If you like the alternative vote (a.k.a. the instant runoff), then you ought to know about the Coombs rule

Bernard Grofman\textsuperscript{a,*}, Scott L. Feld\textsuperscript{b}

\textsuperscript{a} Department of Political Science and Institute for Mathematical Behavioral Science, University of California, Irvine, CA 92697-5100, USA
\textsuperscript{b} Department of Sociology, Louisiana State University, Baton Rouge, LA 70803, USA

Received 1 July 2003

Abstract

We consider four factors relevant to picking a voting rule to be used to select a single candidate from among a set of choices: (1) avoidance of Condorcet losers, (2) choice of Condorcet winners, (3) resistance to manipulability via strategic voting, (4) simplicity. However, we do not try to evaluate all voting rules that might be used to select a single alternative. Rather, our focus is restricted to a comparison between a rule which, under the name ‘instant runoff,’ has recently been pushed by electoral reformers in the US to replace plurality-based elections, and which has been advocated for use in plural societies as a means of mitigating ethnic conflict; and another similar rule, the ‘Coombs rule.’ In both rules, voters are required to rank order candidates. Using the instant runoff, the candidate with the fewest first place votes is eliminated; while under the Coombs rule, the candidate with the most last place votes is eliminated. The instant runoff is familiar to electoral system specialists under the name ‘alternative vote’ (i.e., the single transferable vote restricted to choice of a single candidate). The Coombs rule has gone virtually unmentioned in the electoral systems literature (see, however, Chamberlin et al., 1984). Rather than considering the properties of these two rules in the abstract, we evaluate them in the politically realistic situations where voters are posited to have (at least on balance) single-peaked preferences over alternatives. Evaluating the two rules under this assumption, we argue that the Coombs rule is directly comparable in that Coombs is always as good as AV with respect to two of our four criteria and it is clearly superior to AV with respect to one of the four

\* Corresponding author. Tel.: +1-714-824-6394; fax: +1-714-824-8762.
E-mail address: bgrofman@uci.edu (B. Grofman).
criteria, namely criterion (2), and is potentially inferior only with respect to criterion (3). Key to this argument are two new propositions. The first new result shows that, under the posited assumption, for four alternatives or fewer, AV is always as likely or more likely to select the Condorcet winner than plurality. The second new result shows that, under the same assumptions, the Coombs rule will always select the Condorcet winner regardless of the number of alternatives.

© 2003 Elsevier Ltd. All rights reserved.

1. Introduction

Electoral reformers unhappy with the results of plurality-based elections in single member districts (smds) have largely gone in one of two directions: either proposing to replace smds with multimember districts (mmds) while at the same time replacing plurality with some form of proportional or semi-proportional election method; or accepting the continued use of smds, but proposing to replace plurality-based elections with methods that are viewed as less likely to achieve unrepresentative outcomes in any single seat contest in which more than two candidates might compete. Electoral reformers who focus on single seat offices (such as president or governor), or on individual elections within a legislature elected from single member districts, have offered a number of quite different schemes as improvements over plurality, including approval voting, various Condorcet methods, approval voting, and others.

---

1 In Great Britain, and elsewhere in the British Commonwealth, plurality elections in single member districts are commonly referred to as first-past-the-post voting.

2 There are two main types of proportional representation, the various forms of list PR, and the single transferable vote (STV). From the latter half of the 19th century through most of the 20th century, reformers in Great Britain and the British Commonwealth have tended to prefer the latter (see e.g. Lakeman and Lambert, 1955 and 2nd edition, 1959). Today, electoral reformers whose primary focus is partisan fairness are more likely to advocate so-called mixed methods, involving both a single member district (smd) plurality election component and a PR component—where the PR entitlement may or may not be effected by a party’s success in the single member district component of the system (Shugart and Wattenberg, 2000). In the US, a different set reformers—more concerned about racial than about partisan representation—have advocated so-called semi-proportional methods, such as the limited vote and cumulative voting (Guinier, 1994, 1998; see also Grofman, 1982).

