Soc Choice Welfare (1990) 7:279–317

Social Choice and Welfare

The Borda dictionary*

D.G. Saari

Departments of Mathematics and Economics, Northwestern University, Evanston, IL 60208-2730, USA

Received November 28, 1988/Accepted September 17, 1990

Abstract. For n candidates, a profile of voters defines a unique Borda election ranking for each of the $2^n-(n+1)$ subsets of two or more candidates. The Borda Dictionary is the set of all of these election listings that occur for any choice of a profile. As such, the dictionary contains all positive features, all flaws, and all paradoxes that can occur with single profile, sincere Borda elections. After the Borda Dictionary is characterized, it is used to show in what ways the Borda Count (BC) is an improvement over other positional voting methods and to derive several new BC properties. These properties include several new characterizations of the BC expressed in terms of axiomatic representations of social choice functions, as well as showing, for example, that the BC ranking of n candidates can be uniquely determined by the BC rankings of all sets of $k<n$ candidates for any choice of k between 2 and n.

The Borda Count (BC) is the simple method used to tabulate ballots where, for n candidates, $n-i$ points are assigned to a voter's i^{th} ranked candidate; the candidate with the most points wins. While the BC has both attractive features and flaws, only some of them are known. To redress this situation, I characterize *everything* that could possibly happen when the BC is used. To understand this assertion, note that a profile determines a unique ordinal BC election ranking for each of the possible subsets of candidates; i.e., associated with a profile is a unique listing of election rankings. In this paper I characterize all possible BC election listings that could ever occur. I call this collection of all possible BC election listings a *Borda Dictionary*. By construction, the Borda Dictionary contains everything that could possibly occur with a BC election, so it catalogues all of the BC flaws, all of the "single profile" BC paradoxes, and all of the "single profile" BC positive features.

* This research was supported, in part, by NSF grants IRI8415348, IRI-8803505 and a Fellowship from the Guggenheim Memorial Foundation.

This goal of characterizing the Borda Dictionary continues my program, initiated in [10] (also see [8, 9]), to characterize everything that could possibly occur with positional election procedures.[1] A conclusion of [10], repeated in Sect. 1.2, is that *the Borda method is the unique positional voting method to minimize the kinds and number of paradoxes that can occur.* Thus, this conclusion explains why one should expect to find statements in the literature suggesting that the BC enjoys a favored status. It also follows from this assertion that, in some sense, the BC plays a critical role within the class of positional voting methods. In my development of the Borda Dictionary, given here and in the companion paper [12], I outline the mathematical reasons why the BC has this favored status. Also, I show in what ways the BC is an improvement over other positional voting methods, and I indicate in Sect. 1.3 how these results can be used to understand the class of paradoxes involving changes in the profiles. In doing so, I discuss certain paradoxes involving choice procedures such as tournaments, the Hare method, runoff elections, etc.

Because the Borda Dictionary contains all possible listings of BC election outcomes, one might correctly surmise that it can be used to derive and extend many of the known BC conclusions. To illustrate this, I rederive Smith's important conclusion [15] that the BC ranking of the candidates is related to the majority vote rankings of all pairs of candidates. As I show, Smith's result is just one of many different possibilities; e.g., I show for any value of k satisfying $n > k \geq 2$ that the BC ranking of all n candidates is related to the BC rankings of all sets of k candidates. (There are many other collections of subsets of candidates with the nice property that the BC rankings are related; one of the purposes of the companion paper [12] is to completely characterize these families of subsets.) To indicate another way in which the Borda Dictionary can be applied, I show how to use the dictionary to significantly extend, in several different directions, those characterizations of the BC based on the axiomatic properties of social choice functions. (These results, which start in Sect. 3.2, can be read independently of the technical Sect. 2.) Indeed, if one views the BC Dictionary as a reference tool – a place to start in the analysis of choice methods involving positional voting methods – then a surprisingly large number of other kinds of conclusions can be derived by use of the BC Dictionary. Therefore, rather than trying to provide an exhaustive listing of all of these new results, I adopt the strategy of showing how to use this dictionary to obtain these new statements. In this spirit, several of my examples are designed not only to illustrate particular points, but also to offer new BC conclusions.

In summary, one theme of this paper is that if a collection of subsets of candidates admit relationships among their positional election rankings, then the number and kinds of possible relationships is maximized if the BC is used. The main purpose of this paper is to create the mathematical tools needed to determine the entries in the BC Dictionary and the relationships among BC election rankings. After notation is introduced, the remainder of this introductory section is devoted toward suggesting the advantages of using "dictionaries" and the properties of the BC to analyze social choice issues that involve positional voting rankings. In Sect. 2 the mathematical structures needed to construct the Borda Dictionary are developed. These technical aspects place an emphasis on a vector space interpretation for election outcomes. The purpose of Sect. 3 is to indicate

[1] Although my proofs use the arguments developed in [10], familiarity with [10] is not required for a first reading of this paper.

some of the consequences of the Borda Dictionary; one of the main themes of this section is to introduce several new axiomatic characterizations of the BC. Another theme is to analyze the kinds of relationships admitted by the BC over specified collections of subsets of candidates.

So, if a family of subsets admits relationships among the BC rankings, the techniques of this paper can be used to determine them. The converse issue is to determine what collections of subsets of candidates admit relationships among the BC tallies. For instance, if one is interested in considering only the triplets of candidates, will there be a relationship among the BC rankings? (There will.) This issue of characterizing the collection of sets of candidates that do admit relationships among the BC rankings is the theme of the companion paper [14].

1.1. Notation and dictionaries

To motivate the notation and the idea of a dictionary, consider the following two "paradoxes".

Example 1. a. One of the oldest voting paradoxes, the *Condorcet Cycle*, occurs when, say, 5 voters have the rankings $c_1 > c_2 > c_3$, 5 have $c_2 > c_3 > c_1$, and 5 have $c_3 > c_1 > c_2$. By majority votes of 10 to 5, these voters prefer $c_1 > c_2$, and $c_2 > c_3$. Therefore, one might suspect that the voters prefer c_1 to c_3. The paradox is that, by a vote of 10 to 5, they prefer $c_3 > c_1$; the voters' election rankings create a cycle.
b. A second example [10] has 6 voters with the ranking $c_3 > c_1 > c_2$, 5 with $c_2 > c_1 > c_3$, and 4 with $c_1 > c_2 > c_3$. The plurality ranking is $c_3 > c_2 > c_1$. If candidate c_1 were to withdraw, then it is not uncommon to assume that the electing group's ranking now is given by the truncated ranking of $c_3 > c_2$. However, by majority votes of at least 9 to 6, the majority vote ranking of each pair is the exact opposite of its relative ranking in the plurality outcome. By majority votes, these same, sincere voters prefer $c_1 > c_2, c_2 > c_3$, and $c_1 > c_3$.

Thus, a "paradox", as used here, is where a profile of voters (i.e., choices of complete, binary, transitive rankings without indifference of the candidates for each of the sincere voters) determines election rankings among the subsets of candidates that unveil an unexpected, counter-intuitive outcome. As indicated in (b), paradoxes illustrate that the way we use and interpret election rankings may be incorrect. So, to understand what actually can occur (i.e., to find all possible paradoxes of this kind), we need to determine all listings of election rankings over all possible subsets of candidates that are associated with a profile. To start, list all feasible subsets of candidates. With the $n \geqq 2$ candidates $C^n = \{c_1, \ldots, c_n\}$, there are $2^n - (n+1)$ subsets with enough candidates (at least two) to permit an election. List these subsets as $\{S_1, \ldots, S_{2^n-(n+1)}\}$ where, for convenience, the first $n(n-1)/2$ subsets are the pairs of candidates, the next $n!/3!(n-3)!$ are the sets of three candidates, etc. Also, in each S_j, list the candidates in the lexicographic order determined by the subscripts. For example, with $C^3 = \{c_1, c_2, c_3\}$, the sets could be $S_1 = \{c_1, c_2\}$, $S_2 = \{c_1, c_3\}$, $S_3 = \{c_2, c_3\}$, and $S_4 = \{c_1, c_2, c_3\}$.

For each S_j, let R_j be the set of all complete, binary, reflexive, transitive rankings of S_j. Thus R_j is the listing of all possible ordinal election rankings associated with the S_j candidates. For instance with $S_3 = \{c_2, c_3\}$, $R_3 = \{c_2 > c_3, c_2 = c_3, c_3 > c_2\}$, while R_4 contains all 3! linear rankings of the three candidates in S_4 along with the 7 rankings that have a tie among the candidates.

The *universal space*, $U^n = R_1 \times R_2 \times \ldots \times \mathrm{R}_{2^n-(n+1)}$, is a product space, so an element of U^n is a listing of $2^n-(n+1)$ rankings – there is a ranking assigned to each subset of candidates. (To illustrate, $\{c_2 > c_1, c_1 > c_3, c_3 > c_2, c_1 = c_3 > c_2\}$ specifies a ranking for each subset of candidates, so it is an element of U^3). Thus U^n is the space of all possible, as well as all impossible coordinated (ordinal) election outcomes over the subsets of candidates.

A positional voting system for a set of candidates is where specified weights, $\{w_i\}$, are used to tally the ballots. In tabulating the ballots, w_i points are assigned to a voter's i^{th} ranked candidate, and the election ranking of each candidate is determined by the total number of points she receives. For k candidates, the assigned weights define a *voting vector* $\mathbf{W} = (w_1, w_2, \ldots, w_k)$ where $w_i \geqq w_{i+1}$ and $w_1 > w_k$.[2] Let $\mathbf{W}_j$ designate the voting vector assigned to tally the ballots for the candidates in S_j. The listing of the $2^n-(n+1)$ selected voting vectors, the *system voting vector* is

$$\mathbf{W}^n = (\mathbf{W}_1; \mathbf{W}_2; \ldots; \mathbf{W}_{2^n-(n+1)}) \ . \tag{1.1}$$

The obvious equivalence relationship among the voting vectors can be described with the vector $\mathbf{E}_k$ where, with k candidates, $\mathbf{E}_k = (1, \ldots, 1)$.

Proposition 1. *Let* $\mathbf{W}$ *be a voting vector for a set of k candidates. If* $a > 0$ *and b are scalars, then the ordinal election rankings of an election tallied with the voting vector* $\mathbf{W}$ *and with* $a\mathbf{W} + b\mathbf{E}_k$ *must always be the same.*

This proposition holds because the factor $a\mathbf{W}$ just scales the final tally while $b\mathbf{E}_k$ just adds the same value to the tally of each candidate. Throughout this paper, *I always use equivalence classes of voting vectors.* For instance, for $k=4$, both $(22, 18, 14, 10) = 4(3, 2, 1, 0) + 10(1, 1, 1, 1)$ and $(3, 1, -1, -3) = 2(3, 2, 1, 0) - 3(1, 1, 1, 1)$ are Borda vectors.

Definition. A *Borda voting vector* for $k \geqq 2$ candidates is any voting vector equivalent to the BC vector $(k-1, k-2, \ldots, 0)$. A *Borda system voting vector*, denoted by $\mathbf{B}^n$, is where a Borda voting vector is used to tally the ballots for all subsets of candidates.

In the obvious fashion, once a system voting vector, $\mathbf{W}^n$, is specified, then a given profile of voters, $\mathbf{p}$, uniquely determines a listing of rankings. This listing, denoted by $f(\mathbf{p}; \mathbf{W}^n)$, consists of the election ranking for each subset of candidates.

Example 2. List the subsets of C^3 as $\{\{c_1, c_2\}, \{c_1, c_3\}, \{c_2, c_3\}, \{c_1, c_2, c_3\}\}$.
a. If the system voting vector is $\mathbf{W}^3 = \{1, 0; 1, 0; 1, 0; 1, 0, 0)$, (i.e., a majority election is used for the first three subsets of candidates and a plurality election for S_4) and if $\mathbf{p}_a$ is the profile from Example 1.a, then $f(\mathbf{p}_a, \mathbf{W}^3) = (c_1 > c_2, c_3 > c_1, c_2 > c_3, c_1 = c_2 = c_3) \in U^3$.
b. If $\mathbf{p}_b$ is the profile in Example 1b, then $f(\mathbf{p}_b, \mathbf{W}^3) = (c_1 > c_2, c_1 > c_3, c_2 > c_3, c_3 > c_2 > c_1) \in U^3$.
c. $f(\mathbf{p}_a, \mathbf{B}^3) = (c_1 > c_2, c_3 > c_1, c_2 > c_3, c_1 = c_2 = c_3)$, while $f(\mathbf{p}_b, \mathbf{B}^3) = (c_1 > c_2, c_1 > c_3, c_2 > c_3, c_1 > c_2 > c_3)$.

[2] A *reversed positional voting system* is where $w_j \leqq w_{j+1}$ and where the winning candidate is the one with the lowest total. The main difference is to admit negative values for "*a*" in Proposition 1. With only minor changes, the conclusions of this paper and [10] hold for such systems – some of these extensions are included here with "scoring methods." See [6, 7] for more discussion.

The paradox in Example 1 b disappears when the BC, rather than the plurality vote, is used to tally the ballots. As shown in Sect. 3, this is no coincidence.

Definition. Let $n \geqq 2$ candidates be given. For a given $\mathbf{W}^n$, the *dictionary generated* by $\mathbf{W}^n$ is the set

$$D(\mathbf{W}^n) = \{ f(\mathbf{p}, \mathbf{W}^n) : \mathbf{p} \text{ is a profile for a finite number of voters; the number of voters can change with the profiles} \} \ . \quad (1.2)$$

An entry in the dictionary - the sequence of election rankings over the various subsets of candidates - is called a *word*. Each ranking within a word is called a *symbol*.[3]

In the discussion about axiomatic representations of the BC, we will have need to look at *scoring rules* (Young [16]). This is where the voting vector is replaced by any vector; the only requirement is that all of the components are not the same. Thus, for example, $(-2, 5, 0)$ is a scoring vector where -2 points are assigned to a voter's top ranked candidate, 5 points to the second ranked candidate, and zero points for the bottom ranked candidate. It follows immediately that a positional voting vector is a scoring vector, but a scoring vector need not be a voting vector. A *system scoring vector* is a vector $\mathbf{W}^n = (\mathbf{W}_1, \ldots, \mathbf{W}_{2^n-(n+1)})$ where $\mathbf{W}_j$ is a scoring vector for S_j. In characterizing scoring methods, I will need to use the reversed BC. This is where the value of "a" in Proposition 1 is negative. Thus, let $\mathbf{BC}^n$ be a generic representation for a system scoring vector where each scoring vector entry is either a BC vector, or a negative multiple of a BC vector. Notice that $\mathbf{BC}^n$ can represent a large number of different choices; e.g., for $n=4$ candidates, there are 5 sets of three or more candidates. Thus, there are $2^5 - 1$ different choices of $\mathbf{BC}^n$ that are not system voting vectors.

1.2. Characterization of dictionaries

There is an important difference between a *word* in $D(\mathbf{W}^n)$ and an *element* of U^n. A word is a listing of election outcomes realized by a profile whereas an element of U^n is a sequence of rankings where it may, or may not be possible to achieve them with an election. In other words, if an element of U^n is not in $D(\mathbf{W}^n)$,then it identifies a listing of rankings that never can be attained in an election tallied with the system vector $\mathbf{W}^n$. (For instance, the sequence from Example 1.b, $\{c_1 > c_2, c_1 > c_3, c_2 > c_3; c_3 > c_2 > c_1\}$, is in U^3, but, as shown in Smith [15] and in Corollary 3.1, this listing cannot be in $D(\mathbf{B}^3)$. Thus such a listing of election outcomes never can occur with the BC). Therefore a dictionary, $D(\mathbf{W}^n)$, catalogues all lists of election rankings that ever can happen with $\mathbf{W}^n$, while the complement of $D(\mathbf{W}^n)$ in U^n specifies those listings of rankings that never can occur as election outcomes with $\mathbf{W}^n$.

By construction, $D(\mathbf{W}^n)$ is a subset of U^n. If a system voting vector $\mathbf{W}^n$ admits only a small number of inconsistencies and potentially undesired outcomes, then $D(\mathbf{W}^n)$ is a small subset of U^n. This is not the general situation. The following theorem summarizes those results from [10] basic for my current discussion. Recall

[3] This useful terminology correctly emphasizes my theme that a dictionary serves as a starting point for the analysis of election procedures. Moreover, this term and others reflect the fact that the motivation for this approach came from "Chaos" and "Symbolic Dynamics" in dynamical systems. See [7, 8] for an exposition of the connection.

that $\mathbf{W}^n$ is a vector in a Euclidean space and that an algebraic set is a lower dimensional subset of this Euclidean space determined by the zeros of a given set of polynomials.

