Discussion Paper No. 745

A DICTIONARY FOR VOTING PARADOXES

by

Donald G. Saari
Department of Mathematics
Northwestern University
Evanston, IL 60208

July 1987

Ever since K. Arrow [1] proved it is impossible to construct a voting system that satisfies certain desired properties, a major focus in social choice has been to use an axiomatic formulation to determine what assumptions are, or are not mutually compatible. (See, for example, Sen [22].) In this paper, I introduce a different approach to analyze the important class of positional voting methods, such as the commonly used plurality vote. The idea is to characterize the election outcomes. Namely, for any number of candidates and for any positional voting procedure, I characterize all possible ways election rankings can arise over all possible subsets of candidates. With this calalog, the properties of these voting systems can be determined in a simple, pragmatic fashion - just check the listings to see what can and cannot occur. The conclusions are very disturbing - paradoxes are more plentiful and much more complicated than one might have anticipated. Only **Borda´s Method** avoids many of the potential flaws. Applications of this dictionary of voting outcomes are indicated, in part, by describing all possible plurality election outcomes, by obtaining new results about agendas and runoff elections, and by describing certain strategic situations. Moreover, because I am characterizing *all* possible election outcomes, it follows that all of the election paradoxes in the literature described in terms of ordinal rankings of positional elections must be special cases of this catalog. This is true, and, by using the listings, I show how any such paradox can be extended and generalized in several different ways. Other, quite spectacular paradoxes can be created; indeed, with the dictionary, the kinds of paradoxes that now can be designed are limited only by one´s imagination.

Probably the most widely used voting method is a plurality election, but how should we interpret the election ranking? To see that there is a problem, consider the hypothetical situation where fifteen people select a common luncheon beverage. Six of them have the ranking water (wa) over wine (wi) over beer (be) (i.e., wa>wi>be), 5 have the ranking be>wi>wa, and 4 have the ranking wi>be>wa. The plurality ranking is wa>be>wi with the tally 6:5:4. Nevertheless, these same people prefer the bottom ranked alternative, wine, both to the top ranked water (by 9:6) and to the second ranked beer (by 10:5)! Even beer is preferred to water (by 9:6). Thus wine, the *majority* or *Condorcet winner*, (in any pairwise comparison, it is selected by a majority of the voters) is bottom ranked in the election while water, the *anti-majority alternative*, is top ranked. By using the binary, majority vote comparisons, it is arguable that the "true ranking" is wi>be>wa - the exact reversal of the election ranking. So, which is the correct ranking? In a runoff election, beer would win the runoff between water and beer. Is water, wine, or beer the preferred beverage?

Plurality elections are not the only procedures plagued by paradoxes; they occur with all *positional voting methods*. These are the election procedures that are equivalent to using a *voting vector* $\mathbf{W} = (w_1, \ldots, w_n)$, $w_j \geq w_{j+1}$, $w_1 > w_n$, in the following way. After each voter ranks the n alternatives, the ballots are tabulated by assigning w_j points to a voter's j^{th} ranked alternative, $j=1, \ldots, n$. The group's ranking is determined by assigning higher rankings to alternatives with larger tallies. Thus, a plurality election is identified with the voting vector $(1,0,\ldots,0)$, while the *Borda Count*, BC, is defined by $(n-1, n-2, \ldots, 1, 0)$. For convenience, assume that the weights, w_j, are rational numbers. Clearly, this does not impose any practical restrictions.

It is natural to wonder whether other choices of voting vectors make a difference. Are some vectors better than others? Can a beverage paradox occur if one uses (5,2,0) instead of (1,0,0)? Is there a choice of **W** that avoids the beverage paradox, or maybe some other paradox? What are all possible paradoxes? (Two nice surveys for what currently is in the literature are Niemi and Riker [13] and Moulin [11].) What about those more complicated election procedures that use election rankings as component parts? For instance, the winner of a runoff election, an agenda, or a tournament is determined by the voters' positional rankings of several subsets of the candidates; what can happen here? With our dictionary of election results, we can answer all questions of this kind. To indicate how more complex election methods can be analyzed, in Section 2 some new results about agendas and runoff elections are given. Furthermore, Saari and Van Newenhizen [20] used the techniques derived for this current paper to discover certain new properties of approval voting, cumulative voting, the effects of truncated ballots, and other multiple systems. (Also see the exchange of opinions by Brams, Fishburn, and Merrill [3] and by Saari and Van Newenhizen [21].)

A more subtle reason for studying paradoxes is that positional voting serves as a simple, but important prototype for many kinds of systems. Thus, should something unexpected occur with positional voting, then it probably occurs, for related reasons, elsewhere. To illustrate, positional voting along with probability, statistics, economic indices, etc., are special cases of aggregation procedures. So, are voting paradoxes related to certain difficulties in these areas? Positional voting is a simple "economic message system" of the type introduced by L. Hurwicz [8] (also see [9,14]) where the object is to encode and transmit relevant information about each agent's preferences. In voting, the encoding is given by marking the ballot. Can voting paradoxes suggest hidden flaws in other kinds of message systems? (The answer to these questions is yes.) Because positional voting is a simple, important prototype, it serves as a test case for concepts being developed in decision analysis, the social sciences, and elsewhere. By understanding what "goes right" and what "goes wrong" with voting, insight can be gained about more complex methods as well as other social choice models. The approach developed here extends, in part, to these other systems.

The central theme of this paper is to determine what can go wrong with positional voting and to explain why. To understand what paradoxes can occur and to avoid the standard approach of finding them in a piecemeal fashion, I characterize *all* possible election outcomes over *all* possible subsets of candidates for *all* possible positional voting methods and *all* possible profiles of voters. The reader will recognize the similarity of this goal with the Sonnenschein program [23] where he, Mantel [10], Debreu [4], and others characterized (for the message system of price dynamics) all possible aggregate excess demand functions for all numbers of commodities for all simple trading economies based on neoclassical utility functions. A catalog, or ***dictionary*** for voting outcomes, could be used in much the same way as the Sonnenschein-Mantel-Debreu classification; both serve as a starting point to determine what else can and cannot occur. In this manner, a dictionary serves as the foundation to analyze voting procedures. By using the dictionary, it is easy to create new paradoxes - just check the listing to find what unexpected rankings occur over different subset of candidates with the same sincere voters. We can compare and combine paradoxes into classes - paradoxes that depend on similar dictionary listings probably are related. We can understand strategic voting - just compare the election rankings for nearby profiles of voters. (After a manipulating voter marks the ballot, the actual election is determined by a profile that differs from the sincere one.) All of this is illustrated here.

For reasons explained in a companion paper [15], it is not practical to list the entries of a dictionary. Therefore, two different approaches are developed to determine whether a particular election ranking is admitted. The first one, emphasizing "what can go wrong", is given here. The second, emphasizing "what can go right", is part of a more technical, group theoretic development that is started in [15]. Also, two different kinds of dictionaries are described. Each entry in the one presented here, the *abridged* dictionary, specifies how the election results for the same, sincere votes varies over all subsets of candidates.

One reason I call these catalogs "dictionaries" is to invoke the image of a reference tool.[1] This image underscores that my principal goal is to characterize what can happen, rather than to advocate one system over another. (However, the BC does have properties significantly more favorable than any other system. In [15], the properties of the Borda Dictionary are given.) For instance, if the goal were to promote a system, I should describe how to break tie votes. For a dictionary, this is inappropriate because a tie vote is a possible outcome, so it must be included. Indeed, as shown in Section 5, listings with tie rankings play a major role in the analysis of strategic behavior and related voting issues.

2. THE ABRIDGED DICTIONARY AND APPLICATIONS

To make the notion of a dictionary more precise, note that $n \geq 2$ alternatives $\{a_1, \ldots, a_n\}$ define 2^n different subsets of alternatives. One of them is empty and n of them have only one alternative. This leaves $2^n-(n+1)$ subsets with enough (at least 2) alternatives to be ranked with an election. List the subsets in some manner, and label them as $\{S_1, \ldots, S_{2^n-(n+1)}\}$. For convenience, assume that the first $n(n-1)/2$ subsets are the pairs of alternatives and the last subset, $S_{2^n-(n+1)}$, is the set of all n alternatives. Next, for each $j=1, \ldots, 2^n-(n+1)$, choose a voting vector, $\mathbf{W}_j$, to tally an election for the subset

----

1. Another reason for this terminology comes from the fact that the ideas for these results are derived from concepts in "chaotic dynamical systems". Therefore, I adopted some of the notation from "symbolic dynamics". An expository description of this connection is given in Saari [19].

of candidates S_j. Let the *system vector*, $\mathbf{W}^n = (\mathbf{W}_1, \ldots, \mathbf{W}_{2^n-(n+1)})$, be the listing of these tallying procedures.

Example 1. For n=3, $\{\{a_1, a_2\}, \{a_2, a_3\}, \{a_1, a_3\}, \{a_1, a_2, a_3\}\}$ is a listing of the $2^3-4 = 4$ subsets of two or more candidates. The system vector $\mathbf{W}^3 = (1,0;1,0;1,0;2,1,0)$ signifies that the first three sets of candidates are tallied with a majority vote - the voting vector is (1,0) - while an election for S_4 is tallied with the BC vector (2,1,0).

To describe our results, we need a space of "all possible election outcomes". Toward this end, let R_j be the listing of all possible rankings of the $|S_j|$ alternatives that are generated by complete, transitive, symmetric (to admit tie votes) binary relationships on S_j. For instance, $S_1 = \{a_1, a_2\}$, so $R_1 = \{a_1 > a_2,\ a_1 = a_2,\ a_2 > a_1\}$. If $|S_j| = 3$, then R_j has 13 rankings - 6 of them are without any ties, while each of the other 7 have at least one tie. Let U^n, the *universal set*, be the cartesian product $R_1 x..xR_{2^n-(n+1)}$. An element of U^n is a listing of $2^n-(n+1)$ rankings; there is one for each subset of candidates. The j^{th} ranking, or *symbol*, of this listing is a ranking for S_j.

Example 2. The sequence $\{a_1 > a_2,\ a_2 > a_3,\ a_3 > a_1,\ a_2 > a_1 = a_3\}$ is an element of U^3. Each symbol is the ranking for the appropriate subset of alternatives.

A *profile* is a listing of each voter's linear ranking of the n candidates. Let P^n be the space of all possible profiles of the n alternatives; we impose no restrictions on the (finite) number of voters. Once a profile, **p**, from P^n is given, then, in the obvious manner, $\mathbf{W}^n$ is used to uniquely determine the election rankings for each of the $2^n-(n+1)$ subsets of candidates. This listing of $2^n-(n+1)$ rankings is called a *word*, and a word is an element of U^n. There is an important difference between a word and an element of U^n; an element of U^n might be an arbitrary listing of rankings that has nothing to do with elections, but a *word generated by* $\mathbf{W}^n$ *is a list of election rankings that is attained with a profile of voters.* For instance, in the beverage example, the election results {wi>wa, wi>be, be>wa, wa>be>wi} is a *word* in U^3 generated by the system vector (1,0;1,0;1,0;1,0,0) because these rankings are attained with the specified profile. Let

2.1 $f(-, \mathbf{W}^n): P^n \longrightarrow U^n$

be the mapping that determines the word for a given profile.

Definition. **Let $n \geq 2$ alternatives and the system vector, W^n, be given. Let** 2.2 **$D(W^n) = \{f(p,W^n): p \text{ in } P^n\}$. The subset $D(W^n)$ of U^n is called the (abridged) *dictionary* generated by W^n.**

Each word in a dictionary is a listing of the election results over the $2^n-(n+1)$ subsets that results from the same profile of voters. By considering all possible profiles, a dictionary becomes our catalog of all possible words; i.e., all possible election rankings. There are certain words we want in all dictionaries because they offer no surprises - the election outcome over each subset remains consistent with the ranking of all n candidates.

Definition. **A word is *binarily consistent* iff the ranking for each subset of candidates is generated by the *same* complete, transitive, binary relationship.**

Example 3. For the set in Example 1, $\{a_1=a_2, a_2 \succ a_3, a_1 \succ a_3, a_1=a_2 \succ a_3\}$ is binarily consistent while $\{a_1 \succ a_2, a_2 \succ a_3, a_1 \succ a_3, a_1=a_2 \succ a_3\}$ is not.

Proposition 1. **Let $n \geq 3$ and let W^n be given. If w is a binarily consistent word, then w is in $D(W^n)$.**

A corollary of this proposition is that *all of the remaining words in a dictionary introduce inconsistency in the election results* over the subsets of candidates. These extra words create the unexpected rankings, or paradoxes. Thus, we might hope that $D(W^n)$ is a small subset of U^n clustered around the binarily consistent words. Our first theorem, which completely quenches this hope, gives the generic characterization for the dictionaries. For this statement, recall that the system vector, W^n, is a vector in an Euclidean space. Also recall that an algebraic set is a lower dimensional subset determined by the zeros of a finite number of polynomials.

Theorem 1. **Let $n \geq 3$. There is an algebraic set α^n such that if W^n is not in α^n, then**

2.3 $$D(W^n) = U^n.$$

In particular, if all of the voting vector components of W^n are plurality vectors,

then Eq. 2.3 is satisfied.

So, for almost all choices of $\mathbf{W}^n$, *anything* can happen. This means that the wildest paradox one could possible imagine actually can occur for almost all choices and combinations of voting vectors. Spectacular paradoxes now are eacy to create.

Example 4. 1. There exists a profile of voters so that the plurality outcome is $a_4 \succ a_3 \succ a_2 \succ a_1$, but the *same* voters' election ranking for *all* other subsets of candidates is the exact reverse of this -- the other rankings are generated by the reversed binary relationship $a_j \succ a_k$ iff $j<k$.

2. Let $n>3$. There exists a profile of voters so that their election rankings alternate with the number of candidates -- if a subset has an even number of alternatives, its plurality rankings are generated by $a_j \succ a_k$ iff $j<k$; if a subset has an odd number of candidates, its $(1,1,0,..,0)$ rankings (vote for your two top ranked candidates) is generated by the *reversed* relationship $a_j \succ a_k$ iff $k<j$.

Theorem 1 includes and significantly extends many results in the literature. For instance, a widely quoted example due to P. Fishburn [5] is where the group's plurality ranking is $a_1 \succ a_2 \succ a_3 \succ a_4$, but if a_4 is removed, then the same group's plurality ranking now is $a_3 \succ a_2 \succ a_1$. Saari [16] showed that all results of this kind could be extended in many different ways - there could be any number of candidates, one could use any choice of voting vectors, $\mathbf{W}_n$ and $\mathbf{W}_{n-1}$, the number of candidates that are removed is arbitrary and could involve all sets obtained in this manner, and the rankings could be selected in an arbitrary fashion. But, both Fishburn's and Saari's statements are very special cases of Theorem 1. To see this, note that for four alternatives, there are $2^4-5 = 11$ different subsets of 2 or more alternatives, so each word has 11 different symbols. Fishburn's example specified just 2 of them. According to the theorem, one could fill in the remaining 9 symbols in *any* desired manner, and there is a profile to support it. As an illustration, by choosing appropriate rankings for the pairs of alternatives, Fishburn's example can be extended so that a_2, the middle ranked alternative in the two elections is the Condorcet winner, but the winner of the first election, a_1, is the antimajority candidate while the winner of the second election, a_3, *almost* is the antimajority candidate because it only beats a_1 in the pairwise matches.

My earlier result [16] provides more freedom in the selection of the symbols over more subsets of alternatives. Still, in [16], the subsets must be obtained from other sets by dropping candidates. Thus, for any n, only for n-1 symbols are used. (As we will see starting with Example 5, this limits the applicability of my earlier results.) Theorem 1 asserts that much more is possible; you can select any rankings for the remaining 2^n-2n subsets of candidates, and there is a profile that satisfies all of the conditions.

By using Theorem 1 in the fashion as just described, the election ranking paradoxes described in the Introduction and many of the examples found in the surveys [11,13] can be significantly extended in many different ways, they can involve far more subsets of candidates, with arbitrary selection of rankings, and the conclusion holds for almost any choice of voting vectors. Corollary 1.1 is an extreme case.

Corollary 1.1. For each of the $2^n-(n+1)$ subsets, use a random number generator to determine the ranking. For almost any choice of voting vectors for the subsets of candidates, there is a profile of voters so that, for the *same* voters, their election outcome for each of the subsets coincides with the randomly generated result.

As restated by Corollary 1.1, the conclusion of Theorem 1 is most disturbing! It is commonly assumed that elections extract some kind of aggregated consensus concerning the ranking of the candidates. It is difficult to accept that an election method accomplishes this goal if the outcomes can depend so sensitively upon which subset of candidates just happen to be presented. Theorem 1 asserts that *this negative feature holds for almost all system voting vectors.* Indeed Theorem 1 and Corollary 1,1 have much the same flavor as the Sonnenschein-Mantel-Debreu result asserting that almost any vector field on the price simplex can be an aggregate excess demand function. Namely, for the aggregation procedures of price dynamics and of voting, anything can happen. (Related arguments explain both results.)

