Which voting rule is most likely to choose the best option?

T. Nicolaus Tideman

Department of Economics, Virginia Polytechnic Institute and State University

Blacksburg, VA 24061

ntideman@vt.edu

Florenz Plassmann*

Department of Economics, State University of New York at Binghamton

Binghamton, NY 13902-6000

fplass@binghamton.edu

This version: October 26, 2011

Abstract:

One criterion for determining whether a voting rule is attractive is the frequency with which the rule identifies the best option. We use a spatial model of voting that we calibrate to data from actual elections to simulate elections for which we can define which option is best. We use these simulated data to investigate the frequencies with which eleven voting rules chose this option as their winner. We find that the Black rule tends to perform better than the other rules, especially in elections with few voters. The Bucklin rule and the plurality rule tend to perform worse than the other nine rules, especially when the number of voters gets large.

Journal of Economic Literature Classification Codes: C4, D72

Keywords: spatial model, multinomial-Dirichlet distribution, the Black rule

* Corresponding author. Phone: (607) 777-4304, Fax: (607) 777-2572.

1. INTRODUCTION

When making collective decisions, it is reasonable to suppose that some options are better than others. Otherwise we could make collective decisions by drawing lots. So if we decide to make decisions through voting, it can be presumed that we believe that voting can help us avoid inferior outcomes.[1] Thus a principal purpose of voting is to identify the best option. But which voting rule is most likely to accomplish this task? In this paper we assess the probabilities with which eleven voting rules identify the best option. We undertake our analysis for choices over sets of three options, but in principle it can be extended to choices over sets of more than three options.

It might seem that this effort ought to be futile because it is impossible to determine the true merit of the options under consideration. We resolve this difficulty by evaluating voting rules not on observed voting data but rather on data that we simulate with a statistical model of disagreements among voters about which option is best. In Tideman and Plassmann (2012) we survey, for elections with three options, several statistical models of vote casting that could be used for this task, and we identify one model—a spatial model of voting—that describes observed voting data much better than any of the other models. In Plassmann and Tideman (2011) we compare data that we simulated under the spatial model with voting data from actual elections and find that the simulated data are astonishingly similar to the observed voting data. Thus we are in a position to define the relative merit of three options, use an appropriately calibrated statistical model to simulate ballot data, and use different voting rules on these simulated data to assess how frequently each voting rule declares the best option as the winner. Because our simulated data have the same statistical properties as data from actual elections, we ascribe the success rates of voting rules in finding the best option in our simulations to those voting rules in real voting situations.

In section 2 we discuss the circumstances under which it is meaningful to consider the notion of a best option that can be identified through voting. Here we note that, if one accepts the notion of a best option, then one can interpret voting rules as estimators and winning options as estimates of the best option. This interpretation of voting rules has a long history—for

[1] It is possible to advocate voting on the grounds that, while it would be preferable to make collective decisions in ways other than through voting, abolishing voting would be politically infeasible. We suspect that this is the case because the electorate believes, maybe mistakenly, that voting is likely to result in better outcomes than if we made collective decisions in other ways.

example, Condorcet's goal was to develop a voting rule that was most likely to identify the socially optimal outcome.[2] Young (1986, 1988) shows that Condorcet's voting rule is a maximum likelihood estimator of the best option under specific assumptions regarding the probability structure that describes how voters rank the different options. Young (1986, 1988) also shows that, under different assumptions regarding this probability structure, the voting rule proposed by Jean-Charles de Borda in 1770 is a maximum likelihood estimator of the best option.[3] Conitzer and Sandholm (2005) identify conditions under which various voting rules can be interpreted as maximum likelihood estimators of the best option.

Although many voting rules can be interpreted as estimators, whether a voting rule is a good estimator in a specific election depends first and foremost on conformity between the statistical model to which the voting rule relates and the statistical properties of the distribution of the ballots in question. Statisticians have developed a large number of estimators, and it is understood that these estimators have desirable properties (for example, unbiasedness or efficiency) if they are used on the right data but that they can be highly unreliable if they are used on inappropriate data. For example, a maximum likelihood estimator based on the normal distribution is likely to be a good estimator of, say, the mean if it is used on symmetric data with a range from minus infinity to plus infinity, but a very poor estimator of the mean if it is used on highly skewed data. Thus the question of interest is not whether a voting rule can be interpreted as a maximum likelihood estimator in general but rather whether the ballots on which the rule is used have the statistical structure for which this rule is a good estimator of the best option.[4]

In Tideman and Plassmann (2012) we assess how well different models of vote casting describe the statistical structure of two different data sets that we assembled from actual elections and from surveys. We find that neither the model that supports the Borda rule nor the model that

---

[2] See Condorcet (1785).

[3] See de Borda (1784). Drissi-Bakhkhat and Truchon (2004) and Truchon (2008) offer additional insights regarding the interpretation of the Borda rule as a maximum likelihood estimator.

[4] Young (1986) derives a likelihood function of the relative probability that a particular option is ranked first, which depends on the unknown probability that this option is the best option. Young suggests that one can either estimate the most likely value of this probability from the data or treat all possible probabilities as being equally likely. He considers both strategies to be unpromising. Instead, he shows that the Borda rule is the maximum likelihood estimator of the best option if this probability is small and that the Kemeny rule (the rule Young believes to have been proposed by Condorcet) is the maximum likelihood estimator of the best option if this probability is large. But what *is* the magnitude of this probability and how does it vary across elections? This is the question that needs to be answered before one can recommend a particular voting rule, and estimating this probability from the data (as we do) seems to be the only possible way to answer the question.

supports Condorcet's rule as a maximum likelihood estimator is capable of simulating data with a structure anywhere close to that of the observed data. Instead we find that a spatial model of voting comes closest to describing the structure of observed voting data. By itself this does not imply that either the Borda rule or Condorcet's rule—or any other voting rule for that matter—is a bad estimator of the best option in actual elections. The goal of this paper is to find out precisely how well each rule performs and how likely it is that any of these rules identifies the best option for simulated data that are designed to have the same structure as data from actual elections. Our main conclusion is that the Black rule, which combines the insights of Condorcet and Borda, is the most attractive voting rule in elections with up to a few hundred voters, at least if all voters can be expected to vote sincerely. For elections with more than a few hundred voters, the likelihood of identifying the best option varies very little across most voting rules, although two voting rules—the Bucklin rule and the plurality rule—tend to do notably worse than the other rules that we examine.

The remainder of this paper is organized as follows. In section 2 we discuss different ways in which it is meaningful to identify an option as being best, and we describe our spatial model of voting and our simulation technique in section 3. In section 4 we describe the eleven voting rules and report our results. Section 5 concludes.

2. What does it mean to choose the "best" option?

It is not possible to claim rationally that an option is best on the grounds that the option has been declared the winner by a voting rule, because all reasonable voting rules sometimes produce inconsistencies when the set of available options varies.[5] The notion of a "best option" is meaningful only if the identity of such an option is independent of the outcome of any voting procedure. There are different conceptions of what it means to say that an option is best. For example, there might be an objectively true ranking of options that exists independently of human conceptions of goodness. Moral philosophers have suggested ways of assessing whether

[5] Consider the situation of a majority rule cycle in which each option is beaten by at least one other option in pairwise comparison. All voting rules (augmented perhaps by a tie-breaking procedure) nevertheless identify a winning option when there is a cycle. But if the only options under consideration were the winning option and one to which the winning option loses in the cycle, then none of the voting rules that are consistent with majority rule for pairs of options would still choose the previously winning option. Thus eliminating some of the non-winning options from the set of options might cause a voting rule to declare a different option as winner. The central message of the Arrow theorem is that this inconsistency across available sets of options occurs with all reasonable voting rules. See Tideman (2006), Chapter 10 for details.

some options are objectively better than others, and many philosophers abhor the notion of moral subjectivism. If we believe that some persons—for example, the Pope or Plato's philosopher king—are uniquely qualified to identify the objectively best option, then asking the general public for input is unlikely to provide useful information. But if we believe that all members of the electorate are born with implicit understandings that provide, on average, unbiased estimates of the true ranking of options, then it is reasonable to use voting as a way of estimating the best option.[6]