Theorem 1 [10] **a.** *Let $n \geqq 3$ candidates be given. With the exception of an algebraic subset, a^n, of possible choices for system voting vectors,*

$$D(\mathbf{W}^n) = U^n \ . \tag{1.3}$$

b. *For $n=3$, $\mathbf{B}^3$ is the only system voting vector in a^3. Namely, $D(\mathbf{B}^3)$ is a proper subset of U^3, and if $\mathbf{W}^3 \neq \mathbf{B}^3$, then $D(\mathbf{W}^3) = U^3$.*
c. *For all $n \geqq 3$, if $\mathbf{W}^n \neq \mathbf{B}^n$, then $D(\mathbf{B}^n)$ is a proper subset of $D(\mathbf{W}^n)$.*
d. *Let $n \geqq 3$. Let $\mathbf{w}$ be a word in $D(\mathbf{B}^n)$. There exists a profile of voters, $\mathbf{p}$, so that for any choice of a system voting vector, $\mathbf{W}^n$, $f(\mathbf{p}, \mathbf{W}^n) = \mathbf{w}$.*
e. *Let $\mathbf{W}^n$ be a system scoring vector. With the exception of an algebraic set of choices, $D(\mathbf{W}^n) = U^n$. If $D(\mathbf{W}^n)$ is a proper subset of U^n but where some scoring vector component of $\mathbf{W}^n$ is not a scalar multiple of a* BC *vector, then there is a choice of a $\mathbf{BC}^n$ system vector so that $D(\mathbf{BC}^n)$ is a proper subset of $D(\mathbf{W}^n)$.*

Part a asserts that for almost all system voting vectors, anything can happen. To appreciate the implications of this statement, notice that while many of the surprising, counter-intuitive election examples found in the literature compare election outcomes over different subsets of candidates, most of them use only a few of the symbols of a word. Part a asserts that far more startling paradoxes exist. To create one, just fill in the remaining symbols of the word in *any desired manner*, and Theorem 1 ensures there is a profile to support this conclusion. Thus, by using Theorem 1 it becomes trivial to extend all such examples and paradoxes from the literature in all possible ways. By using a dictionary it is possible to examine the effects of certain kinds of election properties. To illustrate these comments, recall that there are statements asserting that should the majority vote lead to a cycle, then there appears to be no way to select a winner. (For a nice review of this literature, see Nurmi [5]). However, a majority cycle with n candidates specifies only n of the symbols of a word; the remaining symbols can be filled in any desired fashion. For instance, the remaining $2^n - (2n+1)$ symbols could be filled in to be compatible with the ranking $c_1 > c_2 > \ldots > c_n$. In this way, it is clear that even with the majority cycle, the outcomes of all other elections can dramatically support the notion that c_1 is a "natural" selection for the voters.

By using the dictionary it now becomes quite easy to raise questions about almost any procedure. This is because a procedure often uses the rankings of only a limited number of subsets of candidates. These rankings specify the entries for a limited number of the symbols of the word; the rest of the symbols can be filled in any desired manner. In particular, they can be filled in a manner to suggest that a completely different outcome is more appropriate than the one advanced by the procedure. For example, in a standard run-off, the two top ranked candidates from the first stage are advanced to a second stage; here the majority vote winner is declared the winner. This procedure uses only two symbols, so fill them so that c_n and c_{n-1} are advanced to the run-off and c_n wins. Next, let all other symbols so be compatible with the ranking $c_1 > c_2 > \ldots > c_n$. Thus, the winner is c_n, but a very strong argument can be advanced that she should be bottom-ranked, not the selected alternative.

As an extreme case, part a guarantees that the wildest imaginable situations actually occur. As asserted in [10], one could even use a random number generator to determine a ranking for each subset of candidates, and, for almost all system voting vectors, there is a profile so that the randomly selected rankings are the sincere election outcomes of these voters. This conclusion that anything can happen is disturbing; it is difficult to accept that a voting procedure reflects the voters' true wishes when the outcomes can radically change with even minor changes in which group of candidates just happens to be considered. For example, suppose that with only one exceptional case, the ranking of all subsets of candidates is consistent with the ranking $c_1 > c_2 > \ldots > c_n$; the exceptional case is if c_n, and only c_n is absent. In this setting, the election ranking is $c_{n-1} > c_{n-2} > \ldots > c_2 > c_1$. Thus, it is reasonable to believe that c_1 is the top-choice of these voters; but this is not reflected by the election outcome if, for some reason, c_n, and only c_n drops out of the election. By choosing other words from a dictionary, the reader can design many other similar examples.

To avoid these difficulties where the rankings can radically change with minor changes in what candidates are available, it is reasonable to seek procedures that admit relationships among the election rankings of the different sets of candidates. In this way, one has assurances that it is not as likely for radically different outcomes to occur with only minor changes in what candidates are standing for election. Thus the crucial issue is to find a method that permits some consistency among the election rankings. Part b asserts that with three candidates, the only possible way relief can be attained is to use the BC; only the BC offers protection from all imaginable inconsistencies and paradoxes. This is illustrated in Example 2b, c where the plurality ranking is in direct conflict with the same voters' majority vote rankings of the pairs of the candidates while, with the same profile, the BC ranking is consistent with the pairwise rankings.

It turns out that part b of Theorem 1 does not extend; for all $n \geqq 4$, there are other choices of $\mathbf{W}^n \neq \mathbf{B}_n$ where $D(\mathbf{W}^n)$ is a proper subset of U^n. (An argument outlining why this is so is given in the last paragraphs of Sect. 3). However, even if $D(\mathbf{W}^n)$ is a proper subset of U^n, part c asserts that $D(\mathbf{B}^n)$ always is a proper subset of $D(\mathbf{W}^n)$. In other words, $\mathbf{B}^n$ is the unique choice of a positional voting method which minimizes both the number and the kinds of paradoxes. Thus, for instance, whenever one can find a situation (a word) that illustrates an undesirable feature of the BC, it follows from part c that this same feature must also occur with all other choices of positional voting methods. Part d strengthens this statement by asserting that associated with each word in the Borda Dictionary is a profile whereby this word (this same feature) is realized for *all possible positional voting methods.* Part e asserts that this same favored status for the BC extends to the setting where all possible choices of tallying ballots, including scoring and positional voting, are considered. To see the need of using different $\mathbf{BC}^n$ vectors, consider the system scoring vector $\mathbf{W}^3 = (1, 0; 1, 0; 1, 0; 0, 1, 2)$. Here, the symbols in the words from $D(\mathbf{W}^3)$ corresponding to the set of all three candidates is the exact reversal of the BC ranking.

1.3. Social choice procedures and Borda

One way to illustrate the central role played by the BC is to consider some of the properties of those procedures, such as tournaments, and various kinds of runoff elections that are based on the rankings of several subsets of candidates. These procedures include an "agenda", which is a listing of candidates $[c_1, c_2, \ldots, c_n]$. Here a majority vote is held between the first two listed candidates, and the winner is advanced to be compared in a majority vote with the next listed candidate. In general these *social choice* procedures (i.e., instead of ranking the candidates, the procedure finds the "best" candidate or candidates) start by ranking certain specified subsets of candidates. Then, the rankings and the rules of the procedures determine which candidates are ranked at the next stage. For instance, bottom ranked candidates may be dropped from further consideration, as in a runoff election; replaced with other candidates, as with an agenda; or matched against other "winning" candidates, as in certain kinds of tournaments. This process continues until a final set of candidates is ranked. The chosen candidate(s) is determined by the rankings of the final set.

The traditional way to analyze these social choice procedures is to construct a profile to verify that certain suspected properties do occur. However, constructing an appropriate profile can be a difficult combinatorial task. (This is why many of the results in the literature are restricted to small values of n and the plurality vote). Often the only purpose of constructing such a profile is to establish the existence of a particular word. But, the dictionaries catalogue all possible words that ever could occur, so there is no need to continue with this complicated combinatorial step. Indeed, by treating such social choice procedures as a *mapping* from a dictionary to the non-empty subsets of C^n, denoted by $P(C^n)$, the analysis becomes both more complete and much simpler. To provide a distinction, in this section a mapping from $D(\mathbf{W}^n)$ to the set of non-empty subsets of $C^n = \{c_1, \ldots, c_n\}$ is called a social choice mapping. In other words, the dictionary – the domain of the mapping – now becomes a candidate to replace the profiles as the primitive in the analysis. A social choice procedure, then, is the composition of the social choice mapping with the function $f(\mathbf{p})$ used in Eq. (1.2). (However, after this section, I will not make this fine distinction; the word mapping and procedure will be used interchangeably.)

To illustrate an advantage of taking this approach, I offer an important new result. Namely, *the possible outcomes of a given social choice procedure based on the Borda rankings are more restrictive than those based on any other positional voting ranking.* This statement, which is difficult to prove if profiles are the primitives, is an immediate consequence of the assertion that the Borda Dictionary is a proper subset of any other positional voting dictionary; i.e., the associated social choice mapping is restricted to a smaller domain. What this assertion implies is that one can expect more consistency among social choice mappings if their outcomes are based on the BC rankings rather than the rankings of any other positional voting method.

As a first example of this assertion, notice that it is almost an immediate consequence of the properties of the plurality dictionary ($D(\mathbf{W}^n) = U^n$) that the Hare method (see, for instance, Nurmi [5]) need not elect a Condorcet winner. (Recall, Hare's method is based on the plurality rankings. The bottom ranked candidate is eliminated, and the profile is used to rerank the remaining candidates. The procedure continues until only one candidate remains. To prove that a Condorcet winner need not be victorious, choose a word from $D(\mathbf{W}^n)$ where a Condorcet winner exists, but she is bottom ranked in the set of all candidates. According to Theorem 1, such a word exists, so this completes the proof). One might wonder whether Hare's method could be modified to avoid this property by using a positional voting method other than the plurality vote. According to the above assertion, to answer this question one must investigate what happens to Hare's method should it be based on rankings given by the BC; if this Condorcet property cannot happen with the BC, then it cannot occur with any positional voting procedure. However, as shown in Smith [15], if the BC is used, then this modified procedure does select a Condorcet winner when one exists. In other words, when analyzing any choice procedure based on the rankings, one should start with the BC rankings to find what faults and what positive aspects can occur; all faults will be inherited when any other positional voting method is used, but the positive features may be unique to the BC.

As another illustration, consider the IIA property from Arrow's Theorem. (Arrow [1]). Namely, for S_j a subset of S_k and for each fixed profile, the IIA condition requires the ranking of candidates in S_j to mimic the ranking of the candidates in S_k. As we see from the dictionaries for positional voting, this condition is not satisfied. Indeed, one can choose the symbols corresponding to these two subsets of candidates in any desired manner, and, as asserted by Theorem 1, there is a profile leading to this outcome. So, because IIA cannot be satisfied, it is reasonable to seek "relaxed versions" of the IIA condition; conditions that are weak enough so that they admit at least one choice of a positional voting method. According to the above assertion, a necessary and sufficient condition for the new axiom to satisfy this condition is if it is satisfied by the BC. This theme is examined in Sect. 3.

As a third illustration of this new assertion about the BC, we consider the problem of understanding which social choice procedures admit the troubling "abstention paradox."[4] This is where, by abstaining, a voter forces the final result to be personally more favorable than had he voted. This problem seems to have been primarily discussed in the context of run-off elections with the plurality vote (see Smith [15] and Brams and Fishburn [2]). But by using a dictionary, this issue can be analyzed for a large class of social choice procedures – including runoff elections, tournaments, and agendas – rather than just individual procedures as typically is the case when profiles are the primitive.

Definition. A social choice mapping is an assignment of profile to a set in $P(C^n)$; that is, a nonempty subset of candidates from $C^n = \{c_1, c_2, \ldots, c_n\}$. A social choice method based on the $\mathbf{W}^n$ rankings of subsets of the candidates is a mapping

$$f: D(\mathbf{W}^n) \to P(C^n) \ . \tag{1.4}$$

A social choice mapping is *binary susceptible* if

i) The method is based on the positional voting rankings, $\mathbf{W}^n$, of the candidates.
ii) There are $k \geqq 1$ stages. At each stage, the rankings of specified subsets of candidates are examined. After the first stage, the choice of some of the specified subsets of candidates may depend upon the rankings of sets of candidates at earlier stages. At the k^{th} and final stage, there is only one subset of candidates.
iii) There is a subset of candidates, called the *swing set*, and a ranking of this set whereby the reversal of two adjacently ranked candidates changes the choice of the final set of candidates to be ranked.
iv) The selected candidate is based on the ranking of the final set of candidates.
v) The image of f contains at least two different outcomes.

For many of the widely discussed binary susceptible procedures, such as tournaments, agendas, Hare's method, a standard runoff election, etc., any set of candidates, other than the final one, is a "swing set". This is because the

[4] I don't know the history of this paradox, but at least the flovor of it is described in Smith's paper [15]. It may have been discussed much earlier.

reversal of the relative rankings of some two candidates, usually the one just above the cutoff and the one just below, changes which set of candidates are advanced.

In [10] I show *for almost all choices of* $\mathbf{W}^n$ (i.e., those that are not in a^n), that *all binary susceptible social choice methods admit the abstention paradox.* (In [10], the term "disjoint procedure" was used). Namely, there is a profile $\mathbf{p}$ and two additional voters with identical rankings whereby if the two voters vote, the outcome will be personally less favorable than had they abstained. (The proof of this assertion shows how to use dictionaries to construct paradoxes involving a change in profile). Although this conclusion significantly extends the assertions found in the literature, it has the weakness that it does not identify whether a binary susceptible procedure can avoid this negative conclusion by using a $\mathbf{W}^n \in a^n$. To resolve this questions, we need to use the properties of the Borda Dictionary.

Corollary 1.1. *For $n \geqq 3$, if a binary susceptible social choice method based on the $\mathbf{B}^n$ rankings admits the abstention paradox, then the abstention property holds for all choices $\mathbf{W}^n$. In particular, all runoff elections have this property.*

Again, this corollary underscores a basic consequence of Theorem 1. If a "negative" election phenomena occurs with the BC, then it must occur with all positional voting methods. Thus, to understand a social welfare or social choice procedure based on positional voting methods, the analysis must start with the Borda Dictionary to determine what can and cannot occur.

Outline of proof for Corollary 1.1. (The proof is in Sect. 4). To outline the ideas, I show that for $n=3$, the BC runoff election has the abstention property. The basic ideas extend to all $n \geqq 3$. By using results in Sect. 3, it follows that $(c_1 < c_2, c_3 > c_1, c_2 > c_3, c_1 > c_3 = c_2) \in D(\mathbf{B}^3)$. This means there is a profile $\mathbf{p}$ that yields this word, and, of course, any fixed scalar multiple of the number of voters with the same ranking yields the same outcome. As shown in Theorem 6 of [10], this profile can be selected so that at least one of these voters has the ranking $c_2 > c_3 > c_1$. Add another voter with the identical ranking. Choose the scalar multiple determining the replicated number of voters to be large enough so that the outcomes of the pairwise elections are not affected by whether or not these two voters vote. On one hand, if these two voters abstain, then the ranking of the swing set, the set of all candidates, is $c_1 > c_3 > c_2$, and c_3 wins the runoff between c_1 and c_3. On the other hand, should these two voters vote, the outcome is $c_1 > c_2 > c_3$, and the undesired c_1 wins the runoff. This completes the proof. The general proof, in Sect. 4 is similar. It just involves using related words from $D(\mathbf{B}^n)$ where, by changing from one to the other, the final outcome changes.

Notice that this proof involves outcomes based on changes in profiles. It is interesting to note that the approach used in this proof extends to many other settings involving small changes in profiles. For instance, to see how similar ideas can be used to analyze the possibility of manipulating the outcome of a positional voting election, see Saari [13]. As another example, the same ideas can be used to prove that a binary susceptible choice procedure need not be monotonic if it is based on the rankings of positional voting procedures. Thus, for example, we recapture and extend the known result that the Hare method need not be monotonic. (For a definition, discussion and references, see Nurmi [5]).

Corollary 1.2. *A necessary condition for a binary susceptible choice method to be monotonic is if it is monotonic when it is based on the* BC *rankings.*

2. How much of an improvement, and what representation?

Based on the above results, the BC appears to be superior to any other positional voting method – at least with respect to the kinds of issues addressed by Theorem 1. (See [10] for a more detailed discussion). A shortcoming of Theorem 1 is that it does not indicate whether the BC provides only a marginal, or a significant improvement over other positional voting systems. After all, if the difference between $D(\mathbf{W}^n)$ and $D(\mathbf{B}^n)$ is only a small number of words, then the advantage achieved by using Borda's method may not be of importance. This would imply that the BC merely avoids a small number of paradoxes.

The BC offers a significant improvement; I show in this section that the Borda Dictionary has far fewer words than any other dictionary. In fact, for five candidates the general situation is that

$$10^{14}\,|D(\mathbf{B}^5)| \ll |D(\mathbf{W}^5)| \ , \tag{2.1}$$

and for six candidates

$$10^{54}\,|D(\mathbf{B}^6)| \ll |D(\mathbf{W}^6)| \ . \tag{2.2}$$

For instance, suppose for six candidates that each of the $2^6-7=57$ possible subsets of two or more candidates are plurality ranked. It follows from Eq. (2.2) that, on the average, *each* Borda word must be replaced with at least 10^{54} different words to complete the plurality election dictionary. To appreciate the magnitude of this number, recall that the projected supercomputers will perform about 10^{12} operations per second, and that there have been about 10^{20} seconds of time since the "Big Bang". Thus, if such a supercomputer started at the Big Bang to list the words that replace just *one* Borda word from $D(\mathbf{B}^6)$, then, at the very best, the computer would only be $1/10^{22}$ through.

The large multiples in these inequalities not only underscore the point that the BC avoids a shockingly large number of paradoxes[5], but also why it is impossible to list the entries in a dictionary. To circumvent the listing problem, I develop a geometric approach to characterize the words in a dictionary, where I concentrate on the Borda Dictionary. This vector space representation also characterizes all possible tallies for ballots, so it can be used to describe all cardinal relationships relating the election tallies among different subsets of candidates. Consequently, this approach allows one to address issues such as: what profiles define certain specified words in $D(\mathbf{B}^n)$, what percentage of the BC points cast does a candidate need to acquire in order to avoid (or to achieve) certain BC properties, etc., etc. This is illustrated, in part, in Sect. 3.

