Discussion Paper No. 1111

THE COPELAND METHOD I : RELATIONSHIPS AND THE DICTIONARY

by

Donald G. Saari
and
Vincent R. Merlin

November 1994

THE COPELAND METHOD I; RELATIONSHIPS AND THE DICTIONARY

DONALD G. SAARI AND VINCENT R. MERLIN

Northwestern University

ABSTRACT. A central political and decision science issue is to understand how election outcomes can change with the choice of a procedure or the slate of candidates. These questions are answered for the important Copeland method (CM) where, with a geometric approach, we characterize all relationships among the rankings of positional voting methods and the CM. Then, we characterize all ways CM rankings can vary as candidates enter or leave the election. In this manner new CM strengths and flaws are detected.

The Condorcet (or majority) winner [Cn] is the candidate who beats all others (by winning most votes) in pairwise contests. A glaring fault of this widely accepted concept is that it need not exist. Instead, for $n \geq 5$ candidates, c_1 could win all but one pairwise vote while all other candidates lose at least two. Although no one satisfies Condorcet's criterion, c_1 comes the closest, so it is arguable that she is who the voters want. She does win with Copeland's method (CM) – an important, natural extension of the Condorcet winner [C].

More precisely, in a pairwise competition between c_j and c_k let

$$s_{j,k} = \begin{cases} 1 & \text{if } c_j \text{ beats } c_k \\ \frac{1}{2} & \text{if } c_j \text{ and } c_k \text{ are tied} \\ 0 & \text{if } c_k \text{ beats } c_j \end{cases} \tag{1.1}$$

The *Copeland score* for each c_j, defined as

$$C(j) = \sum_{k \neq j} s_{j,k}, \tag{1.2}$$

is used to rank the candidates where more is better. Equivalent to these $(1, \frac{1}{2}, 0)$ weights are the $(\frac{1}{3}, \frac{1}{6}, 0)$ and $(1, 0, -1)$ choices that we use to simplify proofs. Notice that the CM is the method commonly used to rank hockey and other sport teams.

Trivially, the CM ranking is transitive. (The CM score identifies each candidate with a point on the line, so the transitivity of the election ranking is inherited from the transitivity of points on the line.) Equally as trivial, when a Condorcet winner exists, she is CM top-ranked. (Only the Condorcet winner receives a point from each pairwise contest.) Other CM properties and relevant references are in [N], but, in light of its obvious importance,

We would like to thank C. Mount for her careful reading of the manuscript. The work of DG Saari was supported by NSF Grant IRI 9103180 This research was done while V. Merlin was visiting Northwestern University; his permanent address is C.R.E.M.E., Université de Caen, 14032 Caen, France

it is surprising to discover how little is known about this approach. Therefore, a natural objective is to determine the remaining CM properties.

In this and a companion paper [MS], we provide a fairly complete description of the CM properties while emphasizing why they occur. We accomplish this by using a geometric approach ([S1-4]) where the geometry helps to discover and verify new conclusions. Because our arguments outline how to use these geometric techniques for other election and choice procedures, this geometric description may be of independent interest.

Here we examine *single profile concerns;* i.e., the properties, paradoxes, relationships, and perversities of the CM rankings resulting from a fixed profile. We show how to find all possible relationships among the CM and positional rankings of the candidates. (Positional methods extend the plurality vote by giving points to a voter's lower ranked candidates.) Then we show how to find all ways CM rankings can be related when the set of candidates varies because candidates drop out, new ones are added, or comparisons are desired. In this manner new CM faults are discovered and a large set of profiles is identified for which it is arguable that both the CM and Condorcet winners violate the voters' true intent.

Multiple profile concerns, addressed in the companion paper [MS], arise by comparing CM outcomes of two or more profiles. To illustrate, the first profile could be the current sincere preferences of the voters. Options for the second profile include strategic action where certain voters change voter type, or where more voters now support a particular candidate, or where a truncated ballot is cast, or where some voters abstain, or where new voters vote, etc. Another choice has each of two profiles representing different subcommittees while a third is the combined group. As such, the [MS] results describe all CM manipulation, consistency, responsiveness, and monotonicity properties.

A flavor of our single profile assertions (where $c_1 \succ c_2$ means that c_1 beats c_2), comes from profile (i.e., a listin of voters's preferences) $\mathbf{p}^*$ where

Number	Type	Number	Type
8	$c_1 \succ c_2 \succ c_3$	9	$c_3 \succ c_2 \succ c_1$
3	$c_1 \succ c_3 \succ c_2$	3	$c_2 \succ c_3 \succ c_1$
1	$c_3 \succ c_1 \succ c_2$	3	$c_2 \succ c_1 \succ c_3$

has c_2 as the Condorcet and CM winner, while the CM ranking over all subsets

$$(c_2 \succ c_1,\ c_1 \succ c_3,\ c_2 \succ c_3,\ c_2 \succ c_1 \succ c_3) \tag{1.3}$$

conflicts both with the plurality ranking $c_1 \succ c_3 \succ c_2$ and the antiplurality ranking (where each voter votes for his two top-ranked candidates) $c_2 \succ c_3 \succ c_1$. (So, the antiplurality ranking reverses the plurality ranking.) It is not clear who should be selected because the top-ranked candidate is c_1 with the plurality method and c_2 with CM and the antiplurality vote. What else can we say about $\mathbf{p}^*$: for instance, do any or many positional methods agree with the CM ranking?

This example, illustrating how little the outcomes of different procedures can have in common, is typical of our negative results. On the other hand, we introduce a new relationship between the BC (the Borda Count where for n candidates $n-j$ points are assigned to a voter's jth ranked candidate, $j = 1, \ldots, n$) and the CM that leads to positive conclusions about the CM and positional voting rankings.

These assertions raise a natural question: which procedure more accurately reflects the voters' wishes – the CM or, say, the plurality ranking? A tool to analyze this issue is

the *reversal property* introduced in [Sect. 3.1, S1]. To motivate this condition, suppose all voters initially confused Anna with Mary; *neutrality* ensures that we can correct the outcome by interchanging their names in the election ranking. Similarly, suppose after the voters list their rankings of the candidates from top-ranked to bottom, it is discovered that the intended approach is to start with the bottom-ranked candidate. As all voters completely reversed their rankings of the candidates, it is reasonable to expect the election outcome to be similarly reversed. A procedure satisfying this natural condition for all profiles satisfies the *reversal property* [S1].

Theorem 1. *CM satisfies the reversal property*

Although the proof is trivial (once each voter reverses his ranking of the candidates, each pairwise outcome is similarly reversed – the conclusions now follows from Eq. 1.2), what makes this statement important is the surprising conclusion that *the only positional method satisfying this condition is the BC* [S1, 4]. This assertion suggests why the BC and CM rankings are related and that the CM ranking may more accurately reflect the will of the voters than, say, the plurality outcome. The effects of Thm. 1 are felt in Sect. 2 where we use the geometric approach from [S1] to determine all ways 3-candidate positional and CM election rankings can be related (with extensions to $n \geq 3$). It also plays a role in Sect. 3 where, following [S2-4], we characterize all CM ranking properties for $n \geq 3$.

To explain this second project, note that Eq. 1.3 lists $\mathbf{p}^*$'s CM ranking for each subset of two or more candidates. In general for $n \geq 3$ candidates, a profile $\mathbf{p}$ determines the CM election ranking for each of the $2^n - (n+1)$ subsets with two or more candidates. Following [S2-4], call this list of election rankings the *Copeland word defined by* $\mathbf{p}$. The *Copeland dictionary* $\mathcal{D}_C^n$ *is the set of words defined by all possible profiles.*

To appreciate the importance of the CM dictionary, recall that because a Condorcet winner must be CM top-ranked, if a CM word defines a Condorcet winner, then she is top-ranked in all subsets to which she belongs. More generally, any CM ranking relationship (such as the reversal property) restricts how the rankings of the subsets of candidates can differ, so a CM relationship is manifested by the constraints it imposes on the structure of CM words. The converse is more important: if we know all words in a CM dictionary, we also know their structures. In turn, this structure completely identifies all possible CM ranking relationships. As we describe how to characterize the CM dictionary, this approach also characterizes all possible CM ranking relationships.

2. Procedure lines and Copeland outcomes

We start with $n = 3$ candidates so we can geometrically depict all possible relationships among the positional election tallies and the CM ranking. [A positional method is defined by a *voting vector* $\mathbf{w}^3 = (w_1, w_2, w_3)$ where $w_j \geq w_{j+1}$ and $w_1 > w_3 = 0$; w_j points are assigned to a voter's jth ranked candidate and the candidates are ranked according to the total points they receive. Thus, the plurality, BC, antiplurality procedures are defined, respectively, by the voting vectors $(1, 0, 0)$, $(2, 1, 0)$, and $(1, 1, 0)$.] While our first two results assert that a non-BC positional ranking can differ as radically as desired from the CM ranking, the third introduces a new relationship connecting the BC and CM rankings.

Definition 1. Relative to a strict CM ranking, say $\mathcal{A} = c_1 \succ c_2 \succ c_3$, another ranking of these candidates is *BC related to CM-*$\mathcal{A}$ if it can be obtained from $\mathcal{A}$ by altering the ranking of either the top two (c_1, c_2) or the bottom two ranked candidates (c_2, c_3) (but

not both). The new ranking of the chosen pair can be any of the three ways it can be ordered. Should the initial ranking have a tie between two candidates, say $c_1 \succ c_2 \sim c_3$ or $c_1 \sim c_2 \succ c_3$, the BC related rankings are obtained by replacing the tied ranking with any way this pair can be ranked.

To illustrate with the CM ranking $A \succ B \succ C$, the ranking $A \succ C \succ B$ is BC related, but $B \succ C \succ A$ is not. Similarly, with the CM ranking $A \succ B \sim C$, a BC related ranking is $A \succ C \succ B$, but $A \sim B \succ C$ and $A \sim B \sim C$ are not.

Theorem 2. *a. For $n = 3$ candidates, consider a non-BC positional method. For any two rankings of the candidates a profile exists where the positional and CM rankings are, respectively, the first and second selected rankings.*

b. For any CM ranking, there is a supporting profile where the CM ranking disagrees with all positional voting rankings.

c. The Borda ranking is BC related to a CM ranking that is not $c_1 \sim c_2 \sim c_3$; there are no restrictions on the BC ranking with a CM ranking $c_1 \sim c_2 \sim c_3$.

