Public Choice **106:** 275–297, 2001.

President Perot or fundamentals of voting theory illustrated with the 1992 election

ALEXANDER TABARROK*
The Independent Institute, 100 Swan Way, Oakland, CA 94621, U.S.A.;
e-mail: atabarrok@Independent.org

Accepted 4 September 1999

Abstract. Different voting systems can lead to different election outcomes even when voter preferences are held constant. Using the 1992 election as an example, it is shown how the outcome of every positional vote system can be found. Similarly, every possible cumulative and approval vote outcome is shown. Multiple vote systems, like approval and cumulative voting, have disturbing properties. Using the 1992 election as illustration, it is shown how a candidate who wins under every positional vote system, who wins every pairwise vote (i.e. is the Condorcet winner), and who has the most first place and least last place votes may nevertheless lose under approval or cumulative voting. Similarly, it is shown how a candidate who loses under every positional system, who loses every pairwise vote (i.e. is the Condorcet loser), and who has the least first place and most last place votes may nevertheless win under approval or cumulative voting.

1. Introduction

It has long been understood that different voting systems can lead to different election outcomes even when voter preferences are held constant. Voting systems like the Borda Count and Approval Voting have been defended precisely on the grounds that these systems pick better outcomes than the much reviled plurality rule. Unfortunately, the analysis of different voting systems has been mostly restricted to examples, counter-examples, and theorems valid only under restrictive conditions. Recently, however, Saari (1994), has shown how the outcome of *every* positional (also called point-score) vote system can easily be characterized. Similarly, Saari (1994) shows how *every* outcome under approval or cumulative voting can be characterized and how these outcomes relate to those possible under positional vote systems. We use these new techniques to analyze the 1992 US Presidential election. Our analysis has three purposes. First, the analysis illustrates the power and utility of the

* The views expressed here are solely those of the author and do not necessarily reflect the views of The Independent Institute. I thank Don Saari and several referees for useful comments.

new techniques with an interesting and accessible example. Second, the analysis is interesting in its own right. Could Bush have won the election under a different voting system? Could Perot have won? Was Clinton's victory a result of defects in plurality rule or did it reflect the true preferences of the voters? Was Clinton's victory robust or fragile? Third, and most importantly, we use the 1992 election to demonstrate some deep results in voting theory which bear on the desirability of multiple voting systems like approval voting and cumulative voting.

Brams and Merrill (1994) also analyze whether Perot would have won the 1992 election under alternative voting systems. Brams and Merrill, however, look at only two of the infinitely many positional voting systems and only a handful of the many possible outcomes under approval voting. This paper looks at every positional vote system and every possible outcome under approval voting. (This paper also analyzes cumulative voting which Brams and Merrill do not.) Joslyn (1976) analyzes the 1972 Democratic Presidential nomination and finds that the outcome would have been quite different under a number of alternative voting systems. Kellet and Mott (1977) and Kiewet (1979) analyze the difference that approval voting might have made in the 1976 presidential nomination and 1968 presidential election respectively. Tabarrok and Spector (1999) analyze the critical election of 1860.

Throughout this paper we make two simplifying assumptions. First we ignore the electoral college and proceed as if Presidents were elected by popular vote. If we focused on the electoral college the larger points we wish to make about voting systems could become obscured by geographical data peculiar to the United States which have little bearing on elections in general. Second, and more importantly, we will assume for the most part that voters vote sincerely. Because voting systems that perform poorly under sincere voting may perform well when voters vote sophisticatedly, the conclusions made in this paper are conditional. Sophisticated voting, however, relies on an understanding of what is possible under sincere voting. Thus an analysis of sincere voting is important as a precursor to an analysis of sophisticated voting as well as being of independent interest.

The paper proceeds in two sections. In the first section we review the theory of positional, pairwise, approval, and cumulative voting systems.[1] In the second section we apply the theory to the 1992 Presidential election and discuss the bearing of the results on the desirability of different voting systems.

2. Theory of positional, pairwise, approval, and cumulative voting systems

2.1. *Theory of positional voting*

Suppose there are n candidates in an election. We assume that each voter can rank the n candidates from most to least favored. A positional system assigns points to the voter's list with more favored candidates receiving more (or at least not fewer) points than less favored candidates. Plurality rule, is a particularly simple positional system. Plurality rule assigns one point to a voter's top ranked candidate and zero points to every other candidate. The Borda count, a positional voting system devised by the eighteenth century French mathematician Jean Charles Borda, assigns $n - 1$ points to a voter's top ranked candidate, $n - 2$ points to a second ranked candidate and $n - i$ points to an i'th ranked candidate.

Letting $n = 3$, we can write the plurality rule system as $\{1, 0, 0\}$ and the Borda Count as $\{2, 1, 0\}$. Plurality rule and the Borda Count are only two of infinitely many positional voting systems. Another voting system, this one without a particular name, assigns 20 points to the top ranked candidate, 6 points to the second ranked candidate and 3 points to the last ranked candidate or, $\{20, 6, 3\}$.[2] Suppose we modify the $\{20, 6, 3\}$ voting system by subtracting 3 points from each point assignment to get $\{17, 3, 0\}$; this will not change the candidate rankings. The $\{20, 6, 3\}$ and $\{17, 3, 0\}$ voting systems are therefore equivalent. Now divide each point assignment by the sum of all points to get, $\left(\frac{17}{20}, \frac{3}{20}, \frac{0}{20}\right)$. This new voting system is also equivalent to $\{20, 6, 3\}$. By repeating the two processes just described we can change *any* positional voting system into a *standardized* positional system denoted $\{1 - s, s, 0\}$, where $s \in \left[0, \frac{1}{2}\right]$. *Every* positional system is thus associated with a single number, s.[3]

A voter may rank n candidates in any one of n! possible ways. If the candidates are a, b, c, for example, then a voter could rank them (1) abc, (2) acb, (3) cab, (4) cba, (5) bca, or (6) bac where precedence in the list indicates preference. (The numbering of these rankings is arbitrary but later we use the numbers as shorthand.) Note that under vote system s a voter with ranking abc gives $1 - s$ points to a, s points to b and 0 points to c. We can place this information in a matrix.

The vote matrix can be read in two ways. Reading down a particular column we see the number of points given to each candidate from a voter with the ranking indicated by that column. A voter of type cba, for example, gives points of 0, s, and $1 - s$ to candidates a, b, and c respectively. Reading across the rows we see where a candidate's votes come from. Candidate b, for example, gets s votes from each voter of type abc or cba, zero points from

Standardized positional vote matrix						
	abc	acb	cab	cba	bca	bac
a	$1-s$	$1-s$	s	0	0	s
b	s	0	0	s	$1-s$	$1-s$
c	0	s	$1-s$	$1-s$	s	0

each voter of type acb or cab and $1 - s$ points from each voter of type bca or bac.