Many people either reject the notion of an objectively true ranking of options or consider it impossible for humans to identify such an objectively true ranking, if it does exist. In our empirical analysis we do not assume that it is possible to identify an objectively best option. However, by evaluating the attractiveness of options in relation to the voters' subjective preferences, it is nevertheless possible to employ the concept of a best option.[7] We do this by defining the best option as the one that comes closest to the average of the distribution from which everyone's ideal option is drawn. We define "average of the distribution from which ideal options are drawn" and "closest" in the following way. Assume that all options are defined through "attributes" that form a multi-dimensional attribute space. Every option has a unique position in attribute space. It is not necessary to assume that it is possible to measure and specify these attributes objectively but only that voters are aware of them. Each person has an ideal point in attribute space that describes the position that the person's ideal option would possess. A person's preferences define a set of multi-dimensional indifference contours around his ideal point. These contours describe how he values options at other positions. If there are m options, then they span an "option space" that is a subset of attribute space. If attribute space has at least $m - 1$ dimensions and the options are in "general position," so that the dimensionality of the space that they span is constant for slight changes in their positions, then the space they span has $m - 1$ dimensions. The position of each person's ideal point and the associated indifference contours are defined in option space through their definition in attribute space. Throughout this paper we assume that the voters' indifference contours in option space are concentric ($m - 1$)-

[6] For example, voting with the goal of identifying the objectively best option would be meaningful under the Lockean notion that reason can help us to understand natural law or under the sentimentalist notion that humans are born with moral feelings.

[7] More precisely, the best option should be defined in relation to the preferences of all *citizens*. Although one might want to distinguish between citizens and voters conceptually—because not all citizens might decide, be able, or be allowed to vote—we assume in this paper that all citizens are voters to simplify the discussion.

spheres around their ideal points. This assumption is not necessary conceptually, but it simplifies the analysis.

The attractiveness of an available option depends on our understanding of what the ideal points represent. One interpretation is that the ideal points represent the voters' views of the best option from a social perspective. If we assume that voters are on average unbiased when they assess the social merit of attributes, then the best conceivable option is the one at the center of the distribution from which ideal points are drawn.[8] If we also assume that any dependence among their views is not specific to identifiable persons, then the appropriate estimate of the best option among the available options is the one that is closest to the mean of the distribution of the voter ideal points projected orthogonally onto option space.

An alternative interpretation is that the ideal points represent voters' private interests that may be unrelated to the social good. In this case it is meaningful to call the option best that maximizes the aggregate welfare of all persons, which can also be described as the option that minimizes the aggregate loss of all persons from having an option that is not their ideal. To quantify the aggregate loss, one can assume that for every person, the loss from the choice of an option that the person does not regard as ideal is proportional to the distance from this option to the person's ideal point. This assumption implies that the aggregate loss from an option over all persons is proportional to the sum of the distances from the location of the option to the persons' ideal points. In view of the symmetry of the distribution from which voters' ideal points are drawn, in the limit as the population approaches infinity, the option with the smallest aggregate loss is the one at the center of the distribution of the voter ideal points.

For our empirical analysis we adopt the second of these three conceptions of "best"—that the ideal points represent the voters' views of the merit of options from a social perspective and that their views come from a distribution that is on average unbiased and not dependent in any way that is particular to individuals. Thus the appropriate estimate of the option that is best is the one that is closest to the mean of the distribution from which voter ideal points are drawn. In the following section we describe how we use the spatial model to infer this information from

[8] In this context "unbiased" means that voters evaluate the attributes according to what they think would be socially optimal rather than what they prefer privately. For the purpose of defining conceptually which option is best, it is necessary only that voters are honest with themselves when evaluating the attributes, not that they announce their assessments honestly. To identify the option at the center of the distribution of ideal points as best, we use the assumption that voters' indifference contours are concentric circles around the voters' ideal points.

observable rankings and how we use these rankings to calibrate the model's unknown parameters. We do not consider the two other interpretations of "best" because we do not know how to determine the distribution of successes and failures in identifying the best option under these interpretations. The notion that there are objectively true rankings implies that the best option is independent of the voters' conception of what is best, and we find it difficult to defend either a particular philosophical framework that would specify the distribution of objective rankings or the assumptions we would need to make about the voters' knowledge of these rankings that would permit us to use observable ballots to calibrate the distribution of objective rankings. Similarly, applying the notion that the ideal points represent the voters' private interests requires more information about the locations of different options than we can infer from the spatial model. We discuss this issue in more detail in the following section, where we describe the spatial model.

3. A STATISTICAL MODEL OF VOTE CASTING

3. 1. *A spatial model of vote share probabilities*

The spatial model of voting represents a conceptualization of the distribution of voter ideal points and the locations of options in option space. Good and Tideman (1976) describe the philosophy and the technical details of the spatial model that we use; here we summarize its essential elements. Consider the case of m = 3 options. If the positions of these options in attribute space are not collinear, then the three options form a two-dimensional option space, that is, a plane. Label the options *A*, *B*, and *C*, so that the 3! = 6 different strict rankings are *ABC*, *ACB*, *CAB*, *CBA*, *BCA*, and *BAC*.

We assume that the locations of voter ideal points in option space follow a circularly symmetric bivariate normal distribution. The probability that a voter chosen at random will place the three options in a particular order (e.g. *ABC*) is the integral of the density function of voter ideal points over the portion of the plane where the order of the distances to the three options is the order prescribed by the ranking whose probability is being calculated. The portions of the plane assigned to the six rankings of the three options are delineated by the perpendicular bisectors of the line segments connecting pairs of options. Figure 1 shows, for specific locations of the three options, the division of the option plane into the six sectors that define the six rankings. The figure also shows the volume of the triangular wedge (the integral under the spherical multivariate normal distribution centered at *O*) that describes the probability that a randomly chosen voter ranks the three options in the order *ABC*.[9]

While there are six degrees of freedom in the locations of the three options, there are only four degrees of freedom in the probabilities determined by the integrals. One degree of freedom is lost because, if the positions of the three options are rotated around the mode of the distribution of voter ideal points, the pattern of perpendicular bisectors rotates by the same amount, and the integrals are unchanged. A second degree of freedom is lost because, if the locations of the three options are moved the same distance toward or away from the point that is equidistant from them, the perpendicular bisectors and hence the integrals are unchanged. This indeterminacy does not pose any problems if we define the best option as the one closest to *O*, because the spatial model implies that all options are located on rays that emanate from the triangle's circumcenter *T*. Thus the option whose ray is closest to *O* is always closest to *O*, regardless of what (equal) distance from the circumcenter describes the locations of the options along their respective rays.[10]

This indeterminacy prevents us, however, from applying the idea that the voter ideal points represent the voters' private interests. Determining the best option as the one whose location minimizes the sum of the distances to the voters' ideal points, for a finite number of voters, requires knowing not just the ranking of the distances of the options from the mode of the distribution from which the ideal points are drawn, but rather the distances of the options from all voters' actual ideal points. And knowing the number of voters who submit each ranking is not sufficient to determine this, because we need to know the location of each voter's ideal point within the voter's triangular wedge. Thus when the ideal points represent people's private interests we can identify the best option only in the limit when the number of voters is infinite, because in this case the order of the sum of the distances of the options from the individual ideal points is the same as the order of the distances of the options from the mode of the distribution of individual ideal points. Because this is also the order when the ideal points represent people's views of the merit of options from a social perspective, it is not necessary to consider the two cases separately from each other in this limiting case.

---

[9] Owing to the difficulty of having to integrate wedges underneath multivariate normal distributions, we can currently evaluate the spatial model only for elections with three options.

[10] This requires that the options not be located at the triangle's circumcenter where the three rays meet, but our earlier assumption that the positions of the three options are not collinear rules out this possibility.