In the next example, the new results about runoff elections and agendas are meant to suggest how Theorem 1 can be used to analyze more complicated election procedures and other issues raised in the Introduction.

Example 5. 1. One form of a runoff election starts by first ranking the original n candidates with a positional election. The k_1 top ranked candidates are advanced to the next stage to be reranked with another positional election. If $k_1 > 2$, it may be necessary to have still another runoff election with the k_2 top ranked candidates. Indeed, if n is sufficiently large, one could imagine a process involving several elimination stages as characterized by the positive integers $\mathbf{k} = (k_1, \ldots, k_s)$, $k_j > k_{j+1}$, $k_s \geq 2$.

Why use other procedures? Why not just accelerate the process by letting $k_1 = 2$; won't the outcome be the same? The (known) answer is no, not necessarily. Indeed, with Theorem 1, it now is easy to extend the known results by showing that for almost all choices of $\mathbf{W}^n$, there are profiles of voters where different choices of the sequence $\mathbf{k}$ lead to completely different election results. In fact, it now is easy to prove *there is a profile and n-2 choices of runoff procedures so that when the j^{th} procedure is used, a_j is the winner, $j=1,\ldots,n-2$.*

I will illustrate the assertion for n=4; the same proof holds for all values of n. For n=4, there are only two runoffs: (3,2) and (2). First, choose a word with the symbols $a_1 > a_2 > a_3 > a_4$, $a_1 > a_3 > a_2$, $a_1 > a_3$, and $a_2 > a_1$. With these election rankings, the first (3,2) runoff is among $\{a_1, a_2, a_3\}$, and a_1 wins the second runoff between a_1 and a_3. On the other hand, a_2 is the winner of the (2) runoff between a_1 and a_2. This completes the proof because, according to Theorem 1, there are profiles that define this word. (To prove this assertion, we need to select rankings for subsets of candidates that are not admitted by [19].)

The same approach works for all values of n and for almost all voting vectors. The idea is simple: different elimination procedures cause different subsets of candidates to be reranked. But, if there is so much as a one candidate difference between subsets, their rankings can be chosen in any desired manner; there need not be any consistency among them. By choosing the rankings in an appropriate manner, we can prove that radically different outcomes exist. Theorem 1 asserts that, for almost all choices of $\mathbf{W}^n$, a profile exists to support the selected rankings.

2. The runoff example did not use all of the symbols in a word. These extra symbols introduce added flexibility to find even more surprising conclusions. For instance, for all pairs of alternatives $\{a_j, a_4\}$, choose the

ranking $a_4 > a_j$. This illustrates a new feature. *For four candidates, there exists a profile of voters so that i) with the (3,2) runoff, a_1 is elected, ii) with the (2) runoff, a_2 is elected, but iii) the first candidate to be eliminated by either procedure, a_4, is the Condorcet winner.* The same kind of statement holds for all values of n. Indeed, the larger the value of n, the more subsets of candidates there are. Because we do no longer need to find profiles to prove that certain election rankings can occur - Theorem 1 establishes their existence - the larger the value of n, the easier it is to creatively design new examples, conclusions, and paradoxes. This is the exact opposite of the current situation based on finding profiles. Creating a profile with the desired election rankings can be a difficult task. Consequently, it is not surprising that many of the examples in the literature are restricted to n=3,4, and use only plurality voting. Theorem 1 removes all of these restrictions.

3. An *agenda* is a listing of the candidates, say $[a_3, a_1, \ldots, a_n]$. A majority election is held between the first two listed candidates, and the winner is advanced to be compared with the third listed candidate. This iterative procedure is continued, and the final candidate is the *winner*. Can the choice of an agenda affect who is selected as the winner? Yes, and this known result becomes obvious with Theorem 1. The winner at each stage determines the pair of candidates to be matched at the next stage. So, by choosing the rankings of the pairs appropriately, we can rig situations where each alternative wins with some choice of an agenda. This happens with the cycle $a_1 > a_2$, $a_2 > a_3, \ldots, a_{n-1} > a_n$, $a_n > a_1$. Here, a_j wins with the agenda $[a_{j+1}, a_{j+2}, \ldots, a_j]$. (When a subscript j+k exceeds n, replace it with (j+k)-n.)

Cyclic results of this kind are well known, but to create the example, only n of the $2^n-(n+1)$ symbols are specified. According to Theorem 1, the remaining symbols can be specified in any desired manner. For instance, one choice asserts there is a profile manifesting this cyclic agenda property even though it is arguable that a_1 should win because for *all* subsets of three or more candidates the plurality ranking is obtained from $a_1 > a_2 > \ldots > a_n$, and because a_1 wins a majority vote in all pairwise comparisons except when a_1 is compared with a_n (so a_1 *almost* is a Condorcet winner).

A different choice of the symbols demonstrates conflict among the agenda results, runoff elections, etc. After all, the agenda example uses only n of the

$n(n-1)/2$ symbols for the pairs, so the rankings of the remaining pairs could be selected to determine different winners of various runoff procedures, etc. Namely, for $n \geq 4$, *there is a profile of voters, n agendas, and n-2 runoff election procedures so that when the* j^{th} *agenda is used, the outcome is* a_{j+2}, $j=1,\ldots,n$, *and when the* k^{th} *runoff procedure is used, the outcome is* a_k, $k=1,\ldots,n-2$. This example can be enhanced by adding "almost Condorcet winners" to exhibit certain features, etc. Actually, by using Theorem 1, the kinds of examples that can be created are limited only by one´s imagination.

3. ANY GOOD NEWS?

Theorem 1 proves that a lot can go wrong with positional elections, but it also hints there may be good news. It suggests there might be a lower dimensional subset of $\mathbf{W}^n$´s where the outcomes don´t depend so sensitively upon which subset of candidates just happen to be presented. It suggests there are choices of system vectors, $\mathbf{W}^n$ where $D(\mathbf{W}^n)$ is a proper subset of U^n. But, because α^n is an algebraic set, it also means that the component voting vectors of a favorable $\mathbf{W}^n$´s must be carefully coordinated to avoid the negative aspects of Eq. 2.3. Theorem 2 characterizes α^n for $n=3,4$. (An extension to larger values of n requires a different technical development.) First, some preliminaries.

Definition. **A voting vector $W_n = (w_1,\ldots,w_n)$ is a *Borda Vector* iff $w_j - w_{j+1}$ is the same nonzero constant for $j=1,\ldots,n-1$. Let B^n denote the system vector where Borda vectors are used to tally all subsets of three or more alternatives.**

It is easy to see that an election tallied with a Borda Vector and with the BC always have the same ranking. This manifests the fact that an election tallied with $\mathbf{W}$ and with $\mathbf{W}' = a\mathbf{W} + b(1,\ldots,1)$ always agree if $a>0$. These modifications affect the tally, but not the relative rankings of the candidates.

Theorem 2. **1.[17] For $n=3$, $D(\mathbf{W}^3)$ is a proper subset of U^3 iff $\mathbf{W}^3$ is a Borda Vector.**

2. For $n=4$, if $D(\mathbf{W}^4)$ is a proper subset of U^4, then either at least one of the voting vector components of $\mathbf{W}^4$ is a Borda Vector, or the last vector component

$\mathbf{W}_{11} = (w_1, w_2, w_3, w_4)$ satisfies the algebraic condition

3.1 $$w_1 - 3w_2 + 3w_3 - w_4 = 0.$$

In other words, *only* the BC, or some extension of it (Eq. 3.1) can provide any relief from these paradoxes. For n=3, this result is given in Saari [17] with some partial results in Fishburn [6]. The statement for n=4 is new, and a stronger conclusion is in [15]. If Eq. 3.1 applies, then the rankings of the set of all four candidates is related to the rankings of the other sets, but the rankings of these other sets need not have any coordination.

Corollary 2.1. Let 4 candidates be given, and let the four sets of three candidates be tallied with non Borda methods. For any choice of rankings of the four sets of three candidates and for any choice of rankings for the 6 pairs of candidates, there is a profile for which all of these rankings are the election outcomes.

To illustrate Corollary 2.1, generalize the notion of an agenda to have the first three listed candidates ranked in an election, and the top two ranked candidates advanced to be compared in second election with the last listed candidate. If non Borda methods are used, the same cyclic effect that occurs with the usual agendas is obtained. Consequently, because there four possible agendas of this kind (determined by who is the last listed candidate), there is a profile of voters so that when the j^{th} generalized agenda is used, a_j wins. We have not specified any of the rankings for the pairs, so they can be chosen to show that, say, a_1, is a Condorcet winner, or, say, that the pairs have a cycle and that the a_2 almost is a Condorecet winner, etc. In other words, only if the BC is used to rank the elections need there be consistency among the election rankings of the sets.

As it will be clear from the proof of Theorems 2 and 3, for all values of n, paradoxes are avoided iff some voting vector is either a Borda Vector, or closely related to it. If $\mathbf{W}^n$ is in α^n - so its dictionary avoids some paradoxes - then the voting vector components of $\mathbf{W}^n$ must be closely related to BC; if they are not Borda Vectors, then some voting vector component must satisfy an algebraic condition like Eq. 3.1. These extensions of Eq. 3.1 *always* include the BC as a special, singular case. Moreover, it is easy to show that if a system vector has a Borda vector as a component voting vector, then the dictionary avoids certain paradoxes. (The proof of this statement is an elementary extension of the proof of Theorem 2.) In fact, the next statement demonstrates an even stronger, favored feature enjoyed by BC for all values of n.

Theorem 3. **Let $n \geq 3$, and let the system vector $\mathbf{W}^n$ have at least one non Borda vector for a subset of at least three candidates. Then,**

3.2 $$D(\mathbf{B}^n) \subsetneqq D(\mathbf{W}^n).$$

Theorem 3 means that Borda's method admits fewer paradoxes, or words, than any other choice of system vectors. (The exact dictionary, $D(\mathbf{B}^n)$, is characterized in [15].) Eq. 3.2 has many important implications. It means that *any fault or paradox admitted by Borda's method also must be admitted by all other positional voting methods.* For instance, it follows from [15] that Borda's method need not rank a Condorcet winner in first place; e.g., for n=3, the word $\{a_1 \succ a_2, a_2 \succ a_3, a_1 \succ a_3, a_2 \succ a_1 \succ a_3\}$ is in $D(\mathbf{B}^3)$. Consequently, from Theorem 3, this same word *must* be in all other dictionaries, $D(\mathbf{W}^3)$. Thus, *any criticism of Borda's method advanced by means of election rankings also serves as a criticism for all possible voting vectors.* Conversely, all other $\mathbf{W}^n$'s admit words (i.e., paradoxes) that are not permitted by Borda's method. Namely, all other system vectors introduce additional indeterminacies - this means that the resulting election outcomes can be far more sensitive to which subgroup of candidates just happen to be presented. These new electoral difficulties introduced by the other system vectors are not admitted by Borda's Method. These statements serve as strong arguments for using Borda's method over any other choice.

4. WHAT ELSE CAN HAPPEN?

Theorem 1 is not the ultimate description of positional voting paradoxes. To show this, we offer two theorems, both in two parts, to indicate what else can happen; a more complete description is planned for elsewhere. (The proof of Theorem 4 motivates the proofs of Theorems 1-3.) While these theorems are based on the ideas developed for Theorem 1, they are of independent interest because they show how the same voters´ election rankings for the same set of candidates can change along with the voting vector.

Theorem 4. 1. **[16] Let $n \geq 3$ and let $W_1, .., W_{n-1}$ be voting vectors that, along with (1,..,1), span R^n. Choose n-1 rankings of the n candidates. There exists a profile of voters so that the election outcome is the j^{th} selected ranking when W_j is used to tally the ballots, j=1,..,n-1.**

This means that if the voting vectors are linearly independent, there need not be any relationship whatsoever among the resulting election ranking for the same voters. This already can be seen with the beverage example where the Borda ranking is wi>be>wa - the exact reversal of the plurality ranking. Theorem 4 means, for example, that there is a profile of voters so that their plurality outcome is $a_1 > a_2 > a_3 > a_4$, their (1,1,0,0) outcome is $a_4 > a_3 > a_2 > a_1$, while their (1,1,1,0) outcome is $a_4 > a_1 > a_3 > a_2$. Not much consistency here.

Extending Theorem 4 over all $2^n-(n+1)$ subsets would create a "super version" of Theorems 1 and 3. This would allow us to compare not only how one profile effects the rankings of each subset of candidates, but also how this same profile can affect the rankings over each subset as the choice of the voting vector varies. However, we would like to analyze a wider range of situations. To suggest what else is useful, consider the *Coombs runoff system* [11] where, first, the candidates are ranked with a positional election method. Then, some candidates are dropped, and the remaining candidates are reranked with another positional voting election. The Coombs system differs from the runoffs discussed in Section 2 in that the dropped candidate is the one with the largest number of last place votes. (Is this equivalent?) This is the candidate with the largest (0,..,0,1) tally. The second part of Theorem 4 specifies some of the possibilities because voting vectors can be replaced with any vector, including (0,..,0,1). Theorem 4.2 shows that the same voters´ election rankings can vary in an arbitrary fashion not only over different subsets of candidates, but also over any one of the subsets as the choice of the voting vector varies.

Theorem 4. 2. Let $n \geq 2$. Consider the n-1 sets $S_j = \{a_1, .., a_{j+1}\}$. For each j=1,..,n-1, choose j vectors in R^{j+1} that, along with (1,..,1), form a linearly

independent set. for each $j = 1,..,n-1$, choose j rankings of the alternatives in S_j. There exists a profile of voters so that when the j^{th} set is tallied with the k^{th} vector chosen for S_j, the outcome is the k^{th} selected ranking for S_j, $j=1,..,n-1$; $k=1,...,j$.

Example 6. 1. Let n=4. For S_3 choose the vectors (1,0,0,0), (3,2,1,0), (0,0,0,1) and the rankings $a_4 > a_3 > a_2 > a_1$, $a_3 > a_1 > a_2 > a_4$, $a_1 > a_2 > a_3 > a_4$; for S_2 choose (1,0,0) and (0,0,1) with the two identical rankings $a_1 > a_2 > a_3$; and for S_1 choose (1,0) and $a_1 > a_2$. According to Theorem 4.2, there is a profile that realizes all of these outcomes with changes in the choice of the tallying vectors. So, the plurality elections for S_1 and S_2 show some consistency, but the plurality ranking for S_3 is the reverse of what one might expect. On the other hand, the ranking obtained by voting for your bottom ranked alternative, (0,0,0,1), resumes this consistency.

2. This theorem can be used to show how the Coombs, the (3,2), and the (2) runoffs all give different outcomes. Just choose the purality ranking to be $a_1 > a_3 > a_2 > a_4$, the plurality ranking of $a_2 > a_1 > a_3$, and the majority ranking of $a_2 > a_1$. So, the winner of the (3,2) runoff is a_2, while the winner of the (2) runoff is either a_1 or a_3. Now, choose the (0,0,0,1) ranking to be $a_1 = a_2 = a_3 > a_4$. Is only a_4 advanced by the Coombs system?

Theorem 4.2 is not sufficient to completely analyze runoffs and other, more complicated procedures, because it does not admit all subsets of candidates. For instance, in Example 6.1, if the first stage of the election is ranked with the BC, then S_2 is the set of candidates that is to be reranked. However, if a Coombs method is used, then $\{a_2, a_3, a_4\}$ needs to be reranked, and this subset is not admitted by Theorem 4.2. Theorem 5 is a step toward a more general result.

Theorem 5. 1. Let n>3, and let F be the family of subsets of candidates that consists of all n(n-1)/2 pairs of candidates and the set of all n candidates. Choose n-2 vectors in R^n that, along with (1,..,1) form a linearly independent set. Furthermore, suppose the span of the n-2 vectors and (1,1,..,1) do *not* include a Borda Vector. Choose a ranking for each of the pairs and choose n-2 rankings for the set of n alternatives. There is a profile of voters so that for each pair of alternatives, their majority ranking is the selected one. When their

ballots for the set of n candidates is tallied with the j^{th} vector, the outcome is the j^{th} selected ranking.

Again, if we avoid the Borda Vectors, anything can happen. An implication of this statement is that runoff elections can have problems even with the same set of candidates and the same profile of voters if different methods to tabulate the ballots are considered.

Example 7. Let n>4. Consider the (2) runoff system and the voting vectors $\mathbf{V}_j = (1,1,..,1,0,...,0)$, $j = 1,..,n-2$, where j specifies the number of 1´s. There exists a profile of voters so that when $\mathbf{V}_j$ is used to tally the ballots of the n candidates, a_j is the winner of the runoff election, $j=1,..,n-2$, and a_n is the Condorcet winner.

