CONSEQUENCES OF REVERSING PREFERENCES

DONALD G. SAARI AND STEVEN BARNEY

Other than standard election disruptions involving shenanigans, strategic voting, and so forth, it is reasonable to expect that elections are free from difficulties. But this is far from being true; even sincere election outcomes admit all sorts of counterintuitive conclusions.

For instance, suppose after the winner of an important departmental election was announced, it was discovered that *everyone* misunderstood the chair's instructions. When ranking the three candidates, everyone listed their top, middle, and bottom-ranked candidate in the natural order of, respectively, first, second, and third. For reasons only the chair understood, he expected the voters to vote in the opposite way. As such, when tallying the ballots, he treated a first and last listed candidate, respectively, as the voter's last and first choice.[1]

Imagine the outcry if after retallying the ballots the chair reported that the election ranking remained unchanged; in particular, the same person won. Skepticism might be the kindest reaction to greet an announcement that the election ranking for a *profile* — a listing which specifies the number of voters whose preferences are given by each (complete, transitive) ranking of the candidates — is the same for the profile where each voter's preference ordering is reversed. Surprisingly, this seemingly perverse behavior can sincerely occur with most standard election procedures. It is intriguing that this phenomenon can be explained in terms of simple mathematical symmetries. *Of particular interest, the same arguments explain all of the election paradoxes which have perplexed this area for the last two centuries.*

This issue appears to have been first introduced in [Saari 1995] where a section of this book showed that some procedures allow the same election *ranking* to occur with a profile and with its reversal. There is no interest in this phenomenon when the common ranking is a complete tie, but when the common ranking is not a tie this effect is called a "reversal bias." The word "bias" is intended to foreshadow how this anomaly affects election outcomes.

Rather than an election ranking, voters more typically care only about who wins, or who is elected for, say, the departmental budget committee. This raises the question whether an election procedure would allow the same winner, or the same two candidates, . . . , or the same k candidates to be top-ranked with a profile and its reversal. Call this situation a "k-winner reversal bias." Common sense suggests that we should question the reliability of an election procedure if it elects the same committee with a profile and with the profile of reversed preferences; i.e., if the procedure allows a k-winner reversal bias. As one of us (Barney) discovered, an internet discussion group worrying about election methods is particularly concerned about the case $k = 1$ — which we call the "top-winner reversal bias." It should be a concern because, as shown here, rather than a rare and obscure phenomenon, we can expect some sort of reversal behavior about 25% of the time with the standard plurality vote.

Positional methods

Among the widely used election methods are what William Riker [1982] calls *positional* methods. Riker, who was a pioneer in using mathematics to address problems from political

Our thanks to Hannu Nurmi, Tom Ratliff, and two referees for their comments on an earlier version.

[1] Rather than an hypothetical story, this actually occurred in an academic department to which one of us (Saari) belonged. The chair was promoted to a higher administrative position.

science, coined the word "positional" to refer to a method where a ballot for the $n \geq 2$ candidates is tallied by assigning specified weights, $w_1, w_2, \dots, w_n$, respectively, to a voter's first, second, ... , and nth ranked candidate. The candidates are then ranked according to the sum of weights from all ballots. Since the election ranking remains unchanged after adjusting the weights so that $w_n = 0$, assume that this is the case. The *plurality vote* is the commonly used "vote for one" system where $w_1 = 1$ and $w_2 = \cdots = w_n = 0$. The weights for the *Borda Count* (named after Jean Charles de Borda, an eighteenth century French mathematician, inventor, explorer, warrior in the American Revolution, and one of the founders of the metric system) specify the number of candidates ranked below a specified candidate, so $w_1 = n-1$, $w_2 = n-2, \dots, w_n = n-n = 0$. Actually, any choice of weights defines a positional method as long as $w_1 > w_n = 0$ and $w_j \geq w_{j+1}$ for $j = 1, \dots, n-1$. (A positive fixed multiple of the weights scales the tally and yields the same election ranking.)

To demonstrate, we compute each candidate's tally for all positional methods for the profile

Number	Prefer	A	B	C
4	$A \succ C \succ B$	$4w_1$	0	$4w_2$
3	$A \succ B \succ C$	$3w_1$	$3w_2$	0
4	$B \succ C \succ A$	0	$4w_1$	$4w_2$
3	$C \succ B \succ A$	0	$3w_2$	$3w_1$
	Total	$7w_1$	$4w_1 + 6w_2$	$3w_1 + 8w_2$

(1)

Thus the plurality vote, where $w_1 = 1, w_2 = w_3 = 0$, results in the ranking $A \succ B \succ C$ with a 7:4:3 tally. With the antiplurality vote defined by $w_1 = w_2 = 1$, $w_3 = 0$ (called "antiplurality" because by voting for all but one candidate, each voter is effectively voting *against* a candidate; the method is a "negative plurality vote"), the election ranking is $C \succ B \succ A$ with a 11:10:7 tally. Notice the conflict with the plurality outcome.

Now reverse each voter's ranking to obtain the reversed profile

Number	Prefer	A	B	C
4	$B \succ C \succ A$	0	$4w_1$	$4w_2$
3	$C \succ B \succ A$	0	$3w_2$	$3w_1$
4	$A \succ C \succ B$	$4w_1$	0	$4w_2$
3	$A \succ B \succ C$	$3w_1$	$3w_2$	0
	Total	$7w_1$	$4w_1 + 6w_2$	$3w_1 + 8w_2$

(2)

The point to notice is that each candidate's tally for each positional procedure is the same with the Eq. 1 profile as with its Eq. 2 reversal. Thus, unless the outcome is a complete tie, the procedure exhibits a reversal bias. A complete tie requires $7w_1 = 4w_1 + 6w_2 = 3w_1 + 8w_2$, or $w_1 = 2w_2$ — the Borda Count. Consequently, *with the sole exception of the Borda Count, all other positional methods experience a reversal bias with this profile.*

The source of this phenomenon is the considerable symmetry embedded in the profile's two pairs of two rankings. The first pair is the ranking $A \succ C \succ B$ with its reversal $B \succ C \succ A$; each is preferred by four voters. Likewise, with the pair of $A \succ B \succ C$ and its reversal $C \succ B \succ A$, each is supported by three voters. As positional methods respect anonymity (i.e., we do not know who has what preferences), the profile and its reversal are the same. Being indistinguishable, the profile and its reversal must give the same outcome.

To describe this symmetry, first let $\sigma_{i,j}$ be the permutation of candidates that interchanges i and j's names. So, if $\mathbf{p}$ is a profile, then $\sigma_{i,j}(\mathbf{p})$ interchanges each voter's ranking of i and j. It is easy to show that all positional methods f satisfy what is called *neutrality*; namely,

$$f(\sigma_{i,j}(\mathbf{p})) = \sigma_{i,j}(f(\mathbf{p})). \qquad (3)$$

In words, if everyone confused Sue with Mary when marking the ballot (instead of the correct profile $\mathbf{p}$, they used $\sigma_{S,M}(\mathbf{p})$), then the correct outcome is found by exchanging Mary's and Sue's tallies (namely, use $\sigma_{S,M}(f(\sigma_{S,M}(\mathbf{p})))$ as it equals $f(\mathbf{p})$).

Our example involves the symmetry $\mathcal{R}$ which reverses rankings. More precisely, if $\mathcal{R}(\mathbf{p})$ reverses each voter's ranking of the candidates, we want to identify all procedures where

$$f(\mathcal{R}(\mathbf{p})) = \mathcal{R}(f(\mathbf{p})). \tag{4}$$

As reversing a reversal returns to the initial ranking, Eq. 4 means that $\mathcal{R}(f(\mathcal{R}(\mathbf{p}))) = f(\mathbf{p})$. Using the introductory example where everyone marked their ballots in the reversed manner (rather than $\mathbf{p}$, the ballots are marked as $\mathcal{R}(\mathbf{p})$), if the election procedure satisfied Eq. 4, then a way to find the correct $f(\mathbf{p})$ outcome is to reverse the $f(\mathcal{R}(\mathbf{p}))$ ranking. This seemingly natural property can fail with most procedures.

A way to spot the methods with these problems is to mimic the Eq. 1 example by using profiles of the $\mathcal{R}(\mathbf{p}) = \mathbf{p}$ type; i.e., those profiles where each ranking in $\mathbf{p}$ is accompanied by the same number of voters preferring its reversal. With these profiles, $f(\mathcal{R}(\mathbf{p})) = f(\mathbf{p})$. So, if Eq. 4 is true, we have that $\mathcal{R}(f(\mathbf{p})) = f(\mathbf{p})$. But $\mathcal{R}(f(\mathbf{p})) = f(\mathbf{p})$ holds only if $f(\mathbf{p})$ is a complete tie. Thus, we just need to identify those procedures which fail to deliver a complete tie for these special $\mathcal{R}(\mathbf{p}) = \mathbf{p}$ profiles.

Theorem 1. *For three candidate elections, only the Borda Count never exhibits the reversal nor k-winner reversal bias, $k \leq 2$. All other positional methods suffer the reversal, top-winner, and $k = 2$ winner reversal bias. For a procedure to exhibit these effects, a profile must have a sufficiently large component of rankings with their reversal.*

The Borda Count always satisfies Eq. 4 for any $n \geq 3$, so it never has a reversal nor a k-winner reversal bias. Almost all positional methods fail to satisfy the equalities

$$w_1 = w_2 + w_{n-1} = w_3 + w_{n-2} = \cdots = w_{n-1} + w_2 = w_1; \tag{5}$$

methods failing Eq. 5 allow reversal and k-winner reversal biases for any $k \leq n-1$. Indeed, if Eq. 5 fails, then select any ranking; a profile can be constructed where the profile and its reversal support the specified ranking with the same tally.