There are different ways to use the four degrees of freedom to describe the structure of the spatial model. We found that the following parameterization makes it easiest to simulate data that are very similar to data from actual elections. We first place the intersection of the three perpendicular bisectors of the triangle formed by the three options (that is, the triangle's circumcenter T) at the origin of a coordinate system, and we rotate the coordinate system so that the center of voter ideal points O is on its horizontal axis. We then use the first degree of freedom for the horizontal distance OT, and the remaining three degrees of freedom for the angles formed by the line $\overline{OT}$ and the three perpendicular bisectors that describe the boundaries between pairs of the six strict rankings. To simulate data with the spatial model, we adopt distributions for the parameters based on our analysis of real voting data in Plassmann and Tideman (2011). In particular, we assume that $\overline{OT}$ follows a Weibull distribution with scale parameter $\alpha = 0.6858$ and shape parameter $\beta = 2.4608$, and that the three angles follow a Dirichlet distribution with parameters $\delta_1 = 26.47$, $\delta_2 = 23.37$, and $\delta_3 = 23.65$.[11]

3.2 *A model of vote casting*

The spatial model yields a vote share probability p_r for each of the six strict rankings r, with $\Sigma p_r = 1$. To complete the model of vote casting, it is necessary to describe the relationship between the vote share probabilities (six real numbers) and the number of ballots that n voters submit for each of the six rankings (six integers). We assume that the p_rs represent the expected shares of the ballots. To formalize this assumption, let n_r be the number of ballots submitted for ranking r, with $\Sigma n_r = n$ and $r = 1, \ldots, 6$, and let N_r be a random variable that describes the distribution of n_r. If voters cast their ballots independently of each other, then the six N_rs follow a multinomial distribution with $E[N_r] = np_r$ (and thus $E[N_r/n] = p_r$), $Var[N_r] = np_r(1 - p_r)$, and $Cov[N_r, N_s] = -np_rp_s$. One way of accommodating dependent ballots is to assume that the six n_rs follow a multinomial-Dirichlet distribution with $E[N_r] = np_r$, $Var[N_r] = np_r(1 - p_r)\Psi$, and $Cov[N_r, N_s] = -np_rp_s\Psi$, where $\Psi = (\psi + n) / (\psi + 1)$. In Tideman and Plassmann (2011) we show that the multinomial distribution leads to less variation among simulated ballots than what we

---

[11] To match the variance of the angles that we determine from observed election data, we parameterize the Dirichlet distribution so that each of the three shares is multiplied by the common constant 73.5008. Dividing the three values in the text by this number yields three shares that sum to 1.

observe in ballots from actual elections, while the variation of ballots simulated with a multinomial-Dirichlet distribution with $\psi = 330$ is very close to the variation of observed ballots.

In a model of vote casting that uses the multinomial distribution, the variance of each ballot share N_r/n converges to zero as the number of voters becomes large, or

$$\lim_{n\to\infty} Var\left[\frac{N_r}{n}\right] = \lim_{n\to\infty}\left(\frac{1}{n^2}Var[N_r]\right) = \lim_{n\to\infty}\left(\frac{n}{n^2}p_r(1-p_r)\right) = 0. \tag{1}$$

In contrast, this variance does not converge to zero under the multinomial-Dirichlet distribution because

$$\lim_{n\to\infty} Var\left[\frac{N_r}{n}\right] = \lim_{n\to\infty}\left(\frac{n}{n^2}p_r(1-p_r)\frac{\psi+n}{\psi+1}\right) = \frac{1}{\psi+1}p_r(1-p_r). \tag{2}$$

Thus the multinomial-Dirichlet distribution increases the variance of the ballot shares beyond the variance described in the spatial model, by an amount that does not converge to zero as the number of voters becomes infinitely large.

3.3 *Defining the best option*

Which option should be defined as "best" depends on how one interprets the relationship between the spatial model of the vote share probabilities and the multinomial-Dirichlet model of the ballots. One interpretation is that the multinomial-Dirichlet model describes the distribution of and the dependence among the ballots in any given election, while the spatial model specifies the expected shares of the n_rs and describes how these expected shares vary across elections. In section 3.2 we used this relationship to motivate the model of the ballots. Because the spatial model specifies the expected shares, we can use it to determine the best option. To simulate an election whose best option we need to know, we draw a realization of the distance $\overline{OT}$ from the Weibull distribution and realizations of the angles of the three perpendicular bisectors from the Dirichlet distribution, and use the four values to construct the six triangular slices that describe the six rankings. In our parameterization, O lies in the greater of the two wedges with the greatest expected angle.[12] The option ranked highest in this ranking is therefore the best option. We use the six vote share probabilities determined by such a realization of the spatial model to

[12] Any two opposite wedges have the same angle. But because the vote share probability of each wedge is determined as the integral under the bivariate normal distribution over the wedge, two wedges with identical angles describe the same vote share probability unless O and T coincide, that is, unless the center of voter ideal points is located at the triangle's circumcenter.

draw from the multinomial-Dirichlet distribution a set of six integers that represent the total numbers of ballots cast for the six strict rankings. The variation added by the model of the ballots does not change the best option and therefore represents noise. We then apply eleven voting rules to the set of simulated ballots to check which voting rules identify the best option as the winner. Because the multinomial-Dirichlet distribution adds variation even when the number of voters becomes large, the frequency with which any voting rule fails to identify the best option does not converge to 0 as the number of voters becomes infinite.

An alternative interpretation of the relationship between the spatial model of the vote share probabilities and the multinomial-Dirichlet model of the ballots is that the true variation in the vote share probabilities is larger than described by the spatial model that we currently use, and that the multinomial-Dirichlet distribution is needed to remedy this shortcoming. Although we have shown in Plassmann and Tideman (2011) that the combination of the spatial model and the multinomial-Dirichlet distribution is capable of describing the observed distribution of ballots from actual elections, the spatial model alone might not specify the correct distribution of the six vote share probabilities. In this case one cannot identify the best option solely from the spatial model that we currently use, because the spatial model omits relevant variation. Instead, the best option would need to be specified by a revised version of the spatial model that incorporates the additional variation currently supplied by the multinomial-Dirichlet distribution.[13] Because we do not know how to construct this revised spatial model, we use the following method to estimate the best option that such a model would specify. We first draw a set of six values p_r from the spatial model, use the p_rs to draw a set of six n_rs from the multinomial-Dirichlet distribution, and calculate six simulated vote shares as n_r/n. We then use the method described in Good and Tideman (1976) to reconstruct the spatial model from these six vote shares, and we identify the ranking as best whose triangular wedge underneath the bivariate normal density function contains O. Good and Tideman's method of identifying the winning ranking can be interpreted as a voting rule, which Tideman (2006) calls "estimated centrality." Thus we use the estimated centrality winner to approximate the best option that would be specified by a revised spatial model that accounts for the relevant variation among the vote share probabilities. If we were to use the multinomial model to describe the distribution of

[13] Incorporating the additional variation into a revised version of the spatial model would result in different distributions of the distance $\overline{OT}$ and the three angles that describe the positions of the perpendicular bisectors relative to $\overline{OT}$.

the ballots so that the variance of the ballot shares N_r/n converged to zero as the number of voters becomes large, then in the limit as the number of voters approaches infinity, the estimated centrality winner would coincide with the winner suggested by the spatial model. However, because we use the multinomial-Dirichlet model to describe the distribution of the ballots, the estimated centrality winner does not always coincide with the winner suggested by the spatial model, even when the number of voters is large, because the variance of the ballot shares for a given realization of the four parameters of the spatial model does not converge to zero.

Because we have no way of judging which interpretation of the relationship between the spatial model of the vote share probabilities and the multinomial-Dirichlet model of the ballots is more appropriate, we use both methods of identifying the best option in our simulated elections. If the first interpretation is correct, then the absolute frequencies with which different voting rules identify the best option in our simulated data provide information about the absolute accuracy of the different rules in actual elections. If the second interpretation is correct, then the absolute frequencies are less informative because our method of defining the best option is only an approximation. However, the relative frequencies still provide information about the relative attractiveness of the different voting rules, and we find that these relative frequencies are very similar under the two ways of identifying the best option. Thus our analysis provides useful insights irrespective of the true relationship between the spatial model and the multinomial-Dirichlet model.

4. The Frequencies with which Eleven Voting Rules Identify the Best Option

The accuracy of a voting rule increases with the number of voters, and—at least when the number of voters is small—it also depends strongly on whether there is an even or an odd number of voters. We therefore simulate 16 sets of one million elections each with 10, 20, 50, 100, 1000, 10,000, 100,000 and 1,000,000 voters as well as with 11, 21, 51, 101, 1001, 10,001, 100,001, and 1,000,0001 voters. In each set we identify the best option according to the two ways described in section 3.3

One way of classifying voting rules is according to whether or not a voting rule elects the Condorcet winner, if one exists. A Condorcet winner is an option that beats all other options in pairwise comparison, using majority rule. Rules that always elect a Condorcet winner when one exists are called “Condorcet consistent.” Because these rules agree on the winner more often than rules that are not Condorcet consistent, it is reasonable to separate the two types of voting rules to simplify comparison among them. Our first six voting rules with which we evaluate our simulated data are not Condorcet consistent, while the last five voting rules are Condorcet consistent.