To prove this statement, just choose n-2 rankings for the set of n candidates where the j^{th} ranking starts as $a_{n-1} \succ a_j \succ ..$, $j=1,..,n-2$. (Fill in the rest of the ranking in any desired manner.) The j^{th} ranking will be the election outcome for $\mathbf{V}_j$, so, for each j, the runoff is between a_{n-1} and a_j. Choose the rankings $a_j \succ a_{n-1}$ for each j, so that a_j wins the j^{th} runoff. For the pairs (a_n, a_j), choose the rankings $a_n \succ a_j$. By Theorem 5.1, a profile exists that satisfies all of these outcomes, and this completes the proof.

The statement of the theorem prohibits a Borda Vector from even being in the *span of the other vectors.* If it is, then it follows from the froof that many paradoxes no longer are possible. Again, this indicates the power of the BC.

Theorem 3 asserts that Borda´s Method has fewer paradoxes and words than any other choice of a system vector, and that any word in the Borda Dictionary must be in all other dictionaries. It does not follow from Theorem 3 that one profile can give the same word for all dictionaries. The last part of Theorem 5 corrects this. It has much the same effect as Theorem 3 by underscoring the essential, positive role of Borda´s Method.

Theorem 5.2. Let $n \geq 3$. **Select a word from $D(B^n)$. There exists a profile of voters so that for each subset of candidates, the group´s election ranking for this set is the selected Borda ranking. This is true independent of which voting vector is used to tally the voters´ ballots.**

This result generalizes and extends a nice example of Fishburn's [7]. Fishburn created a profile for n=3 to prove that there exist situations where the Condorcet winner never is elected by any positional voting method. *Theorem 5.2 extends this kind of statement in all possible ways.* It asserts that for any n, you can select *any* feature of Borda's method that can be expressed in terms of the rankings over subsets of candidates. This feature defines a word in the Borda dictionary, $D(\mathbf{B}^n)$. Then, according to this corollary, there exists profiles where the same feature holds for all possible positional voting methods. Thus, after $D(\mathbf{B}^n)$ is characterized in [15], all sorts of new examples can be created. Of course, this statement is false if a word is selected from any dictionary other than the Borda Dictionary. This is because if this new word is not admitted by BC then the conclusion cannot possibly hold.

Example 8. There exists a profile so that no matter what positional election method is used, the outcome is $(a_1 \succ a_2,\ a_2 \succ a_3,\ a_1 \succ a_3,\ a_2 \succ a_1 \succ a_3)$.

5. THE GEOMETRY OF THE SPACE OF PROFILES

Before turning to the proofs, it is appropriate to question whether these new results are robust, or whether they depend upon specially constructed examples that disappear with even the slightest perturbation of the profiles. They are robust. This answer, based on the following representation of the space of profiles, P^n, uses the fact there are n! different rankings; i.e., there are n! different types of voters. The basic idea can be seen with the beverage example. There is no qualitative change in the example if I replicated the profile by, say, tripling the number of voters with each ranking. This is because differences in the rankings do not depend on the numbers of voters, but on the *ratios* of the number of voters of each type. So, a profile can be characterized by specifying what *fraction* of all voters have a particular ranking. In this manner, a profile $\mathbf{p}$ can be identified with a vector with n! non-negative components that sum to unity. Namely, $\mathbf{p}$ can be viewed as being a vector on the unit simplex, Si(n!), in the positive orthant of $R^{n!}$.

There is a slight technical difficulty. A profile defines a vector in Si(n!) with rational components, and vice versa. Part of the strength of my approach is to use all of the structure of Si(n!); I embed the discrete problem of voting into the continuous one of analyzing mappings with domain Si(n!). To do this, extend, in the natural manner, the definition of $f(\mathbf{p}, \mathbf{W}^n)$ to all $\mathbf{p}$ in Si(n!), even those with irrational components. The only reservation is that the image of f can be treated as an election ranking for a finite number of voters iff $\mathbf{p}$ has rational components. Such profiles are dense in Si(n!), so there are an infinite number of them in any open set.

This representation of voting as a mapping with domain Si(n!) has many advantages, and most are based on the properties of open sets. For instance, if a paradox is supported by an open set in Si(n!), then it must occur with an infinite number of different profiles (not just replications of the original one). Also, such a word is robust in the sense that a profile can be perturbed and the same word results; so such a paradox, or word, cannot be dismissed as being an anomaly. Another implication is based on the fact that the standard probability measures assign positive values to open sets. Thus, if certain probability measures are introduced on the distribution of profiles, an open set corresponds to a positive probability that the paradox occurs. Because a common denominator of $\mathbf{p}$ is the total number of voters and because the structure of an open set determines which $\mathbf{p}$'s are admitted, by knowing the structure of the open set, programming techniques can be used to determine the minimum number of voters required before a particular paradox can occur. In much the same way, the existence of the open sets can be used to answer "limit" questions about how likely it is for a paradox to occur as the number of voters increase, etc. Thus, we need to understand the structure of the sets in Si(n!) that support each word.

Theorem 6. Let $n \geq 3$ and let $\mathbf{W}^n$ be given. Let w be a word from $D(\mathbf{W}^n)$ and let s_j be the symbol for the subset of candidates S_j. If s_j admits no ties, then the set of profiles supporting s_j is an open set of Si(n!). The boundaries of this open set correspond to tie votes, and they are hyperplanes in Si(n!) determined by $\mathbf{W}_j$. If s_j contains a tie ranking among alternatives, then the set of profiles supporting s_j is in a lower dimensional hyperplane in Si(n!). The set of profiles supporting the word w is the intersection of the sets supporting each of the symbols. So, if none of the symbols of w involve a tie, the supporting set of profiles is a nonempty open set in Si(n!). The profile consisting of an equal number of voters of each type is a boundary point of each region supporting each word.

Thus all paradoxes based on election rankings without ties are supported by open sets, so they are robust. The paradoxes involving tie votes are not robust; a slight change in the profile can alter the outcome. On the surface, Theorem 6 appears to be a technical theorem concerned with the robustness of certain words. In fact, when used with the earlier theorems, Theorem 6 is a "gold mine" for explaining, extending, and describing several other results of current interest. For instance, manipulation and strategic behavior is a topic currently of great concern. But note, if a voter successfully manipulates the outcome of an election with a strategic vote, the sincere profile is on one side of the hyperplane given by a tie vote, and the manipulated profile is on the other. Consequently, the structure of these "tie vote" hyperplanes provide valuable information about how susceptible a system is to being manipulated by individuals or small groups. Such an analysis is in Saari [18]. A similar topic would be the *sensitivity* of a system - small changes in how the voters mark the ballots alters the outcome. Using the techniques of [18], it is easy to determine which systems are more sensitive than others.

Results of a different but related flavor concern those fascinating statements asserting that, by voting, a voter *hurts* his or her interests, so by abstaining the voter is better off. The first result of this type that I am aware of concerns a runoff election with a plurality vote, and it was found by Brams and Fishburn [2]. With Theorems 6 and 1, it is easy to extend and characterize all possible methods that have this behavior. To see how this is done, we offer here a partial result that is easy to prove. Toward this end, call a social choice method that selects a single candidate *disjoint* if i) the outcome is based on the system vector, $\mathbf{W}^n$, positional voting rankings of the subsets of candidates, ii) the ranking of some one subset of candidates determines or affects which one of several subset of candidates will be the final set to be ranked - indeed, just the reversal of the relative rankings of some two adjacently ranked candidates can change the choice of the final set, and iii) the final outcome is based on the ranking of the final subset of candidates, and iv) the method is not constant, at least two different outcomes are possible for each choice of a final set. As an example, all runoff procedures, whether of the Coombs type or the more standard kind defined by the integers $\mathbf{k}$ are disjoint methods. This is because by switching the relative ranking of some two alternatives we can change which candidates are advanced to the next stage, and the winner depends on which candidate is top ranked at the last step. An agenda and most elimination procedures are other disjoint methods.

Corollary 6.1. Let $n \geq 3$, let W^n be a system vector that is not in σ^n, and suppose a disjoint social choice method is given. There exists a profile of voters and two voters of the same type where, by voting, the two voters end up with a personally less favorable outcome than if they had abstained.

The proof is similar to those in Section 2. First, I outline the ideas with a standard (2) runoff election for n=3, and then I describe the steps needed for the more general result. Choose two "final" sets, say $\{a_1, a_2\}$ and $\{a_1, a_3\}$. For these sets, choose the rankings $a_1 \succ a_2$ and $a_3 \succ a_1$. The ranking of a third "swing" set, say $\{a_1, a_2, a_3\}$, determines which of the final subsets is used. Now suppose the two voters have the ranking $a_2 \succ a_3 \succ a_1$. By voting, the two voters could change an outcome from $a_1 \succ a_3 \succ a_2$, where the outcome of the runoff would have been their second ranked a_3, to $a_1 \succ a_3 \succ a_2$, where the outcome is their last ranked a_3. To see that such profiles exist, consider the words with the symbols $a_1 \succ a_2$, $a_3 \succ a_1$, and $a_1 \succ a_2 = a_3$. The first two symbols are supported by open sets in Si(n!), and the third by a hyperplane passing through the intersection of these open sets. Now, positional voting is monotonic, so it is a simple exercise using the fact that vectors with rational components are dense to show there is a profile **p** close enough to the hyperplane but on the $a_1 \succ a_3 \succ a_2$ side so that by adding two voters of the specified type, the new profile is on the other side. Because the other two symbols are supported by open sets, this can be done without changing these rankings. This completes the proof. (If a method specifies how to break a tie, then only one voter of a specified type is necessary. However, without more conditions, we need two or more voters to ensure that the original and the new profile are on different sides of the hyperplane.)

The general proof is much the same. We just need three alternatives, a_1, a_2, and a_3 where a_1 and a_3 are possible outcomes based on the rankings of two final sets. Which final set occurs depends on whether a_2 can be advanced one position in the ranking of a third, swing set. These conditions all occur by the definition of a disjoint method. Start with the ranking of the third set having a_2 tied in the swing position. Construct the ranking for the two voters to have a_2 as top ranked, the outcome of the set they *don't* get by voting and advancing the ranking of a_2 as second ranked, and the outcome of the set they do get by advancing a_2 in third place. The rest of the analysis is the same.

In this proof, there are many other symbols that have not been specified, so they can be assigned in any desired manner to prove other conclusions. For instance, for certain disjoint systems, it is possible to show that, in addition to what already has been proved, the Condorcet winner need not get elected, that the outcome of a runoff election differs from these conclusions, etc. (This generalizes a result of Moulin [12].) As an illustration, we can combine several of the features already described.

Corollary 6.2. Let $n>4$ and suppose the plurality vote is used to rank all subsets of candidates. There are i) $n-2$ runoff procedures, ii) n agendas, iii) a profile of voters and iv) two other voters with the same ranking that has a_2 top ranked, and a_1 bottom ranked, so that a) when the original voters use the j^{th} runoff procedure, a_j is the winner, $j=1,..,n-2$, and b) when the k^{th} agenda is used, a_k wins, $k=1,..,n$. If the two additional voters vote, then outcome of all elections remain the same except for the two procedures where a_2 won. In both of these procedures, the new winner now is a_1.

Other kinds of paradoxes, such as explaining why two subcommittees can independently reach the same conlusion, but when joined as a full committee, they select a different alternative can be based on other geometric structures of Theorem 6. Details will be given elsewhere.

6. PROOFS

The proofs of all of the theorems are based on representing the tally of an election as a mapping from the space Si(n!), described in Section 5, to a cartesian product of simplices. To describe the image space, start with the the set of all n candidates, $S_{2^n-(n+1)}$. In R^n, identify the k^{th} component, x_k, with the k^{th} alternative, a_k. For $\mathbf{x} = (x_1,\ldots,x_n)$, let larger values of x_k denote a "stronger preference for a_k. With this identification, the hyperplane $x_k=x_j$ divides R^n into three regions; the two topologically open regions are identified with the strict ordinal rankings (e.g., $\{\mathbf{x}: x_k>x_j\}$ correspond to $a_k>a_j$), and the hyperplane is identified with indifference between the two candidates. By varying k and j over all pairs of indices, the resulting n(n-1)/2 hyperplanes divide R^n into cones that are in a one to one relationship with the ordinal rankings of the candidates. Call each of these regions a *ranking region*. Call a cone with a nonempty interior an *open ranking region*; it is identified with strict rankings among the alternatives. If a ranking region is in a hyperplane, it corresponds to a ranking with indifference among some of the candidates. For instance, $\{\mathbf{x}$ in $R^5: x_1=x_2>x_3>x_4=x_5\}$ corresponds to the ranking $a_1=a_2>a_3>a_4=a_5$.

In what follows, let A denote the ranking $a_1>a_2>..>a_n$ and let $\mathbf{W}$ be the voting vector for the set of all n alternatives. Because of the monotonicity assumptions on the components of a voting vector, $\mathbf{W}$ is in the closure of the ranking region identified with A. (If at least two of the components of $\mathbf{W}$ agree, then $\mathbf{W}$ is on the boundary; otherwise $\mathbf{W}$ is in the interior.) In fact, $\mathbf{W}$ is the tally of a ballot with the ranking A. When used as a tally, denote the vector as $\mathbf{W}_A$. Any other strict ranking of candidates is a permutation of A; let the generic representation of such a permutation be $\pi(A)$. The tally for a ballot marked $\pi(A)$ is given by a permutation of $\mathbf{W}$, denoted as $\mathbf{W}_{\pi(A)}$. Let $n_{\pi(A)}$ denote the fraction of all voters that have the ranking $\pi(A)$. The tally of an election is given by

6.1 $$\Sigma\, n_{\pi(A)}\mathbf{W}_{\pi(A)}$$

where the summation index, $\pi(A)$, ranges over all n! permutations of A. The sum is in a ranking region, and the ranking associated with this region is the election outcome.

The variables $\{n_{\pi(A)}\}$ are all non-negative and sum to unity. This means a profile **p** is a vector in Si(n!), so the election outcome is in the convex hull of the vectors $\{\mathbf{W}_{\pi(A)}\}$. In turn, this hull is in the affine plane passing through these points and c(1,...,1) where c is the sum of the components of **W**. The analysis is much easier when this plane is a linear subspace of R^n. This motivates the first of the two assumptions I impose on the voting vectors. Because the election outcome for **W** always agrees with the outcome for $\mathbf{W}' = a\mathbf{W} + b(1,\ldots,1)$, these two assumptions only fix the values of a and b. The first one fixes the value of b.

VECTOR NORMALIZATION: The sum of the components of a voting vector equals zero.

Example 9. The voting vector for a plurality election is (1,0,..,0), a vector normalized form is (n-1,-1,..,-1). A vector normalized form of the BC, (n-1,n-2,...,1,0) is (n-1,n-3,...,n+1-2i,...,1-n). So, if n=3, the vector normalized form of a Borda vector is (2,0,-2) while it is (3,1,-1,-3) for n=4.

The vector normalization forces the vectors $\mathbf{W}_{\pi(A)}$ to be orthogonal to (1,...,1), so the vote tally, Eq. 6.1, is in the linear subspace of R^n with the normal vector (1,...,1). Let E^n denote this (n-1) dimensional space; it is the space of interest.

Now, consider all $2^n-(n+1)$ subsets of candidates. Corresponding to each set S_j, there is a division of an Euclidean space of dimension $|S_j|$ into ranking regions. For convenience, assume that the s_j coordinates of this space have the same subscript as the alternatives, and that they are listed in increasing order. For instance, if $S_k=\{a_1,a_4,a_7\}$, the the corresponding *two* dimensional linear subspace for the *three* candidates, E^3, has the coordinates x_1, x_4, and x_7.

Let E(n) be the cartesian product of the $2^n-(n+1)$ linear subspaces E^k. A ranking region in E(n) is obtained for the product of the ranking regions of the component spaces. For instance, $\{x_1=x_2,\ x_2>x_3,\ x_1<x_3,\ x_2>x_1>x_3\}$ is a ranking region in E(4) that corresponds to the element $(a_1=a_2,\ a_2>a_3,\ a_1<a_3,\ a_2>a_1>a_3\}$ in U^4. It is easy to see that there is a one to one correspondence between the ranking regions of E(n) and the entries in U^n.