3 Under approval voting, voters indicate the set of candidates for some office whom they regard as acceptable by placing a check next to their names. Each such approval vote counts as a single point for the candidate receiving it. Summing over all voters, the candidate with the most points is the approval voting winner. Brams and Fishburn (1978, 1983) have been the most important advocates of the use of approval voting (see also Brams and Herschbach, 2001a,b). In large part due to Brams’ efforts, this rule that has now been adopted by a number of private organizations (Brams, 2002).
dorcet extension methods,\textsuperscript{4} the Borda count,\textsuperscript{5} various forms of runoff,\textsuperscript{6} and sequential single ballot methods that resemble runoffs in their effects, such as the alternative vote, \textit{(AV)}\textsuperscript{7} i.e. the single transferable vote restricted to choice of a single candidate.\textsuperscript{8}

In the United States, arguably the most publicly debated proposal for electoral reform for governmental offices in the past several years has been the idea of replacing plurality with AV\textsuperscript{9}—a voting method which, largely for rhetorical reasons, reformers in the US associated with the public interest lobbying organization, the Center for Voting and Democracy, refer to as the \textit{instant runoff} form of voting, often abbreviated as \textit{IRV} (Richie et al., \textit{2001})\textsuperscript{10} In this essay, we wish to call the

\textsuperscript{4} The Condorcet winner (Black, \textit{1958}) is that alternative, if any, who can receive a majority in paired contest against each and every other alternative. Condorcet extension methods are ones that guarantee to pick a Condorcet winner, if one exists. Perhaps the most important contemporary advocate of the use of Condorcet extension methods is Peyton Young (1977; see also Straffin, \textit{1980}). Black (\textit{1958}) proposed a kind of compromise between Borda and Condorcet: using the Borda rule in situations in which there is no Condorcet winner. Perhaps the two best known Condorcet extension methods are the Copeland rule and the Nanson rule. The Copeland rule picks that alternative (or those alternatives) that receive majority support in pairwise contest against the most other alternatives. The Nanson rule operates sequentially by eliminating from among the (still) feasible set of alternatives all alternatives that have less than average Borda scores. Since the Condorcet winner always has a greater than average Borda score (Nurmi, \textit{2002}: 33), this procedure guarantees that a Condorcet winner will be chosen if one exists. (For more formal definitions of Copeland, Nanson, and other Condorcet extension methods see Straffin (\textit{1980}) and Nurmi (\textit{2002}: 32–33.).)

\textsuperscript{5} The Borda count requires voters to rank order candidates and, for each voter’s ranking, assigns a point to an alternative for each alternative to which it is preferred in that voter’s ranking (Black, \textit{1958}). Summed over all voters, the candidate with the highest Borda count is the Borda winner. The most prominent contemporary advocate of the Borda rule, and one who has provided a number of arguments in favor of its use, is Donald Saari (see Saari, \textit{1995}).

\textsuperscript{6} See Colomer (in \textit{press}) for a discussion of runoff methods used in nations with a directly elected president.

\textsuperscript{7} In this article we use AV as the abbreviation for the alternative vote. We would note that, sometimes, AV is used instead as an abbreviation for approval voting.

\textsuperscript{8} This method operates in a fashion that closely resembles runoffs. As in STV, voters submit a list of ranked-ordered preferences. In AV, the electoral ‘quota’ is 50%+; i.e. if no candidate has a majority of first place votes, then the candidate with the fewest first place votes is deleted and the ballots recording first place votes for that candidate are transferred to the candidate on each ballot who is who is next highest in the rankings. If this does not produce a majority winner, we continue in like fashion until we either get a majority winner, or we are down to two candidates, in which case the one with the most votes (including transfers) is chosen. Assuming no change of preferences or of electorate in between rounds of balloting, the AV version of an instant runoff is equivalent to what has been called the sequential elimination procedure (Black, \textit{1958}) when the latter procedure is used to generate a sequence of runoff elections with fewer and fewer candidates, and which has more recently been labeled MRSE, multiple runoff sequential elimination, by Anthony McGann (\textit{2002}) and his colleagues (McGann et al., \textit{2002}).

\textsuperscript{9} In the US the most salient reform for elections to single seat offices within professional societies has been the proposed replacement of plurality with approval voting (see Brams, \textit{2002}).