This geometric representation for election tallies, partially developed in [6, 7, 10], is critical for what follows. To simplify the exposition, first consider the set $S_{2^n-(n+1)}=\{c_1,\ldots,c_n\}$. In the n-dimensional Euclidean space, E^n, identify the k^{th} component, x_k, with the k^{th} candidate, c_k, in the following manner. For $x=(x_1,\ldots,x_n)$, let larger values of x_k denote a "stronger" preference for c_k. With this identification, the hyperplane $x_k=x_j$ divides E^n into three regions; the two

[5] With the large cardinalities of $D(\mathbf{W}^n)$ one might wonder whether most of the words in a dictionary require profiles with more voters than admitted by the population of the world. However, notice that from 30 voters and six candidates $(6!)^{30}=5.24\times10^{85}$ profiles can be constructed; a number that significantly exceeds the cardinality of any $D(\mathbf{W}^6)$. Thus reasonable numbers of voters probably suffice.

half spaces are identified with the strict ordinal rankings (e.g., $x \in E^n$ satisfying $x_k > x_j$ corresponds to the ordinal ranking $c_k > c_j$), and the hyperplane is identified with indifference between the two candidates. By allowing the choices for k and j to vary over all pairs of indices, the resulting $n(n-1)/2$ hyperplanes divide E^n into cones that represent all possible ordinal rankings of the n candidates. Call each of these regions a *ranking region.* In this way, each ranking of the n candidates corresponds to a unique ranking region. For instance, $\{x \in E^5 \mid x_2 > x_4 = x_5 > x_3 > x_1\}$ is identified with the ranking $c_2 > c_4 = c_5 > c_3 > c_1$. The line passing through the origin of E^n and $(1, 1, \ldots, 1)$ represents the ranking of complete indifference among the candidates; this line is the intersection of the $n(n-1)/2$ "indifference" hyperplanes.

For what follows, let A represent the ranking $c_1 > c_2 > \ldots > c_n$. If $\mathbf{W}_{2^n-(n+1)}$ is the voting vector for $S_{2^n-(n+1)}$, then $\mathbf{W}_{2^n-(n+1)}$ is in the closure of the ranking region identified with A. (If two or more of the components, w_j, agree, then $\mathbf{W}_{2^n-(n+1)}$ is on the boundary of the ranking region; otherwise it is in the interior.) This vector serves as the tally of a ballot with the ranking A. When used as a tally, denote it as $\mathbf{W}_{A,2^n-(n+1)}$. Any other strict ranking of the N alternatives is a permutation of A, where I denote the generic representation of a permutation of A as $\pi(A)$. The tally for the ranking $\pi(A)$ is the appropriate permutation of $\mathbf{W}_{2^n-(n+1)}$, denoted by $\mathbf{W}_{\pi(A),2-(n+1)}$. Again, $\mathbf{W}_{\pi(A),2^n-(n+1)}$ is in the closure of the ranking region associated with $\pi(A)$.

Example 3. For $n=3$ and the voting vector (w_1, w_2, w_3), the ranking $c_2 > c_1 > c_3$ is tallied with the vector (w_2, w_1, w_3) to represent that w_2, w_1, w_3 points are assigned, respectively, to c_1, c_2, c_3.

Let $f_{\pi(a)}$ denote the fraction of all the voters with the ranking $\pi(A)$. The tally of an election is given by

$$F(\{f_{\pi(A)}\}, \mathbf{W}_{2^n-(n+1)}) = \Sigma_{\pi(A)} f_{\pi(A)} \mathbf{W}_{\pi(A),2^n-(n+1)} \quad , \tag{2.3}$$

where the summation index varies over all $n!$ permutations of A. The election outcome is determined by the ranking region that contains this vector sum.

The non-negative variables $\{f_{\pi(A)}\}$ sum to unity. Thus the associated profile, $\mathbf{p}$, is identified with a vector in the unit simplex in the positive orthant of $E^{n!}$; i.e., $Si(n!) = \{x \in R^{n!} \mid \Sigma x_j = 1, x_j \geqq 0\}$. Consequently the point defined by Eq. (2.3) is in the convex hull of $\{\mathbf{W}_{\pi(A),2^n-(n+1)}\}_{\pi(A)}$ where, as always, $\pi(A)$ varies over all $n!$ permutations. Thus the set of all possible election outcomes is in this convex hull. (Conversely, any ranking region in this convex hull is an election outcome for some profile of voters.) In turn, this hull is in the affine plane defined by the points $\{\mathbf{W}_{\pi(A),2^n-(n+1)}\}_{\pi(A)}$ and $\tau(1, \ldots, 1)$ where $\tau = \Sigma w_j$. The analysis is considerably simplified when this plane is a linear subspace of E^b. This motivates the first of two assumptions imposed on the voting vectors. The first assumption specifies a value for "b" in the representation of Proposition 1.

Vector normalization. The sum of the components of a voting vector equals zero.

Example 4. 1. The voting vector for a plurality election is $(1, 0, \ldots, 0)$, so a normalized form is $n(1, 0, \ldots, 0) - (1, \ldots, 1) = (n-1, -1, \ldots, -1)$. If $n=3$, all election outcomes are in the convex hull determined by $(1,0,0)$, $(0,1,0)$, and $(0,0,1)$. For the normalized vector, the convex hull is defined by $(2, -1, -1)$, $(-1, 2, -1)$, and $(-1, -1, 2)$.

2. A vector normalized form for the BC vector $(n-1, n-2, \ldots, 1, 0)$ is

$$(n-1, \ldots, n+1-2i, \ldots, 1-n) \ . \tag{2.4}$$

If $n=2$, this vector is $(1, -1)$; if $n=3$, it is $(2, 0, -2)$; and if $n=4$, it is $(3, 1, -1, -3)$.

The vector normalization assumption forces each of the normalized vectors, $\mathbf{W}_{\pi(A), 2^n-(n+1)}$, to be orthogonal to $(1, \ldots, 1)$, so the vote tally, given by Eq. (2.3), also must be orthogonal to $(1, \ldots, 1)$. Let E^{n*} be the linear subspace of E^n defined by the normal vector $(1, \ldots, 1)$. Because this is the subspace E^{n*} of the vote tally vectors, it is the space of interest. For instance, if $n=3$, then E^{3*} is the two dimensional space $x+y+z=0$.

Now consider all $2^n-(n+1)$ subsets of candidates. Corresponding to the set S_j is the division of an Euclidean space of dimension $|S_j|$ into ranking regions. Denote the coordinate functions of this space by $x_{k,j}$ where the first subscript identifies the candidate, c_k, while the second identifies the subset S_j. For example, for $S_j=\{c_1, c_4, c_7\}$ the coordinates of the corresponding E^3 are $(x_{1,j}, x_{4,j}, x_{7,j})$, and the ranking regions are in the two dimensional subspace, E^{3*}, defined by $x_{1,j}+x_{4,j}+x_{7,j}=0$, with $(1, 1, 1)$ as a normal vector.

Let Ω^n be the cartesian product of the $2^n-(n+1)$ linear subspaces E^{k*}. A ranking region in Ω^n is given by the product of ranking regions of the component spaces. For instance, Ω^3 is a five dimensional space where the ranking region $(x_{1,1}=x_{2,1}, x_{2,2}>x_{3,2}, x_{3,3}>x_{1,3}, x_{2,4}>x_{3,4}>x_{1,4})$ corresponds to the ranking $(c_1=c_2, c_2>c_3, c_3>c_1, c_2>c_3>c_1)$. It is important to note that there is a one to one correspondence between the ranking regions of Ω^n and the entries of U^n.

Using the obvious restriction, ranking A defines a ranking for each subset of candidates. If $\mathbf{W}^n$ is a system voting vector, then $\mathbf{W}^n$ is in the closure of the ranking region of Ω^n where each ranking is determined by A. Thus, this system voting vector represents how a voter with ranking A has his ballot tallied over each of the $2^n-(n+1)$ subsets of candidates. When treated as a tally, denote this vector as $\mathbf{W}^n_A$. Any other ranking of the candidates is a permutation of $A, \pi(A)$, so the tally of the ballot for each subset of candidates is given by the appropriate permutation of the vector components of $\mathbf{W}^n$. This permutation, denoted by $\mathbf{W}^n_{\pi(A)}$, is in the closure of the $\pi(A)$ ranking region of Ω^n. For a profile $\{f_{\pi(A)}\}$, the simultaneous voting tally for all subsets of candidates is

$$F(\{f_{\pi(A)}\}, \mathbf{W}^n)=\Sigma_{\pi(A)} f_{\pi(A)} \mathbf{W}^n_{\pi(A)} \ . \tag{2.5}$$

This summation has the same interpretation as Eq. (2.3); it defines a point in the convex hull of $\{\mathbf{W}^n_{\pi(A)}\}_{\pi(A)}$. This point is in one and only one ranking region of Ω^n; the ranking associated with this region is the word in $D(\mathbf{W}^n)$ defined by this profile.

Example 5. Suppose the elections over the sets $\{c_1, c_3, c_4\}$ and $\{c_1, c_2, c_3, c_4)$ are tallied with the voting vectors $(2, 0, -2; 3, 1, -1, -3)$; i.e., both are BC elections. Consider the profile where five people have the ranking $c_1>c_2>c_3>c_4$, three have $c_4>c_1>c_2>c_3$, and two have $c_1>c_4>c_3>c_2$. The tally is $(1/2)(2, 0, -2;\ 3, 1, -1, -3)+(3/10)\ (0, -2, 2;\ 1, -1, -3, 3)+(1/5)\ (0, -2, 2;\ 3, -3, -1, 1)=(1, -1, 0; 2.4, -.04, -1.6, -0.4)$, so it corresponds to the word $(c_1>c_4>c_3; c_1>c_2=c_4>c_3)$.

The key observation used to characterize the dictionaries is that *a word,* $\mathbf{w}$, *is in* $D(\mathbf{W}^n)$ *iff the product regions associated with* $\mathbf{w}$ *intersects the convex hull of* *the vectors* $\{\mathbf{W}^n_{\pi(A)}\}_{\pi(A)}$. This convex hull is in the linear space, $V(\mathbf{W}^n)$, spanned by $\{\mathbf{W}^n_{\pi(A)}\}_{\pi(A)}$. The importance of this observation is based on the fact that it is much easier to analyze a linear space rather than a convex hull.

Proposition 2. *For* $n \geqq 2$, *let* $\mathbf{W}^n$ *be given. A word is in* $D(\mathbf{W}^n)$ *iff the product ranking regions in* Ω^n *associated with this word has a non-empty intersection with* $V(\mathbf{W}^n)$.

This proposition transfers the emphasis of characterizing $D(\mathbf{W}^n)$ to the simpler task of characterizing $V(\mathbf{W}^n)$; e.g., $V(\mathbf{W}^n) = \Omega^n$ iff $D(\mathbf{W}^n) = U^n$.[6] For instance, critical parts of Theorem 1, restated in this framework, assert for $n \geqq 3$ candidate that with the exception of an algebraic subset, a^n, of possible choices for system voting vectors,

$$V(\mathbf{W}^n) = \Omega^n \ . \tag{2.6}$$

Moreover, for $n = 3$, if $\mathbf{W}^3 \neq \mathbf{B}^3$, then $V(\mathbf{W}^3) = \Omega^3$. However, $V(\mathbf{B}^3)$ is a proper linear subspace of Ω^3. Indeed, for all $n \geqq 3$, $V(\mathbf{B}^n)$ is a proper linear subspace of Ω^n.

The added structure obtained by emphasizing the geometry of the vector space $V(\mathbf{W}^n)$, rather than the dictionary $D(\mathbf{W}^n)$, can be exploited in many ways. The first involves using the dimension of $V(\mathbf{W}^n)$. To develop insight, consider the simplified setting of two sets of candidates (c_1, c_2) and (c_1, c_3). The coordinate representation for each pair is in E^{2*}. If $\mathbf{x} = (x_1, x_2)$ and $\mathbf{y} = (y_1, y_3)$ are the coordinates, then Ω is $\{(\mathbf{x}, \mathbf{y}) \mid x_1 = -x_2, y_1 = -y_3\}$. Thus, Ω is a two dimensional vector space that contains 3^2 ranking regions. A one dimensional linear subspace of Ω meets only three regions. So if V_1 is a proper linear subspace of V_2 in Ω, then V_2 meets at least three times as many ranking regions as V_1. This argument generalizes to prove the next statement.

Proposition 3. *Let* $\mathbf{W}_{j^n}$, $j = 1, 2$, *be two system voting vectors where* $V(\mathbf{W}_{1^n})$ *is a proper linear subspace of* $V(\mathbf{W}_{2^n})$ *and the difference in the dimension of the spaces is d. Then*

$$3^d |D(\mathbf{W}_{1^n})| < |D(\mathbf{W}_{2^n}| \ . \tag{2.7}$$

The multiple 3^d is overly conservative because it is based on the assumption that the vector spaces differ in dimension only in those component spaces corresponding to pairs of candidates. This never happens. A more appropriate multiple is obtained by analyzing the geometry of the ranking regions. In particular, this multiple always is larger than 4^d. The estimates in the introductory comments of this section are based on the conservative Eq. (2.7), so the BC actually provides far stronger relief from paradoxes.

It follows from the proposition that $\dim(V(\mathbf{W}^n))$ serves as a crude measure of $|D(\mathbf{W}^n)|$. The next statement imposes a lower bound on the dimension of any voting vector space.

Proposition 4. *For any* $n \geqq 2$

$$\dim(V(\mathbf{W}^n)) \geqq n(n-1)/2 \ . \tag{2.8}$$

[6] An entry in $V(\mathbf{W}^n)$ is a vector whose coefficients are uniquely determined by the profile $\{\mathbf{f}_{\pi(A)}\}$; thus information about which profiles cause what kind of behavior are contained the $V(\mathbf{W}^n)$ representation.

An immediate corollary of Theorem 1 is that all possible rankings can occur with the majority vote rankings of the $n(n-1)/2$ pairs of candidates. The vector space representation for a pair of candidates is one dimensional, so the subspace corresponding to the rankings of all pairs of candidates has dimension $n(n-1)/2$. As this is a subspace of $V(\mathbf{W}^n)$, Inequality (2.8) follows immediately. In other words, it is the dimension of the subspace of the rankings of the pairs of candidates that forces this lower bound on $\dim(V(\mathbf{W}^n))$.

It is clear that if $V(\mathbf{W}_1^n)$ is a proper subspace of $V(\mathbf{W}_2^n)$, then $D(\mathbf{W}_1^n)$ must be a proper subset of $D(\mathbf{W}_2^n)$. However, without imposing additional assumptions on the voting vectors, the converse is false. In fact, even if two system voting vectors are equivalent (so their dictionaries are identical), the vectort spaces need not agree. This can be seen with the two Borda vectors $\mathbf{B}_1^3=(1,0;1,0;1,0;2,0,-2)$ and $\mathbf{B}_2^3=(1,0;1,0;1,0;6,0,-6)$ where the scaling difference in the voting vectors force $V(\mathbf{B}_1^3)\neq V(\mathbf{B}_2^3)$. This example isolates the difficulty. Namely, if a dictionary of one system voting vector is properly contained in the dictionary of another, then the voting vector subspaces can have this same relationship only with an appropriate scaling of the voting vectors. This second normalization specifies the value of "a" from Proposition 1.

Definition. a. Let the system voting vector $\mathbf{W}^n=(\mathbf{W}_1,\ldots,\mathbf{W}_{2^n-(n+1)})$ be given. A *scalar normalization* of $\mathbf{W}^n$ is a choice of $2^n-(n+1)$ positive scalars $\{s_j\}$ used to define the equivalent system voting vector $(s_1\mathbf{W}_1, s_2\mathbf{W}_2,\ldots)$.
b. The *standard scalar normalization* for the Borda system voting vector is where the Borda vectors are given by Eq. (2.4) and the voting vectors for sets of two candidates is $(1,-1)$.
c. Let the system voting vector $\mathbf{W}^n=(\mathbf{W}_1,\ldots,\mathbf{W}_{2^n-(n+1)})$ be given. A *scalar normalization* of $\mathbf{W}^n$ is a choice of $2^n-(n+1)$ non-zero scalars $\{s_j\}$ used to define the equivalent system voting vector $(s_1\mathbf{W}_1, s_2\mathbf{W}_2,\ldots)$.

The next theorem is a much stronger version of Theorem 1. To appreciate the dimension assertions, note that $\dim(\Omega^n)=K(n)=\Sigma_{k=2^n}(k-1)\,n!/(n-k)!k!$. Thus $K(3)=5, K(4)=17, K(5)=49$, and $K(6)=129$.

Theorem 2. a. *Fo* $n\geq 3$, $\dim(V(\mathbf{B}^n))=n(n-1)/2$.
b. *For* $n\geq 3$ *and* $\mathbf{W}^n\neq\mathbf{B}^n$, *there is a scalar normalization of* $\mathbf{W}^n$ *so that* $V(\mathbf{B}^n)$ *is a proper linear subspace of* $V(\mathbf{W}^n)$.
c. *With the exception of a lower dimensional algebraic subset* a^n *of system voting vectors,* $V(\mathbf{W}^n)=\Omega^n$.
d. *If* $n=3$, *and if* $\mathbf{W}^3\neq\mathbf{B}^3$, *then* $V(\mathbf{W}^3)=\Omega^3$.
e. *The assertions of parts a, b, c, d hold for all choices of system scoring vectors* $\mathbf{W}^n$. *(The scalar multiples of some of the scoring vectors may need to be negative).*

An amazing assertion of this theorem is that $\dim(V(\mathbf{B}^n))=n(n-1)/2$; this dimension agrees with the lower bound given in Proposition 4. By use of Proposition 3, this means that $D(\mathbf{B}^n)$ is a very small subset of U^n relative to the size of the dictionaries for most other methods. For instance, if $\mathbf{W}^n$ corresponds to where each subset of candidates is plurality ranked, then the difference in dimension between $V(\mathbf{B}^n)$ and $V(\mathbf{W}^n)$ is $K(n)-n(n-1)/2$. Thus $\dim(V(\mathbf{B}^5))=10$ while $\dim(V(\mathbf{W}^5))=49$ and the dimensional difference is 39; $\dim(V(\mathbf{B}^6))=15$ while $\dim(V(\mathbf{W}^6))=129$ and the dimensional difference is 114. With Proposi-3, one can appreciate the effect the dimensional differences make in the comparative sizes of the dictionaries.