In the obvious manner, the ranking A defines a ranking for each of the subsets of candidates. Let $\mathbf{W}^n$ be the system vector. By the choice of the coordinate axis for each of the component spaces of E(n), $\mathbf{W}^n$ is in the ranking region of E(n) corresponding to A. As above, the different voting vector components of $\mathbf{W}^n$ give the tally for each of the subsets of candidates for a voter with the ranking A. When treated as a tally of a ballot, $\mathbf{W}^n$ is denoted as $\mathbf{W}^n_A$. Any other ranking of A is a permutation of A, $\pi(A)$, and its tally is given by the appropriate permutation of the voting vector components of $\mathbf{W}^n$. This permutation is denoted by $\mathbf{W}^n_{\pi(A)}$. For a profile $\mathbf{p} = \{n_{\pi(A)}\}$, the tally of all subsets of candidates is given by

6.2 $$G(-,\mathbf{W}^n) : Si(n!) \longrightarrow E(n),$$

where

6.3 $$G(\mathbf{p},\mathbf{W}^n) = \Sigma\, n_{\pi(A)}\ \mathbf{W}^n_{\pi(A)}$$

and the summation index, $\pi(A)$, varies over all n! permutations of A. The sum is in a ranking region, and this ranking region defines the word in U^n. This summation has the same interpretation as for Eq. 6.1; it defines a point in the convex hull of the n! vectors $\{\mathbf{W}^n_{\pi(A)}\}$.

The key observation used to characterize the dictionaries is that a word w is in the dictionary $D(\mathbf{W}^n)$ iff the product region associated with w intersects the convex hull of the vectors $\{\mathbf{W}^n_{\pi(A)}\}$. Thus the problem of characterizing a dictionary is equivalent to characterizing how this convex hull intersects the ranking regions in E(n). But, this convex hull is in the linear space, $V(\mathbf{W}^n)$, spanned by the vectors $\{\mathbf{W}^n_{\pi(A)}\}$. Our proofs a based on the following conclusion.

Proposition 2. A word w is in $D(\mathbf{W}^n)$ iff the product ranking region associated with w has a nonempty intersection with $V(\mathbf{W}^n)$.

Proof. First, note that if $\mathbf{p} = (1/n!,\ldots,1/n!)$, then $G(\mathbf{p},\mathbf{W}^n) = \mathbf{0}$; the election ranking for each subset is a complete tie. Next, note that the rank of $G(-,\mathbf{W}^n)$ equals $\dim(V(\mathbf{W}^n))$. This is because $V(\mathbf{W}^n) = \mathrm{span}(\{\mathbf{W}^n_{\pi(A)}\})$, while the column vectors of DG are $\{\mathbf{W}^n_{\pi(A)}\}$). Thus, the rank(DG) = $\dim(V(\mathbf{W}^n))$.

It now follows that G maps an open neighborhood of $\mathbf{p}$ to an open neighborhood, U, of the origin, $\mathbf{0}$, of E(n). But, $\mathbf{0}$ is a boundary point for each product ranking region in E(n). This forces U to meet each of the ranking regions in $V(\mathbf{W}^n)$. Namely, for a ranking region in $V(\mathbf{W}^n)$, there is a point $\mathbf{p}'$ in Si(n!) that is mapped to this particular ranking region. This nearly completes the proof because it means that if $V(\mathbf{W}^n)$ meets a ranking region of E(n), then $G(\mathbf{p}',\mathbf{W}^n)$ is in this same ranking region. All that needs to be proved is that $\mathbf{p}'$ can be found that has rational components. This simple argument is given in the first part of the proofs in [17]. This completes the proof of the Proposition 2.

Proof of Theorem 6. This is a simple exercise using the inverse image of G. If a symbol does not involve any tie, it is in an open ranking region, and the inverse image of an open set is open. Secondly, the structure of these sets with the hyperplanes follows immediately from the linear form of G.

According to the proposition, to prove Theorem 1 we want to show that $V(\mathbf{W}^n) = E(n)$ for most system vectors. My proofs are based on the algebraic, permutation group structure involved in changing a ranking from A to $\pi(A)$. To illustrate the basic idea, I will start with the proof of Theorem 4.1. Here, the group structure is simpler because it involves the permutations of only one set of candidates.

Proof of Theorem 4.1. This theorem involves only the set of all n candidates. The interesting feature is that with the n-1 different tallying processes, the set of outcomes is replicated n-1 times. Therefore the election tally is

6.4 $G(-,\mathbf{W}_1,\ldots,\mathbf{W}_{n-1})\colon Si(n!) \dashrightarrow (E^n)^{n-1}$

where $\{\mathbf{W}_j\}$, $j=1,\ldots,n-1$, are the different voting vectors satisfying the conditions of the theorem and where, if $\mathbf{p} = \{n_{\pi(a)}\}$, then

6.5. $G(\mathbf{p},\mathbf{W}_1,\ldots,\mathbf{W}_{n-1}) = \Sigma\, n_{\pi(A)}\ (\mathbf{W}_{1\pi(A)},\ldots,\mathbf{W}_{n-1,\pi(A)})$

From the assumptions on the voting vectors, for each j, $E^n = \mathrm{span}\{\mathbf{W}_{j,\pi(A)}\}$. To prove the theorem, we need to prove that $\{E^n\}^{n-1}$ agrees with $V^* = \mathrm{span}(\mathbf{W}_{1\pi(A)},\ldots,\mathbf{W}_{n-1,\pi(A)})$.

Let $P'_{jk}(\mathbf{x})\colon E^n \dashrightarrow E^n$ be the permutation mapping that interchanges the j^{th} and the k^{th} components of $\mathbf{x}$, and let $P_{jk}(\mathbf{X}) = (P'_{jk}(\mathbf{x}_1),\ldots,P'_{jk}(\mathbf{x}_{n-1}))$ be a mapping from $\{E^n\}^{n-1}$ back to itself where $\mathbf{X} = (\mathbf{x}_1,\ldots\mathbf{x}_{n-1})$. Let G_p be the group of permutations generated by the n(n-1)/2 permutation mappings $\{P_{jk}\}$, and define $L\{G_p\} = \{V\colon V$ is a linear subspace of $(E^n)^{n-1}$ that is invariant under $G_p\}$. Thus, if V is in $L(G_p)$, and if P is a permutation mapping from G_p, then $P(V)=V$. Such a mapping, P, just permutes the components of the vectors, therefore V^* is in $L(G_p)$. To prove the theorem, I will characterize the elements of $L(G_p)$.

Our characterization of $L(G_p)$ depends on the eigenvalues and eigenspaces of P_{jk}. The eigenvalues are -1 with multiplicity n-1, and 1 with mulitiplicity (n-1)(n-1). It is easy to see that the *-1 eigenspace of (j,k)* is spanned by the $(\mathbf{e}_j-\mathbf{e}_k,\mathbf{0},\ldots,\mathbf{0})$, $\ldots,(\mathbf{0},\ldots,\mathbf{0},\mathbf{e}_j-\mathbf{e}_k)$ where $\mathbf{e}_i$ is the unit vector in R^n with unity in the i^{th} component and zero in all others. That this is the basis follows because only the $\mathbf{e}_j$ and $\mathbf{e}_k$ components are interchanged with the permutation; all others remain the same. Also, elementary arguments prove that the -1 and the +1 eigenspaces are orthogonal to each other.

Claim 1. Let V be in $L(G_p)$, and let V_{jk} be the projection of V into the -1 eigenspace of (j,k). Then, V_{jk} is a linear subspace of V.

Proof of the claim. V_{jk} is a linear subspace, but it is not clear it must be a linear subspace of V. To prove this, let $\mathbf{v}$ be in V and let $\mathbf{v}_1$ be the projection of $\mathbf{v}$. It suffices to show that $\mathbf{v}_1$ is in V. But, by the orthogonality of the two eigenspaces, $\mathbf{v} = \mathbf{v}_1+\mathbf{v}_2$ where $\mathbf{v}_2$ is in the +1 eigenspace. According to the invariance assumption, $P_{jk}(\mathbf{v}) = P_{jk}(\mathbf{v}_1+\mathbf{v}_2) = -\mathbf{v}_1+\mathbf{v}_2$ is in V. Thus,

6.6 $\mathbf{v} - P_{jk}(\mathbf{v}) = 2\mathbf{v}_1$

is in V. This completes the proof.

The next assertion proves that all of the V_{jk}'s are related.

Claim 2. Let V be in $L(G_p)$. Then, $P_{2j}(P_{1k}(V_{jk})) = V_{12}$.

Proof. $P'_{sk}(\mathbf{e}_j-\mathbf{e}_k) = \mathbf{e}_j-\mathbf{e}_s$, so P_{sk} maps the -1 eigenspace of (j,k) onto the -1 eigenspace of (j,s). Thus, $P_{sk}(V_{jk})$ is a linear subspace of the -1 eigenspace of (j,s) with the same vector dimension as V_{jk}. Because V is in $L(G_p)$, the two sets agree.

Claim 3.. Let V be in $L(G_p)$ and suppose $\dim(V_{12}) = j$. V is spanned by $\{V_{1k}\}$ and $\dim(V)=(n-1)j$.

Proof. $E^n = \mathrm{span}\{\mathbf{e}_1-\mathbf{e}_k\}$, k=2,...,n, and this extends, in the natural manner, to define a basis for $\{E^n\}^{n-1}$. So, assume $\dim\{V_{12}\}= j$. According to Claim 2, $\dim(V_{1k}) = j$, k=1,...,n, and V contains the span of these subspaces. Thus, $\dim(V)\geq (n-1)j$. But, if $\dim(V)>(n-1)j$, there is a vector $\mathbf{v}$ in V that cannot be expressed as the linear combination of the vectors from the spaces V_{1k}, k=2,...,n. According to our choice of a basis for $\{E^n\}^{n-1}$, this means there is a k such that the projection of $\mathbf{v}$ into the -1 eigenspace for (1,k) is not in V_{1k}. This contradicts Claim 1 and the definition of V_{1k}, and proves Claim 3.

To complete the proof of Theorem 4.1, it suffices to show that $\dim(V^*_{12}) =$

(n-1) because, from Claim 3, $\dim(V^*) = (n-1)\dim(V^*_{12}) = (n-1)^2 = \dim(\{E^n\}^{n-1})$, so $V^* = \{E^n\}^{n-1}$.

Assume that $\dim(V^*_{12}) = j$ and that a basis for V^*_{12} is $\{\mathbf{c}_i\}$, $i=1,\ldots,j$. Using a standard row reduction argument of the type used to convert a matrix into a diagonal form, we can assume that the basis $\{\mathbf{c}_i\}$ is replaced with the equivalent basis $\{\mathbf{d}^{12}{}_i\}$ where

6.7 $$\mathbf{d}^{12}{}_i = (a^i{}_1(\mathbf{e}_1-\mathbf{e}_2),\ldots,a^i{}_{n-1}(\mathbf{e}_1-\mathbf{e}_2))$$

where $a^i{}_k = 1$ if $i=k$, $i=1,\ldots,j$; $a^i{}_k = 0$ for $k\leq j$ and k different from i; and not all of the remaining $a^i{}_k$ terms can equal zero because this would force V* to be zero for these component spaces. (In turn, this forces the contradiction $\mathbf{W}_s=0$.)

According to Claim 2, a basis for V^*_{1j} is given by Eq. 6.7 where the index 2 is replaced with j; the $a^i{}_k$ terms are the same. This means that the vector $(\mathbf{W}_1,\ldots,\mathbf{W}_{n-1})$ can be expressed as a linear combination of the $\{\mathbf{d}^{1k}{}_i\}$. But if $j=\dim(V^*)<n-1$, it follows that there are j equations, $\mathbf{W}_s = -\Sigma a^s{}_k\mathbf{W}_k$, where the summation is over $k=j+1,\ldots,n-1$, and $s=1,\ldots,j$. This means that the set of vectors $\{\mathbf{W}_j\}$ is linearly dependent. This contradiction proves the theorem.

The proof the Theorem 4.1 is based upon the permutation group structure satisfied by the linear space defined by the vote tally. You can view this in the following manner. If there is a permutation in the voters' preferences, then the new vote tally must also be in V*. The change in the voters' preferences just permuted the entries of the voting vectors. Indeed, the *only* property about a voting vector that was used is that with appropriate choices of P'_{jk}, any $\mathbf{W}_{j\pi(A)}$ can be mapped to some other $\mathbf{W}_{j\pi'(A)}$. Thus, the conclusion holds should voting vectors be replaced with any other vector. The proofs of all of our theorems are proved in much the same fashion; we characterize what happens as the permutation groups act on the domain, or preferences of the voters. The proof of Theorem 4.2 demonstrates there is a difference in the analysis when we have more than one subset of candidates; this difference is basic for all of social choice mappings.

Proof of Theorem 4.2. To demonstrate the basic ideas, I will first prove the theorem where only one voting vector is selected for each set. So, let $\mathbf{W}_{j+1}$ be the voting vector for the set $S_j=\{a_1,\ldots,a_{j+1}\}$, and let $\mathbf{W}_F = (\mathbf{W}_2,\ldots,\mathbf{W}_n)$ be the system vector for this *family of sets*. For each set S_j, there is an associated j dimensional subspace, E^j, of the j+1 candidates. We need to show that the subspace spanned by the system tally vectors, $\{\mathbf{W}_{F,\pi(A)}\}$, is the total space $E^2 x..xE^n$.

Notice that a_n is in only one set, S_{n-1}. Thus, if a_n and an adjacently ranked alternative are interchanged, then the new, permuted ranking only changes the ranking for the set S_{n-1}; the ranking for all other sets are invariant. Now, $\mathbf{W}_n$ is not a multiple of $(1,\ldots,1)$, so there is a value of s where the difference between the s^{th} and the $(s+1)^{th}$ components of $\mathbf{W}_n$ is nonzero. Assume that s is the smallest value for which this is true, and let $w*_n$ be this nonzero difference. (For instance, if $\mathbf{W}=(1,0,\ldots,0)$, then s=1 and $w*_n=1$ while if $\mathbf{W}=(1,1,1,1/2,1/3,0)$, then s=3 and $w*_n=1/2$.) Next, choose a ranking, $\pi(A)$, where a_j and a_n are, respectively, the s^{th} and the $(s+1)^{th}$ ranked candidates. When P_{jn} acts on $\pi(A)$, it interchanges these two alternatives. This permutation creates a new ranking only for the set S_{n-1}; there are no changes in the rankings for any other set. Consequently, $P_{jn}(\mathbf{W}_{F,\pi(A)}) - \mathbf{W}_{F,\pi(A)} = w*_n(\mathbf{0},\ldots,\mathbf{0},\mathbf{e}_n-\mathbf{e}_j)$ Because $w*_n$ is nonzero and because this relationship holds for all choices of j=1,...,n-1, the space spanned by $\{(\mathbf{0},\ldots,\mathbf{0},\mathbf{e}_n-\mathbf{e}_j)\}$ is in the space spanned by $\{\mathbf{W}_{F,\pi(A)}\}$. Namely, 6.8

$0x0x..xE^n$ is in span$\{\mathbf{W}_{F,\pi(A)}\}$.

Next, a_{n-1} is only in the sets S_{n-2} and S_{n-1}. Let $w*_{n-1}$ be the first nonzero difference between weights of $\mathbf{W}_{n-1}$ and assume this is between the s^{th} and the $(s+1)^{th}$ components. For each j, choose a ranking where a_j is s^{th} ranked and a_{n-1} is $(s+1)^{th}$ ranked. When $P_{j,n-1}$ is applied to this ranking, the only changes are in the rankings for S_{n-2} and S_{n-1}. Furthermore, $P_{j,n-1}(\mathbf{W}_{F,\pi(A)}) - \mathbf{W}_{F,\pi(A)} = (\mathbf{0},\ldots,\mathbf{0},w*_{n-1}(\mathbf{e}_{n-1}-\mathbf{e}_j,x_n(\mathbf{e}_{n-1}-\mathbf{e}_j)$ where x_n is some scalar. By use of Eq. 6.8, it follows that the vector $(\mathbf{0},\ldots,\mathbf{0},\mathbf{e}_{n-1}-\mathbf{e}_j,\mathbf{0})$, for each j=1,...,n-2, is in the span$\{\mathbf{W}_{F,\pi(A)}\}$. Thus, $0x..,x0xE^{n-1}x0$ is in span$\{\mathbf{W}_{F,\pi(A)}\}$. Combining this with Eq. 6.8, it follows that $0x..xE^{n-1}xE^n$ is in this span. Continuing in this same fashion, with the obvious induction argument, it follows that $E^2x..xE^n$ is in this span. This completes the proof in the special case.

This proof did not use any properties of voting vectors; it only used that the vectors were not multiples of $(1,\ldots,1)$. Therefore, voting vectors can be replaced with any other vector that is not a multiple of $(1,\ldots,1)$. Next, the algebraic group properties of the subsets and the permutation operators separated the vector spaces for the different sets of alternatives. With only minor modifications, the same argument holds for sets of vectors for each set of candidates. Then, the arguments from the proof of Theorem 4.1 apply to finish the proof.

In the proof of the rest of the theorems, the permutations always effect more than one subset of candidates, so this complicates the proofs. Technically, we are using aspects of the orbit of *the iterated wreath product of permutation groups* to prove the theorem. (It turns out that related arguments can be used to prove and extend Arrow's theorem.) To help with the bookkeeping, we introduce the following definitions.