We write the proportion of voters with ranking abc as p_1 the proportion of voters with ranking acb as p_2 and so forth up until p_6. We call the set $\{p_1, p_2, p_3, p_4, p_5, p_6\}$ a voter profile. We now have all the information we need to find the outcome of any election using a positional voting system. Reading the vote matrix across row a, for example, we see that candidate a receives $p_1 * (1 - s) + p_2 * (1 - s) + p_3 * s + p_4 * 0 + p_5 * 0 + p_6 * s$ votes.

The general positional voting system can therefore be written as:

$$\begin{pmatrix} 1-s & 1-s & s & 0 & 0 & s \\ s & 0 & 0 & s & 1-s & 1-s \\ 0 & s & 1-s & 1-s & s & 0 \end{pmatrix} \begin{pmatrix} p_1 \\ p_2 \\ p_3 \\ p_4 \\ p_5 \\ p_6 \end{pmatrix}$$

$$= \begin{pmatrix} p_1 + p_2 + (-p_1 - p_2 + p_3 + p_6) * s \\ p_6 + p_5 + (p_4 - p_5 + p_1 - p_6) * s \\ p_3 + p_4 + (p_2 - p_3 - p_4 + p_5) * s \end{pmatrix}$$

The equations on the right hand side of the above matrix determine the vote share for a, b, and c respectively. These equations are all linear functions of s, which makes them easy to graph.

By varying s from 0 to $\frac{1}{2}$ we can find, for a given voter profile, what the outcome will be for every possible positional voting system. All three vote shares can be shown on a single diagram. Let the vote share of candidate a be given on the x axis and the vote share of candidate b on the y axis. Since vote shares must add to one the share of c is found implicitly by the distance from the line $a + b = 1$ to the point (a, b) along an orthogonal.

Consider, for example, the profile $\{p_1, p_2, p_3, p_4, p_5, p_6\} = \{0, .419, 0, .258, .322, 0\}$ where .419, for example, indicates that 41.9 percent of the voters are of type 2. We graph the positional vote outcome for

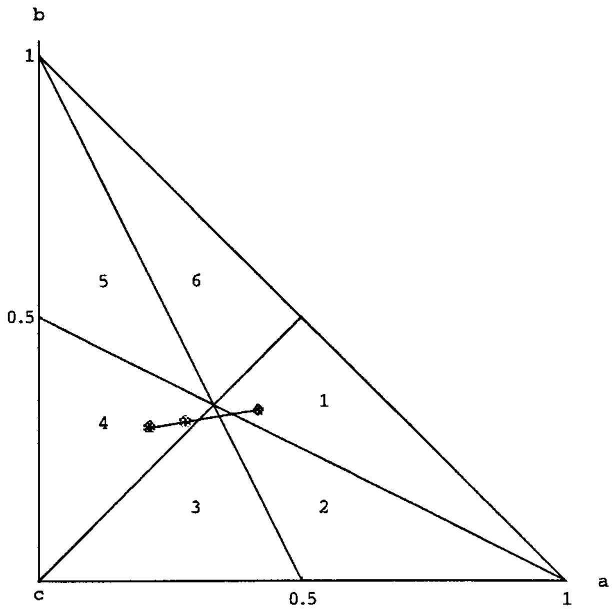

Figure 1. All positional vote outcomes for the profile (0,.419,0,.258,.322,0). The plurality rule outcome is on the right, the Borda Count outcome is in the middle and the anti-plurality rule outcome on the left.

every voting system for this profile of voters in Figure 1. Following Saari (1994), we call the line of vote outcomes the *procedure line*.

In Figure 1 there are six regions labelled (1...6). These six regions correspond to the six possible rankings of candidates, (1) abc, (2) acb, (3) cab, (4) cba, (5) bca, and (6) bac. The center point, where the lines dividing the regions meet, is the outcome where all three candidates are tied for first place.[4] The right-most point along the graphed procedure line is the plurality rule outcome. The left most point is the 'anti- plurality' rule outcome. Anti-Plurality rule assigns one point to each of a voter's top-two ranked candidates (this vote system is called anti-plurality rule because it is equivalent to giving a single negative vote to the last ranked candidate and having the candidate with the fewest negative votes win.) In between the plurality and anti-plurality rule lie the outcomes for all other positional voting systems. The outcome indicated by the middle point, for example, is the Borda Count which is associated with $s = \frac{1}{3}$.[5] The procedure line in Figure 1 tells us that for the profile of voters {0, .419, 0, .258, .322, 0} there are seven possible ordinal rankings (4 strict rankings and 3 rankings involving ties where the procedure line crosses one of the lines separating the regions.)

2.2. *Theory of pairwise voting*

In pairwise voting every candidate is matched against every other candidate and the winner of each contest is decided by majority rule. As is well known, pairwise majority voting can be indeterminate. The pairwise vote matrix is written:

The pairwise vote matrix

	abc	acb	cab	cba	bca	bac
av.b	1	1	1	−1	−1	−1
bv.c	1	−1	−1	−1	1	1
cv.a	−1	−1	1	1	1	−1

Reading across the rows the matrix indicates that in the pairwise election a v. b voters of types 1, 2, and 3 cast votes for a while voters of types 4, 5 and 6 cast their votes against a (for b). If the sum of the votes is positive a wins, if the sum of the votes is negative b wins. An a v. b outcome of $\frac{1}{3}$, for example, indicates that a received a $\frac{1}{3}$ greater share of votes than b. Thus a must have received $\frac{2}{3}$rds of the vote to b′s $\frac{1}{3}$rd. Reading down the columns indicates that a voter of type abc will vote for a in an election of a v. b, b in the election b v. c and a (against c) in the election c v. a. The vote system can then be written:

$$\begin{pmatrix} 1 & 1 & 1 & -1 & -1 & -1 \\ 1 & -1 & -1 & -1 & 1 & 1 \\ -1 & -1 & 1 & 1 & 1 & -1 \end{pmatrix} \begin{pmatrix} p_1 \\ p_2 \\ p_3 \\ p_4 \\ p_5 \\ p_6 \end{pmatrix} = \begin{pmatrix} p_1 + p_2 + p_3 - p_4 - p_5 - p_6 \\ p_1 - p_2 - p_3 - p_4 + p_5 + p_6 \\ -p_1 - p_2 + p_3 + p_4 + p_5 - p_6 \end{pmatrix}$$

2.3. *Theory of approval voting*

In a remarkable group of papers in the late 1970’s, five sets of researchers independently invented a new form of voting now known as Approval Voting (AV). (Steven J. Brams and Peter C. Fishburn have done the most to analyze and popularize this form of voting. See their 1983 book for an overview.) Approval voting lets a voter vote for as many different candidates as she ‘approves.’ If there are 5 candidates, for example, the voter can vote for 1,2,3,4 or even all 5 candidates. (Giving one vote to every candidate is equivalent to not voting so we would not expect to see this happen very often.)