\textsuperscript{10} The Center for Voting and Democracy presumably prefers to talk about AV under the name instant runoff because the concept of runoff elections is already familiar to most US voters, and the adjective ‘instant’ suggests (not unreasonably) that this form of runoff will save voters time (and taxpayers money).
attention of the electoral systems community (especially those engaged in the public debate over electoral reforms) to another runoff-like method, the *Coombs rule* (Coombs, 1964) that, like AV, can be viewed as an alternative to plurality-based elections in single seat districts. Like AV, the Coombs rule requires that voters provide rank ordering, and also involves a sequential elimination procedure in which ballots for candidates who have been eliminated are transferred. The difference between the Coombs rule and the alternative vote is simple. In AV, the candidate with the fewest first place votes is eliminated; in Coombs, the candidate with the most last place votes is eliminated.

Persuaded by arguments of electoral reformers, both local and those associated with the Center for Voting and Democracy, in 2002 the San Francisco City Council agreed to put AV (labeled the instant runoff) on the ballot, and the instant runoff was adopted by referendum later that year for use in future city council elections in San Francisco. Its proposed use for legislative elections in Alaska was the subject of an unsuccessful referendum attempt in 2002, and it was also proposed in 2002 for use in elections in Vermont. Perhaps most importantly, the Center for Voting and Democracy and other reformers have proposed IRV as a way to prevent recurrence of some of the problems found in the 2000 Presidential election, where the popular vote winner, Vice-President Al Gore, was defeated because Ralph Nader siphoned off enough first place votes (about 2% nationwide, primarily from those who preferred Gore to Bush), to deny Gore the plurality victory in key states such as Florida.

But the United States is not the only country where AV either is or might be important. AV has a long history of use in Australia, where it is often argued for in terms of fostering majority rule. In England, the alternative vote has been suggested as an electoral system, which, unlike plurality, would reflect the true extent to which voters support centrist parties and policies. For parliamentary elections to the UK, Patrick Dunleavy and other academics have recently advocated a modified form of AV, AV+ (Dunleavy and Margetts, 1997, 1995, 2001; Dunleavy et al., 1992) and, for a while, there appeared to be some real chance that this hybrid system would be adopted. AV has also been advocated as a tool in mitigating ethnic conflict in deeply divided societies by scholars such as Donald Horowitz (1985, 1991a,b, 1993) and Benjamin Reilly (1999). Horowitz’s main claim is that AV increases the likelihood that candidates would be chosen who are moderate with respect to ethnic issues. Horowitz’s ideas have proved influential in places like Fiji and Papua New Guinea. AV has recently been used for two national elections in

---

11 We should note, however, that the Mayor of San Francisco, the Honorable Willie Brown, is a vociferous opponent of this method (remarks at the University of California, Irvine Center for the Study of Democracy ‘Brown-Bruele Fellowship’ Dinner, Newport Beach, California, January 30, 2003).

12 The Center for Voting and Democracy was involved in each of these campaigns.

13 See various essays in Bowler and Grofman (2000).

14 Here, we should admit to some skepticism about some of the advantages for diffusing ethnic conflict claimed by Donald Horowitz. for AV as compared to plurality (see Fraenkel and Grofman, 2002).

The central point of this essay is the claim that many of the same arguments that have proposed to justify replacing of plurality with AV (a.k.a. instant runoff) apply with greater or equal force to the replacing plurality with the Coombs rule. Indeed, despite the considerable potential attractiveness of other rules for picking a single candidate from among a larger set (e.g. Borda, approval voting, or Condorcet extension methods such as the Copeland rule or the Nanson method, we restrict ourselves in this essay to comparing plurality with AV and with the Coombs rule. We limit ourselves in this way for several reasons, including, perhaps most importantly, that of current policy relevance.

