflect the voters' true views. The generalization of the example is highlighted in the following formal statement. Because we aren't considering <u>all</u> possible subsets of candidates, the condition on the general voting system is relaxed.

Corollary 3.1. Assume there are $N \geq 3$ candidates which are to be ranked with a general voting method. Assume that the general voting system has at least one vector which isn't a Borda vector. There exist examples of voters' profiles so that even though there is a Condorcet winner, the ranking of the N candidates is stochastic.

Again, the introductory example illustrates this result for N=3. Furthermore, this result extends to all subsets of the N candidates. Thus, a general method, such as approval or bullet voting, need not reflect the wishes of the voters over any subset of the candidates.

These negative statements about approval voting can be reconciled with some of the positive ones which appear in the literature. In particular, in certain settings, it has been proved that approval voting can result in a Condorcet winner being ranked in first place [3]. It follows from the above statements that such favorable conclusions can occur, but, they are just one of the many possible stochastic fluxuations of the election! The corollary and the theorem assert that if the voters behave by choosing to tally their choices in certain, specified ways, then these desirable outcomes will result. But if they don't, then anything can occur. In other words, results as undesirable as one may fear are probable.

Another implication of Theorem 3 concerns those procedures used to consider voters' preferences not only over the total set of candidates, but also over subsets. For example, the following is a standard approach: First rank the N candidates, and then drop the candidate who is in last place. Rerank the remaining set and continue this elimination procedure until only the required number of candidates remain. Now, suppose we employ a general method, such as approval voting, to rank the candidates at each step of this elimination procedure. Theorem 3 shows that the stochastic nature of the conclusion could force the final result to have no relationship whatsoever with how the voters really rank the candidates.

Theorem 3 can be extended. For instance, consider a situation where there are 4 candidates, but the election is only closely contested among three of them. Then we might expect that the stochastic effect affects only these three candidates, not the last one. This does happen; results about "partial stochastic" effects can be obtained.

So far we've compared general voting methods only with the rankings of pairs of candidates. Another test is to compare it with plurality voting and other single voting methods. After all, in the introductory example, the conclusions of plurality voting aren't indicative of the voters' preferences as measured by a Condorcet winner. The following statement compares the results of a general voting method with any simple voting system.

Theorem 4. Assume there are $N \geq 4$ candidates. Let $\underline{W}_N$ be a voting vector defining a simple voting method, and assume that a general voting method is given where at least two of the vectors and $\underline{E}_N$ are linearly independent. Then there exist examples of voters' profiles so that

a. There is a Condorcet winner.

b. If $\underline{W}_N$ isn't a Borda vector, then the simple voting method has any previously selected ranking of the candidates.

c. The general method is stochastic.

If the general voting method is either approval or bullet voting, then the conclusion holds for $N \geq 3$. So, with the exception of the Borda Count, Theorem 4 illustrates that examples exist where anything can occur with the simple voting scheme while the general method is stochastic. In particular, this implies the existence of examples of voters' profiles where the plurality outcome does rank the Condorcet winner in first place while approval voting has a stochastic effect. Consequently, it is probable that the plurality election results do reflect the voters' wishes while approval voting does not. Again, this is because the random fluxuations of approval voting allow any type of election result to emerge. A second consequence of this theorem, along with Theorem 2, is that among all single voting systems, only the Borda Count reflects the voters' intent.

A remaining issue is the robustness of these conclusions. That is, can we dismiss these statements because the conclusions occur only with some highly pathological example which is highly unlikely to occur? We show this isn't so; with any fairly general distribution of voters' preferences, these results have a positive probability of occurring. Moreover, as it will become apparent in the proofs, these stochastic outcomes are more likely to occur in those types of situations which have been used to discredit the plurality system. It turns out that these random election outcomes tend to occur when there is a closely contested election among three or more candidates.

To show that these examples are probable, we need to introduce a measure for the distribution of the voters' profiles. If there are N candidates, then there are

N! possible rankings of them. For each possible ranking, A, let n_A be the fraction of the voters with this ranking. The sum of these numbers equals unity, so there are N!-1 degrees of freedom; these numbers define a unit cube, C(N), in the positive orthant of a N!-1 dimensional space. Voters' profiles can be identified with the (rational) points in C(N). Because all of the usual continuous distributions identify a positive probability of occurrance with an open set in C(N), outcomes are probable if they are identified with open sets in C(N). The following theorem asserts that this characterizes our assertions. Also, statements asserting the asymptotic, positive likelihood of these examples, as the number of voters grows, follow immediately. The limits are related to the measure of the open sets in C(N). (See [4,5].)