An important feature of approval voting is that a voter profile is not sufficient to determine a unique election outcome. Consider two voters Joe and

Linda who both have the ranking abc. Joe could vote for his top-two ranked candidates a and b while Linda could vote for a only. The decision to cast a top-two ballot or a top-one ballot will depend on preferences (and also on beliefs about what other voters will do.) If Joe greatly prefers either a or b to c he is likely to cast a top-two ballot, while Linda may cast a top-one ballot if she greatly prefers a to either b or c. A single profile is thus associated with many different outcomes depending on the proportion of each type of voter who casts a top-two or top-one ballot. Let the proportion of type one voters casting top-two ballots be r_1 the proportion of type two voters casting such ballots be r_2 and so forth. We are interested in finding the AV outcome for every possible combination of $\{r_1, r_2, r_3, r_4, r_5, r_6\}$ where the r_i's are independent and each $r_i \in [0, 1]$.

Before we find the AV outcome for every combination of $\{r_1, r_2, r_3, r_4, r_5, r_6\}$ it will be useful to examine a simpler problem. Suppose the proportion of voters casting top-two ballots is the same for every type of voter. Call this proportion r. Assume that $r = 0$, then every voter casts a top-one ballot and approval voting gives the same outcome as plurality rule. Now assume that $r = 1$, then every voter casts a top-two ballot and approval voting gives the same outcome as anti-plurality rule. More generally approval voting with proportion r gives the same outcome as positional voting system $s = \frac{r}{2}$. The procedure line is thus a subset of the total set of possible approval vote outcomes. It follows that if a ranking occurs under some positional vote system then it also occurs under approval voting. Geometrically, if the procedure line crosses a boundary then the convex hull of AV outcomes also crosses that boundary. This leads to the following theorem (Theorem One of Saari and Newenhizen (1988a)):

Theorem 1. A necessary (but *not* sufficient) condition for all the AV outcomes to be within one ranking region is that every positional voting system gives the same ranking.

The bracketed qualifier is the most important aspect of Theorem 1. It is quite possible that every positional vote system ranks the candidates in the same way yet there are multiple rankings possible under AV. Indeed, despite the fact that every positional vote system ranks the candidates the same way, AV may be completely indeterminate, i.e. AV may allow every ranking as a possible outcome. We discuss this possibility further below.

Referring again to Figure 1 we now know that each of the 7 possible rankings under different positional vote systems is also a possible outcome under approval voting. But the procedure line is only a subset of the total approval vote outcomes. To find the other possible approval vote outcomes we need to let r vary independently for each type of voter.

Calculating the approval vote outcome for every combination of $\{r_1, r_2, r_3, r_4, r_5, r_6\}$ would be an impossible task if we could not take advantage of convexity. Fortunately, the space of profiles is convex, the space of outcomes (the representation triangle) is convex and approval voting is a convex (linear) mapping from the space of profiles to the representation triangle. A convex mapping from a convex domain creates a convex image. A vertex in profile space will therefore map to a vertex of the AV outcomes in outcome space. The space bounded by the convex hull of the outcome vertices is the set of all possible approval vote outcomes (for a given profile). A simple example will illustrate. Suppose all voters are of type 1. To find all possible AV outcomes we find the AV outcome where every voter casts a top- one ballot. We then find the AV outcome where every voter casts a top-two ballot. The line connecting these outcomes gives the AV outcome for every other proportion. Suppose all voters are of types one and two. We now have four profile vertices to consider as shown in Table 1. By connecting the outcomes associated with each profile vertex we find the space of possible AV outcomes.

Table 1.

Possible ballots when all voters are of types one or two	
$r_1 = 0, r_2 = 0$	$r_1 = 0, r_2 = 1$
$r_1 = 1, r_2 = 0$	$r_1 = 1, r_2 = 1$

There are six types of voters and each type can cast a top-one or top-two ballot so in general there are $2^6 = 64$ possible ballot vertices. Thus, to find all possible AV outcomes, we compute the outcome for each of the 64 ballot vertices and then plot the convex hull of the AV outcome vertices.

Approval voting has been adopted for use by the US National Academy of the Sciences, the Mathematical Association of America, the Institution of Electronics Engineers (a 300,000 member organization) and many other groups and societies in the United States (Brams and Nagel (1991), Merrill (1988)). Although it has not yet been used in US elections, approval voting has been used for referenda in Oregon and a bill permitting the use of approval voting in public elections has passed the North Dakota senate (Weber, 1995). A modified form of approval voting is used to select the secretary general of the United Nations. Along with cumulative voting (discussed below) approval voting has been forwarded as an alternative solution to the problem of minority vote dilution in the United States (the current solution being racial gerrymandering). The increasing use of approval voting in the US indicates that an investigation of its properties is of more than theoretical interest.

2.4. *Theory of cumulative voting*

Under cumulative voting a voter is given as many votes as there are candidates. Voters are allowed to allocate their votes in any way. If there are three candidates a rational (and sincere) voter will give all 3 votes to her top ranked candidates or 2 votes to her top ranked candidate and 1 to her second ranked candidate (giving 1 vote to all three candidates is equivalent to not voting). Cumulative voting systems may also allow some non-integer allocations of votes. Often, for example, a statement of indifference is allowed so a voter could choose either of the above strategies or she could allocate $\frac{3}{2}$ votes to her first and second ranked candidates.

A single profile can lead to many different outcomes under cumulative voting, just as with approval voting. With three candidates and no indifference ballots, cumulative voting allows $2^6 = 64$ possible ballot vertices. If indifference ballots are allowed there are $3^6 = 729$ possible ballot vertices. To find all possible cumulative voting outcomes we find the outcome under each of these ballot vertices and then plot the convex hull of the outcome vertices.