This theorem ensures, for instance, the existence of a CM ranking $A \succ B \succ C$ although the plurality ranking is the reversed $C \succ B \succ A$. Even more, the CM ranking can be $A \succ B \succ C$ even though not a single positional method admits this ranking! While examples of this type highlight the severe conflict that can exist among procedures, they are highly unsatisfying because they provide no help for the critical issue of identifying which method best represents the voters' interests. This same defect remains even if an example is illustrated with a profile. After all, a specified profile is, at best, anecdotal data; it provides no information about the consequences of other supporting profiles or the typical situation. So, to compare procedures, rather than relying on special cases, we must find and examine all profiles that support contradictory outcomes. This is done below.

While the CM and plurality rankings can reverse one another, the actual situation is more intriguing. This is because part c requires the Borda ranking for any such profile to be BC related to the CM ranking. Even more; although the plurality and CM rankings are at distinct odds, we show next that this CM-BC connection forces over *half* of the remaining positional election rankings – including the antiplurality ranking and many others - to be BC related to the CM ranking. Namely, there exist fascinating, unexpected relationships connecting the CM with "most" positional rankings. All of them are found with the *procedure line* (introduced in [S1, 5]). For the reader's convenience, necessary terms and properties are sketched below (for details, see Sect. 2.4 of [S1]).

2.1. Representation triangle. The normalized election tally for the jth candidate, x_j, is her fractional part of the total vote; i.e., it is her actual tally divided by the sum of points received by all candidates. As the *normalized tally* (x_1, x_2, x_3) has $x_i \geq 0$ and $\sum_{i=1}^{3} x_i = 1$, it is a point on the simplex defined by the vertices $(1, 0, 0)$, $(0, 1, 0)$, $(0, 0, 1)$. Clearly, the election ranking associated with a point in this *representation triangle* is based on the size of each normalized component; e.g., the ranking assigned to $(\frac{1}{3}, \frac{1}{2}, \frac{1}{6})$ is $c_2 \succ c_1 \succ c_3$. This assignment process divides the representation triangle into *ranking regions* with a unique region for each linear ranking of the candidates. In Fig. 1, for instance, each of the six small triangles corresponds to a strict ranking; e.g., the triangle with a dot on the lower edge represents $c_2 \succ c_1 \succ c_3$ because a point in this triangle is closest to the c_2 vertex, next closest to the c_1 vertex, and farthest from the c_3 vertex. The barycenter $(\frac{1}{3}, \frac{1}{3}, \frac{1}{3})$ defines the complete tie $c_1 \sim c_2 \sim c_3$.

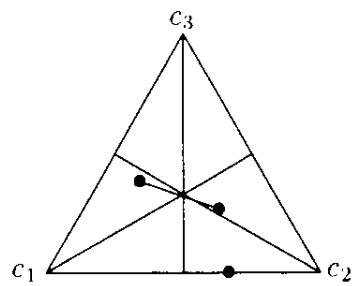


Figure 1. Comparison of the Copeland winner and the procedure line.

To connect elections with points in the representation triangle, use the fact that an equivalent form of $\mathbf{w}^3 = (w_1, w_2, 0)$ (where the election rankings remain the same) can be obtained by multiplying each w_j by a common positive scalar. A natural choice for each vector is $(1, t, 0)$ where $t = 0, \frac{1}{2}, 1$ represent, respectively, the plurality, BC, and antiplurality vote. But, as we want the $\mathbf{w}^3$-outcome for a single voter to be on the representation triangle, the sum of the components must equal unity. Therefore, the desired form is

$$\mathbf{w}^3_s = (1 - s, s, 0), \quad 0 \le s \le \frac{1}{2}, \tag{2.1}$$

where $s = \frac{t}{1+t}$. Thus, the plurality, BC, and antiplurality vectors are defined, respectively, by the values $s = 0, \frac{1}{3}, \frac{1}{2}$.

Any three-candidate, positional voting vector is equivalent to a vector represented by Eq. 2.1. To understand how the vectors are related, notice that the BC, defined by the $s = \frac{1}{3}$ vector $\mathbf{w}^3_{\frac{1}{3}} = (\frac{2}{3}, \frac{1}{3}, 0)$, requires the differences between the first and second, and second and third components to agree. As such, the BC divides the positional vectors into two sets, *the half on the plurality side* (those $\mathbf{w}^3_s$ where $0 \le s \le \frac{1}{3}$) and *the half on the antiplurality side* (where $\frac{1}{3} \le s \le \frac{1}{2}$.) Justification for the word "half" comes from Sect. 3.1 of [S1] and the natural normalization $(1, t, 0)$ where the BC is defined by $t = \frac{1}{2}$. ("Half" has nothing to do with the size of "s" intervals as the normalization distorts the length.) Namely, the BC divides the positional methods into procedures which place less (the plurality half) or greater (the antiplurality side) weight on who is a voter's second-ranked candidate. So, the normalized version of $(6, 1, 0)$ is on the plurality side, while $(4, 3, 0)$ is on the antiplurality side.

Next, replace integer profiles, which list the number of voters of each type, with *normalized profiles* which specify the fraction of all voters that are of each type. The advantage of tallying $\mathbf{w}^3_s$ elections with these fractions is that the outcome is the *normalized election tally* – a point on the representation triangle. To illustrate, the normalized version of profile $\mathbf{p}^*$ is obtained by dividing each integer by 27 – the total number of voters. By reading down the table, this defines the vector $\mathbf{p}^* = (\frac{8}{27}, \frac{3}{27}, \frac{1}{27}, \frac{9}{27}, \frac{3}{27}, \frac{3}{27})$ where the normalized plurality and antiplurality outcomes are, respectively, $(\frac{11}{27}, \frac{6}{27}, \frac{10}{27})$ and $(\frac{15}{54}, \frac{23}{54}, \frac{16}{54})$.

Finally, to compare positional and CM outcomes, choose the CM weights so that the sum of tallies equals unity. We leave it to the reader to show that this requires the CM weights $(\frac{2}{6}, \frac{1}{6}, 0)$ so that, with each candidate's two pairwise elections, a typical CM outcome has the value $(\frac{2}{3}, \frac{1}{3}, 0)$. The extreme CM tallies define points on the boundary of the representation triangle.

2.2. Procedure line. The *procedure line* (for a specified profile) is the straight line segment connecting the plurality and antiplurality normalized tallies. The procedure line for $\mathbf{p}^*$, then, is the line segment in Fig. 1 connecting the computed values of $(\frac{11}{27}, \frac{6}{27}, \frac{10}{27})$ and $(\frac{15}{54}, \frac{23}{54}, \frac{16}{54})$. The importance of this line is that it geometrically displays all positional election outcomes: e.g., the $\mathbf{w}_s^3$ outcome is the point on this line that is $2s$ of the segment length from the plurality endpoint. (This follows from Eq. 2.1.) For instance, because the BC is defined by $s = \frac{1}{3}$, the BC outcome is two-thirds of the way along the procedure line from the plurality endpoint. Reflecting its role as a dividing point, the BC outcome separates the procedure line into the line segment of outcomes from methods that place less weight on a voter's second-ranked candidate (the plurality half) and the segment with increased weight on this identity (the antiplurality side).

The advantages of the procedure line can be illustrated with Fig. 1. It follows from simple algebra that the $(s = \frac{1}{3}, t = \frac{1}{2})$ dividing BC tally is at $(\frac{26}{81}, \frac{29}{81}, \frac{26}{81})$ with ordinal ranking $c_2 \succ c_1 \sim c_3$ and that $s = \frac{4}{15}, t = \frac{4}{19}$, generates the $c_1 \succ c_2 \sim c_3$ ranking. As the dot on the right-hand bottom edge of the representation triangle is the CM outcome, it follows from the simple computation of the procedure line that the $\mathbf{p}^*$ outcomes for commonly mentioned positional methods

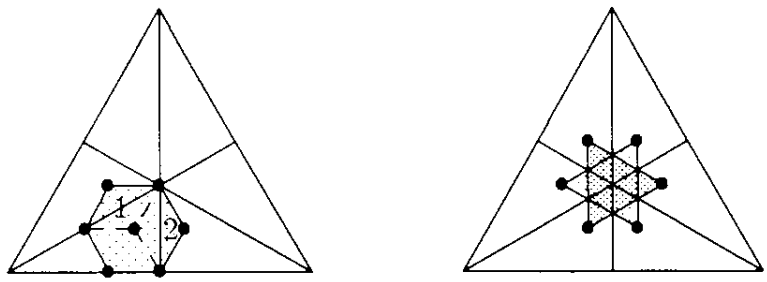
Method	Ranking	Method	Ranking
Plurality	$c_1 \succ c_3 \succ c_2$	BC	$c_2 \succ c_1 \sim c_3$
Antiplurality	$c_2 \succ c_3 \succ c_1$	CM	$c_2 \succ c_1 \succ c_3$

disagree with $\mathbf{p}^*$'s CM ranking. We learn much more: e.g., as a small portion of the procedure line is in the CM ranking region, a small fraction $(\frac{1}{2} - \frac{4}{19} < \frac{1}{3})$ of positional methods agree with the CM ranking. And, because this portion of the procedure line is on the plurality side, agreement (for $\mathbf{p}^*$) is achieved only with procedures that slightly devalue (from the BC) the identity of voters' middle-ranked candidates. In fact, the outcome of a procedure placing more than average weight on a voter's second ranked candidate (the antiplurality half) disagrees with the CM. From a more positive perspective, most $(1 - \frac{4}{19}$ or about 79%) of the rankings (in regions adjacent to the CM ranking region) are BC related to the $\mathbf{p}^*$-CM ranking.

The procedure line provides a powerful, but easily used tool to compare the CM and positional rankings. To use it, we first must find all of its admissible positions relative to a specified CM ranking. Start with a strict CM ranking where, without loss of generality, we use $c_1 \succ c_2 \succ c_3$. One way to determine the admissible procedure line positions is to find the BC scores that accompany a specified CM ranking. This is given by the shaded region of Fig. 2 minus the portion of the boundary that is outside of this ranking region: it is the convex hull of all normalized BC outcomes with a $c_1 \succ c_2 \succ c_3$ ranking along with the normalized BC outcomes $(\frac{1}{3}, \frac{1}{6}, \frac{1}{2})$ and $(\frac{1}{2}, \frac{1}{6}, \frac{1}{3})$. In other words, for any (rational) point in the shaded region there is a profile supporting the selected CM ranking and its BC normalized tally is the specified point! Observe that the two portions extending beyond the $c_1 \succ c_2 \succ c_3$ ranking region identify all conflict between the BC and CM rankings.