While simple, Eq. 5 has surprisingly strong consequences. It means, for example, that all commonly used methods, such as "vote for one," or "vote for two," or methods based on almost any choices of weights are susceptible to the full array of reversal problems. Moreover, since it is arguable that profiles of this $\mathbf{p} = \mathcal{R}(\mathbf{p})$ type should end in a tie, it follows that procedures which fail Eq. 5 bias the outcome; a measure of this bias is the difference in value between the smallest and largest Eq. 5 terms. For three alternatives, then, the bias a positional method introduces into the election outcome is captured by the non-zero difference $w_1 - 2w_2$. We will return to this comment when discussing general election paradoxes.

Once we understand the origin of Eq. 5 and how to construct examples, a formal proof is immediate. To explain Eq. 5, the election tallies for the profile consisting of $c_1 \succ c_2 \succ \cdots \succ c_n$ and its reversal $c_n \succ \cdots \succ c_2 \succ c_1$ are, respectively, $w_1 + w_n, w_2 + w_{n-1}, w_3 + w_{n-2}, \dots, w_1 + w_n$. If these tallies fail to agree, they violate Eq. 5 (remember, $w_n = 0$) and the non-tied outcome means that the procedure suffers a reversal bias.

To construct profiles asserted by Thm. 1 with the election ranking $c_1 \succ c_2 \succ c_3 \succ c_4$, we exploit the bias caused when Eq. 5 is not satisfied. So if $w_1 + w_4 < w_2 + w_3$, exploit the larger $w_2 + w_3$ sum by putting into the profile voters for whom c_1 and c_2 are, respectively, second and third ranked. The two remaining candidates, c_3 and c_4, can be ordered in two ways. Use both orderings to define the two rankings $c_3 \succ c_1 \succ c_2 \succ c_4$ and $c_4 \succ c_1 \succ c_2 \succ c_3$. Include the reversal for each ranking to obtain what we call the "c_2 unit" of $\{c_3 \succ c_1 \succ c_2 \succ c_4,\ c_4 \succ$

$c_2 \succ c_1 \succ c_3\}$ and $\{c_4 \succ c_1 \succ c_2 \succ c_3, c_3 \succ c_2 \succ c_1 \succ c_4\}$. For this c_2 unit, c_1 and c_2 each receive $2(w_2 + w_3)$ points while c_3 and c_4 each receive the smaller $2(w_1 + w_4) = 2w_1$ value.

Replace c_2 with c_j to create a c_3 and a c_4 unit. The number of the c_j units needed to design a profile depends on the desired outcome; e.g., one choice of a $\mathbf{p}$ which generates the specified election ranking consists of two c_2, one c_3, and no c_4 units. The $c_1 \succ c_2 \succ c_3 \succ c_4$ election outcome has the tally $2[3(w_2 + w_3))] > 2[2(w_2 + w_3) + w_1] > 2[(w_2 + w_3) + 2w_1] > 2[3w_1]$. By construction $\mathbf{p} = \mathcal{R}(\mathbf{p})$, so the conclusion of Thm. 1 is satisfied. The formal proof just verifies that this approach extends to any n.

Notice a conspicuous gap: of all positional methods satisfying Eq. 5, Thm. 1 only excuses the Borda Count from these reversal effects for $n \geq 4$ alternatives. For instance, the weights $(2, 1, 1, 0)$, or $(2, 2-y, y, 0)$, $0 \leq y \leq 1$, or $(4, 4-z, 2, z, 0)$, $0 \leq z \leq 2$, define positional methods which satisfy Eq. 5, but Thm. 1 does not state whether they suffer the reversal bias. They do not (the technical proof is omitted), but they have other problems.

These conclusions extend to a much wider class of voting procedures. For instance, Saari and Van Newenhizen [1989] define a *multiple voting procedure* as one which is equivalent to having the voter mark the ballot *and* then select the positional procedure to tally this particular ballot. "Approval Voting" is the multiple voting procedure where a voter can vote for as many candidates as he or she wishes; by voting for one, or two, or say three candidates, the voter is effectively selecting, respectively, the positional methods $(1, 0, \dots, 0)$, or $(1, 1, 0, \dots, 0)$, or $(1, 1, 1, 0, \dots, 0)$. (As one might anticipate from the variability, this procedure, used by both the MAA and AMS, has several serious flaws [Saari, 2001].) Other multiple procedures are "truncated voting" where a voter ignores instructions by voting only for some candidates, and "cumulative voting" where a voter can distribute a specified number of points among the candidates in any desired manner, etc. As these methods clearly fail Eq. 5, the following assertion is immediate.

Theorem 2. *All multiple voting procedures suffer both a reversal and a k-winner reversal bias. In particular, this includes Approval Voting, cumulative voting, and truncated voting when used with any positional method.*

We leave it for the reader to determine (which is not difficult) whether the method of single transferable vote (STV) used by the AMS suffers these problems. In STV, when the goal is to select, say, two of three candidates, as soon as a candidate receives over a third of the vote, she is elected; any remaining ballots that have her top-ranked are reassigned to the second listed candidate.

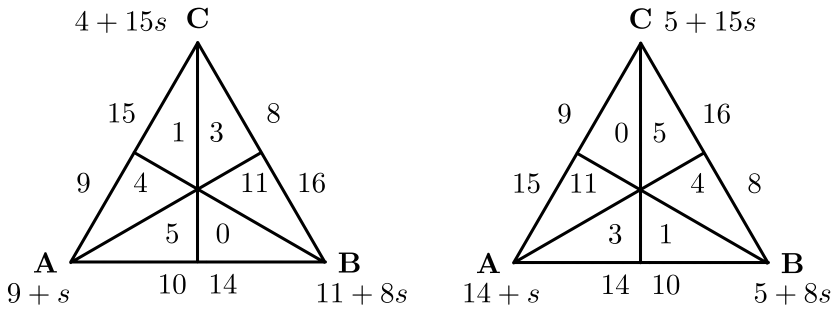
Three candidate positional elections

A more important objective is to understand how reversal effects affect election outcomes. In doing so, we verify our earlier statement about the likelihood of these reversal effects along with the claim that these reversal behaviors explain voting paradoxes. This last theme uses the observation that a positional method which fails to satisfy Eq. 5 can bias the election outcome. Indeed, as we will see, all differences among positional election outcomes reflect the $w_1 - 2w_2$ differences. Along the way, some easily used conditions are developed to identify, for instance, when reversing a profile will not reverse the ranking. In deriving new conclusions while outlining how to find others, the three-candidate setting is emphasized for ease of exposition. Our approach uses the "procedure line" and a geometric representation of profiles introduced in [Saari 1994, 1995] and used in several ways by Nurmi [1999, 2002].

Equilateral triangles, such as Fig. 1a, are useful devices to describe three-candidate election outcomes. Assign a ranking to a point in the triangle according to its distance from each vertex where "closer is better." For instance, any point in the small triangular region of Fig. 1a with "11" is closest to B, next closest to C, and farthest from A, so it is assigned the $B \succ C \succ A$ ranking. Represent a profile by listing the number of voters who have each preference ranking in the appropriate region; e.g., Fig. 1a displays the profile

No.	Ranking	No.	Ranking
5	$A \succ B \succ C$	3	$C \succ B \succ A$
4	$A \succ C \succ B$	11	$B \succ C \succ A$
1	$C \succ A \succ B$	0	$B \succ A \succ C$

(6)



a. Profile **b.** Reversed profile

Fig. 1. Representing profiles and tallies

This geometry simplifies computing election tallies. To see why, notice that all rankings with $A \succ B$ are to the left of the vertical line, so the $5 + 4 + 1 = 10$ sum of the numbers in these three Fig. 1a regions is A's tally in an $\{A, B\}$ pairwise election. All pairwise tallies are similarly computed and listed next to the appropriate triangle edge.

Instead of using $(w_1, w_2, 0)$ for positional method elections, an easier way to compare procedures is to normalize the weights by dividing by w_1; this defines $(1, s, 0)$ where the fixed $s = w_2/w_1$ value, $0 \le s \le 1$, is assigned to a second ranked candidate. In this manner the Borda Count $(2, 1, 0)$ becomes $(1, \frac{1}{2}, 0)$, and the $(7, 5, 0)$ method becomes $(1, \frac{5}{7}, 0)$. To tally positional method ballots, notice from Fig. 1a that A is top-ranked in the two regions with A as a vertex, so add these numbers. Next, A is second ranked in the two adjacent regions; in Fig. 1a these are the two regions containing 1 and 0. Thus, add s times this sum to compute A's final tally of $(5+4)+s(1+0) = 9+s$; this value is placed near the A vertex. The similarly computed tallies for the other two candidates are listed next to the appropriate vertex.

In the three-dimensional space of election tallies, R^3, the A, B, C tallies of $(9 + s, 11 + 8s, 4 + 15s)$ describe a line connecting the plurality tally (where $s = 0$) with the antiplurality outcome (where $s = 1$); this is the *procedure line* [Saari 1992, 1994, 1995]. This line identifies all positional method tallies; the tally for $(1, s, 0)$ is s of the way from the plurality to the antiplurality tally. Since the Borda Count is given by $s = \frac{1}{2}$, the Borda tally is at the midpoint. Procedure lines have proved to be a convenient tool. For instance, by using the procedure line Tabarrok [2001] discovered surprising conclusions about the 1992 presidential election involving Clinton. History buffs will enjoy the Tabarrok and Spector [1999] paper using a natural extension of the procedure line [Saari 1992] to characterize everything that could have happened with the 1860 election involving Abraham Lincoln.