The Coombs rule is in a kind of ‘mirror image’ of AV. In particular, the Coombs rule has exactly the same claim to being called an ‘instant’ runoff as does AV. If we can show that it has a number of desirable properties, then it might well be added to the ‘reform menu’; (1) by at least some of the electoral reformers in the United States, who are presently advocating the use of the AV form of an instant runoff, and who have already succeeded in seeing AV implemented in one major US city (San Francisco, for city council posts), and who have been ardently campaigning for its use at all levels of office; and (2) by electoral reformers elsewhere, including those who have advocated AV as a tool for mitigating ethnic conflict. In contrast, methods such as the Nanson rule, despite how highly thought of they might be by some social choice theorists (ourselves included),15 have never, as far as we are aware, ever actually been adopted for use. Furthermore, to extend our analyses to include the whole range of methods available to pick a single alternative from among many would considerably expand the scope of what is intended to be a short essay focused on two new theorematic results, especially since there is no agreement among social choice theorists as to weight to be given the various proposed criteria that can be used to evaluate alternative rules and no agreement, either, about the extent the debate should be focused on the theoretical properties of voting rules as opposed to more empirically oriented investigations of how they are likely to operate in real world settings.16 Finally, since the Coombs rule is virtually unknown outside the social choice literature17 (and is not even especially well known among social choice theorists),18 we see a sufficient justification of this relatively brief essay in calling the attention of electoral system specialists to this potentially quite significant voting method.

There is one novel feature of the analyses we present in the next section that we regard as especially important to emphasize. We believe that, in many contexts, it is reasonable to assume that candidates or parties can be arrayed along a single continuum, and that (roughly speaking) voters share common perceptions of where

---

15 See the excellent overview of the pluses and minuses of a panoply of voting rules in Nurmi (2002).
16 On this latter point, see Regenwetter et al. (in press).
17 See, however, Chamberlin et al. (1984: 489).
18 See, however, Straffin (1980) and Nurmi (2002: 33).
the parties are along this continuum and vote for that party that is closest to the voter’s own ideal point if they have a single vote to cast (see Downs, 1957). Moreover, in such contexts, it is reasonable for a voter to rank order the candidates in order of proximity to the voter’s ideal point if voters are required to submit a rank-ordered ballot. Such assumptions give rise to what, in the social choice literature, are called single-peaked preferences.19

It is common in the comparative politics literature to plot the location of the parties in the major long-term democracies in a single dimension over which preferences are (at least implicitly) assumed to be single-peaked—although the presence of a second and less important dimension is also often recognized. When there is a single main dimension to political conflict, this dimension is usually found to be an ideological dimension (e.g. a left–right dimension defined by attitudes toward the desirable scope of government), but it need not be. For example, in plural bipolar societies, the principal dimension of cleavage might be defined in ethnic terms, with radical views in favor of special privileges and status for one of the two ethnies defining the two poles of the unidimensional continuum, and moderates of each camp being defined vis-a-vis their more moderate attitudes toward ethnic reconciliation (Fraenkel and Grofman, 2002).20 Because of the empirical importance we attach to single-peakedness, we present comparisons among plurality, AV and Coombs primarily for the case where voters have (at least on balance) single-peaked preferences.21

In comparing plurality, AV and Coombs, we will consider four factors relevant to picking a voting rule to be used to select a single candidate from among a set of choices: (1) simplicity, (2) avoidance of Condorcet losers,22 (3) choice of Condorcet

19 Single-peaked preferences occur when there exists a continuum along which alternatives can be ordered such that the utilities of each voter can be plotted as a single-peaked graph, i.e. a curve which changes its slope at most once, from up to down (Black, 1958; for an alternative definition of single-peakedness in terms of the ‘betweeness’ relation, see Arrow, 1962; see also Sen, 1970 for an alternative definition of single-peakedness in terms of the NW (not worst) condition on a triple of alternatives). If \( k \) is the number of alternatives, then there are \( k! \) (\( k \) factorial) possible (linear) preference rankings if ties are not permitted, but only \( 2(k - 1) \) of these rankings are single-peaked with respect to any given unidimensional continuum.

20 Indeed, in modeling the likely impact of AV on voting support for ethnically and non-ethnically based parties, at least when they focus on bi-racial or bi-ethnic contexts, scholars such as Donald Horowitz are implicitly assuming single-peaked preferences over a continuum that is anchored at each end by the more extreme national nationalists of each of the two groups (Fraenkel and Grofman, 2002).

21 For a precise definition of what it means for voters too, on balance, possess single-peaked preferences, see the definition of partial single-peakedness in Feld and Grofman (1986), and the closely related definition of net single-peakedness (single-peakedness on net preferences) in Regenwetter and Grofman (1998).