For CM tie votes, the region labelled "1" (including the boundaries of the solid and dashed lines but not the barycenter) are the admissible BC scores that accompany a CM ranking $c_1 \succ c_2 \sim c_3$. Similarly, region "2" (with boundaries but not the barycenter) are the BC scores that can accompany a CM ranking $c_1 \sim c_2 \succ c_3$. The BC rankings associated

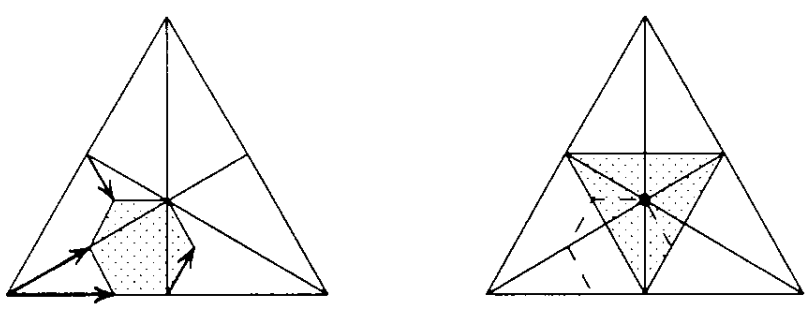
with a CM complete tie are in Fig. 2b. Observe that Thm. 2c follows immediately from Fig. 2. (The proof that this figure represents the BC outcomes is in Sect. 2.6.)



a. BC outcomes for Copeland $c_1 \succ c_2 \succ c_3$ **b.** BC outcomes for Copeland $c_1 \sim c_2 \sim c_3$

Figure 2. Comparing Borda and Copeland rankings.

2.3. Location of procedure line. An alternative way to construct the procedure line is first to connect the plurality and BC outcomes of a profile and then (according to the two-thirds rule) to extend this line segment past the BC score so that it is $\frac{3}{2}$ the original length. Thus, to find all procedure lines, it suffices to find all (normalized) plurality outcomes that can accompany each BC normalized outcome. This is given in Fig. 3 (which is explained in Sect. 2.6).



a. Procedure lines for boundary profiles. **b.** Procedure line for BC tie.

Figure 3. Procedure lines for Copeland $c_1 \succ c_2 \succ c_3$

In Fig. 3-a, the base and arrowhead tip of each vector are, respectively, the plurality and BC outcome. The full procedure line extends the vector so it is half again as long in the indicated direction. Thus, for outcomes on the edges of the representation triangle, the procedure line is easily determined. (They represent profiles where the conclusion is to be expected.) As the CM is intended to handle situations where elections outcomes are not immediate, Fig. 3-b is more important because it indicates what happens when the BC tally approaches a complete tie vote. For the $c_1 \sim c_2 \sim c_3$ BC ranking, the associated plurality outcome can be anywhere in the shaded equilateral triangle [S1: p. 195]. Thus (from continuity considerations), when the BC score approaches the ranking $c_1 \sim c_2 \sim c_3$, the admissible plurality outcomes can be essentially any point in the shaded triangle. (The actual choice is a slight distortion of this triangle; see Chap. 3 of [S1].)

What we learn from Fig. 3b is that as the BC tally approaches the complete tie ranking, there is enormous flexibility in choosing an admissible plurality score; in turn, this offers considerable freedom in choosing the accompanying procedure line. Just choose the plurality endpoint in the shaded region of Fig. 3b and the BC midpoint from the region enclosed by dashed lines but near the triangle barycenter. Then extend the vector so that it is half again as long. With this simple construction all sorts of new assertions can be discovered just by drawing lines according to these rules.

The power of this approach can be illustrated by completing the proof of Thm. 2. For assertion a, recall that $\mathbf{w}_s^3$ defines a specific point on the procedure line. Assume that the CM ranking is $c_1 \succ c_2 \succ c_3$ (or a modification due to a tie for some pair) and choose a ranking for $\mathbf{w}_s^3$. Place the procedure line so that the BC outcome is in the BC related region and the $\mathbf{w}_s^3$ outcome is in the specified ranking region. By choosing these points sufficiently close to the barycenter, this can be done while keeping the plurality end-point in the large shaded triangle. (This ensures the existence of a supporting profile.)

To prove Thm. 2-b, position the procedure line so that it doesn't enter the CM ranking region. This is easy to do; e.g., place the BC and plurality outcomes in the $c_2 \succ c_1 \succ c_3$ region so that the plurality outcome is closer to the $c_1 \sim c_2 \succ c_3$ boundary than the BC. The resulting slope forces the rest of the procedure line away from the $c_1 \succ c_2 \succ c_3$ region. Notice, had we wished, we could use the properties of the procedure line to specify all admissible rankings. Also notice that this line can be chosen so that the antiplurality endpoint is in a region where c_3 is top-ranked. This is part f of the following. An even easier construction applies when the CM ranking is not strict.

Other new conclusions are obtained by learning how the procedure line can cross the regions. Indeed, *all properties relating the CM and positional rankings are obtained in this manner.* Thus, while the following sample of new results are deep, important, and show the wide range between conflict and agreement of rankings, they are trivial to prove. We leave other assertions to the reader interested in experimenting with the procedure line. (By using the profile coordinates developed in [S1], profiles supporting each result are easy to find.)

Theorem 3. *a. For a profile defining a strict CM ranking, over half of the associated positional rankings are BC related.*

b. For x satisfying $0 < x < 1$, there are profiles where the CM ranking has a tie between two candidates, but x of the positional outcomes are not BC related. (So, an x value near unity means that almost all outcomes fail to be BC related.)

c. A CM complete tie imposes no restrictions on the positional rankings.

d. There exist profiles where all positional outcomes have the same normalized election tally, yet the (commonly defined) strict ordinal ranking differs from the strict CM ranking.

e. When the BC and CM rankings disagree, the rankings for over half of the positional methods disagree with the CM ranking.

f. There exist profiles where the top-ranked plurality, CM, and antiplurality winners are three different candidates.

Proof. Part f is proved above. Part c is obvious from Fig. 2. Part a follows from the fact that the BC, the midpoint of the procedure line, must be BC related to the CM strict ranking. From the properties of a straight line, half of the procedure line must be in a BC related region. The conclusion follows. Similarly, for part e, if the BC outcome is in a region different from the CM ranking, then so is some half of the procedure line.

A CM ranking from part b forces the associated BC outcome to be in region "1" or "2" of Fig. 2. Both regions share the barycenter of the representation triangle as a boundary point. Position the procedure line so that a very small portion of it (corresponding to $1-x$ of the positional methods) passes through the BC related regions. The geometry allows the line to be either arbitrarily close to the barycenter or far away, so, from continuity, the line can be positioned to realize the conclusion for any choice of x. While this construction proves that most positional outcomes need not be BC related, at least half of them cannot differ too radically (e.g., reversed).

For part d, choose the BC and plurality outcome to be the same point in the intersection of the shaded portion of the $c_2 \succ c_1 \succ c_3$ region and the dashed lines from Fig. 3b. This forces the procedure line to be this point; namely, all positional methods have the same normalized tally. □

A message of Thm. 3 is that the CM ranking can be at odds with those from commonly used procedures and that this is particularly true when the CM ranking is not a strict one. In other words, precisely when we need guidance from CM, doubt about its reliability is created. This situation becomes worse with more candidates.

Using different choices of lines, we find that some profiles allow procedures on the plurality side to agree with the CM, while other profiles allow CM agreement with the procedures on the antiplurality side – only the BC always admits regularity. This flip-flop behavior suggests that the reversal property plays an important role in these theorems. While true, more valuable than a mathematical explanation is to suggest how the reader can develop insight by experimenting with the procedure line. To start, first select a procedure line by choosing plurality and BC scores according to the rules. Whatever the choice, it follows from Sect. 3.1 of [S1] that there is a supporting profile $\mathbf{p}$ where the reversed profile, $\mathbf{p}^r$, gives the same plurality tally! (So, nothing changes in the plurality tally even after all voters reverse their rankings.) On the other hand, the $\mathbf{p}^r$ CM and BC outcomes *must* be reversed because they satisfy the reversal property. To find the new CM and BC $\mathbf{p}^r$ outcomes, draw a line from the $\mathbf{p}$ choice through the barycenter of the triangle. The associated $\mathbf{p}^r$ score is the exact opposite point on this line. Using this construction, the $\mathbf{p}^r$ procedure line is obtained by connecting the (common $\mathbf{p}$ and $\mathbf{p}^r$) plurality point with the new BC point. Notice how the new procedure line, with the new CM point, reverses the role of each half (relative to the BC) of the positional methods relative to the CM outcome. In words, if $\mathbf{p}$ demonstrates a setting where those procedures on the plurality side disagree with the CM outcome, then perhaps $\mathbf{p}^r$ demonstrates where the other class of procedures disagree with the CM. Only the BC admits balance.

2.4. More than three alternatives. The geometry remains essentially the same for any number of alternatives, but the higher dimensions makes convenient figures impossible. Of more importance, the results change because the above, reasonably comforting statements relating the CM and most positional rankings begin to fade once $n \geq 4$. This definitely occurs when we need the CM to handle the failings of the Condorcet approach. Before describing what happens, it is worth explaining why the three-candidate situation is misleading with respect to the general case.

By ignoring tie votes between pairs, there remain two kinds of CM rankings for $n = 3$: one is a strict ranking of the type $c_1 \succ c_2 \succ c_3$ where c_1 is both a CM and Condorcet winner and c_3 is both the CM and Condorcet loser (c_j is a Condorcet loser if she loses all pairwise elections). Second, when there is no Condorcet winner or loser, the pairwise

rankings create a cycle of the type $c_1 \succ c_2$, $c_2 \succ c_3$, $c_3 \succ c_1$. A cycle always defines the CM ranking $c_1 \sim c_2 \sim c_3$, but the positional method could be anything. (The two possible cycles are manifested by the two triangles in Fig. 2b.) So, where the Condorcet criteria does not apply, the CM evades resolving the issue by declaring a tie.

What relates the CM and BC rankings for $n = 3$, then, is that with only three pairs, the restricted combinatorics of Eq. 1.2 force the CM winner and loser to be, respectively, the Condorcet winner and loser. Completing the connection are the restrictions (for all $n \geq 3$) on how a Condorcet winner and/or loser is BC ranked. (The BC is the only positional method always admitting such relationships [S3].) But $n \geq 4$ candidates introduce so many more pairs that the CM winner no longer must be a Condorcet winner. Once this connection is severed, the general relationships connecting the CM and positional rankings also break. Thus, again, just when we need the CM, its outcomes become questionable.

Should a Condorcet winner and/or loser exist, however, a relationship weaker than, but similar to that for $n = 3$ connects the CM and positional outcomes. In describing this relationship, assume that the normalized voting vectors have the form $(1, w_2, w_3, \ldots, w_{n-1}, 0)$ where $w_j \geq w_{j+1}$ for $j = 2, \ldots, n-1$. Extending the $n = 3$ procedure line, all positional outcomes are in a $n-2$-dimensional *procedural hull* with the BC as the midpoint [S5]. Thus, whenever the BC and CM outcomes are related, the *central BC role forces the outcomes of a fraction of other positional methods to be related to the CM ranking.* The next theorem (proved in Sect. 3.3) specifies the relationship and the minimal fraction of positional outcomes that *must* satisfy it – a fraction that decreases in value with an increase in the number of candidates.