Theorem 5. For each of the above theorems, the set of examples defining the described behavior contains an open set in C(N).

Finally, we are left with the issue of reform. The above demonstrates that a general method does <u>not</u> constitute a reform alternative for plurality voting. But, this doesn't mean we are forced to live in the imperfect world of plurality voting. In Theorems 2 and 4 and in Corollary 3.1, we see that the Borda Count is the unique simple voting scheme which avoids many of the pitfalls associated with plurality voting. This in itself demonstrates that the Borda Count is best candidate, of the voting methods, for reform. For a more detailed analysis of its properties along with a comparison of it with other simple voting methods, see [6].

3. Proofs

Proof of Theorem 1. Assume there are $N \geq 3$ candidates $\{a_1,..,a_N\}$ and that the general voting system consists of the voting vectors $\{\underline{W}_j\}$, j=1,..,s, where

$s \geq 2$. Then, each $\underline{W}_J$ is a vector in the N dimensional space R^N. Let A denote the ranking $a_1 > a_2 > \ldots > a_N$, and let P(A) be a generic representation for the N! permutations of A. For voting vector $\underline{W}_J$, any such permutation P(A) determines how the ballot will be tallied. In fact, this tally can be viewed as being a permutation of the vector $\underline{W}_J$. Denote this permuation by $\underline{W}_{JP(A)}$. For instance, if $\underline{W}=(3,2,1)$, then the standard ranking $a_1 > a_2 > a_3$ defines the vector (3,2,1). The ranking $a_3 > a_1 > a_2$ defines the permutation of $\underline{W}$, (2,1,3), to reflect that for this ranking, two points are tallied for a_1, one for a_2, and three for a_3.

Let $n_{P(A)}$ denote the fraction of the voters with the ranking of the candidates P(A). The tally of a simple election using $\underline{W}_J$ is

4.1 $\sum n_{P(A)} \underline{W}_{JP(A)}$

where the summation is over all N! permutations P(A). The outcome of the election is determined by algebraically ranking the components in this vector sum.

There is a geometric representation for this algebraic ranking. Consider the indifference hyperplane in R^N given by $x_I = x_K$. If the vector sum 4.1 is on the $x_K > x_I$ side of this hyperplane, then a_K ranks higher than a_I, and vice versa. In particular, the N(N-1)/2 possible "indifference hyperplanes" divides R^N into "ranking regions", and the final ranking of the candidates is determined by which ranking region contains the vector sum.

For a general voting system, let $m_{JP(A)}$ denote the fraction of those voters with a P(A) ranking that elect to have their ballots tallied with the j^{TH} voting vector. Then, the fraction of the total number of voters with this tally is $n_{P(A)} m_{JP(A)}$. Consequently, the total tally is given by the double sum

4.2 $\sum_{P(A)} n_{P(A)} [\sum_J m_{JP(A)} \underline{W}_{JP(A)}]$.

Again, the ranking of the candidates is determined by the ranking region of R^N which contains this vector sum.

We represent Eq. 4.2 as a mapping. Toward this end, let

$Si(M) = \{\underline{x}=(x_1,..,x_M) \mid x_K \geq 0, \sum x_K = 1\}$.

Because each term defines a percentage, the set $\{n_{P(A)}\}$ is a (rational) point in the set Si(N!). For each P(A), the set $\{m_{jP(A)}\}$ is in Si(s). (This is because the entries define non-negative fractions which sum to unity.) This means that a domain point is in the $(N!-1)(s-1)^{N!}$ dimensional space

$$T = Si(N!)x(Si(s))^{N!}.$$

Any rational point in T corresponds to an example of voters' profiles along with their individual selection of voting vectors to tally the ballots. Thus, Eq. 4.2 can be viewed as being a mapping from T to R^N

4.3 $$F:T \text{ --------> } R^N,$$

where F is the summation.

Define the "complete indifference" ranking in R^N to be the line given by all scalar multiples of $\underline{E}_N$. The name comes from the fact that this line corresponds to where there is a complete tie in the rankings of all of the candidates. Notice that

a) the complete indifference ranking is the intersection of all of the indifference hyperplanes, and

b) this line is on the boundary of all other ranking regions.

To prove this theorem, we must show the existence of a n* in Si(N!) (a choice of voters' profiles) so that as the variable $m=\{m_{jP(A)}\}$ varies, the image of F((n*,m)) meets all possible ranking regions.