22 The Condorcet loser (Black, 1958) is that alternative, if any, who loses in paired contest against each and every other alternative.
winners, (4) resistance to manipulability via strategic voting. We will, however, make no attempt to attach weights to each of these criteria.\textsuperscript{23}

Our main results have to do with the third criterion, the Condorcet efficiencies of Coombs, AV and plurality. One new result shows that, single-peaked preferences over a single dimension, for four alternatives or fewer, AV is always as likely or more likely to select the Condorcet winner than plurality. However, we also show that this result fails to go through for more than four alternatives. A second new result shows that, for single-peaked preferences, regardless of the number of alternatives, the Coombs rule must always pick the Condorcet winner.

2. Comparisons of the alternative vote, plurality, and the Coombs rule

As noted earlier, we will consider several criteria across which voting methods can be compared: (1) simplicity, (2) avoidance of Condorcet losers, (3) choice of Condorcet winners, and (4) resistance to manipulability via strategic voting;\textsuperscript{24} and we limit ourselves to comparisons among only three voting rules for choosing a single winner: plurality, the Alternative Vote/Instant Runoff, and the Coombs rule.

2.1. Simplicity

Simplicity is a criterion—arguably one of the few—in which plurality comes out higher than most of its alternatives, including AV and Coombs. Plurality only requires the voter to identify the single most preferred candidate, the latter two rules require the voter to present a (full) ordering of alternatives.\textsuperscript{25}

2.2. Avoidance of Condorcet losers

Recall that a Condorcet loser is one who is defeated by each and every other alternative in paired contest. Certainly, it would seem that any voting method that picked such an alternative as the winner was a poor voting method, indeed.

Yet, it is easy to show that plurality can have this defect. In a well known three way US Senatorial contest in the State of New York in 1970, the candi-

\textsuperscript{23} There are numerous other criteria for evaluating voting rules that have been proposed in the social choice literature (see Nurmi, 2002, for a nice overview). In particular, STV has been attacked because it need not be positive responsive (a.k.a. monotonic, a.k.a. non-perversive), i.e., there may be situations in which having a (set of) voter(s) rank a candidate higher (without changing their relative rankings for the other candidates) may actually turn a winning candidate into a losing candidate (Doron and Kronick, 1977)—and this argument carries over to AV. The same potential problem also applies to Coombs. We will consider non-monotonicity effects when we consider strategic voting later in the paper.

\textsuperscript{24} Other forms of manipulability have to do with the ability to change outcomes by the additions or deletions to the candidate set of non-winning candidates (see discussion in Chamberlin et al., 1984: 491–494) or, for procedures involving sequenced comparisons, such as standard amendment procedure, the ability to affect outcomes by delineating the sequencing order (Straffin, 1980).

\textsuperscript{25} When AV is implemented, election rules may impose requirements that voters rank order all (or some given number of) candidates in order for their ballot to be counted as valid. (see discussion in Bowler and Grofman, 2000).
date running on the Conservative party line (Buckley) arguably was a Condorcet loser in that he would have lost in head to head general election contest to with either the Democrat (Otinger) or the liberal Republican (Goodell). Yet it was this Conservative Party candidate who won the general election with a plurality vote despite his more liberal opponents having accumulated about 60% of the vote between them. Although liberal Republicans are a minority among Republican voters, most liberal Republicans preferred Goodell to Buckley, and in a general election pitting Buckley against Goodell they would have been joined by a high proportion of the Democrats who also would have clearly preferred the liberal Goodell to the conservative Buckley. On the other hand, while Otinger, a Democrat, might not have done as well as Goodell among liberal Republicans in a head to head contest with Buckley, he would have made up for that by getting virtually all the Democratic vote. We can see this as a situation involving single-peaked preferences where the two liberal candidates (Goodell and Otinger) split the liberal vote, allowing the least preferred choice among a majority of the voters to win.

In contrast, neither AV nor the Coombs rule can choose a Condorcet loser. The proof is easy. In both AV and Coombs the winner gets a majority of the votes cast at some stage of the balloting and thus must be able to defeat each of the candidates still viable at that stage in head-to-head contest. Hence the winner under these rules must be able to defeat at least one other candidate in paired contest and cannot be a Condorcet loser.