Theorem 4. *For $n \geq 3$ candidates, let $r_n = (\frac{n-2}{n-1})^{n-2}$.*

a. If profile **p** *defines a Condorcet winner, c_1, then the rankings for at least r_n of the positional methods have c_1 ranked strictly above bottom place. Similarly, if* **p** *defines a Condorcet loser, c_n, then at least r_n of the positional methods have c_n ranked strictly below top place. If* **p** *defines both a Condorcet winner and loser, then at least r_n of the positional methods have c_1 ranked strictly above c_n.*

b. The BC is one of the positional methods satisfying part a. For a specified non-BC positional method and any two rankings of the n candidates, there exists a profile where these rankings are, respectively, the CM ranking and the positional ranking.

While not proved here, it follows from [S3] that an admissible BC ranking is any ranking satisfying the conditions of part a. In fact, to connect Thms. 2, 3, 4, observe that our definition of "BC related to an $n = 3$ CM strict ranking" requires the Condorcet winner to be strictly ranked above the Condorcet loser. With this Condorcet restriction (that always is satisfied for $n = 3$), the main difference between $n = 3$ and $n > 3$ is that a smaller fraction of the positional methods are guaranteed to be related to the CM ranking. This result is tight; using higher dimensional methods similar to that employed to prove Thm. 3b, it can be shown that there are profiles that come as close as desired to satisfying these r_n values. (Of course, there also are profiles where all positional rankings agree with the CM ranking.) Because $r_n > r_{n+1}$ and $r_n = (\frac{n-2}{n-1})^{n-2} \to e^{-1}$ as $n \to \infty$, there is a minimal guarantee of compatibility. The real problem is that the Condorcet restriction becomes increasingly unlikely to be satisfied with larger values of n.

So, when are the CM, BC, and other positional rankings related? To analyze this problem it is convenient to use the $(1, 0, -1)$ CM scores for each pairwise election (because the sum of CM points equals zero). With this choice, we exploit the fact that the extreme

CM scores for n candidates, $n-1$ and $-(n-1)$, require, respectively, a Condorcet winner and loser. Even when tie votes are included, a candidate who never loses a pairwise contest and wins at least one is not BC bottom-ranked [S3]. Thus, *a candidate with the CM score of $n-1$ or $n-2$ never can be BC bottom-ranked*, so the positive aspects of Thm. 4 continue to apply. (Conversely, a CM score of $-(n-1)$ or $-(n-2)$ precludes a candidate from being BC top-ranked.) While Thm. 5a appears to offer a more general assertion, it is a restricted case of this observation.

Another way to explain why a Condorcet winner receives preferential BC treatment is to observe that it defines an extreme setting where the pairwise rankings for one set of candidates (the Condorcet winner) dominate all other candidates. This dominated behavior always exists with a CM strict ranking for $n \leq 4$ candidates because there are at most six pairwise rankings. Consequently, the restricted combinatorics of Eq. 1.2 – the number of equations and unknowns – forces a Condorcet winner and loser always to emerge with a strict CM ranking. But, $n \geq 5$ candidates create so many more pairwise elections that a specified CM ranking cannot severely constrain their outcomes. Think of this in terms of the number of equations and unknowns: with $n \geq 5$, there are so many more "unknowns" that the number of equations (a specified CM ranking) cannot determine their values. Thus, we cannot expect the CM and positional ranking connections to continue.

To explore this issue, accept the fact that when the pairwise rankings of one set of candidates dominate all others, the CM and BC rankings are related. (This can be proved using techniques of [S3] and Sec. 3.) To see this with a $n = 6$ example, consider the pairwise rankings that define the two cycles $c_1 \to c_2 \to c_3 \to c_1$ (denoting $c_1 \succ c_2$, $c_2 \succ c_3$, $c_3 \succ c_1$) and $c_4 \to c_5 \to c_6 \to c_4$ where every candidate from the first cycle beats every candidate from the second. Here, the admissible BC rankings are related to the CM ranking $c_1 \sim c_2 \sim c_3 \succ c_4 \sim c_5 \sim c_6$ in the sense that "on the average" the candidates from the first group are BC higher ranked than those from the second. For instance, not all candidates from the second group can be BC ranked above all candidates from the first, or if a candidate from the second group is BC top-ranked (which can occur!), then another candidate from this group must be BC bottom-ranked. However, these pairwise rankings are not the only way to obtain the CM ranking. Some of these choices of supporting pairwise rankings prohibit a set of candidates from dominating all others, so (as we indicate next) no restrictions can be imposed on the associated BC rankings.

Thus, when $n \geq 5$, no longer can we expect the CM rankings to be related with positional rankings. This is because a specified CM ranking cannot ensure that the associated pairwise rankings permit one set of candidates to dominate all others. To define "non-dominated settings" we require the pairwise rankings to admit a cycle involving all candidates. (For the $n = 6$ example, this occurs should even one candidate from the second group beat a candidate from the first. A cycle always is defined for $n = 4, 5$ should each candidate win and lose at least one election and there are no ties.) With a cycle, obvious arguments about how one set of candidates dominates another fail, so we need guidance in how to make a choice. But, as the second and third parts of the next assertion emphasize, it is not clear whether the CM provides reliable help. Of more importance, we observe the wilting of CM-BC connections which, in turn, destroys connections between the CM and positional methods.

Theorem 5. *a. For a strict CM ranking for $n = 3, 4$, or for $n \geq 5$ where there are no pairwise tie votes, the CM winner (loser) is a Condorcet winner (loser). Thus, the BC*

ranking and the rankings of at least r_n of the positional methods have the CM winner ranked strictly above the CM loser.

b. For $n \geq 3$, choose pairwise rankings that admit a cycle and a specified CM ranking. For any ranking of the candidates there exists profiles of voters supporting all specified pairwise election rankings (hence, the related CM ranking) but where the BC outcome is the selected ranking.

c. For $n \geq 5$, choose two rankings of the candidates. There exist profiles where these rankings are, respectively, the CM and the BC rankings.

Part c states that once $n \geq 5$, the CM and BC outcomes can be as different as desired. As this assertion proves that the CM ranking cannot restrict the normalized BC score to be in certain ranking regions, it also follows that we have lost all restrictions on the location of the procedure hull. This permits the CM and positional rankings to differ as radically as desired. The above assertions explain what can happen for all n except $n = 4$ which we leave to the reader as an easy exercise. (As a hint, characterize the CM rankings that do and do not, allow a cycle. Also, see Sect. 3.)

Proof. For $n \geq 3$ and no pairwise ties, the admissible CM scores (when we use $(1, 0, -1)$) are $n-1, n-3, \ldots, -(n-3), -(n-1)$. With precisely n possible scores, a strict CM ranking achieved without the benefit of ties must use each value. Because this requires a Condorcet winner and loser, the assertion follows.

The case of $n = 3$ already is handled. For $n = 4$ where pairwise ties are admitted, the CM scores are $3, 2, 1, 0, -1, -2, -3$ where the sum of the scores for the four candidates equals zero. If the CM strict ranking assigns one point to the top-ranked candidate, then the CM bottom-ranked candidate must receive -1 points: this forces the remaining two candidates to be tied with zero points each. As this is not a strict ranking, the top-ranked candidate must receive at least two points and the bottom-ranked one receives no more than -2 points. The conclusion follows.

The proof of part b is in Sect. 3.3, but we use it now to prove part c. The main complication with $n = 4$ is that, to avoid settings where one candidate dominates the others, we are restricted to fewer scores $(1, 0, -1)$ than candidates; this situation reverses when $n \geq 5$. For $n \geq 5$, start with a cycle; these pairwise rankings contribute a CM score of zero to all candidates so the admissible CM scores, which range from $n-3$ to $-(n-3)$, contain at least n values only if $n \geq 5$. From Eq. 1.2, the CM scores are computed from $n-1$ independent equations (one equation is dropped because the CM scores add to zero) with $\binom{n}{2} - n$ independent variables (the remaining pairwise rankings after the cycle). Notice that the number of independent variables is greater than the number of equations iff $n \geq 5$. Thus, from simple algebra, these equations can be "solved;" i.e., any specified values of the equations admits solutions. The remaining step just involves verifying that integer solutions are possible. As this can be done by substitution, we skip the tedious but straight-forward details. However, we include one such computation because it illustrates the "tightest" case of a CM strict ranking $c_1 \succ c_2 \succ \cdots \succ c_5$ for $n = 5$ with the cycle $c_1 \to c_2 \to \cdots \to c_5 \to c_1$. Here, the remaining five pairwise rankings has c_1 beating all other candidates (other than c_5), c_5 losing to all candidates (other than c_1), and $c_2 \sim c_4$. □

2.5. Profiles of conflict. Now that we have established a central relationship connecting CM and BC rankings, we turn to the critical task of understanding which procedure to trust when a conflict occurs. (The spirit of these results hold for all $n \geq 3$, but to minimize technicalities, we only describe what happens for $n = 3$.) To start, if the voters are evenly

split where five million have the ranking $A \succ B \succ C$ and another five million have the ranking $B \succ C \succ A$, it is arguable that B is the voters' choice. After all, C should be bottom-ranked because nobody ranks her above the middle, and half have her bottom-ranked. A does slightly better by being top- and bottom-ranked by different halves of this population. B, on the other hand, enjoys the enviable status of being the only candidate that never is bottom-ranked while half of the voters have her top-ranked. This natural group ranking $B \succ A \succ C$ is supported by the CM, BC, and all positional methods except the plurality vote (with the outcome $A \sim B \succ C$). In fact, this $B \succ A \succ C$ conclusion is so natural and robust that we should expect it to be preserved even after varying this profile by adding ten thousand or so voters (which is only a 0.1% change). However, just by adding one more voter of the first type, A becomes the Condorcet, CM, and plurality winner. Clearly, at least for this profile, these procedures seriously violate the intent of the voters. (As the procedure line proves, almost all other positional methods, including the BC, have the expected conclusion $B \succ A \succ C$.)

The main result of this section, Thm. 6, asserts that whenever the CM and BC rankings differ, it is because a crucial portion of the supporting profile is of the above type. Therefore, rather than serving as an isolated example, this profile identifies the source of conflict between BC and CM rankings. As this is a setting where the CM ranking fails to reflect the voters' true views, it is arguable that the profiles identified in Thm. 6 – all possible profiles allowing a conflict between the CM and BC outcomes – support the BC ranking rather than the CM ranking (or Condorcet winner). Conversely, this profile set identifies a distinct fault of the CM and Condorcet winners. (Indeed, the arguments in [S1] explaining why the Condorcet winner is a flawed solution concept also apply to the CM.)