A. Six voting rules that are not Condorcet consistent:

1. The Alternative Vote

   An option wins if it is ranked first by a majority of the voters. If there is no such option, then eliminate the option ranked first by the fewest voters and rewrite the ballots without this option. Continue this elimination process until one option has a majority of the first-place votes. That option is declared the winner.

2. Borda rule

   For each ballot, assign m – 1 points to the option ranked highest, m – 2 points to the option ranked second highest, and so on. An option's Borda score is the sum of all points calculated over all ballots. The option with the highest Borda score wins.

3. Bucklin rule

   Each option receives a score equal to the number of times the option is ranked first on all ballots. The option wins that receives the majority of all votes. If there is no such option, then each option receives a score equal to the number of times the option is ranked either first or second on all ballots. Continue the process until an option receives a score that exceeds half the number of voters. If multiple options receive such a score, then the option with the highest score wins.

4. Coombs rule

   An option wins if it is ranked first by a majority of the voters. If there is no such option, then eliminate the option ranked last on the most ballots and rewrite the ballots without this option. Continue this elimination process until one option has a majority of the first-place votes. That option is declared the winner.

5. Estimated Centrality

   Use the spatial model and the observed ballot shares for each of the possible strict rankings to estimate the positions of the perpendicular bisectors of the line segments connecting pairs of candidates, relative to the center of voter ideal points. The ranking wins whose wedge underneath the multivariate normal distribution contains the center of voter ideal points. The option ranked highest in the winning ranking is the winning option.

6. Plurality rule

   The option with the largest number of first ranks wins.

B. Five voting rules that are Condorcet consistent:

7. Black rule

   A Condorcet winner, if it exists, wins the election. If there is no Condorcet winner, then the winner is the option with the highest Borda score.

8. Copeland rule

   An option's Copeland score is the difference between the number of options that this option beats in pairwise comparisons and the number of options by which this option is beaten. The option with the highest Copeland score wins.

9. Dodgson rule

   A Condorcet winner, if it exists, wins the election. If there is no Condorcet winner, then each option is evaluated according to the number of pairs of adjacent options on individual ballots that would need to be reversed to make this option a Condorcet winner. The option wins for which the fewest such reversals are necessary.

10. Kemeny rule

    Consider all pairwise comparisons of the m options. Each strict ranking contains $m(m + 1)/2$ such pairs. For each strict ranking, count the number of pairs of options on a ballot that are ranked in the same order as in this ranking, and record one point for each matching pair. A ranking's Kemeny score is the sum of these points over all ballots. The ranking with the highest Kemeny score is the winning ranking, and the highest ranked option on the winning ranking is the winning option.

11. Nanson rule

    A Condorcet winner, if it exists, wins the election. If there is no Condorcet winner, then calculate each option's Borda score and eliminate all options whose Borda score is less than or equal to the mean Borda score. Compute Borda scores with respect to the reduced set of candidates and continue the prescribed elimination process until there is a Condorcet winner.

For elections with three options, the Kemeny rule chooses the same winner as the Maximin rule, the Ranked-Pairs rule, the Schulze rule, and the Young rule, so we do not need to consider these rules separately.[14] We also do not consider approval voting and range voting. Approval voting permits voters to give a vote to as many options as they choose, while range voting permits voters to assign to each option a score in a specified range. We do not have sufficient data from actual elections that provide information of this kind and that we could use to calibrate statistical models with which to simulate vote-casting processes for such voting rules.

For each simulated election we randomly choose one of the six possible strict rankings as tie-breaker, and we resolve all ties that arise among two or more options according to their order in this ranking. This tie-breaking rule ensures that, across voting rules, we resolve all ties between any two options in the same way. We keep account of whether a voting rule chooses an option as winner with or without employing the tie-breaking rule.

We report our results in tables 1 – 5. Tables 1 – 3 show the results that we obtained under the assumption that the best option is the option ranked highest on the ranking whose wedge contains *O* in the spatial model construct that is determined by the realization of the four parameters that are drawn by a pseudo-random process; tables 4 and 5 show the results when we determine the best option as the estimated centrality winner from an election that is a realization from a pseudo-random process using expected vote share probabilities determined by the previous process. Parts a and b of each table show the results for even and odd numbers of voters, respectively. We first discuss the results in tables 1 – 3, and then comment on the differences between the two ways of defining the best option.

4.1 *Empirical results when the best option is determined by the spatial model*

Tables 1a and 1b show that ten of the eleven voting rules identify the best option with very similar frequencies. The exception is the Bucklin rule, whose accuracy is comparable to that of the other rules when the number of voters is below 20 but is more than six percentage points below that of the other rules for 1,000 and more voters. The second least accurate rule is the plurality rule. The most accurate rule in all cases, except for elections with 10 voters, is the estimated centrality rule, followed by the Black rule. However, the difference in accuracy between estimated centrality and the plurality rule is never greater than 1.3% and converges to

[14] See Tideman (2006) for a description of these voting rules.

about 0.8% with slight fluctuations when then number of voters is 100 or more. Nevertheless, the difference in accuracy is statistically significant.[15]

The accuracy of the Borda rule is comparable to that of the estimated centrality rule and the Black rule in elections with less than 100 voters, but its relative accuracy falls as the number of voters increases, reaching the level of the plurality rule when the number of voters is 1,000 or more. Not surprisingly, in large elections the five Condorcet consistent rules identify the best option with almost identical degrees of accuracy. The variation in accuracy between odd and even numbers of voters is less than two percent in elections with only 10 or 11 voters, and it becomes completely inconsequential for elections with 100 and more voters. From the frequencies in tables 1a and 1b one might conclude that, while it would be prudent to avoid the Bucklin rule, in large elections the choice of voting rules matters very little.

Sometimes voting rules identify multiple options as being equally attractive and require a tie-breaker to select one of them as the winner. Tables 2a and 2b show the percentage of elections in which a voting rule needed a tie-breaker to select a winner. Ties do not matter very much in elections with 1,000 and more voters; the Copeland rule is the only rule for which ties still occur in a meaningful fraction (about 0.1%) of elections with 10,000 and more voters. Ties do occur, however, fairly frequently in elections with fewer voters, especially when the number of voters is even. For example, the Alternative vote employed a tie-breaker in 21% of the elections with 10 voters. As shown in Tables 1a and 1b, the frequency with which the Alternative vote chooses the best option in elections with few voters is comparable to that of the other rules, so this inability to identify a single winning option might not matter much when one considers a large number of elections of equal importance. However, in the 694,845 simulated elections with 10 voters in which the Alternative vote identified the best option, it did so without resorting to a random tie-breaker in only 602,827 elections and it employed a tie-breaker in the remaining 92,018 elections. In 305,155 elections the Alternative vote choose an option other than the best option as winner; 100,288 of these elections satisfied both of the following conditions: (1) the best option was eliminated because of a tie during either the first or the second stage of eliminations and (2) the winning option won only because it was not eliminated

[15] For 1 million Bernoulli trials, the standard error of estimate of the probability of success is 0.05% when the probability is 50%, it is 0.03% when the probability is 10%, and it is 0.01% when the probability is 1%. Using these standard errors as an approximation suggests that the estimated centrality rule is significantly more accurate than any of the other rules in elections with an odd number of voters, as well as in elections with an even number of voters of 50 or larger.

in a tie with either the best option or the third option. These are elections in which the best option failed to win because of ties, which implies that in about 19% of all elections (100,288 + 92,018) with 10 voters, the Alternative vote might not select the best option solely because of ties.[16]