2.3. Choice of Condorcet winners

Supporters of AV (such as Donald Horowitz) argue that, compared to plurality, AV has a greater likelihood of the end result being a victory by a ‘moderate’ candidate (i.e. one who enjoys broad support but is not necessarily the first place choice of many voters) as opposed to an ‘extremist’ candidate (i.e. one who may have first place support from a substantial number of voters but who would lose to one or more of the other candidates in the contest if there were to be a two-candidate head on head contest). In social choice terminology (Black, 1958; Saari, 1995), we may translate this into the claim that AV has a higher probability of choosing the Condorcet winner (when one exists) than does plurality. The basic argument is simply that, under AV, unlike what is true for plurality, preferences for moderate candidates who are Condorcet winners, but who may not be given much first place support, have at least the potential to be important in deciding election outcomes by virtue of the sequential nature of the ballot transfer process.

2.3.1. Condorcet efficiency of the alternative vote compared to plurality

While it is possible to run simulations of the likelihood that AV and plurality will choose the Condorcet winner under different assumptions about the number of
alternatives and the underlying distributions of voter preferences,\textsuperscript{26} e.g. by assuming that all preference orderings are drawn from the so-called impartial culture,\textsuperscript{27} here we prefer to focus on assumptions that are more realistic than the impartial culture\textsuperscript{28} and on analytic rather than simulation results. In particular, as we noted earlier, we believe that (partial) single-peakedness is a reasonable assumption to make about voter preferences in real world political situations, so we will focus on choices over a single dimension over which voters have, on balance, single-peaked preferences. When preferences are, on balance, single-peaked, there always exists a Condorcet winner, and we may identify the Condorcet winner is the alternative which is most preferred by the median voter (Black, 1958).

We now demonstrate that the intuitions offered by a number of authors as to the superiority of AV over plurality with respect to the likelihood of electing ‘moderate’ candidates can be given a firm theoretic underpinning (at least when the set of alternatives is prespecified). In comparing plurality and the alternative vote, we are really only interested in cases in which no candidate has a majority of first place preferences. A candidate with a majority of first place preferences will, of course, be elected under both plurality and AV. We also assume what are called sincere preferences, i.e., that, in every instance when voter submit their ballot they rank order the candidates from top to bottom of their true preference ordering—and, as noted earlier, we assume single-peakedness.

Proposition 1. If we assume single-peaked preferences over a single dimension and no party holding a majority of first place preferences, and posit that voter preferences are sincere, when we have four parties or fewer, the candidate of the median party is more likely (or at least no less likely) to win when voting is conducted under the alternative vote than when voting is conducted under plurality.

Proof. See Appendix.

On the other hand, once we have five or more alternatives, then the superiority of AV over plurality even when preferences are single-peaked, is no longer assured. We demonstrate this fact by providing an example, for five alternatives, where the Condorcet winner is the plurality winner but is eliminated by the alternative vote sequential choice process.

\textsuperscript{26} As noted earlier, for any given voting method used to select a single alternative, the probability that it picks the Condorcet winner when one exists is referred to in the social choice literature as the Condorcet efficiency of the method (Merrill, 1988). However, Condorcet efficiencies are only defined subject to particular assumption about the distribution of voter preferences.

\textsuperscript{27} The impartial culture is one in which all (strict) orderings are equally likely (Gehrlein and Fishburn, 1976a,b; Niemi and Weisberg, 1968).

\textsuperscript{28} Tsetlin et al. (in press) show that the impartial culture involves a set of assumptions about the nature of underlying voter preferences that generates the sampling distribution that is least likely of all possible distributions to generate a Condorcet winner.
Example 1 (a five alternative example where the Condorcet winner is chosen by plurality but not by AV). Suppose that voters are distributed uniformly from 0 to 100 with the median voter being at 50. Suppose that one candidate, C, is located at 50, and that candidates A, B, D and E are located at 9, 29, 71, and 91, respectively.