An advantage of the procedure line for theoretical purposes is that it identifies all positional method outcomes. This suggests that a way to find all consequences of reversing a profile $\mathbf{p}$ is to compare the $\mathbf{p}$ and $\mathcal{R}(\mathbf{p})$ procedure lines. But first we need to represent a reversed profile.

Finding the reversed profile: To find the reversed profile, place the number from each triangular region of the original profile in the diametrically opposite region (relative to the center of the triangle); e.g., the Fig. 1b profile is the reversal of the Fig. 1a profile. Notice from Fig. 1b that while each candidate's plurality tallies (the $s = 0$ values) for the profile and the reversed profile differ, the coefficients of "s" remain the same. This always is true. To explain, when tallying ballots as described above, the s coefficient is the sum of terms in diametrically opposite regions; consequently, reversing a profile preserves the sum.

The s coefficients are the differences between the procedure line's endpoints — the antiplurality and plurality tallies — so they represent the tallies of the voters' *second place vote.* Call this difference vector, or line segment, the *Second Place Tallies* (the SPT). For instance, the Fig. 1a SPT is the vector $(10, 19, 19) - (9, 11, 4) = (1, 8, 15)$.

Theorem 3. *For any three-candidate profile* $\mathbf{p}$*, the SPT for* $\mathbf{p}$ *and* $\mathcal{R}(\mathbf{p})$ *agree. This forces the directions and lengths of their procedure lines to agree; the two lines are parallel.*

Proof. The direction and length of a line are determined by the difference between their endpoints; this difference between a procedure line's antiplurality and plurality tallies is the SPT. For instance, the Fig. 1b SPT of $(15, 13, 20) - (14, 5, 5) = (1, 8, 15)$ is the same as for Fig. 1a. As a profile's SPT is defined by the s coefficients which always agree for $\mathbf{p}$ and $\mathcal{R}(\mathbf{p})$, the theorem is proved. □

Compare a profile's SPT to a straight piece of wire which behaves like a compass needle; when moved, it points in the same direction in the three dimensional space of tallies. So the tally of one positional method and the SPT completely determine the procedure line. As the SPT for $\mathbf{p}$ and $\mathcal{R}(\mathbf{p})$ agree (Thm. 3), knowing how the tallies for a designated procedure change from $\mathbf{p}$ to $\mathcal{R}(\mathbf{p})$ completely determines the $\mathbf{p}$ and $\mathcal{R}(\mathbf{p})$ procedure lines. Since only the Borda Count is immune to reversal effects, it is the designated procedure.

To illustrate this description with the Fig. 1 example, the (normalized) Borda Count tallies for $\mathbf{p}$ and $\mathcal{R}(\mathbf{p})$ are, respectively, $(9.5, 15, 11.5)$ and $(14.5, 9, 12.5)$. The Borda Count defines the midpoint of the procedure line, so $\mathbf{p}$'s procedure line is found by placing the $(1, 8, 15)$ SPT so that its midpoint is at $(9.5, 15, 11.5)$. Similarly, to find $\mathcal{R}(\mathbf{p})$'s procedure line, move the same SPT so that its midpoint now is at $(14.5, 9, 12.5)$.

This description attributes all reversal effects to changes in the Borda tally. These changes involve another symmetry involving how each candidate's Borda vote differs from the average Borda score. To illustrate, Eq. 7 computes these differences for the Fig. 1a profile and its Fig. 1b reversal; the total number of Borda points is 36, so the average of assigned Borda points to the candidates is $36/3 = 12$.

Profile	A	B	C
Original, Fig. 1a	$9.5 - 12 = -2.5$	$15 - 12 = 3$	$11.5 - 12 = -0.5$
Reversed, Fig. 1b	$14.5 - 12 = 2.5$	$9 - 12 = -3$	$12.5 - 12 = 0.5$

(7)

As Eq. 7 suggests, reversing a profile just changes the sign of each candidate's Borda differential from the average Borda score. The reason for this behavior is that each candidate's Borda tally can be computed by adding the points she receives in all pairwise elections. So to compute a candidate's difference from the average Borda score, in each pairwise election add how much the tally differs, either above or below, a complete tie. (To have normalized Borda values, divide by two.) But $\mathcal{R}(\mathbf{p})$ reverses all pairwise tallies — and all differences from the average — so only the sign changes when computing differences from the average Borda score.

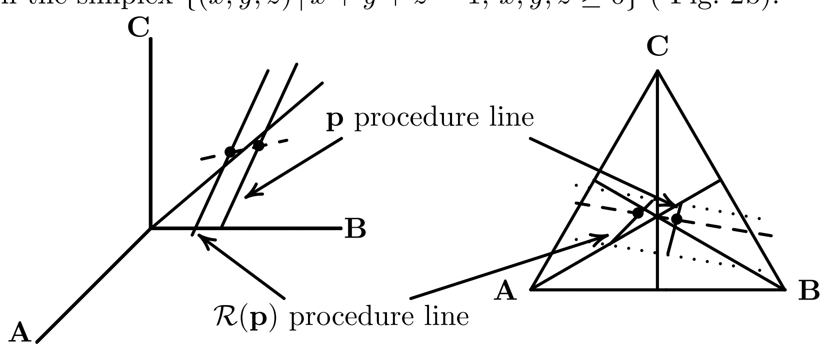
This effect indicates how to find geometrically all three-candidate properties associated with reversing a profile.

Profile $\mathbf{p}$*'s procedure line is determined by its midpoint, the Borda tallies, and its SPT vector. For an* n*-voter profile* $\mathbf{p}$*, the average Borda score is* $(\frac{3}{2}n)/3 = \frac{n}{2}$*, so point* $(\frac{n}{2}, \frac{n}{2}, \frac{n}{2})$ *on the diagonal* $x = y = z$ *indicates each candidate's average Borda score. Construct a line segment where* $\mathbf{p}$*'s Borda tally is one endpoint and segment's midpoint is* $(\frac{n}{2}, \frac{n}{2}, \frac{n}{2})$*; the segment's other endpoint is the* $\mathcal{R}(\mathbf{p})$ *Borda tally so it is the midpoint of the* $\mathcal{R}(\mathbf{p})$ *procedure line. To find the* $\mathcal{R}(\mathbf{p})$ *procedure line, slide* $\mathbf{p}$*'s SPT to this* $\mathcal{R}(\mathbf{p})$ *Borda tally. All consequences of reversing a profile* $\mathbf{p}$ *are found by comparing differences in the* $\mathbf{p}$ *and* $\mathcal{R}(\mathbf{p})$ *procedure lines.*

In Fig. 2a, the solid line on the right represents the Fig. 1a procedure line and the bullet designates the Borda tally in the $B \succ C \succ A$ region; the plurality endpoint is in the $B \succ A \succ C$ region. To find the $\mathcal{R}(\mathbf{p})$ shift, construct a line (the Fig. 2a dashed line) from $\mathbf{p}$'s Borda tally passing orthogonally through the $x = y = z$ diagonal; $\mathcal{R}(\mathbf{p})$'s Borda tally is equidistant on this line on the other side of the diagonal. To find $\mathcal{R}(\mathbf{p})$'s procedure line, which is the slanted line on the left, slide $\mathbf{p}$'s SPT so that its midpoint is at this flipped Borda tally.

Some geometry: We can use this geometry to indicate why the $B \succ A \succ C$ plurality ranking for the Fig. 1a profile $\mathbf{p}$ changes to $A \succ C \sim B$ for $\mathcal{R}(\mathbf{p})$. First, the plurality endpoint of $\mathbf{p}$'s procedure line is close to a $B \sim A$ tie. The SPT remains invariant, so the key is the flipped Borda tally; it favors A, helps C, but hurts B (Eq. 7). In $\mathcal{R}(\mathbf{p})$'s procedure line, this flip pushes the SPT deeper into the region favoring A, somewhat helping C, but hurting B.

For readers comfortable with three-dimensional geometry, this description suffices to explain the new results given below. For most of us, however, the Fig. 2a three-dimensional geometry is difficult to envision. So replace actual tallies with the fraction of the total vote each candidate receives; e.g., replace a $(50, 140, 10)$ tally with $(\frac{50}{200}, \frac{140}{200}, \frac{10}{200})$. Geometrically, the normalized tally is a point in the simplex $\{(x, y, z) \,|\, x + y + z = 1,\ x, y, z \geq 0\}$ (Fig. 2b).



a. Profile line and transfer **b.** Two dimensional projection

Fig. 2. Procedure lines for a profile and its reversal

While Fig. 2b helps to visualize election outcomes, problems arise because the projection distorts geometric properties; this distortion is not dissimilar to the difficulties of observing objects through a convex mirror. Rather than the midpoint, for instance, the normalized Borda tally is two-thirds of the way from the plurality point. As dramatically demonstrated in Fig. 2b, the projected $\mathbf{p}$ and $\mathcal{R}(\mathbf{p})$ parallel procedure lines are skewed. The reason for this distortion is that the plurality tallies are divided by the number of voters while the antiplurality tallies are divided by twice this number. Consequently the two normalized plurality endpoints always are twice as far apart as the antiplurality endpoints.

The two dotted lines in Fig. 2b — the bottom and top one connect, respectively, the $\mathbf{p}$ and $\mathcal{R}(\mathbf{p})$ plurality and antiplurality outcomes — are parallel; this provides an interesting research tool. In the figure, $\mathbf{p}$'s Borda tally (the bullet) is reflected along the dashed line about the center point to identify $\mathcal{R}(\mathbf{p})$'s Borda tally; this flipped Borda position determines $\mathcal{R}(\mathbf{p})$'s procedure line. While the projected procedure lines rarely are parallel, the dashed and dotted lines always are. After formally stating this observation and then indicating which lines in a triangle can be procedure lines, these facts are combined to derive new conclusions.