Definition. Let n alternatives be given. Both the subset of alternatives S_d and its corresponding linear subspace $E^{d'}$ are called (j,..,k) component subspaces iff S_d contains the alternatives $\{a_j,..,a_k\}$.

The next definition corresponds to choosing a value of "a" in the choice of a representation for a voting vector.

Definition. A *scalar normalization* of a system vector W^n is a choice of $2^n-(n+1)$ positive scalars, c_j, used to define the equivalent system vector $(c_1W_1...,c_jW_j,..)$. The *standar scalar normalization* for B^n is where the the Borda vector for k alternatives is given by $(k-1,k-3,..,k+1-2i,..,1-k)$ and the vectors for two alternatives are $(1,-1)$.

First we prove the following.

Proposition 3. $\mathrm{Dim}(V(B^n)) = n(n-1)/2$.

Proof. Consider

6.9 $$Y_{jk} = B^n_{\pi(A)} - P_{jk}(B^n_{\pi(A)}),$$

where $\pi(A)$ ranges through all rankings where a_j is the i^{th} ranked candidate and a_k is the $(i+1)^{th}$ ranked candidate, $i = 1,..,n-1$. This vector difference is 0 in any component space E^d that is not a (j,k) component space, and, in the (j,k) acomponent spaces, it is $2(e_j-e_k)$. Therefore, Y_{jk} is well defined.

There are only n(n-1)/2 distinct veoctors in the set $\{Y_{jk}\}$. What we show is that $V(B^n) = \mathrm{span}(\{Y_{jk}\})$, and that the vectors $\{Y_{jk}\}$ are linearly independent.

To prove the first part, we show that $\mathbf{B}^n_{\pi(A)}$ can be expressed as a linear combination of the vectors $\{\mathbf{Y}_{jk}\}$. But, any $\pi(A)$ can be expressed as the composition of transpositions. So, it suffices to show that if $\mathbf{B}^n_{\pi(A)}$ has such an expression, then so does $\mathbf{B}^n_C$ where C is a ranking obtained by a transposition of some two alternatives adjacent in the ranking $\pi(A)$, and that $\mathbf{B}^n$ has such an expression. The first statement follows from Eq. 6.9 where the choice of j and k are determined by the transposition.

It remains to show that

6.10 $\mathbf{B}^n = \Sigma\, \mathbf{Y}_{jk} \quad 1\le j<k\le n.$

Consider a subset of candidates, S_c, where $|S_c| = s\le n$. Let a_j be in S_c. Then, for precisely s-1 choices of k that differ from j, $\mathbf{Y}_{jk}$ has a nonzero vector component in the space corresponding to the set S_c. The coefficient for $\mathbf{e}_j$ is +1 iff $j<k$, and it is -1 iff $k<j$. Therefore, in the vector sum Eq. 6.10, the vector component corresponding to the set S_c is $(s-1,s-3,\ldots,s+1-2i,\ldots,1-s)$. This sum is the normalized Borda Vector, and it completes the proof of Eq. 6.10.

To complete the proof of the proposition, we must show that $\{\mathbf{Y}_{jk}\}$ is a set of $n(n-1)/2$ linearly indepedent vectors. Note that each is nonzero in only one of the first $n(n-1)/2$ vector components - in the component for the pair $\{a_j, a_k\}$. Here, the component is $2(\mathbf{e}_j-\mathbf{e}_k)$. From this, it is immediate that the vectors are linearly independent. This completes the proof of the proposition.

Proof of Theorem 3. Let the system vector $\mathbf{W}^n$ be given. If each of the voting vector components distinguish between the top and the second ranked alternative, then let the scalar normalization be defined so that the difference between these two weights is 2. Then, for each choice of $j<k$, let $\pi(A)$ be the ranking where a_j is the top ranked alternative and a_k is the second ranked alternative. Thus, the vector

6.11 $\mathbf{W}^n_{\pi(A)} - P_{jk}(\mathbf{W}^n_{\pi(A)}) = \mathbf{Y}_{jk},$

so $V(\mathbf{B}^n)$ is a subspace of $V(\mathbf{W}^n)$.

Next, consider the situation when some of the voting vector components in $\mathbf{W}^n$ do not distinguish between the two top ranked candidates; e.g., one vector component may be $(1,1,-1,-1)$. A voting vector distinguishes between some two rankings of alternatives, so fix j and k, and consider all possible rankings $\pi(A)$ where a_j is the i^{th} ranked alternative while a_k is the $(i+1)^{th}$ ranked alternative, $i=1,\ldots,n-1$. For each such ranking, 6.12 $P_{jk}(\mathbf{W}^n_{\pi(A)}) - \mathbf{W}^N_{\pi(a)}$

has a non-negative multiple of $\mathbf{e}_j-\mathbf{e}_k$ in each (j,k) component space and $\mathbf{0}$ in all others. In each (j,k) component space, there are choices of $\pi(A)$ where the multiple is positive. Therefore, if all of the vectors obtained from Eq. 6.12 are added, the sum has $\mathbf{0}$ for any non-(j,k) compnent space, and a positive multiple of $\mathbf{e}_j-\mathbf{e}_k$ for all (j,k) component spaces. This multiple depends only on the choice of the voting vector for each (j,k) component space, not on the choice of the particular j and k. (This is because the scalar for a particular (j,k) component space depends only on how often two alternatives can be ranked in the i^{th} and $(i+1)^{th}$ position as $\pi(A)$ varies over all specified rankings - it is independent of the particular choice of the alternatives.) Thus the scalar components are independent of j, k and i; they depend only on the choice of the voting vector for the set of candidates. Choose the scalar normalization for $\mathbf{W}^n$ so that, after normalized, all of the scalars in the sum equal the scalar obtained for the binary $\{a_j, a_k\}$. This means that the sum is a multiple of $\mathbf{Y}_{jk}$. This is true for all values of (j,k), so it now follows that $V(\mathbf{B}^n)$ is a linear subspace of $V(\mathbf{W}^n)$.

Finally, we must show that if $\mathbf{W}^n = \mathbf{B}^n$, then $V(\mathbf{B}^n)$ is a proper subspace of $V(\mathbf{W}^n)$. This involves a more careful application of the last argument. Because the choice of j and k determines only which subset of candidates are being considered, we can start with the indices 1 and 2. Again, consider the vector differences

6.13 $\mathbf{W}^n_{\pi(A)} - P_{12}(\mathbf{W}^n_{\pi(A)})$

where only $\pi(A)$ varies. To simplify the notation, let w^*_{cs} denote the difference between the s^{th} and the $(s+1)^{th}$ weights in the voting vector $\mathbf{W}_c$. When $\pi(A) = A$, the vector difference in Eq. 6.13 has $w^*_{c1}(\mathbf{e}_1-\mathbf{e}_2)$ in each (1,2) component space S_c, and $\mathbf{0}$ in all others. Next, consider all choices of $\pi(A)$ where a_1 is the second ranked alternative and a_2 is the third ranked alternative. The only alternatives that concern us are those ranked above a_1. So, consider the n-2 rankings obtained in the following order: The $(j-2)^{th}$ ranking has a_j top ranked, the ranking of the alternatives below a_2 is given in some arbitrary fashion. Here, the vector difference Eq. 6.13 has a scalar multiple w^*_{c2} for the (1,2,j) component spaces, the multiple w^*_{c1} for the rest of the (1,2) component spaces, and $\mathbf{0}$ for all other spaces. This vector is independent of what alternative is ranked in the k^{th} position, k>3.

Continue this construction for i=3,...,n-1. Namely, a_1 is the i^{th} ranked alternative, a_2 is the $(i+1)^{th}$ ranked alternative, and we consider the {(n-2)-(i-1)}!/(i-1)! choices of $\pi(A)$ obtained by choosing all possible sets of (i-1) alternatives ranked above a_1 and a_2. The precise rankings do not matter, it only is important to know which alternatives are ranked above a_1. For each such ranking, the scalars in the vector difference Eq. 6.13, w^*_{cs} in each (1,2) component space S_c depends on how many of the selected (i-1) alternatives are also in this set. The set of all vectors obtained in this manner span a linear subspace of $V(\mathbf{W}^n)$. (This subspace plays a role similar to the "-1 eigenspace for (1,2)" in the proof of Theorem 4.1.) It was shown that a linear combination of these vectors equals $\mathbf{Y}_{jk}$. If any of these vectors differs from a scalar multiple of $\mathbf{Y}_{12}$, then this subspace is at least of dimension 2 and the conclusion that $V(\mathbf{B}^n)$ is a proper subspace of $V(\mathbf{W}^n)$ would follow. So, suppose each such vector is a multiple of $\mathbf{Y}_{12}$; it would be, by comparing the coefficient for the binary $\{a_1, a_2\}$, 2 $\mathbf{Y}_{12}$. Thus, when $\pi(A) = A$, $w^*_{c1} = 2$ for each S_c that is a (1,2) component space. Continuing, because each w^*_{sc} ;occurs in at least one of the vector differences, it follows that the difference between any two successive weights must be 2 for each of the voting vectors. This uniquely determines all of the voting vector components - they all are Borda Vectors. The analysis does not depend on the choice of the indices 1,2; it holds for all $j<k$.

It is a simple exercise to show that if $V(\mathbf{B}^n)$ is a proper linear subspace of $V(\mathbf{W}^n)$, then $D(\mathbf{B}^n)$ is a proper subset of $D(\mathbf{W}^n)$. Indeed, one can show that if the difference in dimension is d, then $3^d|D(\mathbf{B}^n)| < |D(\mathbf{W}^n)|$. This completes the proof of Theorem 3.

Proof of Theorem 1. The proof uses the construction for the proof of Theorem 3, and it is similar to the proof of Theorem 4. What must be shown is that, with the exception of a algebraic set of system vectors, $V(\mathbf{W}^n) = E(n)$. But, $V(\mathbf{W}^n)$ is spanned by $\{\mathbf{W}^n_{\pi(A)}\}$, and the coefficients of these vectors are determined by the weights of the voting vectors. The "open" situation is that for most choices of these coefficients, $V(\mathbf{W}^n)$ will have a fixed, maximal dimension. All we need show is that this maximal dimension is dim(E(n)). This is done by showing that for at least one $\mathbf{W}^n$, dim($V(\mathbf{W}^n)$ = dim(E(n)), so the sets agree. Secondly, it is a simple fact from linear algebra that when the coefficients for the voting vectors are treated as variables, then $V(\mathbf{W}^n)$ will have a dimension lower than the maximal dimension only if the variables are in a particular lower dimensional algebraic set. Namely, the spanning vectors $\{\mathbf{W}^n_{\pi(A)}\}$ have additional dependencies given by the vanishing of certain determinents; these determinent conditions define the algebraic equations.

Suppose $V(\mathbf{W}^n)$ is a proper subspace of $E(n)$. This means it has a normal vector, $\mathbf{N}$, in $E(n)$. By using the basis for $E(n)$, it follows that there is a choice of (j,k) and some component space of $E(n)$ so that $\mathbf{N}$ is nonzero in a $\mathbf{e}_j-\mathbf{e}_k$ direction. Without loss of generality, assume that $(j,k) = (1,2)$. This means that $V(\mathbf{W}^n)$ does not contain the subspace of $E(n)$ generated by the product of $(\mathbf{e}_1-\mathbf{e}_2)$ from each $(1,2)$ component subspace of $E(n)$. The proof of the theorem is based on showing that there is a $\mathbf{W}^n$ where $V(\mathbf{W}^n)$ contains the full subspace generated by the product of the $(\mathbf{e}_1-\mathbf{e}_2)$ subspace. This contradicts the existence of $\mathbf{N}$ and it proves the theorem.

The choice of the system vector, $\mathbf{W}^n$, is where all of the voting vector components are plurality vectors, so $w^*_{c\,1}>0$ for each $(1,2)$ component space S_c, and $w^*_{c\,j}=0$ for $j>1$. To show independence, the rankings described in the proof of Theorem 3 are used, but I will start with the last rankings and word forward. So, if a_1 and a_2 are bottom ranked alternatives $(i=n-1)$, then only in the set $\{a_1,a_2\}$ can they be the two top ranked alternatives. For this $\pi(A)$, the vector difference Eq. 6.13 has $\mathbf{0}$ in all component spaces except the one corresponding to this binary, where the entry is $\mathbf{e}_1-\mathbf{e}_2$.

Next, consider $i=n-2$ where a_j is the bottom ranked alternative. Only in the sets $\{a_1,a_2\}$ and $\{a_1,a_2,a_j\}$ are a_1 and a_2 top ranked. For this $\pi(A)$, the vector difference Eq. 6.13 has nonzero entries only in these two component spaces. The first vector obtained from $i=n-1$ is used to eliminate the $\{a_1,a_2\}$ component of this vector. Thus, $V(\mathbf{W}^n)$ contains both the vector with the only nonzero vector $\mathbf{e}_1-\mathbf{e}_2$ in the $\{a_1,a_2\}$ component space and the vector with the only nonzero component $\mathbf{e}_1-\mathbf{e}_2$ in the $\{a_1,a_2,a_j\}$ component space. The obvious induction argument proves that, for each $(1,2)$ component space of $E(n)$, $E(n)$ has a vector where its only nonzero term is $\mathbf{e}_1-\mathbf{e}_2$ in this component space. This completes the proof of the theorem.

Proof of Theorem 2. Let $n=3$. The proof is much the same as the proof of Theorem 1 except that I now show that the conclusion holds for all non Borda vectors. Again, without loss of generality, we can concentrate on the $(1,2)$

component spaces. These are $S_1 = \{a_1, a_2\}$ and $S_4 = \{a_1, a_2, a_3\}$. The vector differences from Eq. 6.13 involve the coefficients $(2, w*_{41})$ and $(2, w*_{42})$. The vector differences are linearly independent iff the vectors defined by the coefficients are linearly independent. But, these coefficient vectors are dependent iff $w*_{41} = w*_{42} = c$ iff the differences between the weights of $\mathbf{W}_4$ are the same fixed constant c iff $\mathbf{W}_4 = \mathbf{B}_4$. This completes the proof for n=3.

Let n=4. There are four (1,2) component spaces; $S_1 = \{a_1, a_2\}$, $S_7 = \{a_1, a_2, a_3\}$, $S_8 = \{a_1, a_2, a_4\}$ and $S_{11} = \{a_1, a_2, a_3, a_4\}$. The vector differences from Eq. 6.13 define the coefficient vectors

$$(2,\ w*_{7,1},\ w*_{8,1},\ w*_{11,1})$$
$$(2,\ w*_{7,2},\ w*_{8,1},\ w*_{11,2})$$
$$(2,\ w*_{7,1},\ w*_{8,2},\ w*_{11,2})$$
$$(2,\ w*_{7,2},\ w*_{8,2},\ w*_{11,3})$$

from the rankings A, $a_3 \succ a_1 \succ a_2 \succ a_4$, $a_4 \succ a_1 \succ a_2 \succ a_3$, and $a_4 \succ a_3 \succ a_3 \succ a_2$.

If the matrix defined by these four vectors is nonsingular, then the argument for n=3 applies. An equivalent matrix in triangular form is

$$\begin{pmatrix} 2 & w*_{7,1} & w*_{8,1} & w*_{11,1} \\ 0 & w*_{7,2}-w*_{7,1} & 0 & w*_{11,2}-w*_{11,1} \\ 0 & 0 & w*_{8,2}-w*_{8,1} & w*_{11,2}-w*_{11,1} \\ 0 & 0 & 0 & w*_{11,3}-w*_{11,1}-2(w*_{11,2}-w*_{11,1}) \end{pmatrix}$$

If any of the component vectors for $\mathbf{W}^4$ are Borda Vectors, then the corresponding entry in the second to fourth rows of this matrix are zero, and the conclusion follows. If none of the component vectors are Borda Vectors, then a necessary condition for the dictionary not to be all of U^4 is $w*_{11,3}-w*_{11,1}-2(w*_{11,2}-w*_{11,1}) = 0$. When expressed in terms of the weights defining $\mathbf{W}_4$, this is the condition in Theorem 2.

Proof of Theorem 5.1. The matrix associated with Eq. 6.13 when P_{12} is applied to the n-2 rankings with a_1 i^{th} ranked and a_2 $(i+1)^{th}$ ranked is

$$\begin{pmatrix} 2 & w*_{1,1} & w*_{2,1} & \cdots & w*_{n-2,1} \\ 2 & w*_{1,2} & w*_{2,2} & \cdots & w*_{n-2,2} \\ \cdot & \cdot\cdot & \cdot\cdot\cdot\cdot\cdot & & \cdots \\ 2 & w*_{1,n-1} & w*_{2,n-1} & \cdots & w*_{n-2,n-1} \end{pmatrix}$$

The assumption on the voting vectors is that they, and the Borda vector, form n-1 linearly independent vectors. These vectors still are linearly independent if they are replaced by vectors giving the differences between successive weights. This matrix is the one given above. This proves the theorem.