Cumulative voting was used in one form or another to elect members of the Illinois general assembly from 1870 to 1980 (Sawyer and MacRae (1962), Moore (1920)). Recently, the Chicago Tribune (05/30/95) has editorialized in favor of a return to CV in Illinois. Cumulative voting has also been promoted by many civil rights activists, most notably Lani Guinier (1994). Since 1985 nearly one hundred local jurisdictions, small cities, and school districts have adopted cumulative voting in Illinois, New Mexico, South Dakota, Alabama and Texas, mostly in response to lawsuits brought under the Voting Rights Act (Brischetto (1995)). In 1994 a federal judge *ordered* Worcester Count, Md. to adopt cumulative voting to correct problems of minority representation (*Cane v. Worcester County*). In 1995 a bill (HR 2545) was put forward in Congress to allow the use of cumulative voting in congressional elections. (The adoption of CV for such elections would not require a constitutional amendment.) As with AV, the widespread and growing use of cumulative voting makes an analysis of its properties of practical as well as theoretical interest.[6]

3. The 1992 election

3.1. *What would have happened under alternative positional vote systems?*

We use an October 1992 poll of registered voters to classify voters into the six types.[7] In addition to first placed choices, the poll asked "Suppose you could have a second choice, which ticket would be your second choice." Given a three candidate election, a second choice tells us the complete ranking. Of the 3536 registered voters polled, 2489 expressed an opinion as to a second choice.[8] The poll results are given in Table 2.

Table 2.

First and second preferences in the 1992 US presidential election			
	First preference		
Second Preference ⇓	Clinton (1260)	Bush (958)	Perot (271)
Clinton	–	524	145
Bush	519	–	126
Perot	741	434	–

Using the information from Table 2 we arrive at the following profile $\{20.85\%^{CBP}, 29.77\%^{CPB}, 5.83\%^{PCB}, 5.06\%^{PBC}, 17.44\%^{BPC}, 21.05\%^{BCP}\}$, where the superscript indicates the ranking. Given this profile what would have happened under each of the infinite number of positional vote systems? Figure 2 plots the outcome for all positional vote systems.

Surprisingly, the figure indicates that Clinton would have won the 1992 election under every positional vote system. (The right-most point in region one is the plurality rule outcome, the left most point the anti-plurality rule outcome and the inner point the outcome using the Borda Count.) The given profile, however, overestimates Clinton's and Bush's plurality rule vote shares and underestimates Perot's vote share. The actual and predicted vote shares are given in Table 3.

It is evident that the profile over-predicts Clinton's victory mostly at the expense of Perot. We will use the information from the plurality rule outcome to adjust our voter profile for accuracy. The poll over estimates voters with Clinton ranked first and under estimates voters with Perot ranked first, but within these categories (CPB, CBP and PCB, PBC respectively) we don't know exactly which are over and under estimated. Suppose that the poll was accurate at the time it was taken but that some voters changed their minds by the time of the election. It is reasonable to assume that voters with the

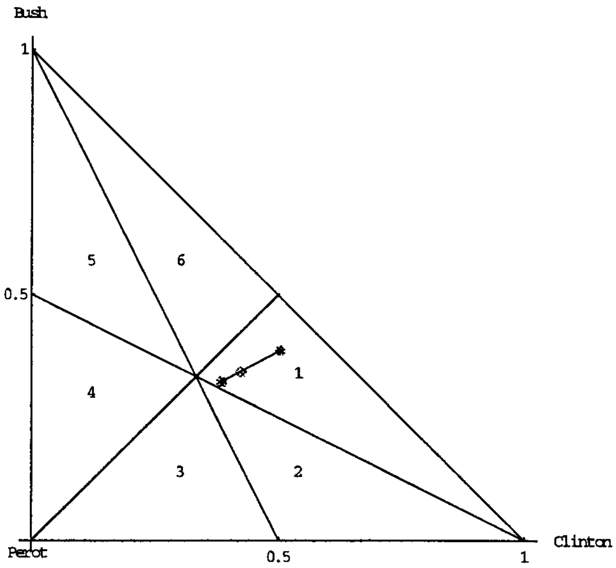

Figure 2. Vote outcome under all positional vote systems, 1992 US Presidential election (initial profile)

Table 3.

Actual and predicted plurality rule vote shares (profile 1)

	Actual	Predicted
Clinton	42.95%	50.62%
Bush	37.40%	38.49%
Perot	18.86%	10.89%

ranking CPB were more likely to switch to Perot (specifically PCB) than voters with the ranking CBP and similarly BPC voters were more likely to switch to Perot than BCP voters. Thus we move 7.67% of the total voters from type CPB to type PCB and we move 1.09% of the total voters from BPC to PBC. The adjusted profile vector is {$20.85\%^{CBP}$, $22.10\%^{CPB}$, $13.5\%^{PCB}$, $6.15\%^{PBC}$, $16.35\%^{BPC}$, $21.05\%^{BCP}$}. We will use this profile vector in further calculations (conclusions are robust to this and similar changes). The adjusted voter profile results in a new plurality outcome given in Table 4. By construction the outcomes for Clinton and Bush are exactly as occurred. Perot's predicted vote share is within 1% of the actual share.

The new procedure line is graphed in Figure 3.

Table 4.

Actual and predicted plurality rule vote shares (profile 2)

	Actual	Predicted
Clinton	42.95%	42.95%
Bush	37.40%	37.40%
Perot	18.86%	19.65%

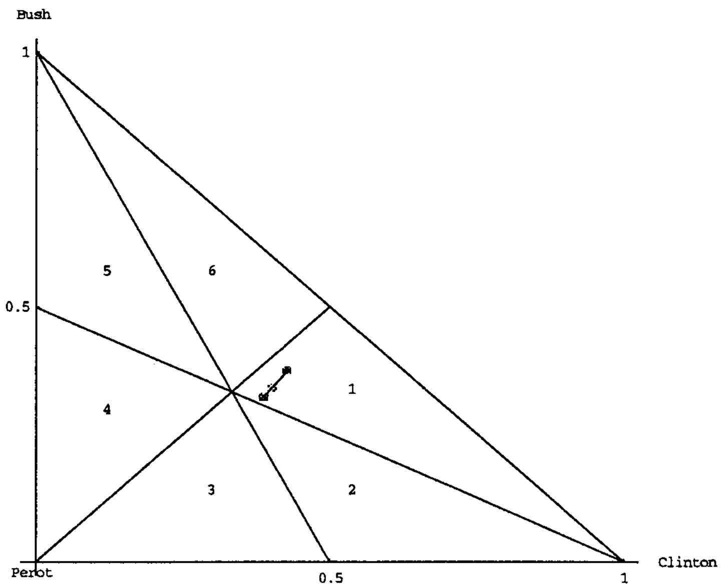

Figure 3. Vote outcome under all positional vote systems, 1992 US Presidential election 1992 (refined voter profile)

Once again, Clinton would have won under any of the infinite number of positional vote systems. Conservative commentators emphasized Clinton's failure to receive more than 50% of the vote in 1992 and thus his failure, in their minds, to achieve a "mandate." The conditions necessary for achieving a mandate are rarely spelled out but a candidate who had enough votes to win under *any* positional vote system surely has a strong claim. Alternatively, the 50% hurdle suggests another condition, a candidate must be the Condorcet winner if he is to achieve a mandate. While it is true that candidates who receive over 50% of the vote are always Condorcet winners, Condorcet winners do not always achieve over 50% of the vote. The 1992 election illustrates.