Theorem 4. *For any* $\mathbf{p}$ *and any two* $(1, s, 0)$ *positional methods, the lines connecting each method's (normalized) tally for* $\mathbf{p}$ *and* $\mathcal{R}(\mathbf{p})$ *are parallel in a Fig. 2b representation. The length of the line connecting the* $(1, s, 0)$ *outcomes is* $\frac{1}{1+s}$ *times the length of the line connecting plurality outcomes, or* $\frac{3}{2(1+s)}$ *the length of the line connecting the Borda outcomes. The* $(1, s, 0)$ *outcome on a procedure line is* $\frac{2s}{1+s}$ *of the distance from the plurality to the antiplurality endpoint.*

The proof is a straightforward exercise in elementary geometry which we leave to the reader.

Procedure lines: Before using Thm. 4, we need a simple way to find all possible positional lines. The surprisingly relaxed rules [Saari 2001] to identify which line segments in a triangle are procedure lines are described in terms of how voters cast their first (the plurality outcome) and second (the SPT) place votes. For any integers selected in the following manner, an unique profile exists with the specified election outcomes and tallies.

- Choose a non-negative integer value for each candidate's plurality tally, the sum determines the total number of voters.
- Select any nonnegative integers to define the SPT where
  - their sum equals the number of voters, and
  - since the sum of the SPT entries for any two candidates includes the plurality tally for the third candidate, it must be at least this large.

A candidate's Borda tally is the average of her assigned plurality and antiplurality values.

To illustrate how easy it is to construct the unique supporting profile, suppose the plurality tallies for A, B, and C are, respectively, 4, 5, and 6 as indicated in Fig. 3, and the SPT is $(10, 5, 0)$. Adding the plurality and SPT tallies determines the antiplurality tallies (in parenthesis) of respectively 14, 10, 6. The zero SPT value requires the diagonal terms defining C's s coefficient to be zero, so the 4 and 5 for A's and B's plurality tallies must be positioned as indicated in Fig. 3a. It is trivial to find the division of C's six plurality votes which allows the correct antiplurality outcomes.

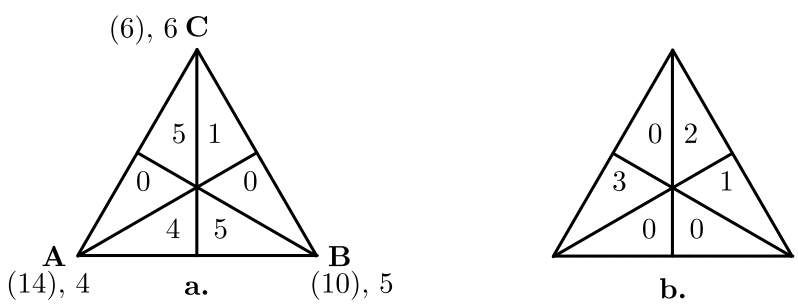


Fig. 3. Creating a profile

For procedure lines on the equilateral triangle, it is easier to describe the plurality and antiplurality endpoints. The following rules follow from properties of the projection.

- Any non-negative rational value can be each candidate's normalized plurality tally as long as the values sum to unity; this defines the plurality endpoint of the procedure line.
- The antiplurality endpoint can be any non-negative rational value which

  - is at least half as large as the assigned plurality value and all values sum to unity,
  - is bounded above by one-half, and
  - for any two candidates, the sum of twice their antiplurality value minus their plurality value is at least as large as the third candidate's normalized plurality value.
- The Borda outcome is two-thirds of the way from the plurality to antiplurality endpoint.

To illustrate, select $(\frac{1}{2}, \frac{1}{6}, \frac{1}{3})$ and $(\frac{1}{4}, \frac{1}{4}, \frac{1}{2})$, respectively, for plurality and antiplurality values. By multiplying the first by 6 and the second by twice this value, integer tallies of $(3, 1, 2)$ and $(3, 3, 6)$ emerge. The corresponding integer profile is in Fig. 3b; fractional values follow by dividing each value by 6. Because fractions are dense, the line segments which depict properties of positional methods can be drawn in almost any way near a complete tie.

Finding new results: It now is easy to find new conclusions about reversal effects. Just draw a line in the triangle — to represent **p**'s procedure line — and use the above structures to compute $\mathcal{R}(\mathbf{p})$'s procedure line. Results follow by comparing differences and similarities of outcomes on the two procedure lines. Thus, all possible results are determined by all possible ways these lines can be drawn. Moreover, a sense about the likelihood of different conclusions is associated with the flexibility in drawing appropriate lines.

We illustrate this approach by using the horizontal procedure line drawn in Fig. 4a which meets seven ranking regions (three regions are lines which represent tie votes). Thus the corresponding profile **p** has seven different election rankings that vary with changes in the positional method; they range from the plurality $A \succ C \succ B$ through $A \succ B \succ C$ and the Borda's $B \succ A \succ C$ to the antiplurality's $B \succ C \succ A$. The Borda tally is identified by the bullet. To find $\mathcal{R}(\mathbf{p})$'s positional line, flip the Borda tally about the center. (Construct a dashed line from Borda tally through the center; $\mathcal{R}(\mathbf{p})$'s Borda tally is equidistant on this line on the other side of the center.) Next, draw a dotted line (on the left) parallel to the dashed line; start it from **p**'s plurality tally. According to Thm. 4, $\mathcal{R}(\mathbf{p})$'s plurality point is on this dotted line and 3/2 as far as the distance between Borda tallies. As these two points define a straight line, they determine $\mathcal{R}(\mathbf{p})$'s procedure line; $\mathcal{R}(\mathbf{p})$'s antiplurality outcome is on the parallel dotted line on the right.

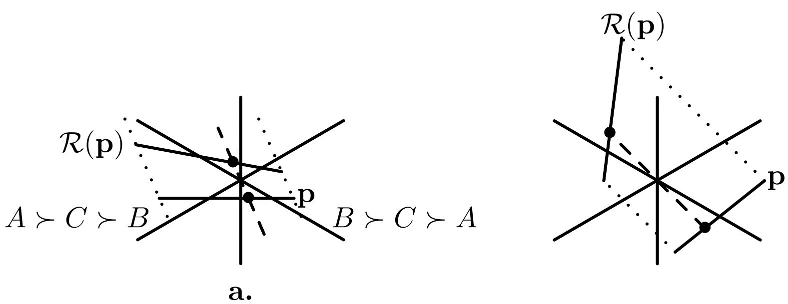


Fig. 4. Finding new results

Figure 4b illustrates this construction with a **p** choice that admits three different election rankings. Notice how the orientation of **p**'s procedure line affects the orientation for $\mathcal{R}(\mathbf{p})$'s line. An amusing example is to choose **p**'s procedure line to be a point. (This profile requires the proportion of the tally assigned to each candidate to be the same for all positional procedures.) The corresponding $\mathcal{R}(\mathbf{p})$ procedure line is a segment (on the line from the point through the center) where all rankings reverse the common **p** ranking.

Conclusions now are apparent. For instance, just by varying the length of the SPT and the location of the Borda tally, the procedure line could allow one, or two, or, ... , or seven different election rankings. An interesting peculiarity of both Fig. 4 diagrams is that the number of rankings allowed by $\mathbf{p}$ and $\mathcal{R}(\mathbf{p})$ agree; this always is the case. Also notice that the closer the Borda tally is to the center — a complete tie — the smaller the changes allowed in election outcomes when reversing the profile.

Theorem 5. *The following statements hold for three-candidate positional method elections.*

1. *For any integer k, $1 \leq k \leq 7$, a profile $\mathbf{p}$ can be found with precisely k different positional method outcomes as the value of s varies; $\mathcal{R}(\mathbf{p})$ also has precisely k different outcomes. (For $k > 1$, some outcomes involve ties).*
2. *All non-Borda positional methods experience the top-two reversal bias; that is, the same two candidates are top-ranked with $\mathbf{p}$ and $\mathcal{R}(\mathbf{p})$, but ranked differently. For instance, $\mathbf{p}$'s plurality tally could be $A \succ B \succ C$ while $\mathcal{R}(\mathbf{p})$'s could be $B \succ A \succ C$.*
3. *All non-Borda positional methods experience a top-winner reversal bias. That is, the same candidate can win with $\mathbf{p}$ and $\mathcal{R}(\mathbf{p})$ but otherwise the rankings differ; e.g., $\mathbf{p}$'s antiplurality ranking could be $A \succ B \succ C$ while $\mathcal{R}(\mathbf{p})$'s could be $A \succ C \succ B$.*
4. *A necessary and sufficient condition for* all *positional methods to have the same ranking and tally for a profile $\mathbf{p}$ and $\mathcal{R}(\mathbf{p})$ is that the Borda ranking is a complete tie.*

The only non-obvious fact (which follows from Thm. 6) is that $\mathbf{p}$ and $\mathcal{R}(\mathbf{p})$ always admit the same number of rankings. The rest of these results can be verified just by drawing lines on the triangle. For instance, no matter how a straight line is drawn, it cannot cross more than seven regions, so the upper bound of part 1 is obvious. Similarly, to create a $\mathbf{p}$ with three, or four, or any other number of outcomes, just draw a line meeting the specified number of regions. To verify the second part, which asserts there is a profile $\mathbf{p}$ with an $A \succ B \succ C$ plurality outcome while $\mathcal{R}(\mathbf{p})$'s plurality outcome is $B \succ A \succ C$, place the plurality endpoint of $\mathbf{p}$'s procedure line in the $A \succ B \succ C$ region near a $A \sim B$ tie; the plurality tip of the $\mathcal{R}(\mathbf{p})$ line will be in the $B \succ A \succ C$ region if you place the Borda point in the $C \succ A \succ B$ region. By using the approach described in the previous section, actual profiles are easy to construct.