Proof of Theorem 5.2. For n alternatives, consider the set of system vectors $\{\mathbf{W}^n_j\}$ that are defined in the following manner. Each component voting vectors of $\mathbf{W}^n_j$ is of the form (1,1,...,1,0,...,0); it assigns one point to a voter's k top ranked candidates and zero for all others. Furthermore, there is precisely one system vector corresponding to any assignment of voting vectors of this kind to the subsets of candidates. (Hence, for n=3, there are only two system vectors: (1,0;1,0;1,0;1,0,0) and (1,0;1,0;1,0;1,1,0). For n=4, there are $2^4(3)$ system vectors. For n alternatives, there are k(n) = $2^{n(n-1)(n-2)/3!}(3^{n(n-1)n-2)(n-3)/4!})\dots(n-1)$ system vectors.) Let $G(-,\mathbf{W}^n_1,\dots,\mathbf{W}^n_{k(n)})$ be the mapping that represents the voting tally for the n system vectors. By using the argument showing that $V(\mathbf{B}^n)$ is a subspace of $V(\mathbf{W}^n)$, it follows that the space spanned by $\{(\mathbf{W}^n_1,\dots,\mathbf{W}^n_{k(n)})_{\pi(A)}\}$ contains the n(n-1)/2 dimensional space given by the span$((\mathbf{Y}_{jk},\mathbf{Y}_{jk},\dots,\mathbf{Y}_{jk}))$ as j<k vary over all pairs. This means that if w is a word in $D(\mathbf{B}^n)$, then there is a profile such that the word for each of the n system vectors is w. Now, let $\mathbf{W}^{n\prime}$ be any system vector. Our choice of k(n) system voting vectors is a basis for the system vectors, and any system vector is in their convex hull. This means that any outcome of this system vector, $\mathbf{W}^{n\prime}$, is in the convex hull of the outcomes for the basis system vectors. Because all of the basis system vectors have the same outcome, and because (by Theorem 6) the regions supporting a word are convex, the outcome for $\mathbf{W}^{n\prime}$ is in the same ranking region. This completes the proof.

Acknowledgement. This research was supported, in part, by NSF grant IST 8415348. Some of this research was done while I was a vistor at CEREMADE, University of Paris, and at the System and Decision Science Program of IIASA in Vienna. I would like to thank my hosts J.P. Aubin and A. Kurzhanski. This paper is based on results in "The optimal ranking method is the Borda count", NU Center for Math Econ Discussion paper #636, Jan., 1985.

References.

1. K. Arrow, *Social Choice and Individual Values*, 2nd ed., New York, Wiley, 1963.

2. S. Brams, P. Fishburn, and S. Merrill, The responsiveness of Approval Voting: comments on Saari and Van Newenhizen, forthcoming in *Public Choice*

3. G. Debreu, *Mathematical Economics*, Paperback version, Econometric Society Monographs #4, Cambridge University Press: Cambridge, 1986.

4. P. Fishburn, "Monotonicity paradoxes in the theory of elections", *Discrete Applied Mathematics* **4** (1982), 119-134.

5. P. Fishburn, "Inverted orders for montone scoring rules", *Discrete Applied Math* **3** (1981), 27-36.

6. P. Fishburn, *The Theory of Social Choice*, Princeton University Press: Princeton, 1973.

7. L. Hurwicz, "Optimality and informational efficiency in resource allocation process", in *Mathematical Models in the Social Sciences, 1959*, ed. Arrow, Karlin, and Suppes, 1960.

8. L. Hurwicz, "On informationally decentralized systems," in *Decision and Organization Volume in Honor of Jacob Marschak*, ed. B. McGuire and R. Radner, North Holland: Amsterdam, 1972.

9. L. Hurwicz, S. Reiter, and D. G. Saari, "On constructing mechanisms with message spaces of minimal dimension for smooth performance functions", Mimeo. (1978)

10. R. Mantel, "On the characterization of aggregate excess demand", JET **7** (1974), 348-353.

11. H. Moulin, "Fairness and strategy in voting", in *Fair Allocation*, ed. H. P. Young, Proceedings of Symposia in Applied Mathematics, vol. 33, American Mathematical Society: Providence, Rhode Island, 1987.

12. H. Moulin, "Condorcet's principle implies the no show paradox", to appear in JET.

13. R. Niemi and W. Riker, "The choice of a voting system", *Sci. Amer.* **234** (1976), 21-27.

14. S. Reiter, "Informational incentive and performance in the $(new)^2$ welfare economics" in *Studies in Mathematical Economics*, ed. S. Reiter, The Mathematical Association of America, 1986.

15. D. G. Saari, "The Borda Dictionary", to appear

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

17. D. G. Saari, "The source of some paradoxes from social choice and probability", *JET*, **41** (1987), 1-22.

18. D. G. Saari, "Susceptibility to manipulation", NU Center for Math. Econ discussion paper, June, 1987.

19. D. G. Saari, "Chaos and the theory of elections", In *Nonlinear Dynamics*, ed. A. Kurzhanski and K. Sigmund, Springer Verlag Series on Mathematical Economics, 1987.

20. D. G. Saari and J. Van Newenhizen, "The problem of indeterminacy in approval, multiple, and truncated voting systems. *Public Choice* forthcoming

21. D. G. Saari and J. Van Newenhizen, "Is Approval Voting an 'unmitigated evil'?", *Public Choice* forthcoming.

22. A. Sen, "Social choice theory", Chp. 22 in *Handbook of Mathematical Economics, Vol. III*, ed. K. J. Arrow and M. D. Intriligator, North-Holland: Amsterdam, 1986.

23. H. Sonnenschein, "Market excess demand functions", *Econometrica* **40** (1972), 549-563

Discussion Paper No. 745

A DICTIONARY FOR VOTING PARADOXES

by

Donald G. Saari
Department of Mathematics
Northwestern University
Evanston, IL 60208

July 1987

Ever since K. Arrow [1] proved it is impossible to construct a voting system that satisfies certain desired properties, a major focus in social choice has been to use an axiomatic formulation to determine what assumptions are, or are not mutually compatible. (See, for example, Sen [22].) In this paper, I introduce a different approach to analyze the important class of positional voting methods, such as the commonly used plurality vote. The idea is to characterize the election outcomes. Namely, for any number of candidates and for any positional voting procedure, I characterize all possible ways election rankings can arise over all possible subsets of candidates. With this catalog, the properties of these voting systems can be determined in a simple, pragmatic fashion - just check the listings to see what can and cannot occur. The conclusions are very disturbing - paradoxes are more plentiful and much more complicated than one might have anticipated. Only **Borda's Method** avoids many of the potential flaws. Applications of this dictionary of voting outcomes are indicated, in part, by describing all possible plurality election outcomes, by obtaining new results about agendas and runoff elections, and by describing certain strategic situations. Moreover, because I am characterizing *all* possible election outcomes, it follows that all of the election paradoxes in the literature described in terms of ordinal rankings of positional elections must be special cases of this catalog. This is true, and, by using the listings, I show how any such paradox can be extended and generalized in several different ways. Other, quite spectacular paradoxes can be created; indeed, with the dictionary, the kinds of paradoxes that now can be designed are limited only by one's imagination.

Probably the most widely used voting method is a plurality election, but how should we interpret the election ranking? To see that there is a problem, consider the hypothetical situation where fifteen people select a common luncheon beverage. Six of them have the ranking water (wa) over wine (wi) over beer (be) (i.e., wa>wi>be), 5 have the ranking be>wi>wa, and 4 have the ranking wi>be>wa. The plurality ranking is wa>be>wi with the tally 6:5:4. Nevertheless, these same people prefer the bottom ranked alternative, wine, both to the top ranked water (by 9:6) and to the second ranked beer (by 10:5)! Even beer is preferred to water (by 9:6). Thus wine, the *majority* or *Condorcet winner*, (in any pairwise comparison, it is selected by a majority of the voters) is bottom ranked in the election while water, the *anti-majority alternative*, is top ranked. By using the binary, majority vote comparisons, it is arguable that the "true ranking" is probability, statistics, economic indices, etc., are special cases of aggregation procedures. So, are voting paradoxes related to certain difficulties in these areas? Positional voting is a simple "economic message system" of the type introduced by L. Hurwicz [8] (also see [9,14]) where the object is to encode and transmit relevant information about each agent's preferences. In voting, the encoding is given by marking the ballot. Can voting paradoxes suggest hidden flaws in other kinds of message systems? (The answer to these questions is yes.) Because positional voting is a simple, important prototype, it serves as a test case for concepts being developed in decision analysis, the social sciences, and elsewhere. By understanding what "goes right" and what "goes wrong" with voting, insight can be gained about more complex methods as well as other social choice models. The approach developed here extends, in part, to these other systems.

The central theme of this paper is to determine what can go wrong with positional voting and to explain why. To understand what paradoxes can occur and to avoid the standard approach of finding them in a piecemeal fashion, I characterize *all* possible election outcomes over *all* possible subsets of candidates for *all* possible positional voting methods and *all* possible profiles of voters. The reader will recognize the similarity of this goal with the Sonnenschein program [23] where he, Mantel [10], Debreu [4], and others characterized (for the message system of price dynamics) all possible aggregate excess demand functions for all numbers of commodities for all simple trading economies based on neoclassical utility functions. A catalog, or ***dictionary*** for voting outcomes, could be used in much the same way as the Sonnenschein-Mantel-Debreu classification; both serve as a starting point to determine what else can and cannot occur. In this manner, a dictionary serves as the foundation to analyze voting procedures. By using the dictionary, it is easy to create new paradoxes - just check the listing to find what unexpected rankings occur over different subset of candidates with the same sincere voters. We can compare and combine paradoxes into classes - paradoxes that depend on similar dictionary listings probably are related. We can understand strategic voting - just compare the election rankings for nearby profiles of voters. (After a manipulating voter marks the ballot, the actual election is determined by a profile that differs from the sincere one.) All of this is illustrated here.

For reasons explained in a companion paper [15], it is not practical to of candidates S_j. Let the *system vector*, $\mathbf{W}^n = (\mathbf{W}_1, \ldots, \mathbf{W}_{2^n-(n+1)})$, be the listing of these tallying procedures.

Example 1. For n=3, $\{\{a_1,a_2\}, \{a_2,a_3\}, \{a_1,a_3\}, \{a_1,a_2,a_3\}\}$ is a listing of the $2^3-4 = 4$ subsets of two or more candidates. The system vector $\mathbf{W}^3 = (1,0;1,0;1,0;2,1,0)$ signifies that the first three sets of candidates are tallied with a majority vote - the voting vector is $(1,0)$ - while an election for S_4 is tallied with the BC vector $(2,1,0)$.

To describe our results, we need a space of "all possible election outcomes". Toward this end, let R_j be the listing of all possible rankings of the $|S_j|$ alternatives that are generated by complete, transitive, symmetric (to admit tie votes) binary relationships on S_j. For instance, $S_1 = \{a_1,a_2\}$, so $R_1 = \{a_1>a_2,\ a_1=a_2,\ a_2>a_1\}$. If $|S_j| = 3$, then R_j has 13 rankings - 6 of them are without any ties, while each of the other 7 have at least one tie. Let U^n, the *universal set*, be the cartesian product $R_1 x..x R_{2^n-(n+1)}$. An element of U^n is a listing of $2^n-(n+1)$ rankings; there is one for each subset of candidates. The j^{th} ranking, or *symbol*, of this listing is a ranking for S_j.

Example 2. The sequence $\{a_1>a_2,\ a_2>a_3,\ a_3>a_1,\ a_2>a_1=a_3\}$ is an element of U^3. Each symbol is the ranking for the appropriate subset of alternatives.

A *profile* is a listing of each voter's linear ranking of the n candidates. Let P^n be the space of all possible profiles of the n alternatives; we impose no restrictions on the (finite) number of voters. Once a profile, $\mathbf{p}$, from P^n is given, then, in the obvious manner, $\mathbf{W}^n$ is used to uniquely determine the election rankings for each of the $2^n-(n+1)$ subsets of candidates. This listing of $2^n-(n+1)$ rankings is called a *word*, and a word is an element of U^n. There is an important difference between a word and an element of U^n; an element of U^n might be an arbitrary listing of rankings that has nothing to do with elections, but a *word generated by* $\mathbf{W}^n$ *is a list of election rankings that is attained with a profile of voters*. For instance, in the beverage example, the election results {wi>wa, wi>be, be>wa, wa>be>wi} is a *word* in U^3 generated by the system vector $(1,0;1,0;1,0;1,0,0)$ because these rankings are attained with the specified profile. Let

2.1 $f(-,\mathbf{W}^n): P^n \text{-----} > U^n$

be the mapping that determines the word for a given profile.

then Eq. 2.3 is satisfied.

So, for almost all choices of $\mathbf{W}^n$, *anything* can happen. This means that the wildest paradox one could possible imagine actually can occur for almost all choices and combinations of voting vectors. Spectacular paradoxes now are eacy to create.

Example 4. 1. There exists a profile of voters so that the plurality outcome is $a_4 > a_3 > a_2 > a_1$, but the *same* voters' election ranking for *all* other subsets of candidates is the exact reverse of this -- the other rankings are generated by the reversed binary relationship $a_j > a_k$ iff $j < k$.

2. Let $n > 3$. There exists a profile of voters so that their election rankings alternate with the number of candidates -- if a subset has an even number of alternatives, its plurality rankings are generated by $a_j > a_k$ iff $j < k$; if a subset has an odd number of candidates, its $(1, 1, 0, \ldots, 0)$ rankings (vote for your two top ranked candidates) is generated by the *reversed* relationship $a_j > a_k$ iff $k < j$.

Theorem 1 includes and significantly extends many results in the literature. For instance, a widely quoted example due to P. Fishburn [5] is where the group's plurality ranking is $a_1 > a_2 > a_3 > a_4$, but if a_4 is removed, then the same group's plurality ranking now is $a_3 > a_2 > a_1$. Saari [16] showed that all results of this kind could be extended in many different ways - there could be any number of candidates, one could use any choice of voting vectors, $\mathbf{W}_n$ and $\mathbf{W}_{n-1}$, the number of candidates that are removed is arbitrary and could involve all sets obtained in this manner, and the rankings could be selected in an arbitrary fashion. But, both Fishburn's and Säari's statements are very special cases of Theorem 1. To see this, note that for four alternatives, there are $2^4 - 5 = 11$ different subsets of 2 or more alternatives, so each word has 11 different symbols. Fishburn's example specified just 2 of them. According to the theorem, one could fill in the remaining 9 symbols in *any* desired manner, and there is a profile to support it. As an illustration, by choosing appropriate rankings for the pairs of alternatives, Fishburn's example can be extended so that a_2, the middle ranked alternative in the two elections is the Condorcet winner, but the winner of the first election, a_1, is the antimajority candidate while the winner of the second election, a_3, *almost* is the antimajority candidate because it only beats a_1 in the election procedures and other issues raised in the Introduction.

Example 5. 1. One form of a runoff election starts by first ranking the original n candidates with a positional election. The k_1 top ranked candidates are advanced to the next stage to be reranked with another positional election. If $k_1>2$, it may be necessary to have still another runoff election with the k_2 top ranked candidates. Indeed, if n is sufficiently large, one could imagine a process involving several elimination stages as characterized by the positive integers $\mathbf{k} = (k_1,\ldots,k_s)$, $k_j>k_{j+1}$, $k_s\geq 2$.

Why use other procedures? Why not just accelerate the process by letting $k_1=2$; won't the outcome be the same? The (known) answer is no, not necessarily. Indeed, with Theorem 1, it now is easy to extend the known results by showing that for almost all choices of $\mathbf{W}^n$, there are profiles of voters where different choices of the sequence $\mathbf{k}$ lead to completely different election results. In fact, it now is easy to prove *there is a profile and n-2 choices of runoff procedures so that when the j^{th} procedure is used, a_j is the winner, $j=1,\ldots,n-2$.*

I will illustrate the assertion for n=4; the same proof holds for all values of n. For n=4, there are only two runoffs: (3,2) and (2). First, choose a word with the symbols $a_1>a_2>a_3>a_4$, $a_1>a_3>a_2$, $a_1>a_3$, and $a_2>a_1$. With these election rankings, the first (3,2) runoff is among $\{a_1,a_2,a_3\}$, and a_1 wins the second runoff between a_1 and a_3. On the other hand, a_2 is the winner of the (2) runoff between a_1 and a_2. This completes the proof because, according to Theorem 1, there are profiles that define this word. (To prove this assertion, we need to select rankings for subsets of candidates that are not admitted by [19].)