Table 5.

Pairwise votes 1992 election	
Contest ⇓	Outcome
Clinton v. Bush	Clinton wins by 12.9 points
Bush v. Perot	Bush wins by 16.5 points
Clinton v. Perot	Clinton wins by 28 points

Using adjusted voter profile we can calculate the pairwise votes, these are given in Table 5.

In 1992 Clinton was the Condorcet winner, easily beating both Bush and Perot in pairwise contests. In a pairwise contest, Bush would also have beaten Perot but by much less than Clinton (Perot was therefore the Condorcet loser). The large difference in the Bush v. Perot and Clinton v. Perot pairwise elections follow from the fact that Clinton voters ranked Perot second more often than they ranked Bush second. In the Bush v. Perot election, therefore, Perot picked up relatively more Clinton voters than did Bush. Bush voters, however, ranked Clinton second more often than they ranked Perot second. Thus, in the Clinton v. Perot election Clinton picked up relatively more voters than did Perot.

Returning to Figure 3 note that as more and more weight is placed on a voter's second ranking the outcome moves towards regions 2 and 3. In region 2 the outcome is CPB and in region 3 the outcome is PCB. Any change in the positional vote system, therefore, would have benefited Perot. Under the Borda Count, for example, the vote outcome would have been as in Table 6. Note that Perot increases his vote share relative to both Clinton and Bush but the largest gain comes at Bush's expense.

Table 6.

Borda count outcome 1992	
Clinton	40.15%
Bush	33.93%
Perot	25.91%

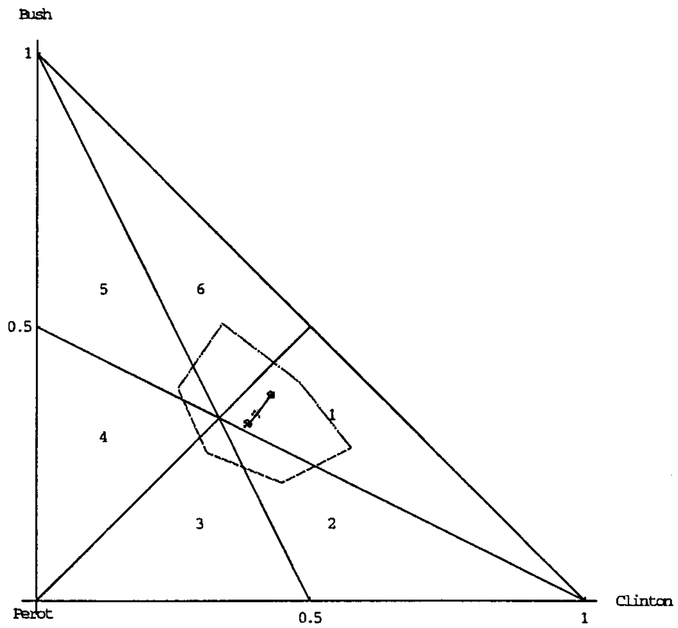

Figure 4. All possible approval and positional vote outcomes, 1992 US Presidential election

3.2. *What could have happened under approval voting?*

In Figure 4 we show every outcome that could have occurred under approval voting (the area encircling the central region) along with the procedure line. Despite the fact that Clinton would have won under every positional vote system and despite the fact that Perot would have lost under every positional vote system, and despite the fact that Clinton was the Condorcet winner and despite the fact that Perot was the Condorcet loser, *Clinton could have come in last and Perot first under approval voting*.[9] (Amazingly, these facts do not exhaust Clinton's advantages and Perot's failings, see further below.)

Under approval voting anything could have happened in 1992! This does not mean, of course, that the 1992 election would have been unpredictable if approval voting were used as a voting system. It's quite probable, for example, that Clinton would still have won. Rather the result indicates something about approval voting in general. Approval voting is inconsistent with *any* standard of voting based on the ordinal voter rankings. AV is inconsistent, for example, with any standard of voting such as Condorcet winners should never come in last, Condorcet losers should never win, candidates with the most last ranked votes should never win etc. All of these things are possible under AV. The analysis of this paper thus supports Saari and Newenhizen's

(1988a,b) contention that the properties of approval voting "appear to be sufficiently bad to disqualify approval voting as a viable reform alternative."[10]

To further illustrate the workings of approval voting we show in some detail *one* of the ways in which Perot could have won the 1992 election. We focus on the AV outcome vertex in region 3. At this vertex, Perot wins the election with 41.96% of the vote, Clinton comes in second with 31% and Bush comes in last with 27%. The region 3 vertex occurs when preferences are given as in profile 2 and voters cast ballots according to Table 7.

Table 7.

Ballot matrix						
	CBP	CPB	PCB	PBC	BPC	BCP
C	1	1	0	0	0	0
B	0	0	0	0	1	1
P	0	1	1	1	1	0

Table 7 shows that if CPB and BPC voters cast top- two ballots while all other voters cast top-one ballots then Perot will win. Intuitively an outcome like this could occur if CPB voters rank Perot higher ('closer' to Clinton) than CBP rank Bush (and similarly for BPC and BCP voters). In an election with multiple issues this is certainly possible. Indeed, for every outcome there is a story about relative intensities of preference which 'justifies' that outcome. Brams, Fishburn, and Merrill (1988), for example, justify a series of odd outcomes using stories about relative intensity and conclude that the 'flexibility' of AV is in fact a virtue. But suppose that preferences were just as given. Would a Perot victory then have been justified? It is doubtful. Clinton, as noted above, was the Condorcet winner and the all positional vote system winner - thus a very strong candidate for best choice. Perot, in contrast, was the Condorcet loser. Furthermore, Perot is ranked *last* by more voters than any other candidate, 41.9% of all voters rank Perot last. Once again, Clinton shows surprising strength (surprising given the conventional wisdom that the 1992 election was 'close'). Clinton is ranked last by only 22.5% of voters. Table 8 gives the proportion of voters ranking each candidate last.

The only area in which Perot exceeded Clinton was in second place votes where Perot edged out Clinton by less than 4% as indicated in Table 9.

Perot has far fewer first place votes than Clinton, far more last place votes, and only a slight edge in second place votes. How is it that Perot can win? The answer lies in the peculiar way in which approval voting counts and does not count second ranked candidates. Approval voting counts some second

Table 8.

Last place votes (profile 2) 1992 election	
Clinton	22.5%
Bush	35.6%
Perot	41.9%

Table 9.