The rule that resorts least to a tie-breaker is the estimated centrality rule. This rule did not require a single tie-breaker in any of our 8 million elections with 1,000 and more voters, and the number of tie-breakers in elections with fewer voters was more than an order of magnitude smaller than the number of the second best rule. For elections with more than 10 voters, the estimated centrality rule also identifies the best option with the highest frequency among the 11 voting rules that we examine. The disadvantage of the estimated centrality rule is that it is rather difficult to compute. The Black rule, which is much easier to compute, has the second lowest frequency of ties in elections with even numbers of voters, and the third lowest frequency, after the estimated centrality rule and the Nanson rule, in elections with odd numbers of voters. The Black rule also has the highest frequency of identifying the best option in elections with 10 voters, the third highest frequency for 21 voters, and the second highest frequency, never more than half a percentage point below the highest, in all other cases. Given its ease of computation, the Black rule is therefore a very attractive voting rule. To our knowledge, voting theorists are dissatisfied with the Black rule mainly because its concatenation of disparate components makes it esthetically less elegant than the other voting rules.[17]

In elections with an even number of voters the Borda rule is a close third. In elections with 101 and fewer voters and an odd number of voters, however, the Borda rule resorts to a tie-breaker significantly more often than the estimated centrality rule, the alternative vote, the Coombs rule, and any of the five Condorcet consistent rules, including the Black rule. Thus unless it is known prior to choosing the voting rule whether the number of voters will be even or

---

[16] When the best option was eliminated by a tie-breaker during the first stage, we considered all elections with second-stage ties between the winning option and the third option (which had survived the tie in the first stage). We did not consider whether there would also have been a tie between the winning option and the best option had the best option not been eliminated in the tie during the first stage. However, because these are elections with only 10 voters, it is likely that the two options that were tied during the first stage would have compared similarly with the winner during the second stage. Thus although the 100,288 elections do not measure exactly how often the best option lost because of ties, this number is likely to be very close to the number in which we are interested.

[17] For example, the mathematician William Zwicker wrote about an experiment in which he and other voting theorists were asked to either approve or disapprove of various voting rules: "Mathematically, Kemeny is beautiful whereas Black is plug-ugly, but I swallowed hard, approved Black, and disapproved Kemeny (because Kemeny winner[s] are rankings, not individual candidates, and I can imagine what would happen the first time some real world election yielded a tie among several rankings)." (See Laslier, 2012, p. 345.)

odd, it might be prudent to use the Black rule rather than the Borda rule in elections with few voters.

Because the spatial model identifies a best ranking, we know the worst option as well as the best in any election. Tables 3a and 3b report the frequencies with which the eleven voting rules declare the worst option as winner. The frequencies are fairly similar across voting rules; they vary from about six to seven percent in elections with 10 or 11 voters to between 0.12 and 0.53 percent in elections with 10,000 and more voters. The Borda rule consistently chooses the worst option with the lowest frequency, while the estimated centrality rule and the Black rule are close behind. The plurality rule chooses the worst option significantly more frequently than the other voting rules do, but the difference between the plurality rule and the Borda rule never exceeds two percent and it is less than half a percent in elections with 1,000 or more voters.

Whether or not one ought to be concerned that a voting rule sometimes declares the worst option as winner depends on the degree of arguable superiority of the best option over the worst in those instances. If each ranking receives about as many votes as the opposite ranking (for example, the ranking ABC receives about as many votes as the ranking CBA), then the voters' preferences are fairly balanced and it should not be a major concern if a voting rule occasionally chooses the worst option. But the less balanced the voters' preferences are, the more troublesome is it that a voting rule sometimes chooses the worst option. In the spatial model the distance $\overline{OT}$ provides a measure of how balanced the voters' preferences are—the larger $\overline{OT}$ is, the less balanced the voters' preferences are. By considering the sampled distance $\overline{OT}$ in any election in which a voting rule declares the worst option as winner, we can assess how balanced the voters' preferences are on average for different voting rules in elections in which they choose the worst option. For such elections, we find a mean distance of $\overline{OT}$ that is between 0.37 and 0.40 in elections with 10 and 11 voters, falling to between 0.09 and 0.14 in elections with one million voters. The parameterization of the Weibull distribution from which we sample $\overline{OT}$ implies $E[\overline{OT}]$ = 0.61. Thus as one would expect, the eleven voting rules tend to choose the worst option when the voter preferences are more balanced than in the average election. Across the different voting rules and elections with different numbers of voters, the relative mean distances $\overline{OT}$ are very similar to the relative percentages in tables 3a and 3b—the mean distance $\overline{OT}$ is shortest for the elections in which the Borda rule identifies the worst option as winner, closely followed by the estimated centrality rule, the Black rule, and the other four Condorcet consistent rules.[18] The Bucklin rule and the plurality rule perform again worst—in elections with one million voters, the mean distances $\overline{OT}$ for these two rules are about 50% larger than those for the voting rules that do best. Thus these two voting rules sometimes choose the worst option when other voting rules are capable of correctly assessing the imbalances among voters' preferences.

4.2 *Empirical results when the best option is specified as the estimated centrality winner*

The results that we obtain under this alternative method of defining the best option are very similar to those that we reported in the previous section. We do not repeat the frequencies of ties because they are identical to those shown in tables 2a and 2b, and we ignore the estimated centrality rule since it does best by design. The lowest error rates are generally attained by either the Black rule or the Borda rule. However, especially in elections with large numbers of voters, the other voting rules are not much worse. The main difference compared to the results in the previous section is that all voting rules have much lower error rates, starting with rates between about 3.5% and 17% in elections with 10 and 11 voters, and converging to about 11% for Bucklin, 4% for plurality, and 2% for all other rules, for elections with many voters (see tables 4a and b). Thus there is further confirmation that Bucklin and Plurality are notably less accurate than other rules. Our analysis of the relative numbers of times that different voting rules choose the worst option yields the same insights as our analysis in section 4.1. The lower error rates in tables 4a and 4b explain why the frequencies of choosing the worst option in tables 5a and b are lower than those in tables 3a and b.

4.3 *Additional insights*

Our simulations revealed several interesting relationships among voting rules. First, inspection of the results of our simulations reveals that, at least for elections with three voters, the Kemeny rule and the Dodgson rule are equivalent when the number of voters is odd but not when the number of voters is even. In all eight sets with odd numbers of voters and one million elections in each set, the two rules always chose the same winner, and they also agreed on the elections in which they considered the best option to be tied with other options. In elections with even

[18] The only notable difference to table 3b is that the mean distance $\overline{OT}$ for the Borda rule is also smallest in elections with an odd number of voters.

numbers of voters and fewer than 10,000 voters, the two rules differed only when the Kemeny rule identified a tie but the Dodgson rule did not.

The two rules lead to different results, however, when it is necessary to establish a winning ranking rather than only a winning option. While the Kemeny rule always selects a winning ranking, the Dodgson rule is designed to select only a winning option. Because the Dodgson rule evaluates each option according to the number of pairs of options on individual ballots that would need to be reversed to make this option a Condorcet winner, it is nevertheless possible to construct a winning ranking under the Dodgson rule by ranking the options in ascending order according to the necessary number of reversals. Because both the spatial model and the estimated centrality rule identify the best option as the option ranked highest in the best ranking, we can evaluate the voting rules according to their ability to identify this ranking. A comparison of the frequencies with which the Kemeny rule and the Dodgson rule identify the best ranking shows that the Dodgson rule is about 4.5% more likely to identify the best ranking than the Kemeny rule, regardless of the number of voters and regardless of whether the number of voters is even or odd. This suggests that, at least in elections with three options, the Dodgson rule is preferable to the Kemeny rule for identifying the best ranking.

The calculation of the Dodgson rule can be extremely tedious. Tideman (1987) identifies a variation on the Dodgson rule that is much easier to calculate and which he calls the "simplified Dodgson rule." We found that the simplified Dodgson rule has many more ties than the Dodgson rule in elections with few voters, but in our 8 million simulations of elections with 1,000 and more voters whose results we report in tables 1 – 3, the two rules chose different winners only twice—in each case the simplified Dodgson rule identified a tie when the Dodgson rule did not. In elections with many voters and without access to a computer program to evaluate the ballots, the simplified Dodgson rule is therefore an attractive alternative to the Dodgson rule.

Second, in elections with 10,000 voters and more, the Nanson rule becomes virtually indistinguishable from the Kemeny rule and the Dodgson rule. In our 6 million simulations of elections with 10,000 and more voters whose results we report in tables 1 – 3, the three rules chose different winners only five times—in each case the Dodgson rule and the Kemeny rule identified ties while the Nanson rule did not. The Nanson rule is slightly more difficult to compute than the simplified Dodgson rule, so the two rules have roughly off-setting advantages.