The last assertion is the easiest to explain. If the Borda Count is a complete tie, then the $\mathbf{p}$ and $\mathcal{R}(\mathbf{p})$ procedure lines coincide. But when the Borda Count is not a complete tie, the flip which determines the $\mathcal{R}(\mathbf{p})$ Borda tally changes the outcomes for all positional procedures which are sufficiently close to the Borda Count.

These results seem to suggest that almost any outcomes can occur, but this is false. As in Fig. 4, a positional procedure's $\mathbf{p}$ and $\mathcal{R}(\mathbf{p})$ outcomes must be on the same side of the dashed line connecting Borda tallies. This geometry restricts the procedure's allowed outcomes.

Does anything reverse? Intuition suggests that *something* must be reversed when a profile is reversed. This is correct; Thm. 6, which slightly generalizes a result in [Saari, 1995] describes a reversal effect which combines reversals of election rankings, profiles, and the choice of a positional method. To explain the notation, let $f(\mathbf{p}, (1, s, 0))$ be the $(1, s, 0)$ tally for profile $\mathbf{p}$, and let $f^N(\mathbf{p}, (1, s, 0))$ be the normalized tally. Recall, the antiplurality vote is the reversal of the plurality vote as it is equivalent to plurality voting *against* somebody; similarly the $(1, 1-s, 0)$ voting method can be viewed as the reversal of $(1, s, 0)$.

Theorem 6. *For any $\mathbf{p}$ involving n voters and for any s, $0 \leq s \leq 1$, the tallies satisfy*

$$f(\mathbf{p}, (1, s, 0)) + f(\mathcal{R}(\mathbf{p}), (1, 1-s, 0)) = (n, n, n). \tag{8}$$

For normalized tallies, the relationship is

$$(1+s)f^N(\mathbf{p}, (1, s, 0)) + (1 + (1-s))f^N(\mathcal{R}(\mathbf{p}), (1, 1-s, 0)) = (1, 1, 1). \tag{9}$$

The $f(\mathbf{p}, (1, s, 0))$ ranking always is the reversal of the $f(\mathcal{R}(\mathbf{p}), (1, 1-s, 0))$ ranking.

Proof. Candidate A's tally is the number of voters who have her top-ranked plus s times the number who have her second ranked. Tallying $\mathcal{R}(\mathbf{p})$ with $(1, 1-s, 0)$ is equivalent to the number of voters with A bottom ranked plus $(1-s)$ times the number who have her second ranked. As the sum is n, Eq. 8 follows. To derive Eq. 9, normalize the tallies.

The same argument generalizes Eq. 8 from three to $c \geq 3$ candidates. After normalizing the positional weights to $(s_1 = 1, s_2, s_3, \dots, s_{c-1}, s_c = 0)$, Eq. 8 extends to

$$f(\mathbf{p}, (1, s_2, \dots, s_{c-1}, 0)) + f(\mathcal{R}(\mathbf{p}), (1, 1-s_{c-1}, \dots, 1-s_2, 0)) = (n, n, \dots, n).$$

This expression allows the above results to be extended to any number of candidates. □

To illustrate Eq. 8, the 24-voter Fig. 1a profile has the plurality tallies of $(9, 11, 4)$ while the antiplurality tallies for $\mathcal{R}(\mathbf{p})$ are $(15, 13, 20)$. It is immediate that

$$(9, 11, 4) + (15, 13, 20) = (24, 24, 24).$$

As required by Thm. 6, $\mathbf{p}$'s plurality ranking of $B \succ A \succ C$ reverses the $\mathcal{R}(\mathbf{p})$ antiplurality ranking of $C \succ A \succ B$. More generally, the $(1, s, 0)$ tallies for Fig. 1a, and the $(1, 1-s, 0)$ for Fig. 1b, are, respectively, $(9+s, 11+8s, 4+15s)$ and $(14+(1-s), 5+8(1-s), 5+15(1-s))$. We find, as required by Eq. 8,

$$(9+s, 11+8s, 4+15s) + (15-s, 13-8s, 20-15s) = (24, 24, 24).$$

As another example, suppose a 30-voter profile $\mathbf{p}$ is constructed to have a plurality ranking of $A \succ C \succ B$ with tallies of $(16, 4, 10)$. It immediately follows that the antiplurality ranking of $\mathcal{R}(\mathbf{p})$ is $B \succ C \succ A$ with tallies $(30, 30, 30) - (16, 4, 10) = (14, 26, 20)$.

The surprising regularity of positional election rankings offered by Thm. 6 makes it easier to determine all relationships between $\mathbf{p}$ and $\mathcal{R}(\mathbf{p})$ outcomes. For instance, to analyze the top-winner reversal bias for the plurality vote, we need to determine all ways to position the procedure line so that the plurality winner is the same for $\mathbf{p}$ and $\mathcal{R}(\mathbf{p})$. But according to Thm. 6, this situation holds if and only if $\mathbf{p}$'s antiplurality ranking has this same candidate *bottom ranked*. So, rather than needing to construct $\mathcal{R}(\mathbf{p})$ to determine whether this behavior occurs, we can concentrate on properties of procedure lines for $\mathbf{p}$. For instance, one such $\mathbf{p}$ with a top-winner bias has a procedure line with a plurality ranking $A \succ B \succ C$ and an antiplurality ranking $B \succ C \succ A$; this line is easy to draw.

Armed with Thm. 6 we can identify all reversal behavior just from the $\mathbf{p}$ election rankings. To illustrate with a plurality $A \succ B \succ C$ ranking, the following lists all possible antiplurality endpoints. The $\mathcal{R}(\mathbf{p})$ plurality ranking, the reversal effects, and the number of positional method outcomes (the number of regions the positional line crosses) are also specified.

Number of outcomes	$\mathbf{p}$ Antiplurality	$\mathcal{R}(\mathbf{p})$ Plurality	Reversal biases
1	$A \succ B \succ C$	$C \succ B \succ A$	no reversal effects
3	$A \succ C \succ B$	$B \succ C \succ A$	no reversal effects
5	$C \succ A \succ B$	$B \succ A \succ C$	two-winner reversal
7	$C \succ B \succ A$	$A \succ B \succ C$	ranking reversal
5	$B \succ C \succ A$	$A \succ C \succ B$	top-winner reversal
3	$B \succ A \succ C$	$C \succ A \succ B$	no reversal effects

As the above demonstrates, some sort of reversal bias occurs for the plurality vote if and only if the procedure line permits five or more rankings. By using results from [Saari and Tataru 1999], which compute the probabilities that positional methods have specified numbers of outcomes[2], we obtain the likelihoods of different reversal behaviors. Incidentally, it also follows from Thm. 6 that if a condition permits one of these reversal phenomena to occur with the plurality method, the same behavior occurs with the antiplurality method.

Theorem 7. *For three candidates, the following probability statements hold for any probability distribution of voter profiles where, as the number of voters grows, the distribution is asymptotically independent with a common variance, and the mean has an equal number of voters of each type.*

1. *A necessary and sufficient condition for a profile's outcomes of all positional method outcomes to be reversed when the profile is reversed is for* $\mathbf{p}$*'s plurality and antiplurality outcomes to agree. The likelihood of such a behavior is* 0.31.
2. *A necessary and sufficient condition for a reversal effect to occur for the plurality outcome is that a profile's antiplurality outcome reverses the plurality outcome. This behavior occurs with probability* 0.06.
3. *A necessary and sufficient condition for a plurality (or antiplurality) top-reversal, or a two-winner reversal effect is for the profile to allow five different election rankings as the positional methods change (and the plurality outcome to be a strict ranking). This occurs with probability* 0.19

According to this theorem, reversal effects are surprisingly likely. Similar results hold for all $(1, s, 0)$ and $(1, 1-s, 0)$ rules, $s \neq \frac{1}{2}$, but with larger likelihoods for the first assertion and smaller likelihoods for the other two. To extend the second statement, notice that a necessary and sufficient condition for a $(1, s, 0)$ reversal effect to occur is for a profile's $(1, 1-s, 0)$ outcome to reverse the $(1, s, 0)$ outcome; this likelihood diminishes to zero as $s \to \frac{1}{2}$.

Constructing examples: When $\mathbf{p} = \mathcal{R}(\mathbf{p})$ components of a profile cause reversal effects, it is reasonable to anticipate that the more these components dominate a profile, the more dramatic the reversal effects. Not only is this true, but all possible three-candidate differences among positional outcomes — the so-called "election paradoxes" — are completely determined by these reversal terms. In other words, we now know that the huge literature characterizing differences among these procedures merely describes consequences of how these procedures are affected by reversal effects. This assertion follows from a convenient decomposition of profiles which allows us to analyze all possible positional and pairwise elections [Saari 1999].