Second place votes (profile 2) 1992 election	
Clinton	34.55%
Bush	27.0%
Perot	38.45%

rankings as fully equivalent to first rankings while it counts other second rankings not at all. Put differently, approval voting misrepresents and ignores vital information. Consider a CPB voter who casts a top-two ballot. By counting each vote equally, AV misrepresents this voter's preferences. The voter ranks Clinton before Perot but is forced to vote as if they were ranked equally. By counting each vote equally, AV throws out the information that Perot is ranked second not first. A CPB voter who casts a top one ballot is equally misunderstood and ignored. By not counting this voter's second ranked candidate at all AV ignores the information that Perot is ranked second. Moreover, AV misunderstand CPB voters because it cannot distinguish between CPB and CBP voters who cast top-one ballots. When so much information is misrepresented or ignored it's not surprising that AV can lead to disturbing results. Thus, Perot wins because AV treats every second place Perot vote as equivalent to a first placed vote and because it throws out everyone else's second placed votes.[11]

Every AV outcome can be justified by some assumption about intensities of preference.[12] But AV provides neither a necessary nor sufficient condition to connect preferences and outcomes. Moreover, to overcome Perot's status as an all positional-system loser, a Condorcet loser etc. requires very strong and precise assumptions about relative intensities of preferences. AV does not respond to relative intensities in anywhere near the precision required to give us confidence in a Perot victory. To *justify* a Perot victory requires that many voters are *very* near indifferent between Clinton and Perot and that many other voters are *very* near indifferent between Bush and Perot. But a

Perot victory could easily occur when voters are *not* nearly as indifferent as required. Because AV counts some second place votes as equal to first place votes and other second place votes not at all, AV gives only a crude and inaccurate measure of relative intensities.

The above argument can be made precise. Relative intensity arguments imply that there is a trade-off between intensity of preference and numbers of voters. If some voters have intense preferences about av.b their preferences may legitimately trump those of a more numerous group whose preferences are less intense. Alternatively stated, a group with less intense preferences may trump those with more intense preferences if the former group is numerous enough. Implicit in every voting system is a *standard of intensity* that determines how intensity and numbers of voters are to be traded.

Suppose that x% of voters are of type bca and the remainder are of type abc, $(0 < x < 100)$. So long as abc voters strictly prefer a to b, the intensity argument asserts that there is some x^* such that if $x > x^*$ then b *should* win and if $x < x^*$ then a should win.[13] x^* is the standard of intensity, it represents a trade-off between numbers of voters and intensity of preference. Plurality rule places zero weight on intensities and defines $x^* = 50$. The Borda Count defines $x^* = 33.33$. Because the Borda Count places positive weight on intensities, fewer bca voters are required in order for b to win, compared with plurality rule. One can argue whether the Borda Count defines the ideal standard but if one makes an intensity argument one is committed to the proposition that an ideal standard exists.

Theorem 2: AV is inconsistent with any standard of intensity. (Under AV, x^* does not exist.)

Proof: Let there be x% of type bca voters who cast $\{1, 0, 0\}$ ballots and $(1 - x)$% of type abc voters who cast $\{1, 1, 0\}$ ballots. Assume that the relative intensities, $U_{bca}(a) - U_{bca}(b)$ and $U_{abc}(a) - U_{abc}(b)$, justify a bca outcome given x.[14] (The relative intensity argument guarantees that some such set of intensities exists.) If a *standard of intensity* exists, then there exists an x^* such that the outcome changes to abc when $x < x^*$. But under AV there is no such x^*, the outcome remains bca regardless of how small the population of bca voters becomes relative to the population of abc voters. ■

Theorem 2 shows that the vaunted ‘flexibility’ of AV cuts both ways. Relative intensity arguments guarantee that a standard can be found that justifies any AV outcome, but it is also guaranteed that some AV outcomes will violate that same standard. AV, therefore, cannot be *systematically* justified by relative intensity arguments.[15,16]

In the above analysis, every possible (sincere) approval vote outcome has been shown. This is not to suggest that in any actual election "anything can happen." Given specific cardinal preferences a specific outcome can be predicted to occur. Following Kiewiet's (1979) "what if" analysis of the 1968 election it has become common to use thermometer scores to estimate which ballots will be cast. In a thermometer poll voters rank candidates on a 0 to 100 thermometer according to how "warmly" they feel towards them. Let a voter rank the candidates abc. It is assumed that the voter always votes for his top ranked candidate, a, and that he votes for his second ranked candidate iff i) $U(b) \geq 50 \geq U(c)$ or ii) $U(b) \geq [U(a) + U(c)]/2$ and $U(c) > 50$. Brams and Merrill (1994) use this method to predict one of the infinitely many AV outcomes in the 92 election. As a way of analyzing what might have happened at a particular time and place this method may sometimes be useful although it is not without problems.[17] The method is misleading, however, if it is used to draw or motivate systematic conclusions (as Brams and Merrill (1994) clearly intend). To properly evaluate AV we need to know that AV is inconsistent with any standard of voting based upon ordinal preferences. We need to know, in other words, that under AV a Condorcet loser can win and a Condorcet winner can lose, that the candidate with the fewest first place and most last placed votes can win and a candidate with the most first place and fewest last placed votes can lose. A complete analysis of the 1992 election illustrates all of these points.[18] The traditional answer to this objection is that AV transcends ordinal preferences to take into account cardinal preferences. Theorem 2 shows that AV cannot be defended on these grounds since it is also inconsistent with any cardinal standard of intensity.

Another defense of AV is that it will perform well under strategic voting. Space precludes a complete analysis of strategic voting in this paper. A few comments, however, are in order. One difficulty with analyzing strategic voting is that the literature offers several different definitions of what it means to vote strategically, especially under AV. Suppose, for example, that a voter has the ranking abc. A "sincere" voter could cast either a top-one or top-two ballot; should the choice of which ballot to cast be considered a strategic choice or should the term strategic be reserved solely for a misrepresentation of ordinal preferences? Furthermore, strategic voting can be defined parametrically as in Myerson and Weber (1993) or in terms closer to that of Nash equilibrium as in Saari (1990). Which of these several definitions is chosen will greatly affect one's conclusions. Thus, Brams and Fishburn (1983) find that AV is among the least manipulable of voting systems while Saari (1990) and Niemi (1984) find that it is among the most manipulable.