These results require describing the space of profiles. (This discussion is technical, so, on a first reading, a reader may prefer to tentatively accept the assertion and skip ahead to Sect. 3.) As a way to introduce notation, we prove that Fig. 2 represents the BC outcomes.

Proof that Fig. 2 represents all BC outcomes. From the geometric approach of Chaps. 2, 3 of [S1], we know that the normalized profiles defining the CM ranking $c_1 \succ c_2 \succ c_3$ is a convex set and that the mapping converting profiles into BC normalized scores is linear. Therefore, after finding the profile vertices of the convex set of supporting profiles (Sect. 2.5 of [S1]), the BC image is the convex hull defined by the BC outcomes of these profile vertices. To describe the profiles in a vector format, assume that $\mathbf{p} = (p_1, \ldots, p_6)$ is a profile where p_j is the fraction of all voters of the jth type as defined next.

Type	Ranking	Type	Ranking
1	$c_1 \succ c_2 \succ c_3$	4	$c_3 \succ c_2 \succ c_1$
2	$c_1 \succ c_3 \succ c_2$	5	$c_2 \succ c_3 \succ c_1$
3	$c_3 \succ c_1 \succ c_2$	6	$c_2 \succ c_1 \succ c_3$

Thus profile $(\frac{1}{3}, 0, \frac{1}{6}, 0, \frac{1}{2}, 0)$ has $\frac{1}{6}$ of the voters with a type-one ranking $c_1 \succ c_2 \succ c_3$, $\frac{1}{6}$ with type-three preference $c_3 \succ c_1 \succ c_2$, while the rest have type-five $c_2 \succ c_3 \succ c_1$. Let $\mathbf{E}_j$ be the unanimity profile where all voters are of the jth type. (So, $\mathbf{E}_j$ is the vector with unity in the jth component and zero in all others.) The next table identifies all profile vertices with the associated normalized BC score – except those which define a complete tie

among all pairwise votes (that are not needed because their CM and BC score is $(\frac{1}{3}, \frac{1}{3}, \frac{1}{3})$).

Profile	BC	Profile	BC
$\mathbf{E}_1$	$(\frac{2}{3}, \frac{1}{3}, 0)$		
$\frac{1}{2}(\mathbf{E}_1 + \mathbf{E}_5)$	$(\frac{1}{3}, \frac{1}{2}, \frac{1}{6})$	$\frac{1}{2}(\mathbf{E}_1 + \mathbf{E}_6)$	$(\frac{1}{2}, \frac{1}{2}, 0)$
$\frac{1}{2}(\mathbf{E}_2 + \mathbf{E}_6)$	$(\frac{1}{2}, \frac{1}{3}, \frac{1}{6})$	$\frac{1}{2}(\mathbf{E}_2 + \mathbf{E}_5)$	$(\frac{1}{3}, \frac{1}{3}, \frac{1}{3})$
$\frac{1}{2}(\mathbf{E}_1 + \mathbf{E}_3)$	$(\frac{1}{2}, \frac{1}{6}, \frac{1}{3})$	$\frac{1}{2}(\mathbf{E}_1 + \mathbf{E}_2)$	$(\frac{2}{3}, \frac{1}{6}, \frac{1}{6})$

Figure 2a is the convex hull of these BC outcomes in the representation triangle. Using the same approach for the CM ranking $c_1 \succ c_2 \sim c_3$, the profile vertices are the last two rows of the table. For $c_1 \sim c_2 \succ c_3$, the profile vertices are those in the second and third rows.

The profile set defining a CM complete tie include the two possible cycles. The profiles supporting each cycle is convex, so the profiles defining a CM complete tie is a nonconvex union of two convex sets. The profile vertices, then, are the profiles leading to tie votes for each pair and the vertices in the following table. The first column are the boundary vertices for the cycle $c_1 \succ c_2$, $c_2 \succ c_3$, $c_3 \succ c_1$, while the second column corresponds to the remaining cycle.

$\frac{1}{2}(\mathbf{E}_1 + \mathbf{E}_3)$	$(\frac{1}{2}, \frac{1}{6}, \frac{1}{3})$	$\frac{1}{2}(\mathbf{E}_2 + \mathbf{E}_4)$	$(\frac{1}{3}, \frac{1}{6}, \frac{1}{2})$
$\frac{1}{2}(\mathbf{E}_3 + \mathbf{E}_5)$	$(\frac{1}{6}, \frac{1}{3}, \frac{1}{2})$	$\frac{1}{2}(\mathbf{E}_4 + \mathbf{E}_6)$	$(\frac{1}{6}, \frac{1}{2}, \frac{1}{3})$
$\frac{1}{2}(\mathbf{E}_1 + \mathbf{E}_5)$	$(\frac{1}{3}, \frac{1}{2}, \frac{1}{6})$	$\frac{1}{2}(\mathbf{E}_2 + \mathbf{E}_6)$	$(\frac{1}{2}, \frac{1}{3}, \frac{1}{6})$

Although we now can find all ways the CM ranking can differ from the positional rankings for $n = 3$ candidates, this does not help us understand which procedure best reflects the true wishes of the voters. These issues require us to identify all profiles with differing outcomes. Clearly (from Thm. 2-a and by experimenting with the procedure line), analyzing where the CM ranking differs from a specified non-BC positional ranking is not useful. Instead, according to Thm. 3, a more valuable exercise is to understand the profile set causing conflict between the CM ranking and over half of the positional methods. This is what we analyze.

The procedure line is centered on the BC outcome, so the sought after profiles are where the CM and BC rankings disagree. (See Thm. 3-e.) It follows from the above tables and figures that this profile set is the union of two convex sets – there is one set for each extension of BC outcomes beyond the $c_1 \succ c_2 \succ c_3$ region. (See Fig. 2.) Both sets contain the profiles leading to a tie for all pairs. The remaining profiles are found by identifying those BC outcomes on the boundary of the extensions of the BC outcomes (outside of the $c_1 \succ c_2 \succ c_3$ region) that come from profile vertices.

The way these profile vertices are found for the extended region to the right is illustrated in Fig. 4. The dashed lines in this figure connect extreme BC outcomes that cross both the shaded region and the boundary of the extended region: that is, they cross the vertical line corresponding to $c_1 \sim c_2 \succ c_3$. (Those lines that fail these conditions are omitted.) The same linear combination describing the connection between BC outcomes defines the combination of profiles.

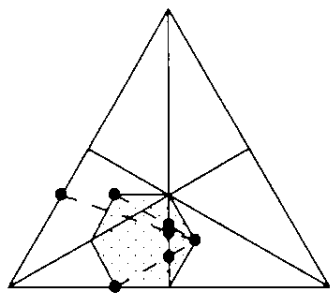


Figure 4. Determining the profile verices.

To illustrate with an example using the lowest dashed line in Fig. 4, because the line connects the extreme BC outcome $(\frac{2}{3}, \frac{1}{3}, 0)$ (the point on the lower left-hand edge) with $(\frac{1}{3}, \frac{1}{2}, \frac{1}{6})$, its equation is $t(\frac{2}{3}, \frac{1}{3}, 0) + (1-t)(\frac{1}{3}, \frac{1}{2}, \frac{1}{6}) = (\frac{1+t}{3}, \frac{1}{2} - \frac{t}{6}, \frac{1-t}{6})$. This line crosses the vertical line (a boundary for this region) where c_1 and c_2 are tied so the BC score is of the form $(x, x, -2x)$ for some value of x. Here, the first and second components of equation for the line agree, so $\frac{1+t}{3} = \frac{1}{2} - \frac{t}{6}$ or $t = \frac{1}{3}$. As the profile defining this outcome satisfies the same linear equation, once the BC outcomes are replaced with the supporting profile we have the profile vertex $\frac{1}{3}\mathbf{E}_1 + \frac{2}{3}(\frac{1}{2}(\mathbf{E}_1 + \mathbf{E}_5))$ which is $\mathbf{v}_3$ in the following table. All other points are found in the same way.

Thus, other than the profiles defining three pairwise ties, the vertex profiles allowing the BC outcome $c_2 \succ c_1 \succ c_3$ with the CM ranking $c_1 \succ c_2 \succ c_3$ are

$$\boxed{\begin{array}{ll} \mathbf{v}_1 = \frac{1}{2}(\mathbf{E}_1 + \mathbf{E}_5) & \mathbf{v}_2 = \frac{1}{2}(\mathbf{E}_1 + \mathbf{E}_6) \\ \mathbf{v}_3 = \frac{2}{3}\mathbf{E}_1 + \frac{1}{3}\mathbf{E}_5 & \mathbf{v}_4 = \frac{1}{5}\mathbf{E}_2 + \frac{2}{5}(\mathbf{E}_1 + \mathbf{E}_5) \\ \mathbf{v}_5 = \frac{1}{2}\mathbf{E}_1 + \frac{1}{6}\mathbf{E}_3 + \frac{1}{3}\mathbf{E}_5 & \end{array}}. \tag{2.2}$$

The profile set supporting this conflict is the convex hull of these vertex profiles; each profile is some mixture of these profile vertices. The effects of various combinations of these profiles can be determined immediately from the geometry and the profile vertices. An important conclusion, evident from Fig. 4, is that the more the BC and CM score differ (that is, the more the BC outcome is in the shaded extension), the more dominant the $\mathbf{v}_1 = \frac{1}{2}(\mathbf{E}_1 + \mathbf{E}_5)$ portion of the profile. (This is the only portion of the profile that can pull the BC outcome into this extended region.) In fact, if the profile does not have a term of this type, then the BC outcome cannot be in the shaded extension. Of importance for understanding the conflict between CM and BC rankings, notice that by changing the c_1, c_2, c_3 names, respectively, to A, B, C, $\mathbf{v}_1$ is the normalized form of the example profile starting this section. In words, whenever there is a conflict between the CM and the BC outcomes (and, hence, over half of the positional outcomes), it is due to preference behavior of the voters of the type introduced at the start of this subsection.