Another surprising assertion is part e. For instance, it follows from part e that the favorable BC properties do not depend on the monotonicity associated with the voting vectors (e.g., $w_i \geqq w_{i+1}$), so, they must be a consequence of a deeper fundamental property of the BC.[7] For instance, it follows from part e that $V(\mathbf{W}^3) = \Omega^3$ for $\mathbf{W}^3 = (1, -; 1, -1; 1, -1; 5, -4, -1)$. $\mathbf{W}^3$ is not a system voting vector because $(5, -4, -1)$ is equivalent to $(10, 1, 4)$ which requires giving 10 point to a top ranked candidate, 4 points to a bottom ranked candidate, but only 1 point to a second ranked candidate. A key point is part e is that the scales may be negative; this eliminates the need to consider the class of vectors $\mathbf{BC}^n$ as needed in Theorem 1.

Part b geometrically extends the assertion that if a word is in $D(\mathbf{B}^n)$, then it is in $D(\mathbf{W}^n)$. Moreover, as it will become clear with the techniques developed in Sect. 3, it also means that if $\mathbf{W}^n \in a^n$, then it must be constructed in terms of the properties of the BC.

From Theorem 2 it follows that the BC is, in certain important ways, a significant improvement over any other choice of a positional voting method. But, what about other kinds of social welfare procedures? For instance, positional voting methods can be viewed as determining certain weighted means over a profile. One could design other voting procedures based on the nonlinear methods commonly used in statistics. In such a manner, or with the use of other techniques, wouldn't it be possible to find a method much better than the BC? How does the BC fare within this larger class of voting procedures?[8]

Definition. A *smooth, majority preserving social welfare mechanism* is a mapping $G^n: Si(n!) \rightarrow \Omega^n$ where the rankings of the pairs of candidates is determined by majority vote. Let $I(G^n)$ be the image set of G^n.

The intersection of $I(G^n)$ with the ranking regions of Ω^n determines what rankings and what paradoxes this social welfare mechanisms admits. With the above argument, it follows that a crude measure is $\dim(I(G^n))$, where smaller values indicate fewer paradoxes. Here, if $I(G^n)$ is a smooth manifold, then $\dim(I(G^n))$ is the dimension of the manifold. If it is not a smooth manifold, then let $\dim(I(G^n)$ be the minimum dimension of all smooth manifolds that contain $I(G^n)$.

Corollary 2.1. *If G^n is a smooth, majority preserving social welfare mechanism, then* $\dim(I(G^n)) \geqq \dim(V(\mathbf{B}^n)$.

In other words, *for the class of smooth, majority preserving, social welfare mechanisms, a class which includes all positional voting procedures and many others, one cannot do better than Borda's method with respect to the dimensional measure.* For related comments concerning social choice mechanisms, see Corollary 3.2. By using the ideas and techniques developed in [11], the smoothness requirement can be dropped.

---

[7] This property is due to the fact the BC creates a singularity in the orbit of the wreath product of permutation groups. This singularity is based on the symmetry created by the requirement that $w_i - w_{i+1}$ is the same constant for all choices of i.

[8] As one might suspect, the ideas developed in [10] and here can be used to analyze certain statistical procedures. Results of this kind are in Haunsperger [4] where she completely analyzes the Kruskal-Wallis Dictionary, etc. It is related to the Borda Dictionary.

3. The Borda dictionary and vector space

To characterize $D(\mathbf{B}^n)$ it suffices to characterize $V(\mathbf{B}^n)$. But $V(\mathbf{B}^n)$ is a linear subspace of Ω^n, so it is uniquely determined by its normal bundle. (This is the set of all vectors that are orthogonal to $V(\mathbf{B}^n)$.) Thus $V(\mathbf{B}^n)$ is characterized once a basis for its normal bundle in Ω^n is determined. To do this, the sets $\{S_1, S_2, \ldots, S_{2^n-(n+1)}\}$ are used to identify the component spaces of Ω^n. Recall that the we are using a lexicographic ordering where the first $n(n-1)/2$ sets S_j are the pairs of alternatives and where the candidates in S_j are listed according to the subscripts. Also, recall that Ω^n is a cartesian product of E^{k*} spaces. Let $\text{Ind}(S_j)$ be the set of the indices (the subscripts) of the candidates in S_j.

Theorem 3, given below, asserts that a basis for the normal bundle of $V(\mathbf{B}^n)$ is given by the vectors $\{\mathbf{Z}_{k,j}\}$, $j=1+n(n-1)/2,\ 2+n(n-1)/2, \ldots, 2^n-(n+1)$, $k \in \text{Ind}(S_J)$, where the first subscript identifies the candidate and the second identifies the subset S_j. Because of the ordering, the second subscript corresponds only to subsets S_j with three or more candidates. The vector $\mathbf{Z}_{k,j}$ is defined in the following manner.

1. In the component subspaces of Ω^n corresponding to $S_i, i > n(n-1)/2$, $\mathbf{Z}_{k,j}$ has only one non-zero vector component, $\mathbf{Y}_{k,j}$. Vector $\mathbf{Y}_{k,j}$, in the component space of Ω^n corresponding to S_j, has $-(|S_j|-1)/|S_j|$ in the $x_{k,j}$ coordinate, and $1/|S_j|$ in all others.

2. For each pair of candidates, S_i, $i \leq n(n-1)/2$, the S_i vector component of $\mathbf{Z}_{k,j}$ is

i. $\mathbf{0}=(0,0)$ iff either $c_k \notin S_i$, or S_i is not a subset of S_j.

ii. $(1/2, -1/2)$ if c_k is the first listed candidate in S_i, $(-1/2, 1/2)$ if c_k is the second listed candidate in S_i.

Example 6. 1. For $n=3$ and $\{\{c_1, c_2\}, \{c_2, c_3\}, \{c_1, c_3\}, \{c_1, c_2, c_3\}\}$, the vector $\mathbf{Z}_{2,4}=(-1/2, 1/2; 1/2, -1/2; 0, 0; 1/3, -2/3, 1/3)$ is identified with c_2 in S_4.This vector corresponds to the S_4 ranking where c_2 is bottom ranked and c_1, c_3, c_4 are tied for first, but where c_2 is top ranked for each of the pairs, S_1, S_2. (The S_3 component is $\mathbf{0}$ because $c_2 \notin S_3$). Because $\mathbf{Z}_{2,4}$ is a normal vector for $V(\mathbf{B}^3)$, this is an impossible outcome for a Borda election.

2. For $n=4$, let $S_1=\{c_1, c_3\}$, $S_2=\{c_2, c_3\}$, $S_3=\{c_3, c_4\}$, $S_7=\{c_2, c_3, c_4\}$, and $S_{11}=\{c_1, c_2, c_3, c_4\}$. The vector $\mathbf{Z}_{3,7}$, which is identified with c_3 in S_7, has $(-1/2, 1/2)$ as the S_2 component, $(1/2, -1/2)$ as the S_3 component, $(1/3, -2/3, 1/3)$ as the S_7 component, and $\mathbf{0}$ for all other components. By being identified with c_3 in S_7, this vector corresponds to the S_7 ranking where c_3 is bottom ranked and the other S_7 candidates are tied for first place, but c_3 is top ranked in the pairwise comparisons with either of these candidates from S_7 (i.e., in the S_2 and S_3 rankings). Thus, such an sequence of election rankings is not a word in $D(\mathbf{B}^4)$; it corresponds to an impossible outcome for a Borda election. The vector $\mathbf{Z}_{3,11}$ also is identified with c_3, but now it relates how c_3 fares among the candidates in S_{11} rather than in S_7. This vector has $(-1/2, 1/2)$ for the S_1 and S_2 components, $(1/2, -1/2)$ for the S_3 component, $(1/4, 1/4, -3/4, 1/4)$ for the S_{11} component, and $\mathbf{0}$ for all other components. Again, this normal vector is in a ranking region that can never be attained by a Borda election. This ranking region has a similar interpretation where c_3 is S_{11} bottom ranked and top ranked in a pairwise comparisons with any candidate from S_{11}.

3.1. The basic result

Theorem 3. a. *For $n \geqq 3$, $V(\mathbf{B}^n)$ is the $n(n-1)/2$ dimensional linear subspace of Ω^n defined by the normal vectors $\{\mathbf{Z}_{k,j}\}$, $j = 1 + \{n(n-1)/2\}, \ldots, 2^n - (n+1)$, $k \in S_j$.*
b. *Suppose $V(\mathbf{W}^n)$ admits a normal vector where the only non-zero components are in those component spaces corresponding to pairs of alternatives and to S_j where $|S_j| \geqq 3$. The voting component of $\mathbf{W}^n$ used to tally the ballots for S_j is a Borda vector.*
c. *Consider a vector space $V(\mathbf{W}^n)$ defined by $F(\mathbf{p}; \mathbf{W}^n)$ where $\mathbf{W}^n$ is a system scoring vector. The assertion in part b holds true for $V(\mathbf{W}^n)$, where the S_j component is either the vector component from $\mathbf{Z}_{k,j}$ or the negative of it.*

This theorem can be used to completely specify the properties of the Borda Count. As I stated, my emphasis is to show how to use this theorem, rather than to provide an extensive listing of new results. To start, consider $n = 3$ and the (normalized) election tally $\mathbf{d} = (10/15, 5/15; 11/15, 4/15; 9/15, 6/15; 6/15, 5/15, 4/15)$ corresponding to the word $(c_1 > c_2, c_2 > c_3, c_1 > c_3, c_1 > c_2 > c_3)$. One might wonder whether $\mathbf{d}$ is a normalized BC election tally. The proof that it is not is simple; if it were based on the BC, then the scalar product of $\mathbf{d}$ with all choice of $\mathbf{Z}_{k,j}$ would be zero. However, the scalar product of $\mathbf{d}$ with $\mathbf{Z}_{2,4}$ (see Example 6) is $2/32 > 0$.

To continue to show how to use Theorem 3, I use it next to recover and extend, in a simple, elementary fashion, several well known conclusions about the BC. To state the first result, recall that a *Condorcet* or *majority* winner is a candidate that wins all pairwise comparisons (by a majority vote), while the *anti-majority* candidate is a candidate that loses all of the pairwise comparisons. (In Example 1, b, c_1 is the Condorcet winner, while c_3 is the anti-majority candidate.) Smith [15] discovered the important relationship, described in Corollary 3.1, a, that a Condorcet winner cannot be BC ranked last. Smith also showed that this particular property is satisfied by no other positional voting method. Combining Smith's assertion with some of Young's results [16, 17], Fishburn and Gehrlein [3] strengthened Smith's statement to obtasin the assertion that the BC is the only positional voting method whereby the winner can be determined "...solely on the basis of the outcome of pairwise votes between candidates."

These two statements are special cases of a more encompassing issue; namely, *for what choices of positional voting mehtods do there exit any relationships, whatsoever, among the rankings of the positional voting method and those of the pairwise comparisons with a majority vote?*[9] For instance, for a positional voting method other than the BC are there pairwise rankings (such as when each pairwise elections ends in a tie, or when both a Condorcet winner and an anti-majority candidate exist, or when the pairwise election rankings define a transitive, binary relationship) that preclude the possibility of at least *one* positional election ranking from occurring? More generally, as scoring methods include positional voting methods, one might wonder from what choices of scoring methods are there relationships, whatsoever, among the rankings of the n candidates and the pairwise majority vote rankings? *This more general issue is answered in Corollary 3.1 b, c.*

[9] There always are some relationships among the *tallies*; e.g., a trivial one is if a particular candidate wins all pairwise elections by unanimous votes, then she is top ranked with any positional method.

Corollary 3.1 a. *Consider $n \geq 3$ candidates. A Condorcet winner can never be* BC *bottom ranked, and an anti-majority candidate can never be* BC *top-ranked.*[10]
b. *The* BC *is the only positional voting method that admits any relationship among the rankings of the positional voting procedure and those of the pairwise votes.*
c. *Consider the class of scoring methods. There is a relationship among the rankings of this method and those of the pairwise votes only if the scoring method is a scalar multiple of the* BC.

With Theorem 3, the proof of Corollary 3.1 reduces to a simple computation. This computation is carried out in full detail to demonstrate the ideas.

Proof. Suppose a profile $\mathbf{p} = \{f_{\pi(A)}\}$ admits c_1 as the Condorcet winner, but c_1 is BC bottom ranked. Let $F(\mathbf{p}, \mathbf{B}^n) = (\mathbf{X}_1, \ldots, \mathbf{X}_{2^n-(n+1)})$, so $\mathbf{X}_k$ is the normalized BC vote tally for S_k. The assumption that c_1 is BC bottom ranked requires the corresponding component of $\mathbf{X}_{2^n-(n+1)}$ to be algebraically the smallest; namely, $x_{1,k} < x_{j,k}$, $j = 2, \ldots, n$, $k = 2^n - (n+1)$. Because $\Sigma_j x_{j,2^n-(n+1)} = 0$ (by the vector normalization assumption), it follows that $x_{1,2^n-(n+1)} < 0$. It follows from a direct computation, using the vector normalization, that

$$(\mathbf{Y}_{1,2^n-(n+1)}, \mathbf{X}_{2^n-(n-1)}) = [-(n-1)x_{1,2^n-(n+1)} + \{\Sigma_{j>1} x_{j,2^n-(n+1)}\}]/n$$
$$= -x_{1,2^n-(n+1)} > 0 \ . \tag{3.1}$$

The assumption that c_1 is a Condorcet winner means that, for each k where $|S_k| = 2$ and $c_1 \in S_k$, the $x_{1,k}$ component of $\mathbf{X}_k$ is positive. For each such choice of k, it follows that

$$((1/2, -1/2), \mathbf{X}_k) = x_{1,k} > 0 \ . \tag{3.2}$$

With Eqs. (3.1), (3.2), the computation of the scalar product is

$$(F(\mathbf{p}, \mathbf{B}^n), \mathbf{Z}_{1,2^n-(n+1)}) = -x_{1,2^n-(n+1)} + \Sigma_k x_{1,k} > 0 \ , \tag{3.3}$$

where the summation is over the values of k selected for Eq. (3.2). But, Theorem 3 requires this scalar product to be zero. This contradiction proves that such a ranking is not a Borda outcome.

The proof that an anti-majority candidate cannot be BC top ranked is essentially the same.

Proof of Part b, c. F is a linear mapping, so if a system voting vector $\mathbf{W}^n$ imposes any relationship, whatsoever, among the outcomes of the pairs and the ranking of a set of candidates, then the image of F lies in a lower dimensional linear subspace of the specified coordinate subspace of Ω^n. This forces $V(\mathbf{W}^n)$ to have a normal vector with its only non-zero vector components in these coordinates subspaces of Ω^n. According to Theorem 3, the voting vector is a Borda Vector. The proof of part c is similar.

[10] A nice, related statement about the BC, by Fishburn and Gehrlein [3], asserts for $n = 3$ candidates and for a specified probabilistic measure over the profiles that, for large numbers of voters, the BC is the unique positional voting method to maximize the *likelihood* that a Condorcet winner is top ranked. Using the mathematical structures introduced in Sect. 2 and in [10], Jill van Newenhizen [18] significantly extends this statement in many different ways. In part, she shows that the same conclusion holds for all $n \geq 3$ and for a much larger class of probability distributions.

3.2. Some axiomatic characterizations of the BC

To illustrate Corollary 3.1b, c, I use it to extend Young's insightful axiomatic characterization of the BC that is based on properties of social choice functions [16]. Say that *two social choice functions, f, g, are equivalent if* $f(\mathbf{p})=g(\mathbf{p})$ *for all* $\mathbf{p}$. To state Young's result, recall that two standard assumptions on social choice functions are that f is *anonymous* if its outcome depends only on the numbers of voters with each preference, and that f is *neutral* if when σ is a permutation of the indices of the candidates, then $f(\sigma(\mathbf{p}))=\sigma f(\mathbf{p})$. (That is, both the different candidates and the different voters are treated equally; the outcome does not depend on their names.) The next assumption is that f is *consistent;* namely, if $\mathbf{p}$ and $\mathbf{p}'$ are profiles for distinct voter sets, then $f(\mathbf{p})\cap f(\mathbf{p}')\neq\phi$ implies that $f(\mathbf{p})\cap f(\mathbf{p}')=f(\mathbf{p}+\mathbf{p}')$. (Suppose a group is subdivided into two subcommittees represented by $\mathbf{p}$ and $\mathbf{p}'$. If the subcommittees agree in that $f(\mathbf{p})\cap f(\mathbf{p}')\neq\phi$, then consistency requires the common set $f(\mathbf{p})\cap f(\mathbf{p}')$ is the choice of the full group $\mathbf{p}+\mathbf{p}'$.) The function f is *faithful* if for a profile of a single voter with c_j as his top ranked candidate, $f(\mathbf{p})=c_j$. Finally, Young states that f has the *cancellation property* if when a profile $\mathbf{p}$ causes all $n(n-1)/2$ pairwise comparisons to result in a tie vote, then $f(\mathbf{p})=C^n$. Young proved *for* $n\geqq 3$ *that if a social choice function is anonymous, neutral, consistent, faithful, and has the cancellation property, then it is equivalent to choosing the top ranked candidates from the* BC *ranking of* C^n.