Let n* correspond to where there is an equal number of voters with each possible ranking of the candidates; i.e., $n^*=(N!)^{-1}(1,1,..,1)$. It follows immediately that if $m_{1P(A)}=1$ for all choices of P(A) (all voters choose the first voting vector), then Eq. 4.2 reduces to Eq. 4.1, and the image of F is on the complete indifference line. The same conclusion holds if all of the $m_{jP(A)}$ are equal. This is because the double summation can be interchanged to obtain separate summations of the type given in Eq. 4.1, each of which yields a point on the complete indifference line. Denote this domain point by (n*,m*).

The idea is the following. Assume that the Jacobian of F at (n^*,m^*) has rank equal to N, where, in the computation of the Jacobian, the $n_{P(A)}$ variables are held fixed. (We treat them as parameters.) This means that there is an open set about the interior point m^* which is mapped to an open set about the image $F((n^*,m^*))$. This open set yields outcomes which can be attained with the same profile of voters (n^*), but where m, which indicates the choice of voting vectors to tally the ballots, varies. Because an open set about any point on the line of complete indifference meets all ranking regions, the conclusion follows. (It is easy to show that there are rational choices of m with this property. For details, see [4,5].)

Thus, the proof is completed if we can determine certain properties about the Jacobian of F at (n^*,m^*). There are two cases to consider, and they are based upon the sum of the components of each voting vector. Either at least two of these sums differ, or they are all the same.

Assume that at least two of the sums differ. For each P(A), eliminate the dependency of the components $\{m_{jP(A)}\}$ by setting $m_{1P(A)}=1-\sum_{j\geq 2} m_{jP(A)}$. Then, the rank of the Jacobian of F is determined by the maximum number of independent vectors in subsets from

4.4 $\{\underline{W}_{jP(A)}-\underline{W}_{1P(A)}\}$,

where P(A) ranges over all N! permutations of A and where j=2,..,s. There is a choice of j where the sum of the components of $\underline{W}_j$ doesn't equal the sum of the components of $\underline{W}_1$, say j=2. Moreover, without loss of generality, we can assume that the sum of the components of $\underline{W}_2$ is larger than the sum of the components of $\underline{W}_1$. What we show is that the set of vectors

4.5 $\{\underline{W}_{2P(A)}-\underline{W}_{1P(A)}\}$

spans R^N. This will complete the proof.

It follows immediately that $\sum_{P(A)} \underline{W}_{jP(A)}$ is a nonzero scalar multiple of $\underline{E}_N$ where the scalar is (N-1)! times the sum of the components of $\underline{W}_j$. Thus,

$\sum_{P(A)}(\underline{W}_{2P(A)}-\underline{W}_{1P(A)})$ is a positive scalar multiple of $\underline{E}_N$, so this vector is in the space spanned by the vectors in Eq. 4.5. The simplex Si(N) has $\underline{E}_N$ as a normal vector, so the theorem is proved if the simplex is spanned by the vectors in Eq. 4.5.

Each vector in Eq. 4.5 can be viewed as being a permutation of the components of $\underline{W}=\underline{W}_2-\underline{W}_1$. Let vector $\underline{V}$ be the permutation of $\underline{W}$ which has the largest value in the first component, the second largest in the second component, etc. (For example, if $\underline{W}_2=(5,4,2,1)$ and $\underline{W}_1=(5,1,1,0)$, then $\underline{W}=(0,3,1,1)$ and $\underline{V}=(3,1,1,0)$.) The set of all possible permutations of $\underline{V}$,

4.6 $\{\underline{V}_{P(A)}\}$,

agrees with the set in Eq. 4.5. Because [illegible] multiple of $\underline{E}_N$, $\underline{V}$ can be viewed as being a voting vector and the vectors in Eq. 4.6 can be viewed as being the various ways to tally ballots. That this set spans Si(N) follows immediately from [7,4].

Suppose the sums of the components for each of the voting vectors are the same. We show that the Jacobian of F has rank N-1 and its image spans a simplex Si(N). This requires the following adjustment in the proof. First, an open set about m* is mapped to an open set about F((n*,m*)) <u>in the simplex</u>. However, such an open set must meet all ranking regions. (The simplex has codimension one, and its normal direction is given by the line of complete indifference.). Thus, all we need to show is that the vectors in 4.6 span the subspace orthogonal to $\underline{E}_N$. This is the same argument given above. This completes the proof.

The proofs of Theorems 3 and 4 depend heavily upon the proofs and results in [6]. Essentially, the idea of the proofs is to use special ways in which the voters choose their voting vectors to obtain a simple voting systems. Then, modifications of the type used in the proof of Theorem 1 and results from [6] lead to a condition of the type where F((n*,m*)) is on the line of complete indifference. The Jacobian condition follows from the analysis given in the proofs of [6].