The other possible CM-BC conflict is where the BC ranking is $c_1 \succ c_3 \succ c_2$. Using the same approach, in addition to those profiles leading to a complete tie vote, the remaining profile vertices are

$$\boxed{\begin{array}{ll} \mathbf{v}_6 = \frac{1}{2}(\mathbf{E}_1 + \mathbf{E}_3) & \mathbf{v}_7 = \frac{1}{2}(\mathbf{E}_1 + \mathbf{E}_2) \\ \mathbf{v}_8 = \frac{2}{3}\mathbf{E}_1 + \frac{1}{3}\mathbf{E}_3 & \mathbf{v}_9 = \frac{1}{5}\mathbf{E}_6 + \frac{2}{5}(\mathbf{E}_1 + \mathbf{E}_6) \\ \mathbf{v}_{10} = \frac{1}{2}\mathbf{E}_1 + \frac{1}{3}\mathbf{E}_3 + \frac{1}{6}\mathbf{E}_5 & \end{array}}.$$

Again, the profile set is the convex hull of these vertices. Again, the deeper the BC normalized score is in the shaded extension, the more dominant the $\mathbf{v}_6 = \frac{1}{2}(\mathbf{E}_1 + \mathbf{E}_3)$ portion of this profile. Again, by the name change c_1, c_2, c_3 to B, C, A, $\mathbf{v}_6$ becomes a normalized form of the example profile: this supports the assertions of the introductory paragraphs of this section. The following summarizes these computations.

Theorem 6. *Consider the set of profiles supporting the CM ranking $c_1 \succ c_2 \succ c_3$. A necessary and sufficient condition for a normalized profile $\mathbf{p}$ to yield this CM ranking and the BC ranking $c_2 \succ c_1 \succ c_3$ is that the profile can be expressed as*

$$\mathbf{p} = t_0\mathbf{p}_0 + \sum_{j=1}^{5} t_j\mathbf{v}_j, \quad \sum_{j=0}^{5} t_j = 1,\ t_j \geq 0.$$

In this expression $t_1 > 0$ and $\mathbf{p}_0$ is any profile where all pairwise votes are tied.

Similarly, a necessary and sufficient condition for profile $\mathbf{p}^*$ to have the CM ranking $c_1 \succ c_2 \succ c_3$ while the BC ranking is $c_1 \succ c_3 \succ c_2$ is if

$$\mathbf{p}^* = \lambda_0\mathbf{p}_0 + \sum_{j=6}^{10} \lambda_j\mathbf{v}_j, \qquad \lambda_0 + \sum_{j=6}^{10} \lambda_j = 1,\ \lambda_j \geq 0$$

where $\lambda_6 > 0$.

3. Copeland and Borda outcomes

To continue our description of the BC-CM relationship, we make precise their similarities and differences. In doing so, it turns out to be more convenient to use the CM scoring values of $(1, 0, -1)$. To start with a similarity, let $n_{j,k}$ be c_j's vote in a pairwise election with c_k. As Borda knew (see [D, S1]), the BC score for c_j is $B(j) = \sum_{k \neq j} n_{j,k}$ where the BC ranking is based on these scores. Alternatively, use $x_{j,k} = \frac{n_{j,k} - n_{k,j}}{n_{j,k} + n_{k,j}}$ where $x_{j,k}$ measures c_j's margin of victory or defeat over c_k. So, $-1 \leq x_{j,k} = -x_{k,j} \leq 1$ where $x_{j,k} > 0$ iff $c_j \succ c_k$ and $x_{j,k} = 0$ iff $c_j \sim c_k$ and the ± 1 values represent a candidate's unanimous vote. The candidates are BC ranked according to their scores

$$b(j) = \sum_{k \neq j} x_{j,k}. \tag{3.1}$$

By comparing Eq. 3.1 with Eqs. 1.1, 1.2 (after replacing the $s_{j,k}$ values with $1, 0, -1$), it becomes obvious why the CM and BC are related and why both satisfy the reversal property: both sum information about pairwise votes. The difference is that BC uses the precise $x_{j,k}$ value, while CM retains only the sign of this term. So, to appreciate differences between the procedures, we only need to understand the implications of CM dropping all information about the precise pairwise tally.

For $n = 3$, all possible pairwise scores can be represented as points in the *representation cube* (Sect. 2.5 of [S1]): this is the shaded region of Fig. 5. (The dots are explained below.) Each $(x_{1,2}, x_{2,3}, x_{3,1})$ score defines a point in the three-dimensional space: e.g., the center dot in the front square is point $(1, 0, 0)$ which represents a unanimous $c_1 \succ c_2$ vote along with $c_2 \sim c_3$, $c_3 \sim c_1$ tie votes.

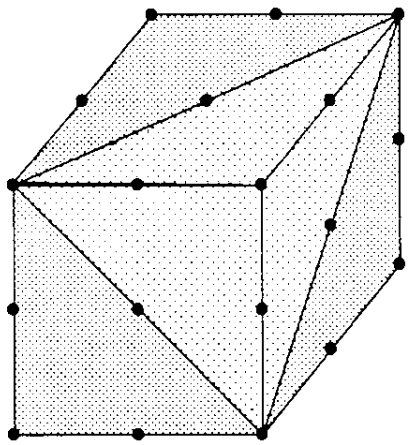


Figure 5. Pairwise vs. Copeland scores.

To explain the representation cube, start with a standard cube where each score can range over all $[-1, 1]$ values. Not all scores in this cube come from a profile: e.g., $(1, 1, 1)$ is impossible as it requires *all* voters to prefer $c_1 \succ c_2$, $c_2 \succ c_3$, $c_3 \succ c_1$, so all voters are irrational with cyclic, nontransitive preferences. To eliminate these perversities the slanted sides of the representation cube (given by $-1 \leq x_{1,2} + x_{2,3} + x_{3,1} \leq 1$) restrict the pairwise outcomes to those supported by transitive preferences (Sect. 2.5, [S1]). In particular, the slanted sides eliminate the troublesome vertex $(1, 1, 1)$. The six vertices of the representation cube represent outcomes for the six unanimity profiles. (To include voters without transitive preferences, the set of pairwise outcomes is the original cube.)

The difference between the BC and CM is that the BC uses the actual $x_{i,j}$ tally, while CM considers only its sign. Because CM treats a tally as though the voters unanimously agreed on the ranking of each pair – either as a strict preference or indifference – we can replace the actual profile with a related "unanimity" one of this type. This is because the CM does not distinguish between the original and the artificial unanimity profile.

For $n = 3$, the 27 artificial unanimity profiles are indicated by the dots in Fig. 5. Eight of them (four are visible), however, are in regions reserved for pairwise outcomes from nontransitive voters. (For larger n values, the nontransitive dots quickly dominate.) Indeed, one dot is the troubling $(1, 1, 1)$ vertex requiring all voters to have the same cyclic preferences. In other words, part of the CM problems arise because the CM cannot distinguish between whether the voters have transitive or cyclic preferences. But, this should be expected: as we know (Sects. 2.5, 4.4 of [S1]), the real difficulty with ordinal pairwise methods (and IIA from Arrow's Theorem) is that they vitiate the critical assumption that voters have transitive preferences.

To summarize, for $n \geq 3$ candidates, *the CM score is a BC score restricted to unanimity profiles but without assumptions about transitive preferences.* The effects of this restriction are displayed in Fig. 6 where the shaded region corresponds to all BC normalized tallies and the dots are the CM tallies. (The interior dots are the unlikely scores where only one pairwise vote is not tied.) This figure makes it clear that all subtle BC relationships resulting from close elections (so, the BC outcome is near the barycenter of the triangle) are lost by CM; again, these are the exact situations the CM was designed to handle.

3.1 Copeland relationships. Recall that the CM word defined by a profile $\mathbf{p}$ is the listing of the $\mathbf{p}$ CM-rankings for each subset of two or more candidates, and the CM dictionary

$\mathcal{D}_C^n$ is the collection of CM words for all possible profiles. We now describe $\mathcal{D}_C^n$ for $n \geq 3$.

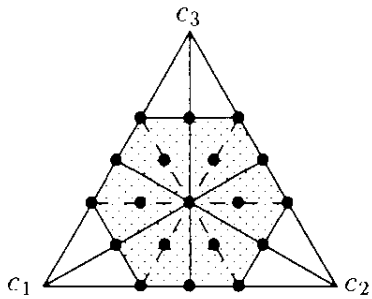


Figure 6. Comparison of Copeland and BC outcomes

The Borda dictionary $\mathcal{D}_B^n$ is characterized [S3] in terms of simple computations. This simplifies the derivation of the CM dictionary because a CM score is a BC score restricted to the artificial unanimity profiles. Therefore, the CM dictionary is found by applying the techniques developed in [S3] to these unanimity profiles. Before introducing the technical conditions, the following theorem provides a flavor of the new results.

Theorem 7. *Assume there are $n \geq 3$ candidates.*

a. If c_1 is CM top-ranked (bottom-ranked) in all $(n-1)$-candidate subsets, then she is CM top-ranked (bottom-ranked) in the full set of n candidates.

b. If the CM ranking has $c_j \succ c_k$ for each $(n-1)$-candidate subset, then the same CM relationship holds for the n-candidate set.

c. The sum of CM points assigned to a candidate over all k-candidate subsets equals $\binom{n-2}{k-2}$ times the points assigned to her in the full n-candidate set. (Recall, $\binom{m}{j} = \frac{m!}{j!(m-j)!}$.)

d. For $n \geq 3$, the set inclusion $\mathcal{D}_C^n \subsetneq \mathcal{D}_B^n$ is satisfied.

e. For $n = 4$, denote the four three-candidate subsets by $S_7 = \{c_1, c_2, c_3\}$, $S_8 = \{c_1, c_2, c_4\}$, $S_9 = \{c_1, c_3, c_4\}$, $S_{10} = \{c_2, c_3, c_4\}$ and c_js CM score in S_k by $C^k(j)$. The CM scores, using the $(1, 0, -1)$ weights, are related by

$$C^7(1) + C^8(2) + C^9(3) + C^{10}(4) = 0$$
$$C^7(2) + C^8(1) + C^9(4) + C^{10}(3) = 0 \tag{3.2}$$

All CM relationships restricted to these four sets are derived by these equations.

So, from a, b, c, we learn that a candidate who fares well in CM elections of subsets of candidates tends to do well in the full set. (This well behaved nature of the CM should be expected from the "unanimity" character of the CM.) However, while it is trivial to prove that a Condorcet winner is CM top-ranked, it is not obvious that a CM top-ranked candidate in all four-candidate subsets must be CM top-ranked in the five-candidate set. Therefore, assertion a generalizes the Condorcet relationship from its restricted setting of $k = 2$ to include $k = n - 1$: b is similar generalization. These statements need not be true for other k values.

Statement c generalizes the computation of a candidate's CM score from the sum of her CM pairwise scores to where it can be expressed as the sum of her CM scores from k-candidate subsets. Assertions a, b are immediate consequences of this new expression. (Even more, with this equation we can extend a and b to show what else can happen for $k \neq 2, n-1$, etc.) In fact, as shown below, all CM relationships are determined by how

a candidate's CM score over certain subsets of candidates uniquely determines her CM score for another subset. Consequently, it becomes important to characterize the family of subsets for which this is true. This is given by Thm. 8.