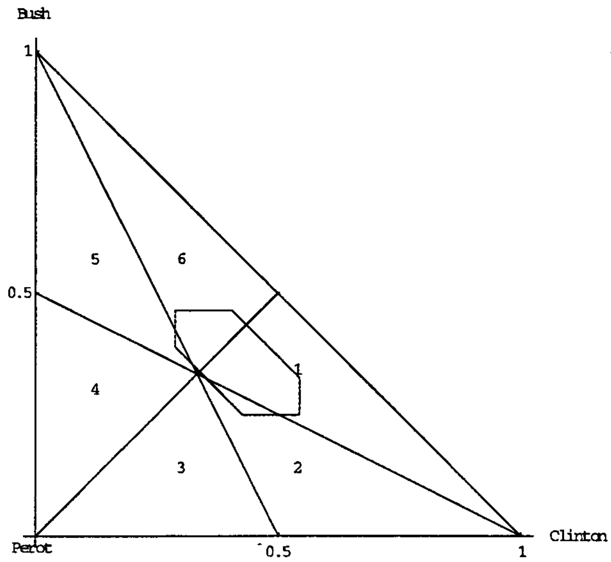

Figure 5. All possible cumulative vote outcomes, 1992 US Presidential election (no indifference ballots allowed)

3.3. *What could have happened under cumulative voting?*

The analysis of cumulative voting is very similar to that of approval voting. In Figure 5 we show every possible cumulative vote outcome when indifference ballots are not allowed. The figure indicates that under cumulative voting Perot could neither have won nor tied for first place in 1992.

The range of possible outcomes increases, however, if we allow voters to cast ballots indicating indifference among candidates. If indifference ballots are allowed then anything can happen under cumulative voting just as with approval voting. In Figure 6 the outermost curve, labelled CV-I, encloses all the outcomes which are possible under CV with indifference ballots allowed. The similarly shaped inner curve, labelled CV, is CV when indifference ballots are not allowed (from Figure 5). For comparison the figure also includes all potential AV outcomes and the procedure line.[19] Notice that CV with indifference and AV have similar potential outcomes but that neither set of outcomes is a subset of the other. Figure 6 summarizes our analysis of the 1992 election.

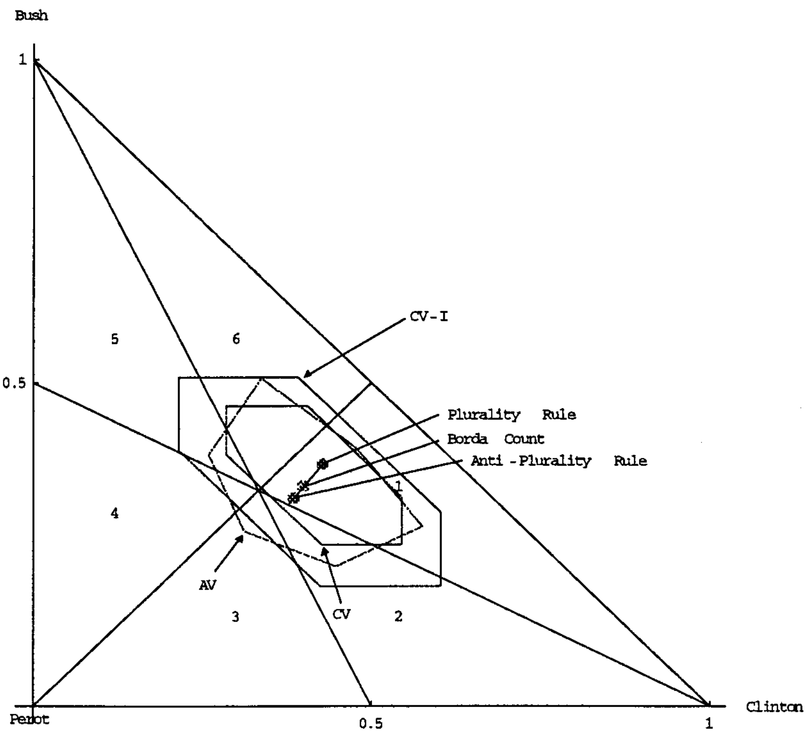

Figure 6. All positional, approval, and cumulative vote outcomes, 1992 US Presidential election

4. Conclusions

Clinton's victory in 1992 was surprisingly strong. Clinton would have won under any positional vote system, Clinton was the Condorcet winner, and he had more first place and fewer last place votes than any other candidate. Perot, in contrast, would have lost under all positional vote systems, was the Condorcet loser and had fewer first place and more last place votes than any other candidate. Plurality rule often hides the strengths of a candidate who loses but it can equally well hide the strengths of a candidate who wins, as it did with Clinton in 1992.

The analysis of the 1992 election illustrates and explains much of fundamental voting theory. Multiple voting systems like approval and cumulative voting are highly indeterminate and highly sensitive to the ways in which voters mark their ballots. Despite Clinton's strengths and Perot's failings, Perot could have won in 1992 and Clinton could have come in last under approval voting and at least one version of cumulative voting. It is extremely difficult to justify such outcomes. Since *every* ranking was a possible outcome under AV or CV with indifference ballots, AV and CV are inconsistent with *any* standard of voting that justifies a specific outcome using information from the voter profile. An extended analysis of one possible Perot AV victory indicates that such disturbing outcomes occur because AV misrepresents and ignores key pieces of information.

Notes

1. The review is based on Saari (1994). A similar review can also be found in Tabarrok and Spector (1999).
2. Positional voting systems other than plurality rule and the Borda Count are sometimes used in practice. The Baseball Writers Association of America, for example, asks two members in each city with a major league baseball team to rank their top ten players. A writer's top ranked player is assigned 14 points, the second ranked player 9 points, the third ranked player 8 points, and so forth all the way down to the last ranked player who is assigned 1 point.
3. This characterization of positional voting systems and the graphical interpretation given below are due to Saari (1994).
4. The three lines separating the regions are lines of equal vote shares. Along the 45 degree line from the origin, for example, candidates a and b have an equal share of votes. At the origin a and b each have 0% of the votes and c has 100%. Moving along the 45 degree line a and b increase their vote shares at the expense of c. At the right-most point along the 45 degree line, therefore, a and b each receive 50% of the votes and c receives 0%. Points below the 45 degree line are areas where $a > b$ and vice versa for points above the 45 degree line. Similar reasoning using the other two lines can be used to deduce the rankings of each region from the graph alone.
5. Plurality rule corresponds to $s = 0$, the Borda Count to $s = \frac{1}{3}$ and anti-plurality rule to $s = \frac{1}{2}$. Since $\frac{1}{3}$ is closer to $\frac{1}{2}$ than to 0 the Borda Count outcome is closer on the procedure line to the anti-plurality rule outcome than to the plurality rule outcome. This makes identifying each outcome simple.
6. Cumulative voting is also used by many firms for corporate governance (see for eg. Bhagat and Brickley (1984)). Homogeneity of interest and ease of exit make it difficult to compare CV in the political and economic contexts.
7. The poll was commissioned by the Times-Mirror Center for the People and the Press. We thank the Pew Research Center, Washington, D.C., for providing this data.
8. The results do not change when various procedures are used to allocate voters who did not indicate a second preference into one of the categories.
9. Perot comes in first and Clinton last in any outcome within region 4.
10. Brams, Fishburn, and Merrill (1988) respond to Saari and Newenhizen (1988) in the same issue.
11. Two simple examples can illustrate these defects. Assume that $100,000$ voters rank the candidates abc and that they rate b nearly as highly as a and so cast top-two ballots. 1 voter ranks the candidates bac and casts a top-one ballot. Amazingly, b wins despite the fact that more than 99.99% of all voters prefer a. Here approval voting fails because it counts the second rankings of abc voters as highly as their first rankings. Now consider a situation where 51% of the voters have the ranking abc and these voters cast top-one ballots. The other 49% have the ranking bca and cast top-one or top-two ballots. In this

case, approval voting ignores the second rankings of the abc voters and so picks a as the winner despite the fact that nearly as many voters rank b first and in addition no voters rank b last (49% of voters rank a last).