Finally, although only the Kemeny rule and the estimated centrality rule are designed to select a winning ranking as well as a winning option, it is possible to use the selection criteria of the other voting rules to identify winning rankings under these rules as well. In all simulations whose results we report in tables 1 – 3, the estimated centrality rule was most likely to identify the best ranking, followed closely by the Black rule and the Copeland rule; these three rules identify the best ranking in more than 89% of elections with 100,000 and more voters. The rules least likely to identify the best ranking in elections with 1,000 and more voters are the Bucklin rule (with a success rate of about 75%), and the plurality rule and the alternative vote (with success rates of about 78%). The Borda rule identified the best ranking in about 87% of the elections with 1,000 and more voters. Young (1986) emphasizes that the choice of a best ranking differs conceptually from the choice of a best option, and that the optimal rules for the two situations might differ. However, our analysis of the frequency with which different voting rules identify the best ranking leads to a very similar assessment of the relative attractiveness of different voting rules as our analysis of their ability to identify the best option.

5. CONCLUSION

Our analysis of the accuracy of eleven voting rules has yielded two notable results. First, when the number of voters exceeds about 1,000, most voting rules identify the best option with comparable frequency. The exceptions are the Bucklin rule and the plurality rule, whose ability to identify the best option is significantly worse than that of the other rules. To the extent that neither of the two rules possesses highly desirable properties that the other voting rules that we examine lack, our analysis suggests that the Bucklin rule and the plurality rule should be set aside in favor of other rules.[19]

Second, among the eleven rules the estimated centrality rule is most likely to identify the best option, and it also resorts to a tie-breaker least frequently. The Black rule has the second highest frequency of identifying the best option and generally the second lowest frequency of resorting to a tie-breaker. The Borda rule comes third in identifying the best option, for elections with 50 votes or fewer, but it becomes significantly less accurate than most other rules as the number of voters increases. The Borda rule also resorts to tie-breakers in elections with an odd

[19] Neither the Bucklin rule nor the plurality rule is Condorcet consistent, and neither rule possesses characteristics that would make the rule particularly attractive when compared to rules that possess Condorcet consistency (see Tideman, 2006, for a comparison of the properties of different voting rules).

number of voters far more frequently than any rule other than the Bucklin rule and the plurality rule. The Black rule selects the Condorcet winner if one exists, and it coincides with the Borda rule if no Condorcet winner exists. Thus the Black rule is only slightly more difficult to calculate than the Borda rule, and it is much easier to calculate than the estimated centrality rule. As long as we can expect voters to vote sincerely, our analysis therefore suggests that the Black rule is the most attractive among the rules that we examine, especially in elections with few voters where ties occur relatively frequently.

We emphasize that the degree to which our simulations results provide information about what we should expect to happen in actual elections depends on the validity of the parameter values that specify our model of vote casting and whose calibration we describe in Plassmann and Tideman (2011). Among the five parameters of our model of vote casting, the parameter ψ of the multinomial-Dirichlet distribution (an inverted measure of dispersion beyond that found in the multinomial distribution) has the strongest influence on our simulations. In Plassmann and Tideman (2011) we calibrated the five parameters to three different data sets; one data set contains ballot data from actual elections while the other two data sets come from survey data. We obtained the value $\psi = 330$ that we use for the simulations in this paper from our analysis of the ballot data from actual elections. The two data sets that came from surveys yielded values of $\psi = 5{,}500$ and $\psi = \infty$, so there is some uncertainty regarding the appropriate value of ψ. However, we consider the value of $\psi = 330$ calibrated with actual election data to be most reliable, partly because our success in using a spatial model to simulate data that resembles real data was greatest when seeking to simulate the actual elections rather than the outcome of the surveys. Our spatial model fits the observed data so well that it does not seem worthwhile to try and refine the spatial model on the basis of the available data. Collecting additional voting data would permit us to calibrate our model with greater accuracy and precision, and simulations with such a revised model of vote casting might yield somewhat different frequencies than the ones that we report here.

We emphasize that our results are valid only for elections with three options. Although it is straightforward to simulate ballots from a spatial model with any number of options, currently we can calibrate the spatial model only to elections with no more than three options, and the main value of our analysis lies in the fact that our simulated ballots have the same statistical properties as ballots in actual elections. Calibrating the spatial model to elections with four options requires numerical integration of non-central triangular wedges of a trivariate normal distribution, and we have not yet developed an algorithm to accomplish this task. Our current results for elections with three candidates suggest that such an extended spatial model would provide additional useful information about voting rules.

Table 1a. Frequencies with which 11 voting rules fail to select the best option when the number of voters is even, assuming that the spatial model provides an adequate description of the distribution of vote share probability vectors

	Number of voters:							
	10	20	50	100	1,000	10,000	100,000	1,000,000
Alternative vote	30.52%	23.05%	15.55%	11.64%	6.42%	5.69%	5.56%	5.59%
Borda	29.17%	22.26%	15.31%	11.65%	6.84%	6.20%	6.07%	6.10%
Bucklin	31.62%	25.44%	19.51%	16.59%	13.19%	12.76%	12.76%	12.74%
Coombs	30.29%	23.16%	15.67%	11.77%	6.53%	5.78%	5.67%	5.69%
Estimated Centrality	29.44%	22.07%	14.73%	11.06%	6.20%	5.37%	5.30%	5.31%
Plurality	30.14%	23.05%	15.68%	11.79%	6.82%	6.20%	6.09%	6.12%
Black	29.06%	22.12%	15.13%	11.41%	6.37%	5.64%	5.51%	5.54%
Copeland	29.67%	22.74%	15.45%	11.60%	6.39%	5.65%	5.52%	5.55%
Dodgson	29.66%	22.70%	15.41%	11.57%	6.37%	5.64%	5.51%	5.54%
Kemeny	30.23%	22.89%	15.44%	11.58%	6.37%	5.64%	5.51%	5.54%
Nanson	30.24%	22.91%	15.45%	11.58%	6.37%	5.64%	5.51%	5.54%

Table 1b. Frequencies with which 11 voting rules fail to select the best option
when the number of voters is odd,
assuming that the spatial model provides an adequate description of the distribution of vote share probability vectors

	Number of voters:							
	11	21	51	101	1,001	10,001	100,001	1,000,001
Alternative vote	28.43%	22.12%	15.20%	11.59%	6.42%	5.68%	5.59%	5.61%
Borda	28.15%	21.83%	15.11%	11.69%	6.84%	6.20%	6.08%	6.10%
Bucklin	29.26%	24.05%	18.81%	16.28%	13.18%	12.81%	12.69%	12.73%
Coombs	28.40%	22.16%	15.32%	11.73%	6.53%	5.78%	5.70%	5.70%
Estimated Centrality	27.93%	21.48%	14.64%	11.09%	6.11%	5.39%	5.32%	5.34%
Plurality	29.13%	22.61%	15.49%	11.80%	6.82%	6.18%	6.13%	6.10%
Black	28.14%	21.86%	15.08%	11.53%	6.38%	5.63%	5.54%	5.55%
Copeland	28.39%	22.09%	15.19%	11.60%	6.40%	5.64%	5.55%	5.57%
Dodgson	28.25%	21.93%	15.11%	11.54%	6.38%	5.63%	5.54%	5.56%
Kemeny	28.25%	21.93%	15.11%	11.54%	6.38%	5.63%	5.54%	5.56%
Nanson	28.34%	22.01%	15.13%	11.55%	6.38%	5.63%	5.54%	5.56%

Table 2a. Frequencies of ties
when the number of voters is even,
assuming that the spatial model provides an adequate description of the distribution of vote share probability vectors

	Number of voters:							
	10	20	50	100	1,000	10,000	100,000	1,000,000
Alternative vote	21.00%	11.46%	4.96%	2.56%	0.25%	0.02%	0.00%	0.00%
Borda	7.61%	4.26%	1.82%	0.92%	0.09%	0.01%	0.00%	0.00%
Bucklin	12.08%	6.11%	3.10%	1.77%	0.19%	0.02%	0.00%	0.00%
Coombs	18.78%	11.34%	4.96%	2.52%	0.25%	0.02%	0.00%	0.00%
Estimated Centrality	0.35%	0.04%	0.01%	0.00%	0.00%	0.00%	0.00%	0.00%
Plurality	11.52%	7.62%	3.04%	1.43%	0.15%	0.01%	0.00%	0.00%
Black	6.48%	2.68%	0.70%	0.23%	0.00%	0.00%	0.00%	0.00%
Copeland	15.06%	9.88%	4.89%	2.67%	0.34%	0.11%	0.08%	0.08%
Dodgson	15.05%	9.74%	4.65%	2.46%	0.25%	0.02%	0.00%	0.00%
Kemeny	20.81%	11.63%	4.99%	2.55%	0.25%	0.02%	0.00%	0.00%
Nanson	20.64%	11.34%	4.82%	2.46%	0.24%	0.02%	0.00%	0.00%