Definition 2 [S6]. A family of subsets $\mathcal{F} = \{S_\alpha\}$, where at least one subset has more than two candidates, has *the cyclic containment property* if it satisfies the following condition. Each triplet of candidates $\{c_i, c_j, c_k\}$ from $U = \cup_\alpha S_\alpha$, the set of all candidates from $\mathcal{F}$, defines three pairs. For each pair, there exits S_α from $\mathcal{F}$ that contains this pair, but not the third candidate.

Theorem 8. *A family of subsets $\mathcal{F} = \{S_\alpha\}$ admits a CM relationship where each candidate's CM tally over the subsets is governed by an equation iff $\mathcal{F}$ has the cyclic containment property.*

To illustrate with $n \geq 4$, as the family $\mathcal{F}$ of all three-candidate subsets satisfies the cyclic containment property, there exist equations relating the CM tallies over these subsets. The exact equations for $n = 4$, given in Eq. 3.2, tightly control the CM scores for different candidates over the sets. One consequence, for instance, is that it is impossible for c_1, c_2, c_3, c_4 to be top-ranked, respectively, in S_7, S_8, S_9, S_{10} (because the first equation would have the sum of positive numbers equal to zero.) While similar expressions are easy to derive, we show how this relationship tells much more: e.g., it can be used to find all CM relationships among the three sets $\mathcal{F}_1 = \{S_7, S_8, S_9\}$. Even though $\mathcal{F}_1$ fails the cyclic containment property, as it is in a family ($\mathcal{F}$) satisfying these conditions, the Eq. 3.2 equalities become inequalities of the form $-2 \leq -C^{10}(4) = C^7(1) + C^8(2) + C^9(3) = -C^{10}(4) \leq 2$, that govern all $\mathcal{F}_1$-CM relationships. To illustrate with CM rankings for S_7 and S_8 of $c_1 \succ c_2 \succ c_3$ and $c_1 \succ c_2 \succ c_4$, if we ignore the unlikely situation of a pairwise tie vote, these inequalities require c_1 to be top-ranked in S^9. Namely, this simple exercise determines the new CM relationship that, with no pairwise ties, a candidate who is top-ranked in two of the three subsets of three candidates is CM top-ranked in the last set. Many other new relationships are found in this same manner.

To further illustrate, the family $\mathcal{F}_2 = \{\{c_1, c_2\}, \{c_1, c_2, c_3\}, \ldots, \{c_1, c_2, \ldots, c_n\}\}$ fails the cyclic containment property because any subset that contains c_2 and c_3 also has c_1. As the theorem asserts there does not exist an equation relating all of the candidates' tallies over these sets, this limits the number and kinds of CM relations admitted by $\mathcal{F}_2$. As true for the above, inequalities (and the resulting CM relationships) are found by treating $\mathcal{F}_2$ as a subfamily of a family that does obey the cyclic conditions. (Just add another subset that doesn't have either c_1 or c_2. The resulting relationships indicate, for instance, that if c_1 does well with most subsets of candidates, she does well with others.)

While Thm. 7d claims that a CM word is a BC word, it also asserts that certain BC words cannot be CM words. In fact, we should suspect – and it is true – that the CM imposes significantly more regularity than the BC upon the rankings of different subsets of candidates. For instance, statement a fails for the BC: all we can assert about a BC top-ranked candidate in all k-candidate subsets is that she is not BC bottom-ranked in larger subsets of candidates. But, because the CM artificially converts the original profile into a related unanimity profile, added regularity must be expected. After all, by treating even close elections as unanimous, the CM loses all ability to retain the subtle distinctions involved in these settings. (And, if the election is not close, the CM is not needed.)

3.2. Copeland Dictionary. These conclusions indicate some of the consequences of the

CM dictionary. The natural desire to list all CM relationships conflicts with reality once $n \geq 5$ – the new relationships would rapidly fill books. Therefore, realism requires us to show how to quickly compute needed relationships while assuring the reader that they all assert that candidates who do well with certain subsets of candidates tend to do well in others. (While this seems obvious, it is false for non-BC positional methods [S4] because they fail to satisfy the reversal property.)

To describe the CM dictionary for the n candidates $\{c_1, c_2, \ldots, c_n\}$, list the subsets of two or more candidates as $S_1, S_2, \ldots, S_{2^n-(n+1)}$. If $|S_j|$ is the number of candidates in S_j, the number of points assigned to a candidate can be represented as a value along an axis in $R^{|S_j|}$. Label the coordinate axis of $R^{|S_j|}$ according to the candidates' names; e.g., if $S_k = \{c_2, c_4, c_5\}$, then the axes of $R^{|S_k}$ are (x_2^k, x_4^k, x_5^k). All of our computations are in the product space

$$\mathcal{R}^n = R^{|S_1|} \times \cdots \times R^{|S_{2^n-(n+1)}|}.$$

$\mathcal{D}_C^n$ is characterized in terms of the following vectors in $\mathcal{R}^n$.

Definition 3. For candidate $c_j \in S_k$ where $|S_k| \geq 3$, let $\mathbf{Z}_{j,k}$, the *c_j normal vector for S_k*, be defined in the following manner.

The $R^{|S_k|}$ vector component of $\mathbf{Z}_{j,k}$ has the value $\frac{1-|S_k|}{|S_k|}$ for the c_j component and $\frac{1}{|S_k|}$ for all other components.

If S_i is a pair of candidates containing c_j and another candidate from S_k, then let the $R^{|S_i|}$ component of $\mathbf{Z}_{j,k}$ have the value $\frac{1}{2}$ for the c_j component, and $-\frac{1}{2}$ for the other component.

For all remaining choices of S_i, the $R^{|S_i|}$ component of $\mathbf{Z}_{j,k}$ is $\mathbf{0}$.

As a $n = 3$ example, let $S_1 = \{c_1, c_2\}$, $S_2 = \{c_1, c_3\}$, $S_3 = \{c_2, c_3\}$, $S_4 = \{c_1, c_2, c_3\}$. The coordinates for $\mathcal{R}^3$ are $((x_1^1, x_2^1), (x_1^2, x_3^2), (x_2^3, x_3^3), (x_1^4, x_2^4, x_3^4))$, so the c_2 normal vector for S_4 is

$$\mathbf{Z}_{2,4} = ((-\frac{1}{2}, \frac{1}{2}), (0,0), (\frac{1}{2}, -\frac{1}{2}), (\frac{1}{3}, -\frac{2}{3}, \frac{1}{3})).$$

In the following statement, let $(-,-)$ be the usual inner product on $\mathcal{R}^n$ where $(\mathbf{v}_1, \mathbf{v}_2)$ is computed by first taking the product of two terms sharing the same component, and then summing these products. For instance, if

$$\mathbf{v}_1 = ((u_1, u_2), (u_3, u_4), (u_5, u_6), (u_7, u_8, u_9)),$$
$$\mathbf{v}_2 = ((v_1, v_2), (v_3, v_4), (v_5, v_6), (v_7, v_8, v_9)),$$

then $(\mathbf{v}_1, \mathbf{v}_2) = \sum_{j=1}^{9} u_j v_j$.

Theorem 9. *For a profile* $\mathbf{p}$, *let* $\mathbf{V} \in R^n$ *be the vector that gives the CM score (using* $(1, 0, -1)$ *weights) for all subsets of candidates. It must be that*

$$(\mathbf{V}, \mathbf{Z}_{j,k}) = 0 \tag{3.3}$$

for all j, k *(and all linear combinations of the* $\mathbf{Z}_{j,k}$ *vectors).*

In mathematical terms, this theorem asserts that the $\{\mathbf{Z}_{j,k}\}$ vectors span normal space for the linear space containing the CM scores. The only difference between Thm. 9 and the characterization of $\mathcal{D}_B^n$ is that the CM scores are integer based. To illustrate with a

trivial $n = 3$ case, we compare all ways the CM and the BC score of $(-1, 0, 1)$ can occur. The scalar product with $\mathbf{Z}_{2,4}$ leads to the equation $\frac{1}{2}(-x_1^1 + x_2^1 + x_2^3 - x_3^3) = 0$, and that with $\mathbf{Z}_{1,4}$ defines $\frac{1}{2}(x_1^1 - x_2^1 + x_1^2 - x_3^2) = -1$. Because the sum of the CM (and BC) scores for a subset of candidates equals zero (so $x_1^1 = -x_2^1$, etc.), we obtain the two equations $x_1^1 = x_2^3$, $x_1^1 + x_1^2 = -1$ in three unknowns. For the BC, these equations admit a line of possible solutions $(-t, t), (t-1, 1-t), (-t, t))$ for $0 \le t \le \frac{2}{3}$. (The $\frac{2}{3}$ bound comes from the transitivity constraint defining the slanted sides of the representation cube.) However, the CM pairwise scores are restricted to the values $-1, 0, 1$, so the only admissible CM solution is the $t = 0$ endpoint of this line with the values $x_1^1 = 0$, $x_1^2 = -1$, $x_2^3 = 0$, and the only possible CM word $(c_1 \sim c_2, c_3 \succ c_1, c_2 \sim c_3, c_3 \succ c_2 \succ c_1)$. In fact, this argument indicates how to prove Thm. 7d: a CM word is a BC word restricted to integer values.

As a second example, we find all BC rankings associated with the CM ranking $c_1 \succ c_2 \succ c_3 \succ c_4$ attained by the pairwise elections values $x_{12} = x_{13} = x_{14} = x_{24} > 0$, $x_{23} = x_{34} = 0$: a setting without a cycle. (Other pairwise rankings can be found to support this CM score of $(3, 0, -1, -2)$.) When pairs and the set of all candidates are compared, Thm. 9 reduces to Eq. 3.1. Thus, the pairwise rankings define the BC score $(x_{12} + x_{13} + x_{14}, -x_{12} + x_{24}, -x_{13}, -x_{14} - x_{24})$. Because the c_1 BC score is positive and the c_3, c_4 BC scores are negative, c_1 always is ranked strictly above c_3 and c_4. If c_1 is top-ranked, then the other three candidates can be ranked in any way. For instance, this can occur if x_{24} is near zero. Here, the BC rankings are $(+, -x_{12}, -x_{13}, -x_{14})$ with c_1 top-ranked and the rest of the candidates ranked in any way desired. If $c_1 \sim c_2$ or $c_2 \succ c_1$, then a simple comparison of the scores show that $c_3 \succ c_4$. This specifies all BC rankings.