12. Without strong utilitarian assumptions it is difficult to justify placing great weight on measures of intensity of preference. If I rank the candidates abc and you rank them acb why do my preferences over the pair $\{a, b\}$ count for less than your preferences? Space precludes an adequate review of this issue. It is worth noting, however, that the Borda Count alone among positional vote systems can be defended in ordinal terms (ie. without reference to intensities of preference). This property follows from the fact that the BC can be understood as a particular method of aggregating the pairwise votes (see Levin and Nalebuff (1995) and Saari (1994).)
13. Candidate c is Pareto dominated by b and so can never win under any reasonable voting system. The handling of the equality case is arbitrary but unimportant. Working in terms of percentages allows us to ignore integer problems, this is a non-essential simplification.
14. I use $U_{abc}(a) - U_{abc}(b)$ to represent the difference in utility levels between outcome a and outcome b when a voter has the ranking abc.
15. Cumulative voting, which we analyze in the next section, lets voters indicate their relative intensities much more clearly than does approval voting. The arguments in the text against AV as a method of measuring relative intensity do not apply to CV. CV, however, may be less robust than AV in the presence of strategic voting. Space does not allow an adequate review of this issue, but see Merrill (1988) for some interesting comparisons.
16. Schwartz (1986) shows more generally that allowing collective choices to depend on preference intensities does not in general eliminate Arrow-like imposibility theorems.
17. Brams and Fishburn (1983) and Merrill (1988), argue that the strategy of voting for all candidates who give the voter greater than average utility will often be optimal. One would have thought, therefore, that a more natural assumption would have been vote for b iff $U(b) \geq \frac{U(a)+U(b)+U(c)}{3}$. That these two behaviour rules sometimes give different answers illustrates only part of the difficulty of prediciting which ballots voters will choose to cast. We have very little information about how voters cast ballots under approval voting. See further below on strategically sincere and strategic voting.
18. Using the thermometer method, Brams and Merrill (1994) estimate that the final tallies under approval voting would have been Clinton (38.37%), Bush (34.49%), and Perot (27.12%), this is quite close to the Borda Count outcome indicated in Figure 4. Note that I have converted Brams and Merrill's numbers into relative shares.
19. A reader pointed out that by dividing the portion of the CV and AV area in regions 3 and 4 with the respective total areas we can find the proportion of outcomes in which Perot wins. Perot wins in 9.65% of all AV outcomes but 0% of CV outcomes with no indifference ballots allowed and 5.69% of CV outcomes when indifference ballots are allowed. Since Perot is the Condorcet loser and the candidate with the least first and most last place votes the relative stinginess of CV is appealing. This computation is loose, however. Since profile space is a 5-dimensional simplex, while outcome space (the triangle) is only a 2-dimensional simplex, outcome areas do not correspond with profile probabilities (not even equally weighted profile probabilities).

References

Bhagat, S. and Brickley, J.A. (1984). Cumulative voting: The value of minority shareholder voting rights. *Journal of Law and Economics* 27: 339–365.

Brams, S.J. and Fishburn, P.C. (1983). *Approval voting.* Boston: Birkhauser.

Brams, S.J., Fishburn, P.C. and Merrill, S. III. (1988). The responsiveness of approval voting: Comments on Saari and Van Newenhizen. *Public Choice* 59: 121–131.

Brams, S.J. and Nagel, J.H. (1991). Approval voting in practice. *Public Choice* 71: 1–17.

Brams, S.J. and Merrill, S. III. (1994). Would Ross Perot have won the 1992 Presidential election under approval voting? *PS: Political Science and Politics* 29: 39–44.

Brischetto, R. (1995). Cumulative voting as an alternative to districting. *National Civic Review* 84: 347–354.

Guinier, L. (1994). *The tyranny of the majority.* New York: Free Press.

Joslyn, R.A. (1976). The impact of decision rules in multi-candidate campaigns: The case of the 1972 democratic Presidential nomination. *Public Choice* 25: 1–17.

Kellett, J. and Mott, K. (1977). Presidential primaries: Measuring popular choice. *Polity* 9: 528–537.

Kiewiet, R.D. (1979). Approval voting: The case of the 1968 Presidential election. *Polity* 12: 170–181.

Levin, J. and Nalebuff, B. (1995). An introduction to vote counting schemes. *Journal of Economic Perspectives* 9: 3–27.

Merrill, S. III. (1988). *Making multicandidate elections more democratic.* Princeton: Princeton University Press.

Moore, B.F. (1920 (1970)). *The history of cumulative voting and minority representation in Illinois 1870–1919.* New York: Johnson Reprint Corporation.

Myerson, R.B. and Weber, R.J. (1993). A theory of voting equilibria. *American Political Science Review* 87: 102–114.

Niemi, R. (1984). The problem of strategic behavior under approval voting. *American Political Science Review* 78: 952–958.

Saari, D.G. (1990). Susceptibility to manipulation. *Public Choice* 64: 21–41.

Saari, D.G. (1994). *Geometry of voting.* Berlin: Springer-Verlag.

Saari, D.G. and Newenhizen, J. van (1988a). Is approval voting an unmitigated evil?: A response to Brams, Fishburn, and Merrill. *Public Choice* 59: 133–147.

Saari, D.G. and Newenhizen, J. van (1988b). The problem of indeterminacy in approval, multiple, and truncated voting systems. *Public Choice* 59: 101–120.

Sawyer, J. and MacRae, D. (1962). Game theory and cumulative voting in Illinois: 1902–1954. *American Political Science Review* 56: 936–946.

Schwartz, T. (1986). *The logic of collective choice.* New York: Columbia University Press.

Tabarrok, A. and Spector, L. (1999). Would the Borda Count have avoided the civil war? *Journal of Theoretical Politics* 11: 261–288.

Weber, R.J. (1995). Approval voting. *Journal of Economic Perspectives* 9: 39–50.