Definition. A social choice procedure for C^n is a *general scoring method* based on the scoring vector $\mathbf{W}^1=(w_1^1,\ldots,w_n^1)$ with tie breaker methods $\mathbf{W}^j=(w_1^j,\ldots,w_n^j)$, $j=2,\ldots,s$, if the following conditions are satisfied.
1. Not all of the w_i^j's in $\mathbf{W}^j$ have the same value.
2. A voter's k^{th} ranked candidate receives w_k^1 points.
3. The candidate, or candidates with the largest point total are selected. If a tie breaker is used to determine among several candidates with the largest point total, then, inductively, at the i^{th} stage *all* candidates are reranked with $\mathbf{W}^i$. Those candidates with the largest point total that *also* are selected at the $(i-1)^{\text{th}}$ stage are selected for the i^{th} stage.

As already emphasized, the weights for a scoring vector need not satisfy any monotonicity condition; e.g., the vector $(-1,-1,4,-3)$, where the third ranked candidate gets 4 points, the first and second get -1 points, and the last ranked candidate gets -3 points, is a scoring method. Young [17] found an interesting axiomatic characterization of general scoring methods.

Proposition 5 [17]. *For* $n\geqq 3$, *if a social choice function is anonymous, neutral consistent, and does not have a fixed value, then it is equivalent to a general scoring method.*[11]

[11] Young incorrectly asserts that the procedure is a scoring rule, not that it is only equivalent to one. To see that the condition of equivalence is needed, consider the procedure that chooses the winner of a plurality election should she receive no more than 65% of the vote. Otherwise, choose the top ranked candidate based on the voting vector (1, 1/40, 0). This procedure satisfies all of Young's conditions, it is not a scoring rule, but it is equivalent to the scoring rule based on the plurality vote. However, as stated in Proposition 5, the assertion is correct. These same comments apply to Young's characterization of the BC; his conditions define a procedure that is *equivalent* to the BC, but not necessarily the BC. While this point is not of importance for the kinds of issues raised here, it is of importance for other kinds of social choice issues. See, for instance, Saari [13].

To understand the tie breaking scheme, suppose for $n=5$ that the set is initially ranked with the BC, and that the two successive tie breakers are $(1,0,0,0,0)$ and $(1,1,1,1,0)$. So, according to this procedure, first find the BC ranking of the candidates. If there is a tie for top place, than choose the tied candidates that have the most first place votes. If there still is a tie, of the remaining candidates, choose the ones with the least number of last place votes. The dictionaries can be used to illustrate some of the properties of this procedure. For instance, there is a profile [10] so that the rankings of these three procedures are, respectively, $c_1=c_2=c_3>c_4>c_5, c_5>c_4=c_3=c_2>c_1$, and $c_1>c_5>c_4>c_3>c_2$. From, the first election, c_1, c_2, c_3 are advanced to the first tie breaker scheme. In the first tie breaker, the fact that c_5 is top-ranked plays no role in the process; the only relevant fact is that of the three remaining candidates, c_1 is bottom ranked. Thus, c_2 and c_3 are advanced to the second tie-breaker. In the second tie-breaker, it just so happens that all of the eliminated candidates are ranked above the remaining two candidates, but this ranking plays no role in the procedure. The important fact is that c_2 is ranked below c_3, so c_3 wins. Notice how sensitive this procedure is to the order in which the tie-breakers are used. For instance, if $(1,1,1,1,0)$ is used for the first tie-breaker, rather than for the second one, then c_1 wins rather than c_3.

One might wish to streamline the tie breaking procedure by using the first runoff among only $\{c_1, c_2, c_3\}$, rather than reconsidering c_4 and c_5; after all, c_4 and c_5 are eliminated from further consideration. However, *any such procedure violates consistency*. The proof of this fact again illustrates an application of the dictionaries. For example, consider the symbols $\{c_1=c_2=c_3>c_4, c_1=c_2>c_3, c_1=c_2, c_3>c_2\}$ and $\{c_2=c_3=c_4>c_1, c_2=c_4>c_3, c_2=c_4, c_3>c_2\}$. It follows from Theorem 2 that for any choice of scoring methods, satisfying part e, there are profiles $\mathbf{p}$ and $\mathbf{p}'$ defining the above two listings of election outcomes. Thus with a tie breaker for a set of three and a set of two candidates, $f(\mathbf{p})=\{c_1, c_2\}$ while $f(\mathbf{p}')=\{c_2, c_4\}$, so $f(\mathbf{p})\cap f(\mathbf{p}')=\{c_2\}$. However, $f(\mathbf{p}+\mathbf{p}')=\{c_3\}$. (This is because the linearity of summing tallies forces the top ranked candidates for $\mathbf{p}+\mathbf{p}'$ to be $\{c_2, c_3\}$.) Again, from the linearity of summation processes, c_3 is the winner of the run-off. Thus, while consistency may appear to be a natural, fairly weak requirement, in fact it imposes a very strong condition on the social choice procedure. (Weaker versions of consistency are discussed and characterized in Saari [13].)

"Faithfulness" in Young's Theorem plays two roles. The first is to impose a monotonicity on the choice of a scoring method; it requires $w_1>w_j$ for $j\geqq 2$. As such, this condition prohibits $\mathbf{W}^1$ from being $(1,1,0,0,0)$, or any other voting vector where the weights assigned to a voter's first and second ranked candidates agree. I use a weaker condition that includes "faithfulness" as a special case. The new condition admits all positional voting methods because it only requires that w_n, the weight assigned to the bottom ranked candidate, must have a smaller value than at least one other weight. The second role of "faithfulness" is to ensure that the image of f contains singletons; i.e., there are situations where only one candidate is selected. It turns out that this property is not needed for what follows. On the other hand, this property does significantly simplify the proofs of the results, so, for convenience, I include it as part of the following definition.

Definition. A social choice function, f, is *somewhat faithful* if the image of f contains a singleton and if for a profile $\mathbf{p}$ of a single voter, his bottom ranked candidate is not in the set $f(\mathbf{p})$.

To illustrate the difference between faithfulness and somewhat faithful, consider the single voter profile **p** where the voter has the ranking $c_1 > c_2 > c_3 > c_4$. If a procedure is faithful, then $f(\mathbf{p}) = \{c_1\}$. If it is somewhat faithful, then there are 2^3 ways to define $f(\mathbf{p})$; this image set can be any set of candidates that does not contain c_4. Thus, the requirement of being somewhat faithful admits far more procedures. As specific examples for $n = 4$, the positional voting methods $(1, 1, 0, 0)$ and $(1, 1, 1, 0)$ and the scoring methods $(1, 0, 1, 0)$ and $(0, 0, 1, 0)$ all define somewhat faithful procedures that are not faithful.

In light of Proposition 5, the role of the cancellation property in Young's result is to impose a pairwise ranking restriction on the choice of the quasi-positional voting method. It now follows immediately from Corollary 3.1 b, c, that the only possible choice is the BC. Indeed, in light of Corollary 3.1 b, c, it follows that *Young's result can be extended in many different directions simply by replacing the cancellation property with any other BC property.* For instance, the same conclusion holds when the cancellation property is replaced with the much weaker requirement of *non-determinancy* whereby if **p** creates a tie vote in all pairwise elections, then $f(\mathbf{p})$ is not a singleton. Another natural choice is to replace the cancellation property with the desirable requirement that if c_j is an anti-majority candidate, then $c_j \notin f(\mathbf{p})$. This substitution of the cancellation property does admit tie breakers. Either of these appealing, substitute axiomatic representations of the BC appears to be difficult to prove directly, but they are immediate consequences of Theorem 3 and Corollary 3.1. Indeed, most of the BC properties derived in this essay serve as substitute conditions for the cancellation property. This fact is emphasized in the first part of the following statement.

Corallary 3.2. a. *For $n \geqq 3$, suppose an anonymous, neutral, consistent, and somewhat faithful social choice function, f, satisfies another specified condition whereby the pairwise rankings of the candidates imposes a (non-trivial) constraint on the image of f. If this specified condition is not satisfied by the* BC, *then no such f exists. If the condition is satisfied by the* BC, *then f is equivalent to first choosing the* BC *top ranked candidates of C^n where possible ties may be broken by tie breakers.*
b. *For $n \geqq 3$, there does not exist an anonymous, neutral, consistent, somewhat faithful social choice function that always selects the Condorcet winner.*
c. *For $n \geqq 3$, if an anonymous, neutral, consistent, and somewhat faithful social choice function satisfies the non-determinacy property, then it is equivalent to the* BC *with no tie breakers.*

By a "non-trivial constraint", in part a, I mean that when the designated pairwise rankings occur for a profile **p**, then there is a subset of $P(C^n)$ that cannot be the image for $f(\mathbf{p})$. The proof of part b, which serves as an example of a condition that leads to an impossibility theorem, follows from the observation that the Condorcet winner need not be BC top ranked. Part c is a direct extension of Young's statement, but it is not obvious whether Young's techniques could be used to prove this stronger assertion. One purpose for including this result is to indicate in the proof (Sect. 4) why tie breakers are not admitted. Finally, the only role played by the "somewhat faithful" condition is to impose enough monotonicity on the procedure to outlaw the reversed Borda Count – this is equivalent to choosing the BC bottom ranked, rather than top ranked candidate. One could replace this condition with one of many different substitute monotonicity conditions.

3.2. The Borda rankings

To extend Corollary 3.1a, note that the proof just uses the fact that $F(\{f_{\pi(A)}\}, \mathbf{B}^n)$ is orthogonal to each $\mathbf{Z}_{k,j}$. In the proof I separately computed the contribution of the scalar product due to the outcomes of the pairwise elections and the contribution due to the Borda tally. This same computation generalizes Corollary 3.1a from an assertion about the ranking of a Condorcet winner to a statement about any candidate who fares well in pairwise comparisons. More precisely, *if the pairwise election outcomes for candidate c_i satisfy $\Sigma_k x_{i,k} > 0$, where the summation index k is over the pairs of candidates S_k that include c_i, then c_i cannot be* BC *bottom ranked. Equivalently, if this summation is negative, then c_i cannot be* BC *top ranked.* This condition permits candidate c_i to lose several of the pairwise majority vote comparisons as long as she wins other elections by a sufficiently large margin. To further appreciate the flexibility offered by this condition $\Sigma_k x_{i,k} > 0$, note that if $S_j = \{c_i, c_a\}$, then $(x_{i,j} + 1)/2$ is the fraction of the voters that prefer c_k to c_a. Thus, this generalization of Corollary 3.1a asserts that if a candidate receives enough votes over all of the pairwise comparisons to be, on the average, over 50% (i.e., if the sum of the fractions of votes from each of the $n-1$ pairwise elections exceeds $(n-1)/2$), then she cannot be BC ranked last. Therefore, it is possible for her to lose all but one of the pairwise election and still satisfy the inequality!

This assertion, which involves the use of all of the vectors $\mathbf{Z}_{i,2^n-(n+1)}$, holds because in light of Theorem 3 the assumption about the pairwise elections forces the scalar product $(\mathbf{Y}_{i,2^n-(n+1)}, \mathbf{X}_{2^n-(n-1)})$ to be negative. In turn, this negative value forces the angle between $\mathbf{X}_{2^n-(n-1)})$ and $\mathbf{Y}_{i,2^n-(n+1)}$ to be greater than $\pi/2$. It now follows from the position of $\mathbf{Y}_{i,2^n-(n+1)}$ and the geometry of the ranking regions that, not only is it impossible for c_i to be BC bottom ranked, but *it is impossible for her even to be* BC *tied for last.*

Example 7. a. What words are admitted in $D(\mathbf{B}^n)$? To illustrate both the geometry and how algebraic relationships are found, I'll show that $(c_1 > c_2;\ c_1 > c_3;\ c_2 > c_3;\ c_2 = c_1 > c_3) \in D(\mathbf{B}^3)$. This word corresponds to the components $(x_{1,1}, x_{1,1};\ x_{1,2}, -x_{1,2};\ x_{2,3}, -x_{2,3};\ x_{1,4}, x_{1,4}, -2x_{1,4})$ where $x_{1,j} > 0$ for $j = 1, 2, 4$, and $x_{2,3} > 0$. The orthogonal scalar product of this vector with $\mathbf{Z}_{1,4}$ yields $x_{1,1} + x_{1,2} = x_{1,4}$, while with $\mathbf{Z}_{2,4}$ it is $-x_{1,2} + x_{2,3} = x_{1,4}$. ($\mathbf{Z}_{3,4} = -(\mathbf{Z}_{1,4} + \mathbf{Z}_{2,4})$, so it provides no new information.) Thus, the word is admissible in $D(\mathbf{B}^3)$ because it is possible to satisfy the equation

$$x_{1,1} + 2x_{1,2} = x_{2,3}\ . \tag{3.4}$$

Therefore, not only is such a word in $D(\mathbf{B}^3)$, but *Eq.* (3.4) *is a necessary and sufficient condition for a Condorcet winner to be* BC *tied for first place.* A necessary and sufficient condition for c_2 to be BC top ranked is if $x_{2,3}$ has a larger value. Notice that this imposes a significant burden on her winning margin over c_3, particulrly if c_1 does not just barely win both pairwise elections. Alternatively, a sufficient condition for c_1, the Condorcet winner, to be BC top ranked is $x_{1,1} + 2x_{1,2} > 1/2\,(\geqq x_{2,3})$. As extreme examples, this happens if she just barely beats $c_2\,(x_{1,1} > 0)$ but gets at least 62.5% of the vote against $c_3\,(x_{1,2} > 1/4)$, or if she beats c_2 with a 75% vote $(x_{1,1} = 1/2, x_{1,2} > 0)$. The main purpose for these new conclusions is to indicate how similar supporting relationships for any word in $D(\mathbf{B}^n)$ and for any $n \geqq 3$ can be found.

b. If the BC is used, what words can accompany the symbols $c_1 > c_2$, $c_2 > c_3$, $c_1 > c_3$? The choice of the one remaining symbol is governed by the above equations for $x_{j,4}$. It follows immediately that this symbol can be filled with any ranking that has c_1, the Condorcet winner, ranked strictly above c_3, the Condorcet loser.
c. What words are in the dictionary for the plurality ranking, but not in $D(\mathbf{B}^n)$? If $S_1 = \{c_1, c_2\}$, $S_2 = \{c_1, c_3\}$ and $S_7 = \{c_1, c_2, c_3\}$, then, by using $\mathbf{Z}_{1,7}$, it follows that a word with the three symbols $c_1 > c_2$, $c_1 > c_3$, $c_2 > c_3 > c_1$ is not in $D(\mathbf{B}^4)$. On the other hand, all such words are in $D(\mathbf{W}^4)$ if the system vector is based on plurality votes. So, by permuting the indices, by filling in the other symbols in an arbitrary fashion, etc., this simple assertion accounts for 172, 974, 204 of the words that are in the plurality dictionary but not in $D(\mathbf{B}^4)$.

In his paper, Smith reminds the reader that Black questioned the naturalness of the Condorcet winner. For instance, the Condorcet winner, c_1, may barely win each pairwise election, while c_2 barely loses to c_1 and then wins all other pairwise elections by substantial margins. It is reasonable to feel that c_2 should be the winner. Smith notes that "it would be interesting to try to formulate this feeling as a precise property; it may be that a suitable formulation is a necessary and sufficient condition for a [positional voting] system." To continue this line of thought, note that it is possible for a Condorcet winner to achieve her status not throught excellence, but rather through mediocrity by serving as the compromise – the second choice candidate – for most voters. For instance, suppose three voters have the ranking $c_1 > c_2 > c_3$, 50 have $c_3 > c_1 > c_2$, 50 have $c_2 > c_1 > c_3$, and two have $c_2 > c_3 > c_1$. Here, because 52 voters have c_2 as top choice, 50 have c_3, and only three have c_1, it is reasonable to believe that the true choice is between c_2 and c_3. Nevertheless, because 100 of the voters rank c_1 in second place, c_1 is the Condorcet winner, and c_3 is the *anti-majority candidate.*

Examples and criticisms of this kind do raise concern about the virtue of the Condorcet winner, so they add emphasis to Black's and Smith's obxervations. Actually, Smith's question is easy to answer because, according to Proposition 5, the imposition of natural assumptions force the social choice function to be equivalent to a scoring method. If the social choice procedure is to be related in any manner whatsoever with the pairwise election outcomes – and this is mandatory if one is to follow Smith's suggestion – then, according to Corollary 3.1 b, c, *the procedure can only be the* BC. Thus, to avoid an impossibility conclusion, the appropriate condition must be based on BC properties. Example 7a illustrates the sensitivity of the BC to the pairwise tallies, while Eq. (3.4) illustrates that when the BC does allow a non-Condorcet winner to be BC top ranked, she can do so only by overcoming a significant burden. For instance, in the example of the previous paragraph, c_2 is the Borda winner with 107 points, while the Condorcet winner, c_1, is Borda second ranked with 106 points. (The interested reader may wish to expand these comments to address other issues surrounding the conflict between the choice of the Condorcet winner or the BC. For example, see Nurmi [5] for more details.)

3.4 Other subsets and feasible sets

The above argument just uses the vectors $\{\mathbf{Z}_{i,2^n-(n+1)}\}$. Many other conclusion follow by using all of the vectors $\{\mathbf{Z}_{i,j}\}$. A first step is Corollary 3.3 given below. Part c is included to handle an obvious gap of Corollary 3.1a; namely, is it possible for a Condorcet winner be BC ranked below an anti-majority candidate?