This decomposition expresses any profile as a union of profiles of four types. To start, a "Neutral" configuration has the same number of voters assigned to each of the six rankings. Second, a "Condorcet" profile configuration affects only pairwise rankings; it is given by Condorcet triplets such as $A \succ B \succ C$, $B \succ C \succ A$, $C \succ A \succ B$, or its reversal. While such a triplet (a Z_3 orbit of a ranking) has no effect on positional rankings (because each candidate is ranked first, second, and third once), it causes pairwise cycles; in part these cycles arise because Z_3 does not admit symmetries (that is, a subgroup) of order two. More is said about this below. The "Reversal" configuration is created by using pairs consisting of a ranking and its reversal; these profile components (the Z_2 orbit of a ranking) do not effect pairwise rankings but, as demonstrated above, they affect positional outcomes. The remaining "Basic" configuration requires all positional and pairwise rankings and tallies to agree.

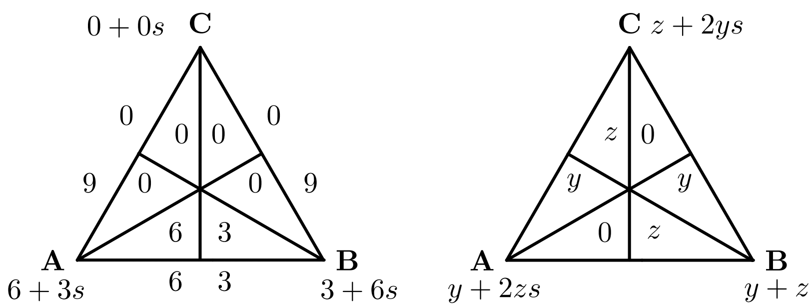
Using a vector approach, all three-candidate profiles can be decomposed into components of these four types where the coefficients may be fractions. As demonstrated by Eq. 5, only the Reversal and Basic directions affect positional outcomes; only the Basic directions affect

[2]The Saari-Tataru approach using procedure lines and differential geometry has subsequently been used by others, primarily various combinations of M. Tataru, V. Merlin, and F. Valgones, to obtain several fascinating results; e.g., see [Tataru, Merlin 1999], [Merlin, Tataru, Valgones 2000]).

the Borda Count. Consequently all differences between the Borda and any other positional ranking are strictly due to Reversal terms [Saari 1999].

The new twist added here is a description how these components geometrically affect the positioning of the procedure line. (Details can be verified by using [Saari 1999].) The procedure line (in a three dimensional space) generated by the Basic portion of a profile is parallel to the $x = y = z$ diagonal. So, for all positional methods the difference of each candidate's Basic tally from the average number of assigned points is the same. All differences in election outcomes, then, are introduced by Reversal terms; they pivot the procedure line about the Borda outcome. If a strong Reversal component creates a $B \succ C \succ A$ plurality outcome, for instance, then the pivoting of the procedure line creates a tendency for the other endpoint — the antiplurality outcome — to define the opposite $A \succ C \succ B$ outcome.

To illustrate by creating examples, start with the Fig. 5a profile. The average number of assigned points per candidate for $(1, s, 0)$ is $[(6+3s)+(3+6s)]/3 = 3+3s$, so each candidate's tally minus this average is $(3, 3s, -3-3s)$. Because these differences change with the s value, it means that the profile has a Reversal component. (This is not obvious; it illustrates a case with a fractional coefficient.) By comparing the differences for the antiplurality ($s = 1$) with the Borda ($s = \frac{1}{2}$) pivot point, the $(3, 3, -6) - (3, 1.5, -4.5) = (0, 1.5, -1.5)$ components show that the effect of this hidden Reversal term is to create a bias for the antiplurality outcome favoring B at the expense of A and C. Indeed, this distortion causes the antiplurality ranking of $A \sim B \succ C$ to conflict with the $A \succ B \succ C$ conclusion for all other positional procedures.



a. Starting profile

b. Reversal terms

Fig. 5. Creating examples

To make these statements more concrete, we modify the Fig. 5a profile to create a $\mathbf{p}$ with a plurality top-winner reversal bias; $\mathbf{p}$'s plurality ranking will be $B \succ A \succ C$ and $\mathcal{R}(\mathbf{p})$'s plurality ranking will be $B \succ C \succ A$. To achieve this goal, we need to add reversal terms. That is, select y and z values from Fig. 5b so that adjoining it to the profile of Fig. 5a gives the plurality $B \succ A \succ C$ and antiplurality $A \succ C \succ B$ rankings.

Adding each candidate's tallies from the Figs. 5a, b triangles, the desired $B \succ A \succ C$ plurality outcome and $A \succ C \succ B$ antiplurality outcomes occur, respectively, if and only if

$$z + y + 3 > y + 6 > z,$$

or

$$6 + y > z > 3, \tag{10}$$

and

$$9 + y + 2z > z + 2y > 9 + y + z,$$

or

$$9 + z > y > 9. \tag{11}$$

The simplest choice of $y = 10, z = 4$ defines the $\mathbf{p}$ in Fig. 6a; $\mathcal{R}(\mathbf{p})$ is in Fig. 6b. While the Fig. 6a plurality ranking changes from that of Fig. 5a, the Borda ranking remains the same reflecting Borda's immunity to reversal terms.

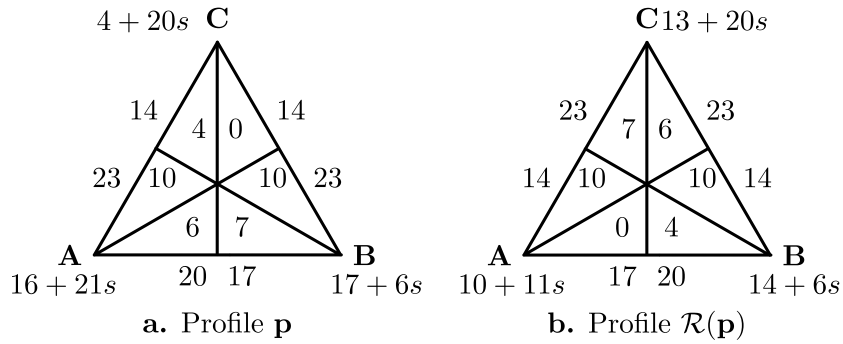


Fig. 6. Final example

This construction can also be used to demonstrate how the "size" of the Reversal term affects the $\mathcal{R}(\mathbf{p})$ outcomes. While all Eq. 10 choices define a $\mathbf{p}$ with a plurality $B \succ A \succ C$ ranking, we know from the properties of the procedure line that different Reversal components generate different $\mathcal{R}(\mathbf{p})$ plurality rankings. To illustrate with the smallest value of $z = 4$, observe in Eq. 12 how the $\mathcal{R}(\mathbf{p})$ plurality ranking changes as the y value increases; the $\mathcal{R}(\mathbf{p})$ outcome moves through five different rankings until the $y \geq 14$ values require the plurality ranking to be the same for $\mathbf{p}$ and $\mathcal{R}(\mathbf{p})$. As $\mathbf{p}$'s Borda ranking is immune to Reversal terms, it remains A $\succ B \succ C$; the $\mathcal{R}(\mathbf{p})$ Borda ranking reverses $\mathbf{p}$'s Borda outcome to become $C \succ B \succ A$.

y **value,** $z = 4$	$\mathcal{R}(\mathbf{p})$ **plurality ranking**	**Reversal behavior**
$0 \leq y \leq 8$	$C \succ B \succ A$	–
$y = 9$	$C \sim B \succ A$	–
$10 \leq y \leq 12$	$B \succ C \succ A$	Top-winner
$y = 13$	$B \succ C \sim A$	Top-winner
$y \geq 14$	$B \succ A \succ C$	Reversal

(12)

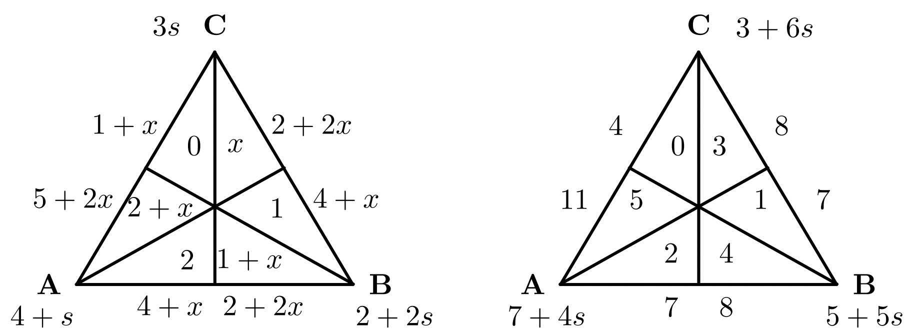
These Reversal terms provide a tool which now makes it trivial to create paradoxical examples. Of interest for our earlier claim, because these terms are fully responsible for all possible differences among three-candidate positional method election outcomes, they explain this two century mathematical mystery about election procedures.

Using parts

Now consider those election methods which are based on pairwise majority votes. As illustrated by the pairwise tallies in Figs. 1 and 6, $\mathcal{R}(\mathbf{p})$ always reverses $\mathbf{p}$'s pairwise rankings and tallies; this suggests that maybe pairwise procedures never suffer reversal problems. After all, should the pairwise rankings form a transitive ranking, then any reasonable procedure will select the top-ranked candidate. But $\mathcal{R}(\mathbf{p})$ reverses the rankings, so $\mathbf{p}$'s bottom ranked candidate becomes $\mathcal{R}(\mathbf{p})$'s top ranked candidate; reversal biases cannot occur.