The power of Thm. 9 arises when comparing several sets of three or more candidates. This is illustrated with a proof of Thm. 7e which requires finding normal vectors (linear combinations of the $\mathbf{Z}_{j,k}$ vectors) where the only non-zero components are in the spaces associated with $S_j, j = 7, \ldots, 10$. Now, as the $\mathbf{Z}_{j,7}$ components for pairs has c_j beating the other candidate, to eliminate these terms, we need to use other $\mathbf{Z}_{i,k}$ vectors where c_j loses to these candidates. By adding, the normal vectors are $\mathbf{N}_1 = ((-2, 1, 1), (1, -2, 1), (1, -2, 1), (1, 1, -2))$, $\mathbf{N}_2 = ((1, -2, 1), (-2, 1, 1), (1, 1, -2), (1, -2, 1))$. (For instance, $3\mathbf{N}_2 = \mathbf{Z}_{1,8} + \mathbf{Z}_{2,7} + \mathbf{Z}_{3,10} + \mathbf{Z}_{4,9}$.) Thm. 7e follows by using these vectors in Eq. 3.3.

Similarly, the proofs of Thm. 7a,b,c involve finding normal vectors with non-zero components only in the component spaces of $\mathcal{R}^n$ corresponding to the k-candidate subsets and the full subset. If $\{S_\alpha\}$ represents all k-candidate subsets and $S_{2^n-(n+1)}$ the n-candidate set, then we need linear combinations of $\{\mathbf{Z}_{j,\alpha}\}$ and $\mathbf{Z}_{j,2^n-(n+1)}$ to eliminate the coordinates corresponding to pairs. Here, because each pair appears in $\binom{n-2}{k-2}$ different k-candidate subsets, a normal vector for c_j is where if c_j is in a k-candidate subset, then the c_j coordinate of this vector component is $\frac{k-1}{k}$ while all other coordinates are $-\frac{1}{k}$, the $S_{2^n-(n+1)}$ component is a $\binom{n-2}{k-2}$ multiple of the vector with $\frac{1-n}{n}$ for the c_j coordinate and $\frac{1}{n}$ for all others; all other components are zero. Using these normal vectors in Eq. 3.3 leads to Thm. 7c. Relationships for other families of subsets are found in the same way. To complete the discussion, we need to know when a linear combination of the $\mathbf{Z}_{j,k}$ vectors can be found where all non-zero components correspond to the subsets in the specified family. This is Thm. 8; the proof (for the BC) is in [S6].

3.3. Remaining proofs. The proof of Thm. 9 is a minor modification of the argument in [S3]; the proof of Thm. 8 follows directly from [S6]. All that remains is Thm. 4 and the

second part of Thm. 5. To prove the second part of Thm. 5, the basic argument of [S7] is used twice. This approach uses a smooth mapping, $F : M \to N$, between spaces where the domain has at least as large dimension as the range. If the Jacobean of F at an interior point p has maximal rank, then any sufficiently small open set about p is mapped to an open set about $F(p)$. (So, these equations can be solved for an open set of specified values.) After specifying pairwise election rankings that define a cycle, if a profile $\mathbf{p}$ can be found where its pairwise elections rankings are the specified ones and where the BC outcome is the complete tie $c_1 \sim \cdots \sim c_n$, the open mapping theorem ensures that small changes can be made in the profile to preserve the pairwise election rankings (and the CM score), but the BC ranking can be anything desired. The difficult task is to show that such a profile $\mathbf{p}$ can be found for specified pairwise election rankings.

Instead of the base profile, it follows from [S3] that we can use pairwise election outcomes: as with the representation cube, this restricts attention to a region in a $\binom{n}{2}$ dimensional space with variables $x_{ij} \in [-1, 1]$ where $i < j$ and a positive value indicates that c_i beats c_j. Let $\mathbf{X}$ represent a vector in this $\binom{n}{2}$ dimensional region M. The mapping $F(\mathbf{X})$ converts these pairwise election tallies into the BC scores as mandated by Eq. 3.1. Thus, the space N, of BC election outcomes, is $n-1$ dimensional (as the sum of BC scores equals zero), and the goal is to find $\mathbf{X}^*$ where the pairwise rankings are as specified and $F(\mathbf{X}^*) = \mathbf{0}$.

Assume, without loss of generality, that the pairwise rankings define the cycle $c_1 \to c_2 \to \cdots \to c_n \to c_1$. If $x_{1,2} = x_{2,3} = \cdots = x_{n-1,n} = -x_{n,1} = t > 0$, then the pairwise rankings for the cycle are satisfied and the BC outcome is a complete tie *should all other pairwise elections be tied so these* $x_{j,k} = 0$. (Choose the value of t so that it is in the open region of the pairwise outcomes M.) Call this list of pairwise outcomes $\mathbf{X}_t$; we have that $F(\mathbf{X}_t) = \mathbf{0}$. Defining F_{cycle} to be F restricted to the n dimensional subspace, M_t, of M spanned by these n nonzero variables, then $F_{cycle}(\mathbf{X}_t) = \mathbf{0}$. With a cycle, it follows from the diagonal form that the Jacobean, DF_{cycle}, has maximal rank of $n-1$.

Clearly, the variables off of M_t can be chosen so that each of the remaining pairwise elections satisfy the specified ranking and the sum of these contributions to each BC score can be made arbitrarily small; let this vector of BC scores be q. Allowing the variables on M_t to be free, the problem of finding $\mathbf{X}^*$ is the same as finding $\mathbf{Y} \in M_t$ near $\mathbf{X}_t$ so that $F_{cycle}(\mathbf{Y}) = -q$. Thanks to the open mapping approach, this can be done should the magnitude of q be sufficiently small. Thus, $\mathbf{X}^*$ exists and the theorem holds. □

Proof of Theorem 4. Extensions to $n \geq 3$ require generalizing the procedure line to a *procedure hull* (supporting details are in [S5]): the convex hull of the normalized outcomes for voting vectors $\mathbf{W}_1^n = (1, 0, \ldots, 0)$, $\mathbf{W}_2^n = (1, 1, 0, \ldots, 0), \ldots, \mathbf{W}_{n-1}^n = (1, 1, \ldots, 1, 0)$. All positional methods can be expressed as a linear combination of the $\{\mathbf{W}_k^n\}_{k=1}^{n-1}$. In fact, because the BC vector $\mathbf{B}^n = (n-1, n-2, \ldots, 0)$ has the representation $\mathbf{B}^n = \frac{1}{n-1}\sum_{j=1}^{n-1} \mathbf{W}_j^n$, it is the barycenter of the positional voting methods. The procedure hull, one dimension less than the space of normalized outcomes, is uniquely determined by the scores of the BC and $n-2$ of the $\{\mathbf{W}_k^n\}$ methods. As restrictions on the BC normalized score constrain the position of the hull, they require the rankings for a portion of positional methods to be similar to the BC. These restrictions (see Borda [D], Nanson [Na] and [S1] for a simple proof) are that a Condorcet winner (loser) can never be BC bottom-ranked (top-ranked), and a Condorcet winner is BC strictly ranked over the Condorcet loser. (No other positional method satisfies any relationship based on pairwise comparisons [S2,3].) As the assumptions on profile $\mathbf{p}$ require the BC – the midpoint of the positional procedures – to satisfy

the ranking conditions of part a, the procedure hull geometry forces these relationships on a portion of the positional methods.

To find the portion of the positional methods related to the CM ranking, note that the specified restrictions of the theorem admit at least half of all possible strict rankings. (The strongest restriction, where $c_1 \succ c_n$, admits precisely half of all $n!$ strict rankings.) In the space of normalized outcomes, the admissible rankings are in a half-space. (For the $c_1 \succ c_n$ restriction, this half-space $x_1 > x_n$ is defined by the plane $x_1 = x_n$.) Therefore, the r_n value is determined by finding the smallest possible portion of the procedure hull that can be in a half-space along with the barycenter point. Geometrically, the problem corresponds to dividing the simplex of the positional methods (defined by the vertices $\{\mathbf{W}_k^n\}_{k=1}^{n-1}$) into two parts with a plane passing through the barycenter. The object is to find where one portion has the minimum volume. For $n = 3$, the hull is a line divided into two equal parts, so the answer is the $\frac{1}{2}$ value specified earlier. For $n > 3$, it follows from calculus methods that the minimal situation has the dividing plane parallel to one of the sides. Because the components of the BC (in this normalized sense) are $(1, \frac{n-2}{n-1}, \ldots, \frac{1}{n-1}, 0)$, the conclusion follows from a direct computation. (However, a geometric approach using "similar" triangles defined by the dividing plane in the hull leads to an elementary computation.)

Part b requires finding the region that assumes the role of the shaded triangle in Fig. 3b. Again (by use of the techniques of [S5]), when the BC is the complete tie ranking, the admissible accompanying $\mathbf{W}_k^n$ choices are contained in an open set that includes the barycenter. As such, accompanying a BC ranking near the complete tie is considerable flexibility in choosing the remaining vertices for the procedure hull. In the same manner as for $n = 3$, the assertion follows. □

References

[Cn] Condorcet, Marquis de, *Essai sur l'application à la possibilitité des décisions rendues à la pluralité de voix*, Paris, 1785.

[C] Copeland, A.H., *A 'reasonable' social welfare function*, Notes from a seminar on applications of mathematics to the social sciences; University of Michigan, 1951.

[D] De Grazia, A., *Mathematical derivation of an election system.*, Isis **44** (1953), 42-51.

[MS] Merlin, V., and D.G. Saari, *Copeland method II: Manipulation, monotonicity, and paradoxes* (1994).

[Na] Nanson, E.J., *Methods of election*, Trans. Proc. R. Soc. Victoria **18** (1882), 197-240.

[N] Nurmi, H., *Comparing voting systems*, Reidel, 1987.

[S1] Saari, D.G., *Geometry of Voting*, Springer-Verlag, 1994.

[S2] Saari, D.G., *A dictionary for voting paradoxes*, Jour. Econ. Theory **48** (1989), 443-475.

[S3] Saari, D.G., *The Borda Dictionary*, Soc. Choice Welfare **7** (1990), 279-317.

[S4] Saari, D.G., *Symmetry extensions of "neutrality," I Advantage to the Condorcet loser*, Soc. Choice Welfare **9** (1992), 307-336.

[S5] Saari, D.G., *Millions of election rankings from a single profile*, Soc. Choice Welfare **9** (1992), 277-306.

[S6] Saari, D.G., *Relationship admitting families of candidates*, Soc. Choice Welfare **8** (1991), 21-50.

[S7] Saari, D.G., *The source of some paradoxes from social choice and statistics*, Jour. Econ. Theory **41** (1987), 1 - 22.

Department of Mathematics, Northwestern University, Evanston, Illinois, 60208-2730