Corollary 3.3. a. *Let $n \geqq 3$, and consider the subset S_j where $|S_j| \geqq 3$. Each component of the* BC *tally for S_j, $\mathbf{X}_j$ is given by*

$$x_{i,j} = \Sigma_k x_{i,k} \tag{3.5}$$

where the summation is over all pairs of candidates, S_k, where $c_i \in S_k$ and S_k is a subset of S_j.
b. *Suppose for a set S_j, $\Sigma_k x_{i,k} > 0$, where the summation is as described in part a. Then, candidate c_i cannot be* BC *ranked or tied for last in S_j. If this summation equals zero, then c_i can be* BC *bottom ranked or top ranked iff the election ranking has all candidates in a tie.*
c. *If c_1 is the Condorcet winner and c_n is the anti-majority candidate, then the* BC *ranking for any subset of candidates that includes c_1 and c_n ranks c_1 strictly above c_n. The* BC *is the only positional voting method for which this is true.*

Example 8. a. To illustrate Eq. (3.5), return to the introductory beverage example where c_1 corresponds to water, c_2 to beer, and c_3 to wine. This means that $x_{1,1} = -x_{2,1} = (6/15) - (9/15) = -3/15$, $x_{1,2} = -x_{3,2} = -3/15$, and $x_{2,3} = -x_{3,3} = (5/15) - (9/15) = -4/15$. Therefore, according to Eq. (3.5), these pairwise tallies uniquely define the BC tally to be $x_{1,4} = -6/15$, $x_{2,4} = -1/15$, and $x_{3,4} = 7/15$. This leads to the BC ranking $c_3 > c_2 > c_1$, or wine > beer > water.
b. Suppose the runoff election is being designed so that if a Condorcet winner exists, she will be elected. If the elimination procedure is based on ordinal rankings, then as Smith noted, each set of candidates must be BC ranked and, at each stage, only the bottom ranked candidate is dropped. Notice that this procedure is the generalization of the Hare method by using the BC (see Sect. 1). This procedure can be generalized by using the normalized BC tally. At each stage, drop all candidates with a negative (vector normalized) BC tally, or, if at least one candidate has a positive tally, then drop all candidates with a non-positive normalized tally. As the Condorcet winner always has a positive normalized tally, she is advanced to all stages and she becomes the winner. The advantage of this procedure is that it accelerates the process because more than one candidate can be eliminated at each stage. A similar procedure holds for other social choice methods, such as a "generalized agenda" where the candidates are listed in some order and then the first $k > 2$ candidates are ranked. The idea is to replace those candidates dropped from further consideration by the next listed candidates. Again, to ensure a Condorcet winner, at each stage, the candidates with a negative (or non-positive) normalized BC tally should be dropped.

Equation (3.5) follows from the requirement that $(F(\{f_{\pi(A)}\}, \mathbf{B}^n), \mathbf{Z}_{i,j}) = 0$ for all i and j. Part c follows because if $c_1 \in S_j$, then, according to Eq. (3.5), $x_{1,j} > 0$. Similarly, if $c_n \in S_j$, $x_{n,j} < 0$. Part b is a simple illustration. Corollary 3.3b significantly extends Corollary 3.1 to all candidates (not just the Condorcet winner) and to all subsets of candidates (not just the set of all candidates). Based on the approach used by Borda, it wouldn't surprise me to discover that he already knew of a result of the general nature of Eq. (3.5) for $n=3$. Versions of it do appear in Smith [15], and then later in Young [16], but only for the set of all candidates. By using this relationship for all sets of candidates, one can address issues such as whether there is a relationship between the BC ranking of a candidate in the set of all candidates and her BC ranking in the set where the bottom ranked candidate is removed. This question is easily answered below; the basis for the solution is stated for all subsets in Corollary 3.4a [by use of Eq. (3.5)], and it is illustrated, with one simple example, in Corollary 3.4b. Corollary 3.4a also shows that only the BC admits such relationships. As such, Corollary 3.4 extends Corollary 3.1, and it can be used, in the obvious fashion, to extend Corollary 3.2. But, before the corollary is stated, a related issue, coming from social choice theory, is introduced.

An important theme from social choice is to understand the appropriate restrictions on the "feasible" sets of candidates that leads to a possibility theorem. (Social choice functions are extended to subset of C^n by asserting that for S_j, $f(\mathbf{p}, S_j)$ is a subset of S_j.) A natural condition, introduced by Arrow [1], is based on the idea that if a candidate is top ranked in S_j and she belongs to S_k, a subset of S_j, then she should remain top ranked when consideration is restricted to S_k. Namely, if $f(\mathbf{p}, S_j)$ is a subset of S_k, where S_k is a subset of S_j, then $f(\mathbf{p}, S_k)=f(\mathbf{p}, S_j)$. However, when combined with some other natural assumptions, this condition turns out to be too strong to permit the existence of any social choice function. (For a geometric explanation that is related to the approach used in this paper, see [11].) Therefore, it is worth searching for alternative restrictions. As argued in Sect. 1, if such a restriction is to be successful for procedures based on positional voting rankings, the restriction must be based on the properties of the BC.

An Arrow type condition leads to an impossibility conclusion. So it is reasonable to suspect that the appropriate substitute condition is based on the notion that the selection of who is the "best" candidate depends on how she compares with the other available candidates. To see some of the possibilities, suppose that Ann, Rose and Kay are three of several candidates running for President where Rose may, or may not decide to run. Suppose if Ann (c_3) and Rose (c_2) were compared in a pairwise majority election, Rose would win. Nevertheless, should Rose choose to join the other candidates, Ann would win. It is reasonable to assume that if Rose decides not to run for office, then Ann's comparative advantage could only improve by the absence of Rose – Ann would have an even better chance to win. Namely, it is reasonable to expect that Ann still would be top ranked whether or not Rose stands as a candidate. However, such an axiom, which is closely related to Arrow's condition, can lead to impossibility assertions.

As an alternative, instead of designing a substitute condition for IIA in terms of who should win with the truncated set of candidates, we might examine those conditions specifying who does not win should Rose choose not to run. So, suppose that Kay (c_1) would beat Rose in a pairwise election. This implies that part of Kay's appeal in a general election is derived from her favorable comparison with Rose. (We do not specify whether Kay is, or is not a winner should Rose decide to run.) Therefore, should Rose decide not to run, then Kay loses some of her comparative advantage. Thus, it is reasonable to suspect that Kay will not win should Rose decide not to be a candidate. However, when the natural extensions of this scenario are converted into an axiom, it can still lead to impossibility assertions.

What does lead to a possibility conclusion is if these two scenarios are combined. In other words, while we can't say that Ann would win should Rose decide not to run, we can say that because Ann is a candidate and because Ann stands to benefit more from Rose's withdrawal than Kay, Kay will not win. Thus, by combining these two scenarios based on Kay's (c_1) and Ann's (c_3) different kinds of comparative advantage with respect to Rose (c_2), we end up with a constrained Arrow type condition that does lead to a possibility assertion, but only for certain classes of procedures.

Definition. a. A social choice function satisfies the *comparative advantage* property if for any subset of candidates, S_j, the following hold: Suppose $c_2 \notin S_j$ and that $c_1 \in S_j$ beats c_2 in a pairwise election. If there is a candidate $c_3 \in f(\mathbf{p}, S_j U\{c_2\})$ that loses to c_2 in a pairwise election, then $c_1 \notin f\mathbf{p}, S_j)$.
b. A social choice function satisfies the condition of *independence of missing alternatives* (IMA) if the following condition holds for every feasible set S_j. If $\mathbf{p}$ and $\mathbf{p}'$ are two profiles where the relative rankings of the candidates in S_j are the same, then $f(\mathbf{p}, S_j) = f(\mathbf{p}', S_j)$.

The IMA condition is, in fact, a form if the IIA condition; it asserts that the rankings of the candidates depends only on what candidates are being considered, not on what candidates may be admitted. I choose to give it a different name to underscore the fact that no assumption is made about how $f(\mathbf{p}, S_k)$ may compare with $f(\mathbf{p}, S_j)$.

Corollary 3.4. a. *Let S_j be a set of candidates and S_k a proper subset of at least three of the candidates in S_j. The* BC *is the only positional voting method that admits a relationship among the tallies or rankings of S_j, S_k, and the pairwise comparisons of the candidates in S_j. If $S_a = S_j - S_k$, then the* BC *relationship for each candidate $c_i \in S_k$ is*

$$x_{i,j} = x_{i,k} + \Sigma_\tau x_{i,\tau} \tag{3.6}$$

where the summation is over all pairs $S_\tau = \{c_i, c_\tau\}$ where $\tau \in \mathrm{Ind}(S_a)$.
b. *Suppose $|S_j| \geqq 3$ and that $S_k = S_j / \{c_a\}$. For $c_i \in S_k$, denote the pair $\{c_i, c_a\}$ by S_a. Then, $x_{i,j} = x_{i,k} + x_{i,a}$. In particular, if c_i has over half of the* BC *total vote tally in S_j and if c_a beats c_i in a pairwise election, then c_i cannot be bottom ranked, nor tied for bottom ranked in S_k. Such an assertion holds only if the* BC *is used to rank both sets.*
c. *If for $n \geqq 3$, an anonymous, neutral, consistent, somewhat faithful social choice function f that satisfies the* IMA *and the comparative advantage properties, then f is equivalent to the* BC *where $f(\mathbf{p}, S_j)$ is the top ranked candidate from the* BC *ranking of S_j.*

It follows from part c that if Arrow's IIA condition is replaced by the comparative advantage axiom, then a possibility theorem emerges. The proofs for parts a and b involve direct computations using the $\mathbf{Z}_{k,j}$ vectors. The assertion that part b holds only for the BC is a consequence of Corollary 3.1. The combination of the IMA along with the other assumptions implies, from Corollary 3.2, that the choice method associated with each set S_j is equivalent to using a positional voting method to select the winner. As the rankings are based on pairwise comparisons, this positional method must be the BC. That the BC satisfies these conditions follows from parts a and b. Incidentally, comparative advantage property can be made simpler if the majority vote tallies, rather than just the rankings, are used.

For instance, if Ann is the candidate to most benefit from Rose not being a candidate (i.e., in a majority vote, Rose does better against Ann than against any other candidate), then the condition can read that Ann will win should Rose not run. There are many other options; e.g., if Ann is the only person to lose to Rose in a pairwise election, then she wins whether or not Rose is a candidate.

Example 9. As an example capturing much of the flavor of the comparative advantage axiom as well as parts a and b of the corollary, note that if c_1 is BC bottom ranked in $S_j=\{c_1,c_2,c_3,c_4,c_5,c_6\}$ (so $x_{i,j}<0$) and BC top ranked in $S_k=\{c_1,c_2,c_3\}$ (so $x_{1,k}>0$), then it seems reasonable to expect that c_1 must have lost at least one of the pairwise comparisons with c_4, c_5, and/or c_6. After all, she must suffer in the voters' comparisons of her with respect to these added candidates if she dropped from top-rank in S_k to bottom-rank in S_j. While this assertion may seem to be obvious, it is surprising to learn that only the BC always satisfies it.

3.5. *A sufficient condition*

According to Eq. (3.5), and the various other expressions deriving from the $\mathbf{Z}_{k,j}$ vectors, all sorts of results can be derived just by determining the tallies for the majority vote elections. Once these tallies are known, then the BC tallies for all other sets are uniquely determined. On the other hand, a difficulty with this approach is that it useful only *if* the tallies for all of the pairwise elections are known. This suggests that in order to demonstrate various BC properties, examples need to be constructed so that the right hand side has the appropriate value. For normative or theoretical studies we need much more. We need to know what values of $\{x_{j,k}\}$, $k\leqq n(n-1)/2$, are admissible. With the vector space representation used here, the answer to this seemingly difficult combinatorial problems is immediate, and the solution is given in Corollary 3.5. To state the results, notice that the vector $\mathbf{M}^n_A=(1/2,-1/2;1/2,-1/2;,\ldots,;1/2,-1/2)$ with $n(n-1)/2$ vector components corresponds to the tally of a ballot of the pairwise comparisons for a voter with the ranking A. Similarly, $\mathbf{M}^n_{\pi(A)}$ corresponds to the tally for a voter with the ranking $\pi(A)$.

Corollary 3.5. *The vector $\mathbf{X}=(\mathbf{X}_1,\ldots,\mathbf{X}_{n(n-1)/2})$ can be realized by a profile of voters iff $\mathbf{X}\in\mathscr{CH}^n$, the convex hull of $\{\mathbf{M}^n_{\pi(A)}\}$, and the components of X are rationals.*

Example 10. a. While the normalized tally $\mathbf{t}=(1/2,-1/2;1/2;-1/2;-1/2,1/2)$ corresponds to the admissible ranking of the cycle $c_1>c_2$, $c_2>c_3$, $c_3>c_1$, $\mathbf{t}$ is not an admissible outcome. Here, the six vectors defining $\mathbf{M}^3_{\pi(A)}$ are $\mathbf{b}_1=(1/2,-1/2;\ 1/2;\ -1/2;\ 1/2,-1/2)$, $\mathbf{b}_2=(1/2,-1/2;\ -1/2,1/2;\ 1/2,-1/2)$, $\mathbf{b}_3=(1/2,-1/2;\ -1/2,1/2;\ -1/2,1/2)$, $\mathbf{b}_4=(-1/2,1/2;\ -1/2,1/2;\ -1/2,1/2)$, $\mathbf{b}_5=(-1/2,1/2;1/2,-1/2;\ -1/2,1/2)$, and $\mathbf{b}_6=(-1/2,1/2;1/2;\ -1/2;1/2,-1/2)$. To see this, note that $\mathbf{t}=\Sigma_j s_j\mathbf{b}_j$ where $\Sigma s_j=1$ and $s_j\geqq 0$. It follows that in order to satisfy both this summation expression for $\mathbf{t}$ and the values of the first two vector components of $\mathbf{t}$, if must be that $s_1=1$; $s_j=0$ for $j\geqq 2$. The contradiction is that the third vector component of $\mathbf{t}$ has the incorrect sign.

b. Next I exploit the geometry of the hull formed by the majority vote tallies. By use of Corollary 3.5 and Eq. (3.5). one can show, for example, that there are profiles where c_1 is the Condorcet winner, but BC ranked in $(n-1)^{\text{th}}$ place, c_2 only loses to c_1 in the pairwise comparisons, but she is BC ranked in $(n-2)^{\text{th}}$ position,..., c_n loses a pairwise comparisons to c_j iff $j < i < n$ but she is BC ranked in the $(n-i)^{\text{th}}$ position, and c_n, the anti-majority candidate is BC bottom ranked. To prove this, note that $\mathbf{0}$ is an interior point of the convex hull specified in Corollary 3.5. Thus the $x_{1,j}$'s, for the pairwise comparisons, can be chosen to have arbitrarily small positive values. For candidate c_2, only the value for pairwise comparison with c_1 is specified, so the remaining $x_{2,j}$'s can be selected as arbitrarily small positive values that satisfy the requirement $0 < x_{1,2^n-(n+1)} < x_{2,2^n-(n+1)}$. Continuing, for candidate c_i, only the values for the pairwise comparisons with c_k, $k < i < n$, have been specified. The remaining values of $x_{i,a}$, for the pairwise comparisons, can be selected so that they are positive and so that $x_{i-1,2^n-(n+1)} < x_{i,2^n-(n+1)}$. All of the values for the pairwise comparisons with c_n have been specified, and they are all negative. Thus, $x_{n,2^n-(n+1)} < 0$, and the conclusion holds. (I used the fact that all of the hyperplanes associated with a tie majority vote pass through $\mathbf{0}$. It is immediate that the approxiate values are admitted because the relevant hyperplanes obviously have independent normal vectors.)

With a similar argument, it is easy to show that if we only know that c_1 *is the Condorcet winner and* c_n *is the antimajority candidate, then the only restrictions on the* BC *ranking is that* c_1 *is strictly ranked above* c_n (Corollary 3.3c); all rankings satisfying this condition are admitted. Related statements hold for the various subsets S_j.

c. By use of Theorem 1,d and using a special case of the above, it follows that there is a profile of voters, $\mathbf{p}$, so that *for all choices of positional voting methods, the election ranking is* $c_{n-1} > c_{n-2} > \ldots > c_3 > c_1 > c_n > c_2$ *while the pairwise rankings of the pairwise votes are* $c_i > c_j$ *iff* $i < j$. Therefore c_1 is the Condorcet winner, c_2 almost is the Condorcet winner (it only loses to c_1), and c_n is the anti-majority winner.

d. Because I show in [10] how to extend voting paradoxes from the literature by use of the dictionaries, I do not emphasize this approach here. However, when this approach is used with the BC, complications arise because $D(\mathbf{B}^n)$ is a subset of U^n. This means that the existence of certain symbols precludes the possibility that other symbols can occur. To find what symbols are admitted, note that Theorem 3 and Corollary 3.5 imply that they are determined by the image of a mapping from $\mathscr{CH}^n$ to the various component spaces of Ω^n; the components of the mapping are given by Eq. (3.5).

To illustrate this with a simple example, consider all possible words in $D(\mathbf{B}^3)$ that have admit the Condorcet cycle $c_1 > c_2, c_3 > c_1$, and $c_2 > c_3$. This corresponds to the region in $\mathscr{CH}^3$ defined by positive values for $x_{1,1}, x_{3,2}$, and $x_{2,3}$. The easiest way is to find whether this region admits any symbols in S_4 with a tie vote. But, according to Eq. (3.5), if all three variable are equal to $\varepsilon > 0$, then the S_4 ranking is $x_{1,4} = x_{2,4} = x_{3,4}$, or $c_1 = c_2 = c_3$. Because $\mathbf{0}$ is an interior point of $\mathscr{CH}^3$, the value of ε can be chosen so that $x_{1,2} = x_{2,3} = \varepsilon$ is an interior point. Thus, by perturbing these values, it follows that the S_4 symbol can be anything. This same approach holds for all values of n.