The reason difficulties arise is that there are $2^{\binom{n}{2}}$ ways to rank the $\binom{n}{2}$ pairs of the n candidates. Consequently, reversal problems may be created by the way these procedures handle the $2^{\binom{n}{2}} - n!$ *non-transitive* pairwise outcomes. For n=3, there are only $2^3 - 6 = 2$ possibilities, but for $n = 4$ there are $2^6 - 24 = 40$ such situations. Since the non-transitive settings significantly outnumber the transitive ones once $n \geq 4$, plenty of opportunities exist for unexpected behavior. What helps in our analysis is that we now know [Saari 1999, 2000] that all non-transitive settings for n candidates — hence all possible reversal problems — are caused by the "Z_n cyclic symmetry orbits" of the n alternatives; these are natural generalizations of the Condorcet triplets. To construct such a profile component, start with any n-candidate ranking such as $A \succ B \succ C \succ \cdots \succ Z$. For the second ranking, move the top-ranked candidate to the bottom to have $B \succ C \succ \cdots \succ Z \succ A$. Continue until there are n rankings. With n candidates, this *Condorcet n-tuple* creates the cyclic outcomes $A \succ B, B \succ C, \dots Z \succ A$, each with $n-1:1$ tallies.

To see how reversal problems can occur, consider an *agenda.* This is a form of a tournament where candidates are compared with a pairwise vote in a specified manner; after each comparison the winner is advanced to be compared with the next specified candidate. One example, then, is where the winner of an A and B pairwise vote is compared with C. With a Condorcet triplet $\{A \succ B \succ C,\ B \succ C \succ A,\ C \succ A \succ B\}$, A beats B to advance to a vote with C; C wins by a 2:1 vote. $\mathcal{R}(\mathbf{p})$ is the reversed Condorcet triplet $\{C \succ B \succ A,\ B \succ A \succ C,\ A \succ C \succ B\}$ with the opposite pairwise cycle of $A \succ C$, $C \succ B$, and $B \succ A$ with the 2:1 tallies. With $\mathcal{R}(\mathbf{p})$, B beats A in the first comparison, but loses to C in the second election. Since C is the winner with both $\mathbf{p}$ and $\mathcal{R}(\mathbf{p})$, an agenda admits the top-winner reversal bias.



a. Profile $\mathbf{p}_1$ is where $x = 0$ **b.** Profile $\mathbf{p} = \mathbf{p}_1 + \mathbf{p}_2$

Fig. 7. Adding cycles

This cyclic effect for the agenda example suggests that *all positional method runoff procedures* — where the top two candidates in a positional election are advanced to a majority vote runoff — allow a top-winner reversal bias. To explain, if profile $\mathbf{p}$ satisfies Eq. 4 with an $A \succ B \succ C$ outcome, then A and B are advanced to the runoff while the $\mathcal{R}(\mathbf{p})$ outcome of $C \succ B \succ A$ advances B and C to the runoff. To have a top-winner reversal bias, the profile needs to have a cyclic effect where B beats A with $\mathbf{p}$ and B beats C with $\mathcal{R}(\mathbf{p})$. Again, what simplifies the construction is that such an example requires C to beat B with $\mathbf{p}$.

To construct illustrating examples, notice that a Condorcet n-tuple does not effect positional method election rankings (as each candidate is ranked in each position once). So, as illustrated in Fig. 7, create a $\mathbf{p} = \mathbf{p}_1 + \mathbf{p}_2$ where (according to Thm. 7) $\mathbf{p}_1$ has the same plurality and antiplurality ranking and $\mathbf{p}_2$ defines an appropriate cycle. Profile $\mathbf{p}_1$ is given in Fig. 7a for $x = 0$; the positional method outcomes are $A \succ B \succ C$ with $4 + s : 2 + 2s : 3s$ tallies.

The $\mathbf{p}_2$ portion is the Condorcet triplet given by the x's in Fig. 7a. As these terms add the same $x + xs$ value to each candidate's positional method tally, they do not affect the positional method rankings. But as indicated in the figure, the x terms can change the pairwise rankings. In particular, for B to beat A, and C to beat B with $\mathbf{p}_1 + \mathbf{p}_2$, select x where $2 + 2x > 4 + x$; i.e., $x \geq 3$. The $x = 3$ choice in Fig. 7b defines a $\mathbf{p}$ with the top-winner bias for any positional method runoff.

A word of caution; not all elimination procedures suffer these reversal problems. An example is Nanson's method [Nanson 1882] which, at each stage, drops all candidates who fail to receive more than the average Borda score; the remaining candidates are reranked with the Borda method and the process continues until a single candidate remains. Since the Nanson winner survives the first cut with $\mathbf{p}$, when the average Borda score is subtracted from her Borda tally, it must be positive. But, as demonstrated earlier, with $\mathcal{R}(\mathbf{p})$ this difference is negative. Because the Nanson winner with $\mathbf{p}$ is dropped at the first stage with $\mathcal{R}(\mathbf{p})$, there is no reversal problem. So while a Borda runoff can suffer a top-reversal bias, Nanson's approach never does.

To obtain a general result for the $n \geq 3$ alternatives $\{a_1, a_2, \dots, a_n\}$, represent the tallies for the $\binom{n}{2}$ pairs with a point in $R^{\binom{n}{2}}$. To do so assign an axis for each $\{a_j, a_k\}$ pair. The value used for a $\{a_j, a_k\}$ tally is the difference between a_j's and a_k's vote divided by the number of voters. Thus, the outcomes are on the $[1, -1]$ interval of this axis, where 1 means that a_j wins unanimously, 0 means a tie, and -1 means that a_k wins unanimously. All pairwise outcomes, then, are in a cube of $R^{\binom{n}{2}}$ centered at the origin $\mathbf{0}$ called the *representation cube.*[3] The coordinate planes define $2^{\binom{n}{2}}$ orthants in the representation cube; each orthant contains all pairwise tallies supporting a specific choice of pairwise rankings.

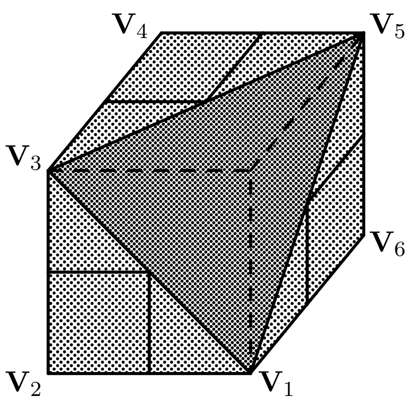


Fig. 8. Representation Cube

To illustrate with $n = 3$ and Fig. 8, let the x, y, z coordinates represent, respectively, the rankings $A \succ B, B \succ C, C \succ A$. While the cube $[-1, 1]^3$ has eight vertices, only six of them can be identified with the six transitive rankings. It is not difficult to show that the labeled vertices in Fig. 8 correspond to transitive rankings; for instance $\mathbf{V}_1$ corresponds to $A \succ B$, $B \succ C$, $A \succ C$ or $A \succ B \succ C$. The two remaining vertices, $(1, 1, 1)$ and $(-1, -1, -1)$, correspond to cyclic rankings.

The representation cube is the convex hull of the six labelled vertices; it turns out [Saari 1995] that the rational points in this hull represent all possible pairwise election outcomes. Notice that this hull meets the positive and negative orthants; the points in these two orthants are the pairwise cyclic outcomes that can cause problems. Indeed, the 15-voter Fig. 7b choice of $\mathbf{p}$ defines the point $(\frac{7-8}{15}, \frac{7-8}{15}, \frac{4-11}{15})$; as all components are negative the election rankings form the cycle $B \succ A, C \succ B, A \succ C$. The $\mathcal{R}(\mathbf{p})$ point reverses each sign; it is $(\frac{8-7}{15}, \frac{8-7}{15}, \frac{11-4}{15})$. For any n, the $\mathbf{p}$ and $\mathcal{R}(\mathbf{p})$ tallies differ only by the sign of each component, so they are endpoints of a line segment with $\mathbf{0}$ as the midpoint. This statement introduces a geometric test for a top-winner reversal bias. (Procedures mentioned in this theorem which have not been introduced are described below.)

[3]The set of all admissible pairwise tallies is a subset which can be determined; this is done in [Saari 1995] for $n = 3$, and a similar approach holds for all n.

Theorem 8. *Suppose a specified election method using pairwise votes is given. For each candidate, find all points in the representation cube which elect that candidate. If a line segment of positive length centered at* $\mathbf{0}$ *with the* $\mathbf{p}$ *and* $\mathcal{R}(\mathbf{p})$ *pairwise outcome as endpoints has both endpoints in the same candidate's region, then the procedure has a top-winner Reversal bias. Thus, for instance, agendas and Dodgson's method (for* $n \geq 4$*) have the top-winner Reversal bias. If all such line segments have the endpoints in regions for different candidates, then the method never has a top-winner Reversal bias. As examples, Copeland's, Borda's and Kemeny's methods never experience a top-winner Reversal bias.*

To illustrate Thm. 8, consider the agenda where the winner of an A and B pairwise vote is compared with C. Each orthant of the representation cube in $R^{\binom{3}{2}} = R^3$ determines a specific agenda winner. But $3! = 6$ of these eight orthants represent transitive rankings where the top-ranked candidate is the agenda winner. Since reversing a transitive ranking makes the previously bottom-ranked candidate top-ranked, none of these six regions pass the line segment test. It remains to examine the two remaining orthants where the pairwise rankings define cycles. Both orthants elect C and they are diametrically opposite one another, so a top-winner reversal bias must occur with any profile which allows a cycle.

An intriguing election approach was introduced by the mathematician Charles Dodgson who is better known as Lewis Carroll of *Alice in Wonderland* fame. Dodgson's method selects the Condorcet winner — the candidate who beats all others in pairwise comparisons. If a Condorcet winner does not exist, replace the actual rankings with the "closest" set of rankings which have a Condorcet winner. For Dodgson, "closest" is the minimum number of adjacent changes in individual rankings which create a new profile with a Condorcet winner. Ratliff [2001, 2002, 2003] has discovered a surprising array of unexpected behaviors allowed by this procedure.