Proof of Theorem 4. The following is a consequence of Theorems 5 and 7 in [6].

Lemma. Suppose there are two simple voting vectors, $\underline{V}_1$ and $\underline{V}_2$ which,
a) form a linearly independent set along with $\underline{E}_N$ and
b) a Borda vector isn't in the span of these three vectors.
Rank the pairs of candidates in any way and rank the N candidates in any two ways. Then, there exist profiles of voters for which when the same voters consider each pair of candidates, a majority prefer the designated one. When these same voters rank the N candidates by the simple voting system $\underline{V}_j$, the outcome is the j^{TH} ranking of the candidates, j=1,2.

According to the statement of the theorem, the range space containing the tally of the various subsets of candidates is given by $S=(R^N)x(R^N)x(R^2)^p$ where $p=N(N-1)/2$. The first component space is the tally of the simple voting system, the second is the tally of the general voting system, and the last p components contain the tally of the binary comparisons. The domain is T. Thus, the obvious summations define the mapping

$$F*: T \text{--------}> S.$$

The proof follows much as in that of Theorem 1. We show the existence of a set of profiles n" for which the rankings of the simple system and the rankings of the pairs of candidates are as specified. Moreover, n" is such that there is an interior point, m", in the product of the simplices which designate how the voters select their tallying vectors so that F((n",m")) is complete indifference for the general system. Then, the above argument concerning the Jacobian of F*, is repeated. The main difference is that it is evaluated at (n",m") rather than at (n*,m*).

First we find m". To do this, we choose the $m_{jP(A)}$'s to depend on j but not on P(A). This defines a convex combination of the voting vectors which are available to tally the rankings. In turn, this defines a new voting vector; indeed, it defines a continuum of them where the dimension of the continuum depends upon the number of linearly independent vectors in the general voting system. Since we have a continuum of them available, and since at least two vectors in this system define a three dimensional space with $\underline{E}_N$, we can choose the m_j's to obtain a voting vector $\underline{V}_N$ which, along with $\underline{W}_N$, satisfies the condition of the lemma. To use the lemma, choose the ranking corresponding to $\underline{V}_N$ to be complete indifference, the ranking corresponding to $\underline{W}_N$ to be as specified in the theorem, and the pairwise rankings as specified in the theorem. The conclusion then follows from the above and the lemma.

Proof of Theorem 3. In this setting, the domain and the image of F changes drastically from that given above. Here we have $2^N-(N+1)$ different subsets with at least two candidates. Thus, the range space is the cartesian product over all of these sets of Euclidean spaces of the same dimension as the number of candidates in the subset. The domain also is increased significantly. For each subset, there is a general voting method. Thus, for each ranking of the candidates in each subset, the domain is increased by another product of a simplex reflecting the various choices the voters have to tally their ballots. Let the new domain, which is a much larger product space of simplices, be given by T′, and let the larger image space be given by R′. The tally of the ballots still is given by summations of the type found in Eq. 4.2. They define a mapping

$$F':T' \text{--------}> R'.$$

As in the statement of the theorem, designate for each pair of candidates which one is to be preferred by a majority of the voters. We now appeal to Theorem 6 in [6]. A consequence of this result is that for "most" simple voting systems, there exist profiles of voters so that for each of the pairs of the candidates, a majority of them favor the designated candidate. Yet, their rankings of all subsets with three or more candidates is complete indifference. "Most" replaces the linear independence condition in the lemma, and it means that the voting vectors for various subsets don't make a certain determinant vanish. For our purposes it suffices to note that for this condition is satisfied for any voting vector which is "plurality like". To use this theorem, we find a special case of the general voting system which is a simple voting system satisfying the above.

For each subset of more than three candidates, choose $m_{JP(A)}=m_J$. This, then, defines a convex combination of the voting vectors which are available to tally the rankings. Now, choose the m_J's so that $\sum m_J \underline{W}_J$ defines a vector of the type in Definition 3. Such a vector doesn't make the determinant condition vanish. If not all of the m_J's are positive, then they can be perturbed so that all are positive and the sum is still a vector which satisfies the non-vanishing of the determinant. (This is because the determinant condition is an open condition.) Thus, for each subset, a choice of the $\{m_J\}$ can be made so that the resulting vectors over all subsets do not satisfy the vanishing determinant condition.