3.6. *More general* BC *characterizations*

So far I have emphasized the relationship among the BC ranking of a set S_j and the rankings of the pairs of candidates in S_j. This is not necessary; as already suggested by Eq. (3.6), there are relationships among the BC rankings of other subsets of candidates. The idea developed here is the following: Suppose for some reason we are not interested in the relationships among the BC rankings of all subsets of candidates, but rather just the relationships admitted by the BC rankings over a certain collection of specified subsets of candidates. To illustrate, suppose $n=5$, and we are interested in learning whether there is a relationship among the BC rankings of the five subsets of four candidates. Call F, the specified collection of subsets of candidates, a *family*. Notice that a family F defines a linear subspace, Ω_F, of the space Ω^n. Namely, Ω_F is the product of only those component spaces corresponding to subsets of candidates in F. To find what relationships are admitted by the BC over this family F, one only needs to find the normal bundle corresponding to the BC. This construction reduces to nothing more than a linear algebra problem. Namely, a new basis is found for the vector space spanned by the vectors $\{\mathbf{Z}_{j,k}\}$ given by Theorem 3. This basis is chosen so that all of the basis vectors are either in Ω_F or orthogonal to Ω_F. (That this can be done follows from the theory of elementary vector analysis.) Then, in the same way that the $\{\mathbf{Z}_{j,k}\}$ vectors are used to determine the relationships among the BC tallies over all subsets of candidates, the basis vectors that are in Ω_F determine the relationships among the BC tallies of the subsets of candidates in F. This approach is demonstrated in this last subsection. Thus, the only remaining problem is to determine what are the families of subsets of candidates that admit relationships among the BC rankings. This issue is solved in the companion paper [14].

To start this discussion, note that the above characterizations of the BC are based on election outcomes of the $n(n-1)/2$ pairs of candidates. It is natural to wonder whether the BC rankings also can be characterized strictly in terms of the BC rankings of all triplets, or of all sets of four candidates. It can. Corollary 3.6 asserts that the BC ranking for the $|S_j|=k$ candidates in S_j can be determined by the BC rankings for all possible subsets of m candidates from S_j for $2 \leq m < k$. As described above, the explicit relationships are easily determined with elementary linear algebra techniques where the final result is of the form of Eq. (3.5) or of the form described in Example 11,d.

For a different kind of motivation for Corollary 3.6, recall the "money pump" argument occasionally used by economists to dismiss intransitivities. The argument is that a person with intransitive preferences soon would change his rankings to avoid being exploited. To see the idea, suppose a person has the rankings $c_1 > c_2, c_3 > c_1, c_2 > c_3$. If he is to choose a candidate from $\{c_1, c_2\}$, he selects c_1. Presumably, he would be willing to pay (a bribe?) to change to set $\{c_3, c_1\}$ so he could select the personally more desirable outcome of c_3. For similar reasons, after paying, presumably he will pay another sum to choose from the set $\{c_2, c_3\}$ rather than from $\{c_3, c_1\}$. Now, faced with selecting from $\{c_2, c_3\}$, presumably he will pay to obtain the personally more favorable situation of choosing from $\{c_1, c_2\}$. So, after paying, our victim returns to the original situation; unless he changes his rankings, his cyclic preferences provide a never ending opportunity to pump money out of him.

Can a money pump be applied to an organization? The same argument applies if there are subsets of candidates where an organization's top choices form a cycle. For instance, suppose a group's election rankings over triplets from C^4 are $c_2 > c_4 > c_1$, $c_3 > c_1 > c_2$, $c_4 > c_2 > c_3$, and $c_1 > c_3 > c_4$. In such a situation, presumably, when faced with selecting from a specified set of three candidates, the organization would pay a sum of money for the privilege of selecting from the next triplet of candidates in order to obtain a "significantly better" candidate. (The first set of candidates "follows" the last set.) After all, the top ranked candidate in any set is bottom ranked in the following set of candidates. Thus, such an organization is a "money pump" target. In light of this example, it is reasonable to question if positional voting methods avoid such cyclic rankings. It turns out that no positional voting method - not even the BC - can avoid all of them. However, only the BC gives partial relief from such a scam because the BC does avoid many of these kinds of cycles. This is becausae, for the family of triplets from C^4, there are relationships among the BC rankings. For instance, it follows as a consequence of Corollary 3.6 that the BC does not admit situations where the triplets would have the above choices of rankings. (This follows from the first specified normal vector. The normal vectors stated in Corollary 3.6 are in the subspace of Ω_F identified with the family F of triplets.)

Corollary 3.6. a. *Let $n=4$. The* BC *is the only positional voting method that admits relationships among the tallies of the four triplets of candidates. If the sets are listed as* $\{\{c_1, c_2, c_3\}, \{c_1, c_2, c_4\}, \{c_1, c_3, c_4\}, \{c_2, c_3, c_4\}\}$, *then all possible relationships among the* BC *tallies for these four sets of three candidates are characterized by the normal spanned by the two normal vectors* $\mathbf{N}_1 = (1, -2, 1; -2, 1, 1; 1, 1, -2; 1, -2, 1)$ *and* $\mathbf{N}_2 = (-2, 1, 1; 1, -2, 1; 1, -2, 1; 1, 1, -2)$.
b. *Let S_j be a subset of $k > 3$ candidates, and consider all subsets of m candidates, $k > m \geqq 3$. If all sets are* BC *ranked, then there is a relationship among the rankings that is uniquely determined by a set of normal vectors of co-dimension $k(k-1)/2$.*
c. *Let S_j be a subset of candidates with $k \geqq 3$ candidates, and let m be an integer such that $2 \leqq m < k$. The* BC *imposes a relationship among the S_j rankings and the $k!/m!(k-m)!$ subsets of m candidates from S_j. These relationships are uniquely determined by the vectors in a normal bundle. A basis of this normal bundle of c-dimension $k(k-1)/2$ is given by the normal vectors determining the* BC *relationships among the sets of m candidates and by $k-1$ additional vectors of the following form: in the following manner: For $c_a \in S_j$, the S_j vector component is $[(k-2)!/(m-2)!(k-m)!]\mathbf{Y}_{a,j}$, while the S_i component, for $|S_i| = m, c_a \in S_i$,* is $-\mathbf{Y}_{a,i}$. *All other components are zero. If $k=4, m=3$, and none of the sets are* BC *ranked, then there is no relationship among the admitted rankings.*

In part c, the new normal vector associated with c_a is where she is top-ranked in all sets of m candidates that she belongs to while the remaining candidates are tied for bottom ranked, yet she is bottom ranked in the set of k candidates while the rest of the candidates are tied for top rank. This is a normal vector so such a ranking is not an admissible BC outcome. On the other hand, this ranking is an admissible election outcome for any other choice of a system voting vector. Along with Proposition 3, the dimension of the normal bundle indicates the number of cycles and other phenomena avoided by the BC. The uniqueness assertion from part a extends to part c for $k=4$, $m=3$, and to other values of k and m. But, the approach used here to prove part a becomes overly complicated to prove part c. Therefore, a different kind of argument (this is not developed here) is required. Incidentally, this uniqueness assertion of part a is strong; if even one of the sets is not BC ranked while the other three are, then there is no relationships among the set of all possible election outcomes.

Proof of Part a. Suppose the sets listed in part a are S_7, S_8, S_9, and S_{10}. To prove the corollary, it suffices to show that the vectors $\{\mathbf{Z}_{k,j}\}, j=7,8,9,10, k \in \text{Ind}(S_j)$, determine the specified normal bundle over the specified linear subspace of Ω^n. This is a straightforward computation. The uniqueness assertion is proved in the next section.

Example 11. a. Because $F(\mathbf{p}, \mathbf{B}^n)$ satisfies the neutrality condition, other choices of normal vectors can be determined from the first specified one by using the permutation group theoretic structure. A quick way to find other normal vectors is to list the candidates in a repeating *chain*, such as $--c_1--c_2--c_3--c_4--c_1--c_2--$. Each candidate c_j is the *central candidate* of that subset consisting of c_j and the two candidates on either side of her in the chain. A normal vector for Corollary 3.6a is determined in the following manner; for each subset, assign the value -2 for the central element and 1 for the other two candidates. The first listed normal vector in the corollary comes from the above chain, while the second one comes from the chain $--c_1--c_2--c_4--c_3---c_1--c_2---$. Another normal vector (but linearly dependent) comes from the chain $--c_1--c_4--c_2--c_3---c_4--$, etc. One immediate consequence is that *it is impossible for all four of the central candidates to be* BC *bottom ranked, or to be* BC *top ranked.* The selection of normal vectors for other sets can be determined by a graph theoretic argument.

b. The two specified normal vectors determine the relationships

$$x_{2,7}+x_{1,8}+x_{4,9}+x_{3,10}=0 \tag{3.7}$$

$$x_{1,7}+x_{2,8}+{}_{3,9}+x_{4,10}=0\ . \tag{3.8}$$

These are the only (independent) restrictions imposed on the Borda outcomes. Consequently, in either expression, it is impossible for the four variables all to be of the same sign. This implies, as already noted in part a, that the central candidates cannot all be top (or bottom) ranked. In particular, it implies that the example demonstrating an organizational "money pump" never can occur with the BC. On the other hand, these expressions do permit the rankings $c_1 > c_2 > c_3, c_4 > c_1 > c_2, c_3 > c_4 > c_1$, and $c_2 > c_3 > c_4$; rankings that admit cycles. This is because these rankings require $x_{1,7}, x_{4,8}, x_{3,9}$, and $x_{2,10}$ to be positive. A simple way to show these choices are compatible with Eqs. (3.7), (3.8) is to set all of the variables in Eq. (3.7) equal to zero. This forces $x_{2,8}$ and $x_{4,10}$ to be negative. So, values exist to satisfy Eq. (3.8).

c. These results provide another axiomatic representation of the BC. For instance, for $n=4$ candidates, *if a social choice function f satisfying the* IMA *is defined over all subsets of three candidates, if over each set it is anonymous, neutral, consistent, somewhat faithful, and if it never selects the central candidates for any chain, then f is equivalent to the* BC *(without tie breakers).* This serves as another alternative to Arrow's condition discussed in Corollary 3.4c.

d. Part c asserts that the BC admits a relationship among the rankings of the four candidates and those of the four subsets of three candidates. The relationship is given by the specified normal vectors. Therefore, the new equations relating the outcomes of the four sets are of the form

$$x_{1,7}+x_{1,8}+x_{1,9}=2x_{1,11}\ . \tag{3.9}$$

From this equation, all sorts of new conclusions, such as the impossibility of c_i being BC bottom ranked in the three subsets of three candidates (which forces the left hand side of Eq. (3.9) to be negative), yet BC top ranked in S_{11} (so $x_{1,11}>0$) follow immediately.

In a similar manner, it now is easy to show for n candidates and for any m where $n>m\geqq 2$ that c_i cannot be BC top (bottom) ranked in all possible sets of m candidates while being BC bottom (top) ranked in the set of all n candidates. The BC is the only positional voting vector with this property. (If another voting vector had this same property, then the linear space it defines would have the same dimension as the one for the BC. As the vector space for the BC is contained in this new vector space, they must agree. Thus the new vector is a BC.) When $m=2$, this becomes a restatement of Corollary 3.1a. Moreover, by following the earlier arguments given in this paper, all of the results given in terms of pairs and the ranking of n candidates extend to this new situation.

3.7. Other system voting vectors

A remaining issue is to gain some insight into the structure and the properties of the other choices of system voting vectors that belong to a^n. As shown here, the structure of these system voting vectors is governed by the structure of the BC. To see why this is so, recall that if $\mathbf{W}^n\in a^n$ where $\mathbf{W}^n\neq\mathbf{B}^n$, then there is a scaling so that $V(\mathbf{B}^n)$ is a proper linear subspace of $V(\mathbf{W}^n)$. In turn, this means that the normal bundle of $V(\mathbf{W}^n)$ (in Ω^n) is a proper linear subspace of the normal bundle of $V(\mathbf{B}^n)$. Therefore, a normal vector for $V(\mathbf{W}^n)$ is given by a linear combination of the normal vectors for $V(\mathbf{B}^n)$. The importance of this observation derives from the fact that these normal vectors determine what kind of relationships exist among the rankings (and tallies) of the different subsets of candidates. Therefore, it follows that any relationship which is admitted by the election rankings of $\mathbf{W}^n$ also is admitted by the BC election rankings.

This linear combination assertion provides a stronger statement; the conditions admitted by other choices of $\mathbf{W}^n\in a^n$ can be viewed as coming from linear combinations of the relationships admitted by the BC. Thus, while it is an interesting issue to characterize all choices of $\mathbf{W}^n\in a^n$, it follows from the above argument that these other choices of vectors in a^n do not add anything new to our understanding of the kinds of relationships that can be admitted among positional election rankings.

Corollary 3.7. *Let* $\mathbf{W}^n\in a^n$.
a. *A normal vector for* $V(\mathbf{W}^n)$ *can be expressed as a linear combination of the vectors* $\{\mathbf{Z}_{k,j}\}$.
b. *If there exists a relationship among the election rankings of* $\mathbf{W}^n$ *describing what kinds of rankings of the subsets of candidates cannot occur with the same profile, then the* BC *admits the same, or a more strict relationship among the* BC *election rankings.*

The task of finding, in a direct fashion, an example of $\mathbf{W}^N\in a^n, \mathbf{W}^n\neq\mathbf{B}^n$, appears to be difficult. However, by use of part a of the corollary, a more tractable, indirect approach can be fashioned to determine classes of such vectors. As an outline of how this is done, the emphasis is changed from finding an $\mathbf{W}^n$

to finding a normal bundle for a vector space of Ω^n that can be identified with a space $V(\mathbf{W}^n)$ for some choice of $\mathbf{W}^n$. In this manner, once the space $V(\mathbf{W}^n)$ is found, we can determine a choice for $\mathbf{W}^n$.

What makes this approach simpler is that, according to part a of the corollary, the normal bundle for any such space $V(\mathbf{W}^n)$ must be based on linear combinations of the normal vectors $\{\mathbf{Z}_{k,j}\}$. Thus, by use of Corollary 3.7, we have a place to start our search – we consider all linear subspaces of the normal bundle to $V(\mathbf{B}^n)$. However, not all such linear subspaces serve as the normal bundle for some space $V(\mathbf{W}^n)$. This is because such a space $V(\mathbf{W}^n)$ must satisfy certain orientation properties to reflect that $\mathbf{W}^n$ is a system voting vector and that positional voting methods are neutral processes. These properties for $V(\mathbf{W}^n)$ force any such linear subspace to assume a particular orientation in Ω^n, which in turn forces its normal bundle to have a particular orientation. (This symmetry constraint on the normal bundle is already exhibited by the obvious symmetry properties exhibited by the $\mathbf{Z}_{k,j}$ vectors – for example, certain obvious permutation of indices map these vectors back into the set of these vectors. In more technical terms, these vectors are in the orbit of particular algebraic group. Also, in more technical terms, the analysis of these linear spaces is conceptually easier if it is done with Grassmanians. However, a straight forward approach does provide answers.) This outline provides a systematic method to determine other choices of $\mathbf{W}^n \in a^n$.

To see the kind of outcomes that can occur, it follows that there exist choices of $\mathbf{W}^5$ so that if c_j is top-ranked in all subsets of four or fewer candidates, then c_j cannot be bottom-ranked in the set of all five candidates. Notice that this is a much weaker assertion than that holding for the BC; for the BC the same assertion holds but where c_j is top-ranked for, say, all subsets of k candidates where k assumes any value $2 \leqq k \leqq 4$. Therefore, the conditions for this choice of $\mathbf{W}^5$, then, need to combine the various conditions for the BC. This illustrates the above assertion that the kinds of relationships that emerge are weaker and are based on linear combinations of the assertions for the BC. Incidentally, the normal vector defining such a process is given by an appropriate linear combination of *all* $\mathbf{Z}_{j,k}$ vectors related to c_j. The scalars in this linear combination determine the choices of the $\mathbf{W}^5$. In this manner, the stratified structure of a^n can be determined. The main point for the current paper is to indicate that all of these vectors along with the structure of a^n are determined by the BC. The BC does play a fundamental role in understanding the properties of positional voting processes.

4. Proofs

The proofs of most of the assertions either are contained in the body of this paper, or they are immediate. The proofs of the certain of the remaining statements follow directly from arguments developed in [10]. For instance, with the exception of the last part of Theorem 1, both Proposition 2 and Theorem 1 are proved in [10]. The proof of Theorem 2 also is given in the last section of [10]; indeed, while Theorem 2 is not formally stated in [10], the proof of this statement is the manner in which Theorem 1 is proved. As an outline, recall that a basis for $V(\mathbf{B}^n)$ is determined in the following manner. Starting with a ranking $\pi(A)$, a new ranking $\pi(A)'$ is determined by transposing some two adjacently ranked candidates. The basis is found by computing

$$\mathbf{B}^n_{\pi(A)} - \mathbf{B}^n_{\pi(A)'} \ , \tag{4.1}$$

where the system vector $\mathbf{B}^n$ is expressed in a vector normalized form. The resulting vector has a fixed vector form in each subset of candidates containing these two candidates. By carefully choosing the two candidates that are to be transposed, it is shown that $\dim(V(\mathbf{B}^n)) = n(n-1)/2$. Moreover, as also shown in [10], each of the above basis vectors are in $V(\mathbf{W}^n)$ for an appropriately scalar normalization of $\mathbf{W}^n$. To prove this, the vector from $V(\mathbf{W}^n)$ involves the tally from more than one voter, and then an expression of the form given in (4.1) is used. This computation uses nothing more than the fact that when $\mathbf{W}^n$ is expressed in a vector normalized form, each voting vector component is non-zero. Thus the last part of Theorem 2 holds.