The same approach works for all values of n and for almost all voting vectors. The idea is simple: different elimination procedures cause different subsets of candidates to be reranked. But, if there is so much as a one candidate difference between subsets, their rankings can be chosen in any desired manner; there need not be any consistency among them. By choosing the rankings in an appropriate manner, we can prove that radically different outcomes exist. Theorem 1 asserts that, for almost all choices of $\mathbf{W}^n$, a profile exists to support the selected rankings.

2. The runoff example did not use all of the symbols in a word. These extra symbols introduce added flexibility to find even more surprising conclusions. For instance, for all pairs of alternatives $\{a_j,a_4\}$, choose the

$n(n-1)/2$ symbols for the pairs, so the rankings of the remaining pairs could be selected to determine different winners of various runoff procedures, etc. Namely, for $n \geq 4$, *there is a profile of voters, n agendas, and n-2 runoff election procedures so that when the* j^{th} *agenda is used, the outcome is* a_{j+2}, $j=1,\ldots,n$, *and when the* k^{th} *runoff procedure is used, the outcome is* a_k, $k=1,\ldots,n-2$. This example can be enhanced by adding "almost Condorcet winners" to exhibit certain features, etc. Actually, by using Theorem 1, the kinds of examples that can be created are limited only by one's imagination.

3. ANY GOOD NEWS?

Theorem 1 proves that a lot can go wrong with positional elections, but it also hints there may be good news. It suggests there might be a lower dimensional subset of $\mathbf{W}^n$'s where the outcomes don't depend so sensitively upon which subset of candidates just happen to be presented. It suggests there are choices of system vectors, $\mathbf{W}^n$ where $D(\mathbf{W}^n)$ is a proper subset of U^n. But, because α^n is an algebraic set, it also means that the component voting vectors of a favorable $\mathbf{W}^n$'s must be carefully coordinated to avoid the negative aspects of Eq. 2.3. Theorem 2 characterizes α^n for $n=3,4$. (An extension to larger values of n requires a different technical development.) First, some preliminaries.

Definition. **A voting vector** $\mathbf{W}_n = (w_1,\ldots,w_n)$ **is a** ***Borda Vector*** **iff** $w_j - w_{j+1}$ **is the same nonzero constant for** $j=1,\ldots,n-1$. **Let** $\mathbf{B}^n$ **denote the system vector where Borda vectors are used to tally all subsets of three or more alternatives.**

It is easy to see that an election tallied with a Borda Vector and with the BC always have the same ranking. This manifests the fact that an election tallied with $\mathbf{W}$ and with $\mathbf{W}' = a\mathbf{W} + b(1,\ldots,1)$ always agree if $a>0$. These modifications affect the tally, but not the relative rankings of the candidates.

Theorem 2. **1.[17] For** $n=3$, $D(\mathbf{W}^3)$ **is a proper subset of** U^3 **iff** $\mathbf{W}^3$ **is a Borda Vector.**

2. For $n=4$, **if** $D(\mathbf{W}^4)$ **is a proper subset of** U^4, **then either at least one of the voting vector components of** $\mathbf{W}^4$ **is a Borda Vector, or the last vector component**

special, singular case. Moreover, it is easy to show that if a system vector has a Borda vector as a component voting vector, then the dictionary avoids certain paradoxes. (The proof of this statement is an elementary extension of the proof of Theorem 2.) In fact, the next statement demonstrates an even stronger, favored feature enjoyed by BC for all values of n.

Theorem 3. Let $n \geq 3$, and let the system vector $\mathbf{W}^n$ have at least one non Borda vector for a subset of at least three candidates. Then,

3.2 $$D(\mathbf{B}^n) \subsetneqq D(\mathbf{W}^n).$$

Theorem 3 means that Borda's method admits fewer paradoxes, or words, than any other choice of system vectors. (The exact dictionary, $D(\mathbf{B}^n)$, is characterized in [15].) Eq. 3.2 has many important implications. It means that *any fault or paradox admitted by Borda's method also must be admitted by all other positional voting methods*. For instance, it follows from [15] that Borda's method need not rank a Condorcet winner in first place; e.g., for n=3, the word $\{a_1 > a_2, a_2 > a_3, a_1 > a_3, a_2 > a_1 > a_3\}$ is in $D(\mathbf{B}^3)$. Consequently, from Theorem 3, this same word *must* be in all other dictionaries, $D(\mathbf{W}^3)$. Thus, *any criticism of Borda's method advanced by means of election rankings also serves as a criticism for all possible voting vectors.* Conversely, all other $\mathbf{W}^n$'s admit words (i.e., paradoxes) that are not permitted by Borda's method. Namely, all other system vectors introduce additional indeterminacies - this means that the resulting election outcomes can be far more sensitive to which subgroup of candidates just happen to be presented. These new electoral difficulties introduced by the other system vectors are not admitted by Borda's Method. These statements serve as strong arguments for using Borda's method over any other choice.

4. WHAT ELSE CAN HAPPEN?

Theorem 1 is not the ultimate description of positional voting paradoxes. To show this, we offer two theorems, both in two parts, to indicate what else can happen; a more complete description is planned for elsewhere. (The proof of Theorem 4 motivates the proofs of Theorems 1-3.) While these theorems are based on the ideas developed for Theorem 1, they are of independent interest because independent set. for each $j = 1,..,n-1$, choose j rankings of the alternatives in S_j. There exists a profile of voters so that when the j^{th} set is tallied with the k^{th} vector chosen for S_j, the outcome is the k^{th} selected ranking for S_j, $j=1,..,n-1$; $k=1,..,j$.

Example 6. 1. Let n=4. For S_3 choose the vectors (1,0,0,0), (3,2,1,0), (0,0,0,1) and the rankings $a_4 > a_3 > a_2 > a_1$, $a_3 > a_1 > a_2 > a_4$, $a_1 > a_2 > a_3 > a_4$; for S_2 choose (1,0,0) and (0,0,1) with the two identical rankings $a_1 > a_2 > a_3$; and for S_1 choose (1,0) and $a_1 > a_2$. According to Theorem 4.2, there is a profile that realizes all of these outcomes with changes in the choice of the tallying vectors. So, the plurality elections for S_1 and S_2 show some consistency, but the plurality ranking for S_3 is the reverse of what one might expect. On the other hand, the ranking obtained by voting for your bottom ranked alternative, (0,0,0,1), resumes this consistency.

2. This theorem can be used to show how the Coombs, the (3,2), and the (2) runoffs all give different outcomes. Just choose the purality ranking to be $a_1 > a_3 > a_2 > a_4$, the plurality ranking of $a_2 > a_1 > a_3$, and the majority ranking of $a_2 > a_1$. So, the winner of the (3,2) runoff is a_2, while the winner of the (2) runoff is either a_1 or a_3. Now, choose the (0,0,0,1) ranking to be $a_1 = a_2 = a_3 > a_4$. Is only a_4 advanced by the Coombs system?

Theorem 4.2 is not sufficient to completely analyze runoffs and other, more complicated procedures, because it does not admit all subsets of candidates. For instance, in Example 6.1, if the first stage of the election is ranked with the BC, then S_2 is the set of candidates that is to be reranked. However, if a Coombs method is used, then $\{a_2, a_3, a_4\}$ needs to be reranked, and this subset is not admitted by Theorem 4.2. Theorem 5 is a step toward a more general result.

Theorem 5. 1. Let n>3, and let F be the family of subsets of candidates that consists of all n(n-1)/2 pairs of candidates and the set of all n candidates. Choose n-2 vectors in R^n that, along with (1,..,1) form a linearly independent set. Furthermore, suppose the span of the n-2 vectors and (1,1,..,1) do *not* include a Borda Vector. Choose a ranking for each of the pairs and choose n-2 rankings for the set of n alternatives. There is a profile of voters so that for each pair of alternatives, their majority ranking is the selected one. When their

This result generalizes and extends a nice example of Fishburn's [7]. Fishburn created a profile for n=3 to prove that there exist situations where the Condorcet winner never is elected by any positional voting method. *Theorem 5.2 extends this kind of statement in all possible ways.* It asserts that for any n, you can select *any* feature of Borda's method that can be expressed in terms of the rankings over subsets of candidates. This feature defines a word in the Borda dictionary, $D(\mathbf{B}^n)$. Then, according to this corollary, there exists profiles where the same feature holds for all possible positional voting methods. Thus, after $D(\mathbf{B}^n)$ is characterized in [15], all sorts of new examples can be created. Of course, this statement is false if a word is selected from any dictionary other than the Borda Dictionary. This is because if this new word is not admitted by BC then the conclusion cannot possibly hold.

Example 8. There exists a profile so that no matter what positional election method is used, the outcome is $(a_1 \succ a_2, a_2 \succ a_3, a_1 \succ a_3, a_2 \succ a_1 \succ a_3)$.

5. THE GEOMETRY OF THE SPACE OF PROFILES

Before turning to the proofs, it is appropriate to question whether these new results are robust, or whether they depend upon specially constructed examples that disappear with even the slightest perturbation of the profiles. They are robust. This answer, based on the following representation of the space of profiles, P^n, uses the fact there are n! different rankings; i.e., there are n! different types of voters. The basic idea can be seen with the beverage example. There is no qualitative change in the example if I replicated the profile by, say, tripling the number of voters with each ranking. This is because differences in the rankings do not depend on the numbers of voters, but on the *ratios* of the number of voters of each type. So, a profile can be characterized by specifying what *fraction* of all voters have a particular ranking. In this manner, a profile **p** can be identified with a vector with n! non-negative components that sum to unity. Namely, **p** can be viewed as being a vector on the unit simplex, Si(n!), in the positive orthant of $R^{n!}$.

There is a slight technical difficulty. A profile defines a vector in Si(n!) with rational components, and vice versa. Part of the strength of my approach is to use all of the structure of Si(n!); I embed the discrete problem of

Thus all paradoxes based on election rankings without ties are supported by open sets, so they are robust. The paradoxes involving tie votes are not robust; a slight change in the profile can alter the outcome. On the surface, Theorem 6 appears to be a technical theorem concerned with the robustness of certain words. In fact, when used with the earlier theorems, Theorem 6 is a "gold mine" for explaining, extending, and describing several other results of current interest. For instance, manipulation and strategic behavior is a topic currently of great concern. But note, if a voter successfully manipulates the outcome of an election with a strategic vote, the sincere profile is on one side of the hyperplane given by a tie vote, and the manipulated profile is on the other. Consequently, the structure of these "tie vote" hyperplanes provide valuable information about how susceptible a system is to being manipulated by individuals or small groups. Such an analysis is in Saari [18]. A similar topic would be the *sensitivity* of a system - small changes in how the voters mark the ballots alters the outcome. Using the techniques of [18], it is easy to determine which systems are more sensitive than others.

Results of a different but related flavor concern those fascinating statements asserting that, by voting, a voter *hurts* his or her interests, so by abstaining the voter is better off. The first result of this type that I am aware of concerns a runoff election with a plurality vote, and it was found by Brams and Fishburn [2]. With Theorems 6 and 1, it is easy to extend and characterize all possible methods that have this behavior. To see how this is done, we offer here a partial result that is easy to prove. Toward this end, call a social choice method that selects a single candidate *disjoint* if i) the outcome is based on the system vector, $\mathbf{W}^n$, positional voting rankings of the subsets of candidates, ii) the ranking of some one subset of candidates determines or affects which one of several subset of candidates will be the final set to be ranked - indeed, just the reversal of the relative rankings of some two adjacently ranked candidates can change the choice of the final set, and iii) the final outcome is based on the ranking of the final subset of candidates, and iv) the method is not constant, at least two different outcomes are possible for each choice of a final set. As an example, all runoff procedures, whether of the Coombs type or the more standard kind defined by the integers **k** are disjoint methods. This is because by switching definition of a disjoint method. Start with the ranking of the third set having a_2 tied in the swing position. Construct the ranking for the two voters to have a_2 as top ranked, the outcome of the set they *don't* get by voting and advancing the ranking of a_2 as second ranked, and the outcome of the set they do get by advancing a_2 in third place. The rest of the analysis is the same.

In this proof, there are many other symbols that have not been specified, so they can be assigned in any desired manner to prove other conclusions. For instance, for certain disjoint systems, it is possible to show that, in addition to what already has been proved, the Condorcet winner need not get elected, that the outcome of a runoff election differs from these conclusions, etc. (This generalizes a result of Moulin [12].) As an illustration, we can combine several of the features already described.

Corollary 6.2. Let $n>4$ and suppose the plurality vote is used to rank all subsets of candidates. There are i) n-2 runoff procedures, ii) n agendas, iii) a profile of voters and iv) two other voters with the same ranking that has a_2 top ranked, and a_1 bottom ranked, so that a) when the original voters use the j^{th} runoff procedure, a_j is the winner, $j=1,..,n-2$, and b) when the k^{th} agenda is used, a_k wins, $k=1,..,n$. If the two additional voters vote, then outcome of all elections remain the same except for the two procedures where a_2 won. In both of these procedures, the new winner now is a_1.

Other kinds of paradoxes, such as explaining why two subcommittees can independently reach the same conlusion, but when joined as a full committee, they select a different alternative can be based on other geometric structures of Theorem 6. Details will be given elsewhere.

profile **p** is a vector in Si(n!), so the election outcome is in the convex hull of the vectors $\{\mathbf{W}_{\pi(A)}\}$. In turn, this hull is in the affine plane passing through these points and c(1,...,1) where c is the sum of the components of **W**. The analysis is much easier when this plane is a linear subspace of R^n. This motivates the first of the two assumptions I impose on the voting vectors. Because the election outcome for **W** always agrees with the outcome for $\mathbf{W}' = a\mathbf{W} + b(1,\ldots,1)$, these two assumptions only fix the values of a and b. The first one fixes the value of b.

VECTOR NORMALIZATION: The sum of the components of a voting vector equals zero.

Example 9. The voting vector for a plurality election is (1,0,...,0), a vector normalized form is (n-1,-1,...,-1). A vector normalized form of the BC, (n-1,n-2,...,1,0) is (n-1,n-3,...,n+1-2i,...,1-n). So, if n=3, the vector normalized form of a Borda vector is (2,0,-2) while it is (3,1,-1,-3) for n=4.

The vector normalization forces the vectors $\mathbf{W}_{\pi(A)}$ to be orthogonal to (1,...,1), so the vote tally, Eq. 6.1, is in the linear subspace of R^n with the normal vector (1,...,1). Let E^n denote this (n-1) dimensional space; it is the space of interest.

Now, consider all $2^n-(n+1)$ subsets of candidates. Corresponding to each set S_j, there is a division of an Euclidean space of dimension $|S_j|$ into ranking regions. For convenience, assume that the s_j coordinates of this space have the same subscript as the alternatives, and that they are listed in increasing order. For instance, if $S_k=\{a_1,a_4,a_7\}$, the the corresponding *two* dimensional linear subspace for the *three* candidates, E^3, has the coordinates x_1, x_4, and x_7.

Let E(n) be the cartesian product of the $2^n-(n+1)$ linear subspaces E^k. A ranking region in E(n) is obtained for the product of the ranking regions of the component spaces. For instance, $\{x_1=x_2,\ x_2>x_3,\ x_1<x_3,\ x_2>x_1>x_3\}$ is a ranking region in E(4) that corresponds to the element $(a_1=a_2,\ a_2>a_3,\ a_1<a_3,\ a_2>a_1>a_3\}$ in U^4. It is easy to see that there is a one to one correspondence between the ranking regions of E(n) and the entries in U^n.

In the obvious manner, the ranking A defines a ranking for each of the subsets of candidates. Let $\mathbf{W}^n$ be the system vector. By the choice of the coordinate axis for each of the component spaces of E(n), $\mathbf{W}^n$ is in the ranking this same ranking region. All that needs to be proved is that $\mathbf{p}'$ can be found that has rational components. This simple argument is given in the first part of the proofs in [17]. This completes the proof of the Proposition 2.

Proof of Theorem 6. This is a simple exercise using the inverse image of G. If a symbol does not involve any tie, it is in an open ranking region, and the inverse image of an open set is open. Secondly, the structure of these sets with the hyperplanes follows immediately from the linear form of G.

According to the proposition, to prove Theorem 1 we want to show that $V(\mathbf{W}^n) = E(n)$ for most system vectors. My proofs are based on the algebraic, permutation group structure involved in changing a ranking from A to $\pi(A)$. To illustrate the basic idea, I will start with the proof of Theorem 4.1. Here, the group structure is simpler because it involves the permutations of only one set of candidates.