Dodgson's method selects the top-ranked candidate from a transitive ranking, so ignore the $n!$ orthants with transitive outcomes. Similarly, suppose the non-transitive rankings for $\mathbf{p}$ define a Condorcet winner and a Condorcet loser (a candidate who loses to each of the other candidates). Since the reversal converts $\mathbf{p}$'s Condorcet loser into $\mathcal{R}(\mathbf{p})$'s Condorcet winner, no reversal bias occurs. More generally, any $\mathbf{p}$ which defines a Condorcet loser which differs from the Dodgson winner cannot have the top-winner reversal bias.

Next consider profiles with Condorcet winner A but no Condorcet loser. Because A is the $\mathcal{R}(\mathbf{p})$ Condorcet loser, it is reasonable to suspect that nothing can go wrong. What makes the actual story more complicated is that $\mathcal{R}(\mathbf{p})$ has no Condorcet winner, so we need to invoke Dodgson's metric. The problem arises if A barely is a Condorcet winner with $\mathbf{p}$ — so she barely is a Condorcet loser with $\mathcal{R}(\mathbf{p})$ — and the tallies for all other pairwise rankings involve substantial differences. Such a situation requires cyclic symmetries [Saari 2000a]. Combining these two notions, examples are immediate; e.g., the next profile repeatedly uses the Condorcet $\{B, C, D\}$ triplet to create sizeable differences in their pairwise tallies. To ensure that A barely is the Condorcet winner, she is top-ranked in slightly over half of the preferences, and she is bottom ranked in the others.

Number	**Ranking**	**Number**	**Ranking**
10	$A \succ B \succ C \succ D$	9	$B \succ C \succ D \succ A$
10	$A \succ C \succ D \succ B$	9	$C \succ D \succ B \succ A$
10	$A \succ D \succ B \succ C$	9	$D \succ B \succ C \succ A$

(13)

A is the Condorcet winner by beating the other candidates with a 30:27 tally. Here, $\mathcal{R}(\mathbf{p})$ is

Number	Ranking	Number	Ranking
10	$D \succ C \succ B \succ A$	9	$A \succ D \succ C \succ B$
10	$B \succ D \succ C \succ A$	9	$A \succ B \succ D \succ C$
10	$C \succ B \succ D \succ A$	9	$A \succ C \succ B \succ D$

(14)

where A is the Condorcet loser since she loses to each opponent with a 30:27 tally. The remaining rankings define the $B \succ D, D \succ C, C \succ B$ cycle with 38:19 tallies. Without a $\mathcal{R}(\mathbf{p})$ Condorcet winner, we need to invoke Dodgson's metric; the Dodgson winner is A. Indeed, interchange the last pair for two individuals in each ranking on the left of Eq. 14, the revised rankings allow A to beat each of the other candidates by 29:28 to become the Condorcet winner. Thus, Dodgson's method admits a top-winner reversal bias.

Other methods, such as the one developed by the mathematicians Borda [1781], Copeland [1951] and Kemeny [1957], and Dodgson's method for $n = 3$ do *not* have a reversal bias, because these methods replace the actual pairwise rankings with the "nearest" *transitive* ranking. For instance, Saari and Merlin [2000] showed that the Kemeny method can be viewed as finding the nearest transitive ranking with a l_1 metric — the sum of the difference between coordinates. With $n = 3$ and Dodgson's metric, the nearest region with a Condorcet winner is either transitive, or on the boundary of a transitive orthant. The other two methods use the *transitivity plane* introduced in [Saari 1999, 2000b]; it is a lower-dimensional plane symmetrically positioned in the representation cube passing through the origin and transitive orthants. Borda's method can be viewed as replacing a point in the representation cube with the nearest (l_2 or Euclidean distance) point on the transitivity plane. Copeland's method converts each pairwise tally into a 1 or -1, indicating who won or lost, and sums the tallies; i.e., it replaces a point in an orthant of the representation cube with the outside vertex of that orthant. Then, Copeland's method replaces the vertex with the nearest (l_2) point on the transitivity plane.

It is easy to show that the distance from a point in the representation cube defined by $\mathbf{p}$ to one of these regions is the same as the distance from the point defined by $\mathcal{R}(\mathbf{p})$ to the reversal of these regions. But these reversed regions define reversed transitive rankings, so the ranking of $\mathcal{R}(\mathbf{p})$ reverses that given by $\mathbf{p}$.

Final comments

On first glance the study of elections seems to be trivial because, seemingly, only counting is involved. From a mathematical perspective, however, everything become delightfully complex. As we have recently learned, an important source of the mathematical complexity is that profiles can be full of hidden symmetries from higher dimensional spaces; symmetries which cause all sorts of unanticipated problems and difficulties for election procedures. The reversal problems identify only a small portion of the tip of a very big iceberg of complexity. Of interest, this structure extends to problems from statistics, probability, and other aggregation methods; different symmetry groups are needed, but the ideas are similar.

As an illustration of related issues, consider strategic voting — something all of us have done. For instance, if you have $A \succ B \succ C$ preferences in a close election between A and B, you might be tempted to mark your ballot as $A \succ C \succ B$ to increase A's point spread over B. More generally, Gibbard [1973] and Satterthwaite [1975] proved the amazing result that all reasonable election procedures for three or more candidates admit situations where some voter, by voting strategically, gets a better election outcome. But if all methods admit strategic options, the next natural question is to determine which $(1, s, 0)$ method is least susceptible to a small number of strategic voters being successful. The answer [Saari 1995] is Borda's method; the level of susceptibility decreases as $s \to \frac{1}{2}$. (According to this theorem, the plurality vote is highly susceptible to strategic behavior. We know this; just recall those "Don't waste your vote" calls for strategic action voiced during close elections involving more than two candidates.) But as $s \to \frac{1}{2}$ a procedure becomes less susceptible to the Reversal components. Is there a connection? Probably, but it has not been established.

References

[1] Borda, J. C. 1781, *Mémoire sur les élections au scrutin*, Histoire de l'Académie Royale des Sciences, Paris.
[2] Condorcet, M. 1785, *Éssai sur l'application de l'analyse à la probabilité des décisions rendues à la pluralité des voix*, Paris.
[3] Copeland, A.H. 1951, A reasonable social welfare function. *Mimeo,* University of Michigan.
[4] Gibbard, A., 1973, Manipulation of voting schemes: a general result, *Econometrica* **41**. 587-601.
[5] Kemeny J., 1959, Mathematics without numbers. *Daedalus* **88**, 571-591.
[6] Merlin, V., M. Tataru, and F. Valognes 2000, On the probability that all decision rules select the same winner, *Journal of Mathematical Economics* **33**, 183-208.
[7] Nanson, E. J., 1882, Methods of elections, *Trans. Proc. R. Soc. Victoria* **18**, 197-240.
[8] Nurmi, H., 1999, *Voting Paradoxes and How to Deal with Them*, Springer-Verlag, NY.
[9] Nurmi, H., 2002, *Voting Procedures under Uncertainty* Springer, Heidelberg.
[10] Ratliff, T. 2001, A comparison of Dodgson's method and Kemeny's rule, *Social Choice & Welfare* **18**, 79-89.
[11] Ratliff, T. 2002, A comparison of Dodgson's Method and the Borda Count, *Economic Theory* **20**, 357-372.
[12] Ratliff, T. 2003, Some startling paradoxes when electing committees, to appear *Social Choice & Welfare*
[13] Riker, W. H., 1982, *Liberalism Against Populism*, W. H. Freeman, San Francisco.
[14] Saari, D. G., 1992, Millions of election rankings from a single profile, *Soc. Choice & Welfare* (1992) **9**, 277-306
[15] Saari, D. G. 1994, *Geometry of Voting*, Springer-Verlag, New York.
[16] Saari, D. G. 1995, *Basic Geometry of Voting*, Springer-Verlag, New York.
[17] Saari, D. G. 1999, Explaining all three-alternative voting outcomes, *Journal of Economic Theory* **87**, 313-355.
[18] Saari, D. G. 2000a, Mathematical structure of voting paradoxes 1; pairwise vote. *Economic Theory* **15**, 1-53.
[19] Saari, D. G. 2000b, Mathematical structure of voting paradoxes 2: positional voting. *Economic Theory* **15**, 55-101
[20] Saari, D. G. 2001, *Chaotic Elections! A Mathematician Looks at Voting*, American Mathematical Society, Providence, RI.
[21] Saari, D. G., and V. Merlin 2000, A geometric examination of Kemeny's rule, *Social Choice & Welfare* **17**, 403-438.
[22] Saari, D. G., and J. Van Newenhizen, 1988, Is Approval Voting an "unmitigated evil?" *Public Choice* **59**, 133-147.
[23] Saari, D. G. and M. Tataru 1999, The likelihood of dubious election outcomes, *Economic Theory* **13**, 345-363.
[24] Satterthwaite, M., 1975, Strategyproofness and Arrow's conditions, *Jour. Econ. Theory* **10**, 187-217.
[25] Tabarrok, A., 2001, Fundamentals of voting theory illustrated with the 1992 election, or could Perot have won in 1992? *Public Choice* **106**, 275-297.
[26] Tabarrok, A. and L. Spector 1999, Would the Borda Count have avoided the civil war? *Journal of Theoretical Politics* **11**, 261-288.
[27] Tataru, M., and V. Merlin 1997, On the relationships of the Condorcet winner and positional voting rules, *Mathematical Social Sciences* **34**, 81-90.