The same argument extends to scoring vectors. The only difference is that there exist situations (where the scoring vector is in a lower dimensional algebraic set) where the scale multiple must be negative. This can be handled in either one of two ways. The first is to change the appropriate vector component of $\mathbf{B}^n$ to be the reversed BC. This is the approach used in Theorem 1. The second is to allow negative scale multiples. This is the approach used in Theorem 2. Both approaches are indicated because there exist settings where one approach is preferred to the other.

Proof of Corollary 1.1. In the proof (Sect. 1.3) that the BC run-off election has the abstention paradox, the key step involved using three words that differed only in one symbols. The proof for the general case is much the same; again I use three words from $D(\mathbf{B}^n)$ that differ only in a specified symbol. One choice for this symbol has the relative rankings of two candidates tied, a second choice has one candidate ranked above the other, and the third choice reverses this relative ranking. To do this, I first show that such rankings can be found; namely, if two words in $D(\mathbf{B}^n)$ differ only in one symbol and if the only difference in that symbol has two adjacently ranked candidates in reversed positions, then the word where this symbol has a tie vote between these candidates also is in $D(\mathbf{B}^n)$. For instance, if $(c_1 > c_2,\ c_1 > c_3,\ c_2 > c_3,\ c_1 > c_2 > c_3)$ and $(c_1 > c_2,\ c_1 > c_3,\ c_2 > c_3,\ c_2 > c_1 > c_3)$ are in $D(\mathbf{B}^3)$, then so is $(c_1 > c_2,\ c_1 > c_3,\ c_2 > c_3,\ c_1 = c_2 > c_3)$.

This assertion is immediate. There are several vectors in $V(\mathbf{W}^n)$ corresponding to each of the first two words; let $\mathbf{v}_1$ be a vector corresponding to the first word, $\mathbf{v}_2$ to the second, and $\mathbf{v}_t = t\mathbf{v}_1 + (1-t)\mathbf{v}_2$ to a point on the line segment between them. Of course, because $V(\mathbf{W}^n)$ is a vector space, $\mathbf{v}_t \in V(\mathbf{W}^n)$.

From Theorem 6 in [10], the region of profiles corresponding to a fixed symbol is a convex region, therefore all but one of the symbols of $\mathbf{v}_t$, $\mathbf{v}_1$, $\mathbf{v}_2$ must agree. The remaining symbol is where the rankings of the two candidates are reversed. Assume without loss of generality that this is for the ranking of the set S_j and that the candidates are $\{c_1, c_2\}$. This means that in the appropriate component space of Ω^n, one ranking has $x_{1,j} > x_{2,j}$ and the other has the reversed inequality. For all other values of $x_{k,j}$, the same ordering applies. It now is immediate that there is a value of t so that the c_1, c_2 coordinates for $\mathbf{v}_t$ are $x_{1,j} = x_{2,j}$, while the remaining coordinate orderings are preserved. This completes the proof of the assertion.

If a social choice method based on $\mathbf{B}^n$ admits the abstention paradox, then there are three words in $D(\mathbf{B}^n)$, based on three candidates, c_1, c_2, c_3 where c_1 and c_3 are possible outcomes depending on the rankings of two possible final sets. The choice of which final set occurs depends on whether c_2 can be advanced one position in the ranking of a swing set. Start with the ranking of the swing set where c_2 is tied in the swing position. Choose the rest of the symbols of the word in $D(\mathbf{B}^n)$ so that if c_2 is lowered from the tie position, then c_3 is the outcome, yet if c_2 is advanced from the tie, then c_1 is the final outcome. (Again, this is possible because the binary susceptible method has the abstention property.) This identifies the words that changes the outcomes.

The selected profile will be associated with the word with the tie vote in the swing symbol. Construct the rankings for the two voters with c_2 as top ranked, c_1 is bottom ranked, and c_3 is ranked somewhere in the middle. By use of Theorem 6 of [10], $\mathbf{p}$ can be selected so that it includes one of these voters. So, if both vote, the outcome is the undesired c_1, if they abstain, the outcome is the more desired c_3. Because this argument is based on a particular word being in $D(\mathbf{B}^n)$ and because this same word is in all choices of $D(\mathbf{W}^n)$, the same argument proves that this phenomenon occurs holds for all choices of positional voting vectors. (A slightly more complicated argument extends the conclusion to a much wider class of procedures.)

Proof of Theorem 3, a. This is a simple computation using the basis for $V(\mathbf{B}^n)$ derived in [10].

Proof of part b. This follows immediately from Theorem 5.1 in [10]. An alternative proof can be derived by using the scheme developed to prove Corollary 3.6a, c which is given below.

Proof of part c. This follows from Theorem 2 and the fact, pointed out above, that the main component of the proofs for other choices of $\mathbf{W}^n$ is that when $\mathbf{W}^n$ is expressed in a vector normalized form, each voting vector component is non-zero.

Proof of Corollary 3.2., a. First I show that a somewhat faithful social choice function f cannot be single valued. If it could be, then $f(\mathbf{p})$ is the same set from $P(C^n)$ for all choices of $\mathbf{p}$. Let $c_1 \in f(\mathbf{p})$, and let $\mathbf{p}_1$ be a profile of a single voter where c_1 is bottom ranked. Then, according to the assumption of being somewhat faithful, $c_1 \notin f(\mathbf{p}_1)$. This contradicts the assumption that f has only one value and proves the assertion.

According to assumption and the above, f is an anonymous, neutral, consistent social choice function that takes more than one value. Thus, according to Propostion 5, f is equivalent to a general scoring method. It now follows from Corollary 3.1 b that the added assumption about f satisfying a condition involving pairwise rankings forces f to be equivalent either to the BC or to the reversed BC. However, because f is somewhat faithful, the weight assigned to the bottom ranked candidate must be less than that for some other candidate, so f must be equivalent to choosing the top ranked candidates from the BC ranking. This completes the proof.

Incidentally, this assertion does not preclude the possibility of tie breakers. For instance, suppose the binary condition is that the anti-majority candidate is not elected and that a single winner is being selected, then tie breakers are admitted. This is because, at the end of the first stage, the anti-majority candidate cannot be tied for first place. Thus, as the tie breaker is applied to those candidates that are top ranked, this does not involve the anti-majority candidate.

Proof of part b. The BC need not have the Condorcet winner as top ranked, so, by part a, an impossibility assertion follows.

Proof of part c. According to part a, the method at the first stage must be the BC. Thus, the first stage is equivalent to selecting the top ranked candidate(s) from the BC ranking. Now, suppose a tie breaker $\mathbf{W}_2$ is admitted where $\mathbf{W}_2$ is not a Borda Vector. According to Theorem 2, a profile can be found so that the pairwise votes are all ties (which forces the BC outcome to have all candidates tied for first place), but c_1 is the top ranked candidate of the $\mathbf{W}_2$ tally. This means that the social choice mechanism selects c_1. This choice of a singleton violates the assumption of non-determinancy. This completes the proof.

Incidentally, this is one of the few places I use the fact that the social choice function has a singleton in its image. On the other hand, if f does not admit any singleton sets in the image, only sets of at least $k>1$ candidates, then the social choice method is equivalent to selecting the k top ranked candidates from the election ranking. For instance, this may correspond to selecting a committee of k candidates. In this case the BC ranking of C^n *can be reduced by a tie breaker* to the top k candidates without coming into conflict with non-determinacy or consistency. To regain the same conclusion of a BC election without use of a tie breaker, one just strengthens the non-determinacy condition to require more than k candidates in the set of f whenever all of the pairwise elections result in tie voters.

Proof of Corollary 3.6b. Let $k=n$ and let Ω^n_m be the subspace of Ω^n corresponding to the entries in the sets of m candidates. Let $Pr: \Omega^n \to \Omega^n_m$ be the obvious projection mapping. Part b holds iff the dimension of the vector space $Pr(V(\mathbf{B}^n))$ is $n(n-1)/2$. Because $\dim(V(\mathbf{B}^n))=n(n-1)/2$, we have that $\dim(Pr(V(\mathbf{B}^n))) \leq n(n-1)/2$. It remains to show that the dimension cannot be smaller than $n(n-1)/2$.

The space $Pr(V(\mathbf{B}^n))$ can be viewed (Corollary 3.5) as the image of a mapping F: $\mathscr{CH}^n \to \Omega^n_m$ where each coordinate of F is given by Eq. (3.5). If the dimension of the image is less than $n(n-1)/2$, then, by the linearity of F, there is a point $\mathbf{p} \in \mathscr{CH}^n$ and a direction $\mathbf{a}$ so that $F(\mathbf{p}+\mathbf{a}t)=F(\mathbf{p})$. By the Taylor series, this means that DF does not have rank $n(n-1)/2$.

The matrix DF is $\dim(\Omega^n_m) \times n(n-1)/2$ where each row corresponds to the value of $x_{j,k}$ while each column corresponds to a pair of candidates. It follows from Eq. (3.5) that the entries in this matrix are either unity or zero. To show that the rank of DF is $n(n-1)/2$, it suffices to show that the span of the row vectors includes the $n(n-1)/2$ coordinate vectors in $R^{n(n-1)/2}$, $\mathbf{e}_i$, with unity in the i^{th} component and zero in all others. To show this, first consider the coordinates associated with c_1 (i.e., $x_{i,j}$) and assume that the $(j-1)^{\text{th}}$ columns of DF represents the pair $\{c_1, c_j\}, j=2, \ldots, n$. List all candidates other than c_1 according to subscipt as $c_2, c_3, \ldots, c_n, c_2, c_3, \ldots$ From this list form the set of candidates $S_k, k=1, \ldots, n-1$, where c_1 is in S_k along with c_{k+1} and the $m-2$ candidates listed to the right of her. For instance, $S_1=\{c_1, c_2, \ldots, c_m\}$, $S_3=\{c_1, c_4, \ldots, c_{m+2}\}$, and $S_{n-1}=\{c_1, c_n, c_2, \ldots, c_{m-1}\}$. Assume that the first $n-1$ rows of DF are listed in the order $x_{1,j}$ where j corresponds to the sets S_j defined above. With this notation, the first row of DF (for $x_{1,1}$) has unity in the first $m-1$ columns, and zero in all others. In general, the j^{th} row has unity on the main diagonal, and unity in the next $m-2$ columns. If there are not $m-2$ columns remaining in this

$(n-1)\times(n-1)$ block, then the remaining number of ones are placed starting from the right hand side. All other entries in this row are zero.

Either by noting that this matrix is a circulant matrix (see [7] and the listed references) whose determinant is non-zero, or by elementary row reduction, it follows that the resulting $(n-1)\times(n-1)$ block has non-zero determinant. This means that the row vectors with unity in a component corresponding to a pair $\{c_1, c_j\}$ and zero in all others is in the span of the row vectors of DF. The same construction is continued with all other choices of c_j, and the conclusion follows.

Part c. Again, let $k=n$. For this part, we are considering the projection of $V(\mathbf{B}^n)$ into the product space of Ω_m^n with the vector space representing the outcomes of $S_{2^n-(n+1)}$. As this projection has a subspace of dimension $n(n-1)/2$ (by part b), the dimension of the projection is no less than $n(n-1)/2$. As $\dim(V(\mathbf{B}^n))=n(n-1)/2$, the projection has exactly this dimension. Thus, it remains to find the $n-1$ normal vectors that were not in Part b.

Start with candidate c_1. For each term in the sum $\Sigma_j \mathbf{Z}_{1,j}$, where the summation is over the $(n-1)!/(m-1)!(n-m)!$ of the sets S_j containing c_1 and $m-1$ other candidates, the vector component for each pair is either $(1/2, -1/2)$ (c_1 is top ranked) or $\mathbf{0}$. Each pair gets counted in a non-zero faction $(n-2)!/(m-2)!(n-m)!$ times. Thus, if $-[(n-2)!/(m-2)!(n-m)!]\mathbf{Z}_{1,2^n-(n+1)}$ is added to the sum, all of the components for pairs become zero. This is the vector described in the statement of the corollary. Clearly, by changing the choice of the candidate from c_1 to c_a, $n-1$ normal vectors are defined.

What remains is to show the uniqueness assertion of part a. Namely, I prove that if any of the four sets of three candidates is not ranked with a BC, then all possible election rankings occur. Let Ω' denote the product space of the vector ranking regions for the four sets. Suppose there exist choices of voting vectors, $\mathbf{W}_i$, that admit a relationship among the ordinal election rankings for the four sets of three candidates specified in part a. This means that the span of the election tallies define a lower dimensional, linear subspace of Ω', so this linear subspace has a non-zero normal vector, $\mathbf{N}$, in Ω'.

My first assertion is that $\mathbf{N}$ must be a linear combination of the two normal vectors specified in part of the corollary. Extend the above choices of $\mathbf{W}_i$'s in any desired manner to obtain a $\mathbf{W}^n$. Recall from Theorem 2 that $V(\mathbf{B}^n)$ is a proper subset of $V(\mathbf{W}^n)$ for any $\mathbf{W}^n$ that is not a Borda system vector. (Of course, this requires the proper scalar normalization of the components of $\mathbf{W}^n$.) Therefore, a normal vector for $V(\mathbf{W}^n)$ must be in the span of the normal vectors for $V(\mathbf{B}^n)$. Thus, $\mathbf{N}$ must be in the span of the two vectors specified in part a of the corollary. Namely, $\mathbf{N}=a_1\mathbf{N}_1+a_2\mathbf{N}_2$ because we assume that $\mathbf{N}\neq\mathbf{0}$, either a_1 or a_2 is non-zero. Without loss of generality, assume that $a_1\neq 0$. Thus, $\mathbf{N}$ can be expressed as $\mathbf{N}_1+b\mathbf{N}_2$.

A scalar normalization of the $\mathbf{W}_i$'s that suffices for Theorem 2 is to assume that $\mathbf{W}_i=(2, x_i, -2-x_i)$, $i=1,2,3,4$, where $-2\leqq x_i\leqq 2$. Thus, each $\mathbf{W}_i$ becomes a Borda Count iff $x_i=0$. As $\mathbf{N}$ is a normal vector, there are $4!=24$ resulting equations corresponding to the scalar product of $\mathbf{N}$ with the vote tally assigned to a profile consisting of a single voter where the voter's preferences vary over each of the possible permutations of $A=c_1>c_2>c_3>c_4$. This gives rise to a system of 24 equations with the four unknowns x_i. Because $\mathbf{N}$ is a linear combination of the normal vectors for the BC tally, each equation has at least one solution – the BC. It remains to show that this is the unique solution. This requires finding four equations where the coefficients, determined by the entries of **N**, are linearly independent.

By looking at the equations for the cyclic permutations of the ranking $c_1 > c_2 > c_3 > c_4$, (i.e., $c_1 > c_2 > c_3 > c_4$, $c_2 > c_3 > c_4 > c_1$, $c_3 > c_4 > c_1 > c_2$, and $c_4 > c_1 > c_2 > c_3$), the equations for $\mathbf{X} = (x_1, x_2, .x_3, x_4)$ becomes $B\mathbf{X}' = \mathbf{0}$, where B is a 4×4 matrix with entries that are scalars or scalar multiples of b (the multiple of $\mathbf{N}_2$ in the definition of **N**). The unique solution for this system is the BC as long as the matrix B is non-singular. The determinant of B equals $(-5+5b+b^2)(1-b-b^2)$, so $b = (-5 \pm 3[5]^{1/2})/2$ are the only real values of b that cause this matrix to be singular. On the other hand, if one looks at the equations resulting from the cyclic permutations of the ranking $c_1 > c_2 > c_4 > c_3$, a similar matrix equation results where the determinant of the coefficient matrix has the value $-1-4b+9b^2-10b^3+5b^4$. This new polynomial has its real zeros at the values $b = (-5 \pm 3[5]^{1/2})/10$. Thus, no matter what is the choice of b, there exist some selection of four independent equations that forces the only solution of the system to be the BC. This completes the proof.

References

1. Arrow KJ (1952) Social choice and individual values. Wiley, New York
2. Brams S, Fishburn P (1983) Paradoxes of preferential voting. Math Mag 56: 207–214
3. Fishburn P, Gehrlein W (1976) Borda's rule, positional voting, and Condorcet's simple majority principle. Public Choice 28: 79–88
4. Haunsperger D (1991). PhD dissertation, Northwestern University, Department of Mathematics
5. Nurmi H (1987) Comparing voting systems. Reidel, Dordrecht
6. Saari DG (1982) Inconsistencies of weighted summation voting systems. Math Oper Res 7: 479–490
7. Saari DG (1984) The ultimate of chaos resulting from weighted voting systems. Adv Appl Math 5: 286–308
8. Saari, DG (1987) Chaos and the theory of elections. In: Kurzhanski AB, Sigmund K (eds) Dynamical systems. (Lect Notes Econ Math Systems, vol 287) Springer, Berlin Heidelberg New York, pp. 179–188
9. Saari DG (1988) Symmetry, voting, and social choice. Math. Intell 10: (No 3): 32–42
10. Saari DG (1989a) A dictionary for voting paradoxes. J Econ Theory 48: 443–475
11. Saari DG (1989b) Calculus and extensions of Arrow's theorem. J Math Econ
12. Saari DG (1990a) Susceptibility to manipulation. Public Choice 64: 21–41
13. Saari DG (1990b) Consistency of decision processes. Ann Oper Res 23: 103–137
14. Saari DG (1991) Relationship admitting families of candidates. Soc Choice Welfare 8
15. Smith JH (1973) Aggregation of preferences with variable electorate. Econometrica 41: 1027–1041
16. Young HP (1974) An axiomatization of Borda's rule. J Econ Theory 9: 43–52
17. Young HP (1975) Social choice scoring functions. SIAM J Appl Math 28: 824–838
18. Van Newenhizen J (1988) A geometric investigation of voting techniques: a comparison of approval voting, positional voting techniques and the Borda count. Northwestern University PhD dissertation, 1988