Proof of Theorem 4.1. This theorem involves only the set of all n candidates. The interesting feature is that with the n-1 different tallying processes, the set of outcomes is replicated n-1 times. Therefore the election tally is

6.4 $\quad G(-,\mathbf{W}_1,\ldots,\mathbf{W}_{n-1})\colon Si(n!) \longrightarrow (E^n)^{n-1}$

where $\{\mathbf{W}_j\}$, $j=1,\ldots,n-1$, are the different voting vectors satisfying the conditions of the theorem and where, if $\mathbf{p} = \{n_{\pi(a)}\}$, then

6.5. $\quad G(\mathbf{p},\mathbf{W}_1,\ldots,\mathbf{W}_{n-1}) = \Sigma\, n_{\pi(A)}\ (\mathbf{W}_{1\pi(A)},\ldots,\mathbf{W}_{n-1,\pi(A)})$

From the assumptions on the voting vectors, for each j, $E^n = \text{span}\{\mathbf{W}_{j,\pi(A)}\}$. To prove the theorem, we need to prove that $\{E^n\}^{n-1}$ agrees with $V* = \text{span}(\mathbf{W}_{1\pi(A)},\ldots,\mathbf{W}_{n-1,\pi(A)})$.

Let $P'_{jk}(\mathbf{x})\colon E^n \longrightarrow E^n$ be the permutation mapping that interchanges the j^{th} and the k^{th} components of $\mathbf{x}$, and let $P_{jk}(\mathbf{X}) = (P'_{jk}(\mathbf{x}_1),\ldots,P'_{jk}(\mathbf{x}_{n-1}))$ be a mapping from $\{E^n\}^{n-1}$ back to itself where $\mathbf{X} = (\mathbf{x}_1,\ldots,\mathbf{x}_{n-1})$. Let G_p be the group of permutations generated by the $n(n-1)/2$ permutation mappings $\{P_{jk}\}$, and define $L\{G_p\} = \{V\colon V$ is a linear subspace of $(E^n)^{n-1}$ that is invariant under $G_p\}$. Thus, if V is in $L(G_p)$, and if P is a permutation mapping from G_p, then $P(V)=V$. Such a mapping, P, just permutes the components of the vectors, therefore $V*$ is in $L(G_p)$. To prove the theorem, I will characterize the elements of $L(G_p)$.

(n-1) because, from Claim 3, $\dim(V^*) = (n-1)\dim(V^*_{12}) = (n-1)^2 = \dim(\{E^n\}^{n-1})$, so $V^* = \{E^n\}^{n-1}$.

Assume that $\dim(V^*_{12}) = j$ and that a basis for V^*_{12} is $\{\mathbf{c}_i\}$, $i=1,\ldots,j$. Using a standard row reduction argument of the type used to convert a matrix into a diagonal form, we can assume that the basis $\{\mathbf{c}_i\}$ is replaced with the equivalent basis $\{\mathbf{d}^{12}{}_i\}$ where

6.7 $$\mathbf{d}^{12}{}_i = (a^i{}_1(\mathbf{e}_1-\mathbf{e}_2),\ldots,a^i{}_{n-1}(\mathbf{e}_1-\mathbf{e}_2))$$

where $a^i{}_k = 1$ if $i=k$, $i=1,\ldots,j$; $a^i{}_k = 0$ for $k\leq j$ and k different from i; and not all of the remaining $a^i{}_k$ terms can equal zero because this would force V* to be zero for these component spaces. (In turn, this forces the contradiction $\mathbf{W}_s=0$.)

According to Claim 2, a basis for V^*_{1j} is given by Eq. 6.7 where the index 2 is replaced with j; the $a^i{}_k$ terms are the same. This means that the vector $(\mathbf{W}_1,\ldots,\mathbf{W}_{n-1})$ can be expressed as a linear combination of the $\{\mathbf{d}^{1k}{}_i\}$. But if $j=\dim(V^*)<n-1$, it follows that there are j equations, $\mathbf{W}_s = -\Sigma a^s{}_k\mathbf{W}_k$, where the summation is over $k=j+1,\ldots,n-1$, and $s=1,\ldots,j$. This means that the set of vectors $\{\mathbf{W}_j\}$ is linearly dependent. This contradiction proves the theorem.

The proof the Theorem 4.1 is based upon the permutation group structure satisfied by the linear space defined by the vote tally. You can view this in the following manner. If there is a permutation in the voters' preferences, then the new vote tally must also be in V*. The change in the voters' preferences just permuted the entries of the voting vectors. Indeed, the *only* property about a voting vector that was used is that with appropriate choices of P'_{jk}, any $\mathbf{W}_{j\pi(A)}$ can be mapped to some other $\mathbf{W}_{j\pi'(A)}$. Thus, the conclusion holds should voting vectors be replaced with any other vector. The proofs of all of our theorems are proved in much the same fashion; we characterize what happens as the permutation groups act on the domain, or preferences of the voters. The proof of Theorem 4.2 demonstrates there is a difference in the analysis when we have more than one subset of candidates; this difference is basic for all of social choice mappings.

Proof of Theorem 4.2. To demonstrate the basic ideas, I will first prove the theorem where only one voting vector is selected for each set. So, let $\mathbf{W}_{j+1}$ be the voting vector for the set $S_j=\{a_1,\ldots,a_{j+1}\}$, and let $\mathbf{W}_F = (\mathbf{W}_2,\ldots,\mathbf{W}_n)$ be the system vector for this *family of sets*. For each set S_j, there is an associated j dimensional subspace, E^j, of the j+1 candidates. We need to show that the candidates. Then, the arguments from the proof of Theorem 4.1 apply to finish the proof.

In the proof of the rest of the theorems, the permutations always effect more than one subset of candidates., so this complicates the proofs. Technically, we are using aspects of the orbit of *the iterated wreath product of permutation groups* to prove the theorem. (It turns out that related arguments can be used to prove and extend Arrow´s theorem.) To help with the bookkeeping, we introduce the following definitions.

Definition. Let n alternatives be given. Both the subset of alternatives S_d and its corresponding linear subspace E^{d}' are called (j,..,k) component subspaces iff S_d contains the alternatives $\{a_j,..,a_k\}$.

The next definition corresponds to choosing a value of "a" in the choice of a representation for a voting vector.

Definition. A *scalar normalization* of a system vector W^n is a choice of $2^n-(n+1)$ positive scalars, c_j, used to define the equivalent system vector $(c_1W_1,\ldots,c_jW_j,..)$. The *standar scalar normalization* for B^n is where the the Borda vector for k alternatives is given by (k-1,k-3,..,k+1-2i,..,1-k) and the vectors for two alternatives are (1,-1).

First we prove the following.

Proposition 3. $\mathrm{Dim}(V(B^n)) = n(n-1)/2$.

Proof. Consider

6.9 $$Y_{jk} = B^n_{\pi(A)} - P_{jk}(B^n_{\pi(A)}),$$

where $\pi(A)$ ranges through all rankings where a_j is the i^{th} ranked candidate and a_k is the $(i+1)^{th}$ ranked candidate, i = 1,..,n-1. This vector difference is 0 in any component space E^d that is not a (j,k) component space, and, in the (j,k) acomponent spaces, it is $2(e_j-e_k)$. Therefore, Y_{jk} is well defined.

There are only n(n-1)/2 distinct veectors in the set $\{Y_{jk}\}$. What we show is that $V(B^n) = \mathrm{span}\{\{Y_{jk}\}\}$, and that the vectors $\{Y_{jk}\}$ are linearly independent.

6.12 $P_{jk}(\mathbf{W}^n_{\pi(A)}) - \mathbf{W}^N_{\pi(a)}$

has a non-negative multiple of $\mathbf{e}_j-\mathbf{e}_k$ in each (j,k) component space and 0 in all others. In each (j,k) component space, there are choices of $\pi(A)$ where the multiple is positive. Therefore, if all of the vectors obtained from Eq. 6.12 are added, the sum has 0 for any non-(j,k) compnent space, and a positive multiple of $\mathbf{e}_j-\mathbf{e}_k$ for all (j,k) component spaces. This multiple depends only on the choice of the voting vector for each (j,k) component space, not on the choice of the particular j and k. (This is because the scalar for a particular (j,k) component space depends only on how often two alternatives can be ranked in the i^{th} and $(i+1)^{th}$ position as $\pi(A)$ varies over all specified rankings - it is independent of the particular choice of the alternatives.) Thus the scalar components are independent of j, k and i; they depend only on the choice of the voting vector for the set of candidates. Choose the scalar normalization for $\mathbf{W}^n$ so that, after normalized, all of the scalars in the sum equal the scalar obtained for the binary $\{a_j, a_k\}$. This means that the sum is a multiple of $\mathbf{Y}_{jk}$. This is true for all values of (j,k), so it now follows that $V(\mathbf{B}^n)$ is a linear subspace of $V(\mathbf{W}^n)$.

Finally, we must show that if $\mathbf{W}^n = \mathbf{B}^n$, then $V(\mathbf{B}^n)$ is a proper subspace of $V(\mathbf{W}^n)$. This involves a more careful application of the last argument. Because the choice of j and k determines only which subset of candidates are being considered, we can start with the indices 1 and 2. Again, consider the vector differences

6.13 $\mathbf{W}^n_{\pi(A)} - P_{12}(\mathbf{W}^n_{\pi(A)})$

where only $\pi(A)$ varies. To simplify the notation, let w^*_{cs} denote the difference between the s^{th} and the $(s+1)^{th}$ weights in the voting vector $\mathbf{W}_c$. When $\pi(A) = A$, the vector difference in Eq. 6.13 has $w^*_{c1}(\mathbf{e}_1-\mathbf{e}_2)$ in each (1,2) component space S_c, and 0 in all others. Next, consider all choices of $\pi(A)$ where a_1 is the second ranked alternative and a_2 is the third ranked alternative. The only alternatives that concern us are those ranked above a_1. So, consider the n-2 rankings obtained in the following order: The $(j-2)^{th}$ ranking has a_j top ranked, the ranking of the alternatives below a_2 is given in some arbitrary fashion. Here, the vector difference Eq. 6.13 has a scalar multiple w^*_{c2} for the (1,2,j) component spaces, the multiple w^*_{c1} for the rest of the (1,2) component spaces, and 0 for all other spaces. This vector is independent of what alternative is ranked in the k^{th} position, k>3.

maximal dimension only if the variables are in a particular lower dimensional algebraic set. Namely, the spanning vectors $\{\mathbf{W}^n_{\pi(A)}\}$ have additional dependencies given by the vanishing of certain determinents; these determinent conditions define the algebraic equations.

Suppose $V(\mathbf{W}^n)$ is a proper subspace of $E(n)$. This means it has a normal vector, $\mathbf{N}$, in $E(n)$. By using the basis for $E(n)$, it follows that there is a choice of (j,k) and some component space of $E(n)$ so that $\mathbf{N}$ is nonzero in a $\mathbf{e}_j-\mathbf{e}_k$ direction. Without loss of generality, assume that $(j,k) = (1,2)$. This means that $V(\mathbf{W}^n)$ does not contain the subspace of $E(n)$ generated by the product of $(\mathbf{e}_1-\mathbf{e}_2)$ from each $(1,2)$ component subspace of $E(n)$. The proof of the theorem is based on showing that there is a $\mathbf{W}^n$ where $V(\mathbf{W}^n)$ contains the full subspace generated by the product of the $(\mathbf{e}_1-\mathbf{e}_2)$ subspace. This contradicts the existence of $\mathbf{N}$ and it proves the theorem.

The choice of the system vector, $\mathbf{W}^n$, is where all of the voting vector components are plurality vectors, so $w^*_{c,1}>0$ for each $(1,2)$ component space S_c, and $w^*_{c,j}=0$ for $j>1$. To show independence, the rankings described in the proof of Theorem 3 are used, but I will start with the last rankings and word forward. So, if a_1 and a_2 are bottom ranked alternatives $(i=n-1)$, then only in the set $\{a_1,a_2\}$ can they be the two top ranked alternatives. For this $\pi(A)$, the vector difference Eq. 6.13 has $\mathbf{0}$ in all component spaces except the one corresponding to this binary, where the entry is $\mathbf{e}_1-\mathbf{e}_2$.

Next, consider $i=n-2$ where a_j is the bottom ranked alternative. Only in the sets $\{a_1,a_2\}$ and $\{a_1,a_2,a_j\}$ are a_1 and a_2 top ranked. For this $\pi(A)$, the vector difference Eq. 6.13 has nonzero entries only in these two component spaces. The first vector obtained from $i=n-1$ is used to eliminate the $\{a_1,a_2\}$ component of this vector. Thus, $V(\mathbf{W}^n)$ contains both the vector with the only nonzero vector $\mathbf{e}_1-\mathbf{e}_2$ in the $\{a_1,a_2\}$ component space and the vector with the only nonzero component $\mathbf{e}_1-\mathbf{e}_2$ in the $\{a_1,a_2,a_j\}$ component space. The obvious induction argument proves that, for each $(1,2)$ component space of $E(n)$, $E(n)$ has a vector where its only nonzero term is $\mathbf{e}_1-\mathbf{e}_2$ in this component space. This completes the proof of the theorem.

Proof of Theorem 2. Let $n=3$. The proof is much the same as the proof of Theorem 1 except that I now show that the conclusion holds for all non Borda vectors. Again, without loss of generality, we can concentrate on the $(1,2)$

The assumption on the voting vectors is that they, and the Borda vector, form n-1 linearly independent vectors. These vectors still are linearly independent if they are replaced by vectors giving the differences between successive weights. This matrix is the one given above. This proves the theorem.

Proof of Theorem 5.2. For n alternatives, consider the set of system vectors $\{\mathbf{W}^n_j\}$ that are defined in the following manner. Each component voting vectors of $\mathbf{W}^n_j$ is of the form $(1,1,\ldots,1,0,\ldots,0)$: it assigns one point to a voter's k top ranked candidates and zero for all others. Furthermore, there is precisely one system vector corresponding to any assignment of voting vectors of this kind to the subsets of candidates. (Hence, for n=3, there are only two system vectors: $(1,0;1,0;1,0;1,0,0)$ and $(1,0;1,0;1,0;1,1,0)$. For n=4, there are $2^4(3)$ system vectors. For n alternatives, there are $k(n) = 2^{n(n-1)(n-2)/3!}(3^{n(n-1)n-2)(n-3)/4!})\ldots(n-1)$ system vectors.) Let $G(-,\mathbf{W}^n_1,\ldots,\mathbf{W}^n_{k(n)})$ be the mapping that represents the voting tally for the n system vectors. By using the argument showing that $V(\mathbf{B}^n)$ is a subspace of $V(\mathbf{W}^n)$, it follows that the space spanned by $\{(\mathbf{W}^n_1,\ldots,\mathbf{W}^n_{k(n)})_{\pi(A)}\}$ contains the $n(n-1)/2$ dimensional space given by the span$((\mathbf{Y}_{jk},\mathbf{Y}_{jk},\ldots,\mathbf{Y}_{jk}))$ as j<k vary over all pairs. This means that if w is a word in $D(\mathbf{B}^n)$, then there is a profile such that the word for each of the n system vectors is w. Now, let $\mathbf{W}^{n\prime}$ be any system vector. Our choice of k(n) system voting vectors is a basis for the system vectors, and any system vector is in their convex hull. This means that any outcome of this system vector, $\mathbf{W}^{n\prime}$, is in the convex hull of the outcomes for the basis system vectors. Because all of the basis system vectors have the same outcome, and because (by Theorem 6) the regions supporting a word are convex, the outcome for $\mathbf{W}^{n\prime}$ is in the same ranking region. This completes the proof.

Acknowledgement. This research was supported, in part, by NSF grant IST 8415348. Some of this research was done while I was a vistor at CEREMADE, University of Paris, and at the System and Decision Science Program of IIASA in Vienna. I would like to thank my hosts J.P. Aubin and A. Kurzhanski. This paper is based on results in "The optimal ranking method is the Borda count", NU Center for Math Econ Discussion paper #636, Jan., 1985.

15. D. G. Saari, "The Borda Dictionary", to appear

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

17. D. G. Saari, "The source of some paradoxes from social choice and probability", *JET*, **41** (1987), 1-22.

18. D. G. Saari, "Susceptibility to manipulation", NU Center for Math. Econ discussion paper, June, 1987.

19. D. G. Saari, "Chaos and the theory of elections", In *Nonlinear Dynamics*, ed. A. Kurzhanski and K. Sigmund, Springer Verlag Series on Mathematical Economics, 1987.

20. D. G. Saari and J. Van Newenhizen, "The problem of indeterminacy in approval, multiple, and truncated voting systems. *Public Choice* forthcoming

21. D. G. Saari and J. Van Newenhizen, "Is Approval Voting an 'unmitigated evil'?", *Public Choice* forthcoming.

22. A. Sen, "Social choice theory", Chp. 22 in *Handbook of Mathematical Economics*, *Vol. III* ed. K. J. Arrow and M. D. Intriligator, North-Holland: Amsterdam, 1986.

23. H. Sonnenschein, "Market excess demand functions", *Econometrica* **40** (1972), 549-563