Let m′ correspond to these choices of $\{m_{JP(A)}\}$ over all subsets. Then we have from Theorem 6 in [6] that there exist profiles of voters, n′, so that the various components of F((n′,m′)) are on the line of complete indifference, yet the ranking of the pairs is as designated. What remains to be shown is that the rank of the Jacobian of F′, where n is held fixed, is of the rank of the the dimension of the range. But, with the modifications of the type given in the proof of Theorem 3, this follows from the proof of Theorem 6 in [6]. Indeed, the proof of Theorem 6 is based upon this independence condition holding.

Proof of Corollary 3.1. Theorem 4 in [6] asserts that if the voting vector isn't a Borda Vector, then for any rankings of the pairs of alternatives and for any ranking of the N alternatives, there is a profile of voters which will realize all of these outcomes simultaneously. In particular, the rankings of the pairs can be chosen to define a Condorcet winner, and the ranking of the N candidates can be chosen to define the complete indifference ranking. The hypothesis of the corollary asserts that there is at least one voting vector in the general methods which isn't a Borda Vector. Thus, m′ can be chosen in a manner similar to the above so that the resulting voting vector isn't Borda. Then, the same type of proof goes through.

Proof of Theorem 5 for hypothesis of Theorem 1. The basic ideas are demonstrated for Theorem 1; since the ideas extend immediately for the other theorems, we only prove this case.

To prove the theorem, all that is necessary is to show that there is an open set U about n* in Si(N) (the space of voters′ profiles) so that if n′ is in U, then there is an interior point m′ in $(Si(s))^{N!}$ such that F((n′,m′)) is on the line of complete indifference. To do this, we give a geometric interpretation of Eq. 4.2.

For each ranking of the candidates, P(A), the bracketed term in the double summation Eq. 4.2 defines the convex hull of the vectors $\{\underline{W}_{JP(A)}\}$. Thus, this means that the double summation yields the convex hull of the N! convex hulls. The fact that the image of F((n*,m)) contains an open set means that this particular convex combination of the convex hulls contains a open set around the point F((n*,m*)) on the line of complete indifference. It now follows from continuity considerations that the conclusion holds.

We conclude with an observation which indicates that there is a large likelihood of a stochastic effect for a general voting system. According to the above proof, a measure of this is the abundance of the points (n′,m′) such that F((n′,m′)) is on the line of complete indifference. But, according to the above independence argument, it follows from the implicit function theorem that the inverse image of the complete indifference line in T is an affine space of codimension N-1. To obtain some feeling for the size of this space, consider the setting where N=4 and where we are using an approval voting system (s=3). Then, the base points define a 58 dimensional linear subspace in a 61 dimensional space. Incidently, these large dimensions already for only 4 candidates indicates 1) why standard methods won't suffice in the analysis of such voting schemes, 2) why we used a convexity argument to prove Theorem 5 instead of an implicit function argument (which would have involved a massive linear independence argument), 3) why the stochastic effect occurs (F is trying to force the input from a 61 dimensional space into a 3 dimensional space, so we must expect such results), 4) why simple voting systems don't have as adverse effects (the subspace is 20 dimensional in a 23 dimensional space), and 5) that there are many examples other than those suggested by the proof of the theorems. (Because the 58 dimensional space is affine, it must intersect the boundaries of T. The boundaries correspond to examples of voters' profiles where there are no voters which have certain rankings of the N candidates.)

Acknowledgements

The work of both authors was supported, in part, by NSF Grant IST 8415348. As mentioned in the introduction, this work was partially stimulated by questions and comments made by S. Brams, Carl Simon, and the other participants at the July, 1984, "Workshop on Mathematical Models in Political Science" which was organized by John Aldrich and Carl Simon. This conference was jointly sponsored by the NSF Institute for Mathematics and its Applications and the Department of Political Science, both at the University of Minnesota, Minneapolis, Minnesota.

References

1. Nagel, J., A debut for approval voting, PS Winter 1984, 17, 62-65.

2. Brams, S., and P. Fishburn, America's unfair elections, The Sciences, Nov. 1983, reprinted in Eastern Airlines' Review Magazine, Feb. 1984.

3. Brams, S., and P. Fishburn, Approval Voting, Birkhauser, Boston, 1982.

4. Saari, D. G., The ultimate of chaos resulting from weighted voting systems, Advances in Applied Mathematics 5 (1984), 286-308.

5. Saari, D. G., The source of some paradoxes from social choice and probability, NU Center for Math Econ Discussion paper #609, June, 1984.

6. Saari, D. G., The optimal ranking method is the Borda Count, NU Discussion paper, Jan. 1985.

7. Saari, D. G., Inconsistencies of weighted voting systems, Math of Operations Research 7 (1982), 479-490.