Table 2b. Frequencies of ties
when the number of voters is odd,
assuming that the spatial model provides an adequate description of the distribution of vote share probability vectors

	Number of voters:							
	11	21	51	101	1,001	10,001	100,001	1,000,001
Alternative vote	1.05%	1.19%	0.32%	0.10%	0.01%	0.00%	0.00%	0.00%
Borda	7.01%	4.08%	1.79%	0.92%	0.09%	0.01%	0.00%	0.00%
Bucklin	6.23%	4.92%	2.82%	1.64%	0.19%	0.02%	0.00%	0.00%
Coombs	1.50%	0.97%	0.32%	0.12%	0.01%	0.00%	0.00%	0.00%
Estimated Centrality	0.03%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%
Plurality	13.10%	6.33%	2.82%	1.47%	0.14%	0.01%	0.00%	0.00%
Black	0.85%	0.36%	0.11%	0.04%	0.00%	0.00%	0.00%	0.00%
Copeland	2.79%	1.83%	0.87%	0.47%	0.11%	0.09%	0.09%	0.09%
Dodgson	2.27%	1.12%	0.34%	0.13%	0.01%	0.00%	0.00%	0.00%
Kemeny	2.27%	1.12%	0.34%	0.13%	0.01%	0.00%	0.00%	0.00%
Nanson	0.78%	0.24%	0.04%	0.01%	0.00%	0.00%	0.00%	0.00%

Table 3a. Frequencies with which 11 voting rules select the worst option
when the number of voters is even,
assuming that the spatial model provides an adequate description of the distribution of vote share probability vectors

	Number of voters:							
	10	20	50	100	1,000	10,000	100,000	1,000,000
Alternative vote	7.59%	4.24%	1.84%	1.01%	0.27%	0.21%	0.20%	0.20%
Borda	6.34%	3.47%	1.45%	0.76%	0.18%	0.14%	0.12%	0.12%
Bucklin	6.98%	4.02%	1.85%	1.09%	0.39%	0.33%	0.34%	0.34%
Coombs	6.95%	3.91%	1.65%	0.90%	0.23%	0.18%	0.16%	0.16%
Estimated Centrality	6.92%	3.64%	1.48%	0.77%	0.14%	0.14%	0.13%	0.13%
Plurality	8.07%	5.01%	2.48%	1.52%	0.62%	0.55%	0.54%	0.53%
Black	6.36%	3.51%	1.50%	0.80%	0.20%	0.15%	0.14%	0.14%
Copeland	6.64%	3.76%	1.62%	0.88%	0.22%	0.15%	0.15%	0.15%
Dodgson	6.63%	3.73%	1.58%	0.85%	0.20%	0.15%	0.14%	0.14%
Kemeny	7.03%	3.85%	1.60%	0.86%	0.20%	0.15%	0.14%	0.14%
Nanson	7.04%	3.87%	1.61%	0.86%	0.20%	0.15%	0.14%	0.14%

Table 3b. Frequencies with which 11 voting rules select the worst option
when the number of voters is odd,
assuming that the spatial model provides an adequate description of the distribution of vote share probability vectors

	Number of voters:							
	11	21	51	101	1,001	10,001	100,001	1,000,001
Alternative vote	6.57%	3.94%	1.76%	0.99%	0.26%	0.21%	0.21%	0.21%
Borda	5.82%	3.32%	1.40%	0.75%	0.18%	0.14%	0.14%	0.14%
Bucklin	6.20%	3.75%	1.79%	1.06%	0.40%	0.34%	0.33%	0.33%
Coombs	6.20%	3.62%	1.58%	0.87%	0.22%	0.18%	0.18%	0.18%
Estimated Centrality	6.02%	3.46%	1.46%	0.79%	0.14%	0.14%	0.13%	0.13%
Plurality	7.62%	4.83%	2.41%	1.52%	0.62%	0.54%	0.54%	0.53%
Black	6.07%	3.52%	1.51%	0.82%	0.19%	0.15%	0.15%	0.15%
Copeland	6.33%	3.73%	1.62%	0.89%	0.16%	0.16%	0.16%	0.16%
Dodgson	6.19%	3.57%	1.54%	0.83%	0.19%	0.15%	0.15%	0.15%
Kemeny	6.19%	3.57%	1.54%	0.83%	0.19%	0.15%	0.15%	0.15%
Nanson	6.27%	3.64%	1.56%	0.84%	0.20%	0.15%	0.15%	0.15%

Table 4a. Frequencies with which 10 voting rules fail to select the Estimated Centrality winner (an approximation of the best option when vote share probabilities have more variation than the spatial model allows) when the number of voters is even

	Number of voters:							
	10	20	50	100	1,000	10,000	100,000	1,000,000
Alternative vote	12.57%	8.25%	5.15%	3.75%	2.09%	1.83%	1.82%	1.80%
Borda	7.58%	4.46%	2.85%	2.36%	1.93%	1.91%	1.89%	1.91%
Bucklin	16.90%	14.27%	12.42%	11.62%	10.79%	10.83%	10.74%	10.78%
Coombs	11.94%	8.02%	4.99%	3.65%	2.06%	1.81%	1.79%	1.77%
Plurality	10.03%	8.28%	6.60%	5.54%	4.33%	4.15%	4.13%	4.13%
Black	6.81%	3.64%	2.68%	2.39%	1.72%	1.55%	1.55%	1.53%
Copeland	8.66%	6.30%	4.40%	3.31%	1.86%	1.61%	1.60%	1.59%
Dodgson	8.64%	6.19%	4.26%	3.22%	1.82%	1.58%	1.57%	1.56%
Kemeny	11.20%	7.12%	4.43%	3.26%	1.82%	1.58%	1.57%	1.56%
Nanson	11.27%	7.24%	4.51%	3.30%	1.83%	1.58%	1.57%	1.56%

Table 4b. Frequencies with which 10 voting rules fail to select the Estimated Centrality winner
(an approximation of the best option when vote share probabilities have more variation than the spatial model allows)
when the number of voters is odd

	Number of voters:							
	11	21	51	101	1,001	10,001	100,001	1,000,001
Alternative vote	6.41%	6.28%	4.67%	3.60%	2.09%	1.84%	1.82%	1.81%
Borda	5.27%	4.14%	2.84%	2.36%	1.93%	1.91%	1.92%	1.90%
Bucklin	8.97%	11.22%	11.40%	11.20%	10.83%	10.81%	10.82%	10.81%
Coombs	5.77%	5.93%	4.48%	3.50%	2.06%	1.79%	1.79%	1.80%
Plurality	10.40%	8.32%	6.50%	5.54%	4.36%	4.20%	4.10%	4.12%
Black	3.48%	4.36%	3.66%	2.96%	1.79%	1.56%	1.55%	1.54%
Copeland	4.76%	5.33%	4.16%	3.24%	1.86%	1.61%	1.60%	1.60%
Dodgson	4.22%	4.82%	3.87%	3.09%	1.82%	1.58%	1.58%	1.57%
Kemeny	4.22%	4.82%	3.87%	3.09%	1.82%	1.58%	1.58%	1.57%
Nanson	4.92%	5.26%	4.02%	3.15%	1.82%	1.58%	1.58%	1.57%

Table 5a. Frequencies with which 10 voting rules select the Estimated Centrality worst option (an approximation of the worst option when vote share probabilities have more variation than the spatial model allows) when the number of voters is even

	Number of voters:							
	10	20	50	100	1,000	10,000	100,000	1,000,000
Alternative vote	1.90%	0.92%	0.36%	0.19%	0.06%	0.05%	0.05%	0.05%
Borda	0.34%	0.09%	0.01%	0.00%	0.00%	0.00%	0.00%	0.00%
Bucklin	1.68%	1.11%	0.72%	0.52%	0.31%	0.28%	0.28%	0.28%
Coombs	0.72%	0.41%	0.18%	0.10%	0.04%	0.03%	0.03%	0.03%
Plurality	2.44%	1.64%	0.87%	0.59%	0.34%	0.32%	0.32%	0.31%
Black	0.34%	0.09%	0.01%	0.00%	0.00%	0.00%	0.00%	0.00%
Copeland	0.40%	0.23%	0.15%	0.10%	0.04%	0.03%	0.03%	0.03%
Dodgson	0.34%	0.10%	0.02%	0.00%	0.00%	0.00%	0.00%	0.00%
Kemeny	0.51%	0.10%	0.02%	0.00%	0.00%	0.00%	0.00%	0.00%
Nanson	0.51%	0.10%	0.02%	0.00%	0.00%	0.00%	0.00%	0.00%

Table 5b. Frequencies with which 10 voting rules select the Estimated Centrality worst option (an approximation of the worst option when vote share probabilities have more variation than the spatial model allows) when the number of voters is odd

	Number of voters:							
	11	21	51	101	1,001	10,001	100,001	1,000,001
Alternative vote	1.28%	0.75%	0.34%	0.19%	0.06%	0.05%	0.05%	0.05%
Borda	0.28%	0.08%	0.01%	0.00%	0.00%	0.00%	0.00%	0.00%
Bucklin	0.94%	0.98%	0.70%	0.54%	0.31%	0.28%	0.29%	0.29%
Coombs	0.23%	0.24%	0.15%	0.10%	0.04%	0.03%	0.03%	0.03%
Plurality	2.29%	1.54%	0.87%	0.58%	0.35%	0.32%	0.31%	0.31%
Black	0.28%	0.08%	0.01%	0.00%	0.00%	0.00%	0.00%	0.00%
Copeland	0.97%	0.62%	0.29%	0.15%	0.04%	0.03%	0.03%	0.03%
Dodgson	0.30%	0.09%	0.02%	0.00%	0.00%	0.00%	0.00%	0.00%
Kemeny	0.30%	0.09%	0.02%	0.00%	0.00%	0.00%	0.00%	0.00%
Nanson	0.31%	0.09%	0.02%	0.01%	0.00%	0.00%	0.00%	0.00%

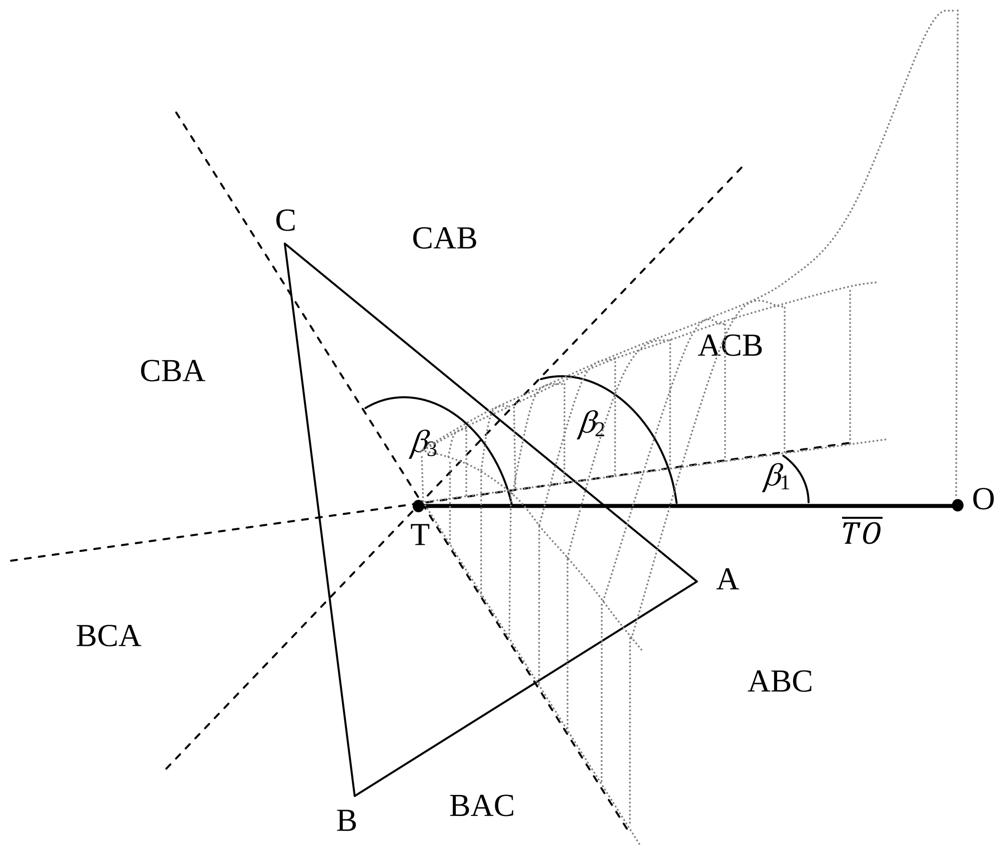

Figure 1. The division of the option plane into the six sectors that define the six rankings, and the volume of the triangular wedge (the integral under the spherical multivariate normal distribution centered at *O*) that describes the probability that a randomly chosen voter ranks the three options in the order *ABC*

References:

Condorcet, Marie Jean Antoine Nicolas de Caritat, Marquis de. 1785. *Essai sur l'application de l'analyse à la probabilité des décisions rendues à la pluralité des voix* (Paris).

Conitzer, Vincent and Tuomas Sandholm. 2005. "Common voting rules as maximum likelihood estimators." In: *Proceedings of the 21st Annual Conference on Uncertainty in Artificial Intelligence* (UAI-05) (pp. 145-152). Edinburgh, UK.

de Borda, Jean-Charles. 1784. "Mémoire sur les Elections au Scrutin." Reprinted in: Alfred de Grazia. 1953. "Mathematical derivation of an election system." *Isis* **44**, 42 – 51.

Drissi-Bakhkhat, Mohamed, and Michel Truchon. 2004. "Maximum likelihood approach to vote aggregation with variable probabilities." *Social Choice and Welfare* **23**, 161 – 185.

Good, Jack, & T. Nicolaus Tideman. 1976. "From individual to collective ordering through multidimensional attribute space." *Proceedings of the Royal Society of London (Series A)* **347**, 371 – 385.

Kemeny, John. 1959. "Mathematics without numbers." *Daedalus* **88**, 571 – 91.

Laslier, Jean-François. 2012. "And the loser is… plurality voting." In Dan Felsenthal and Moshé Machover (eds.) *Electoral systems: Paradoxes, assumptions, and procedures*, Springer (forthcoming).

Tideman, T. Nicolaus. 1987. "Independence of Clones as a Criterion for Voting Rules." *Social Choice and Welfare* **4**, 185-206.

Tideman, T. Nicolaus. 2006. *Collective Decisions and Voting*. Burlington, VT: Ashgate.

Plassmann, Florenz and T. Nicolaus Tideman. 2011. "How to assess the frequency of voting paradoxes and strategic voting opportunities in actual elections." Mimeo.
A copy is available at http://papers.ssrn.com/sol3/cf_dev/AbsByAuth.cfm?per_id=250705.

Tideman, T. Nicolaus and Florenz Plassmann. 2012. "Modeling the outcomes of vote-casting in actual elections." In: Dan Felsenthal and Moshé Machover (eds.) *Electoral systems: Paradoxes, assumptions, and procedures*, Springer (forthcoming).
A copy is available at http://papers.ssrn.com/sol3/cf_dev/AbsByAuth.cfm?per_id=250705.

Truchon, Michel. 2008. "Borda and the maximum likelihood approach to vote aggregation." *Mathematical Social Sciences* **55**:1, 96 – 102.

Young, H. Peyton. 1986. "Optimal ranking and choice from pairwise comparisons." In: B. Grofman and G. Owen (eds.). *Information Pooling and Group Decision Making*. Greenwich, CN: Jai Press.

Young, H. Peyton. 1988. "Condorcet's theory of voting." *The American Political Science Review* **82**:4, 1231 – 1244.

Young, H. Peyton. 1995. "Optimal voting rules." *Journal of Economic Perspectives* **9**:1, 51 – 64.