We show this situation below:

<table>
<thead>
<tr>
<th></th>
<th>A</th>
<th>B</th>
<th>C</th>
<th>D</th>
<th>E</th>
</tr>
</thead>
<tbody>
<tr>
<td>9</td>
<td>29</td>
<td>50</td>
<td>79</td>
<td>91</td>
<td></td>
</tr>
</tbody>
</table>

Here, of course, C is the Condorcet winner; and we further note that C is also the plurality winner. The first place votes go roughly 19.5 to A, 20 to B, 21 to C, 20 to D, and 19.5 to E. But what happens under AV? Under AV, we would, say, first eliminate A. After the resulting vote transfers, B now gets 39.5. So, then we would eliminate E, and D gets 39.5 votes. Now we are down to the set {B, C, D}. But now AV will eliminate C, since C has only 21 first place choices, fewer than either B or D. So even though the alternative supported by the median voter (the Condorcet winner) was also the plurality winner, it would not the alternative vote winner.

Nonetheless, particular counterexamples notwithstanding, simulation work by Anthony McGann (2002; see also McGann et al., 2002) has shown that, in general AV (or its runoff equivalent, MRSE, multiple runoffs with sequential elimination) has a higher Condorcet efficiency than plurality under single-peakedness over a single dimension almost regardless of what assumptions we make as to the underlying distribution of voter ideal points.

2.3.2. Condorcet efficiency of the alternative vote compared to the Coombs rule

We have previously demonstrated that, when preferences are single-peaked, for a fixed set of four or fewer alternatives, we can show the superiority of AV over plurality in terms of the greater likelihood that the former method will choose ‘moderates.’ Moreover, as noted earlier, it is easy to demonstrate that the alternative vote never picks a Condorcet loser, while plurality can do so. But, if a greater likelihood of avoiding the selection of extremist candidates with substantial first round but little overall support is a central merit of AV as compared to plurality, if you like AV then you should like the Coombs method even more—as the following theorem demonstrates.

Proposition 2. If voters have single-peaked preferences over a single dimension, and voter preferences are sincere, then the Coombs rule always selects the Condorcet winner (i.e., the alternative supported by the median voter).\(^{29}\)

\(^{29}\) Also, even if voter preferences are not completely single-peaked but only net single-peaked (see Feld and Grofman, 1986; Regenwetter and Grofman, 1998) we would still expect that Coombs would pick the Condorcet winner. Because the introduction of the concept of net single-peakedness would take into technical complexities we do not wish to pursue here we refer the reader to the above references. We also omit considerations of a further complication, systematic bias in voter perceptions of candidate/party locations (see Merrill et al., 2001).
Proof. See Appendix.

We can make the intuitive plausibility of this result clearer by returning to our earlier five-candidate example. When we have five candidates and single-peaked preferences, only the ‘extremist’ candidates (A and E in the above example) might be last place choices. Thus, one of these two alternatives is sure to be eliminated under Coombs. Without loss of generality we may let that one be A. Now, on the next round, C still will not be eliminated. Either B or E must be the candidate with the most last place choices. If E is eliminated on this round it is easy to see that we have no problems and that C will ultimately win; but imagine that, instead, B is eliminated on this round. That leaves the choice among the set \{C, D, E\}. But, given its location, C must have gotten the preference transfers from the supporters of the previously eliminated A and B. But, since C is by assumption a Condorcet winner, that means that on the third round, C will win by virtue of getting an absolute majority of the votes.

Since the reader may wonder about the realism of the assumptions we are using it is useful to provide a check with some real world data. Chamberlin et al. (1984) have run mock elections on ballots from the American Psychological Association under a variety of rules including AV, Coombs and plurality. In the data they examine (five contests, under two different assumptions about how full preference orderings are generated, for a total of ten hypothetical elections), there is always a Condorcet winner and Coombs always chooses that winner.\(^{30}\) In contrast, AV fails to pick the Condorcet winner two out of ten times, and plurality fails to do so five out of ten times.

2.4. Resistance to manipulability via strategic voting

Compared to plurality, AV is supposed to offer a reduced incentive for voters to vote ‘strategically’ (e.g. to fail to mark their ballot for, or put at the top of their ranking, the candidate who they actually most want) because of their fear of casting a ‘wasted’ vote. In AV, unlike plurality, it is argued, voters should have much less worry that failing to give their first place support to a desirable candidate with a greater chance of winning the election than the candidate they sincerely most prefer might lead to the election of a candidate whom the voter truly did not want. The need to consider the strategic voting option (a) imposes cognitive problems for voters in deciding how to vote to best achieve their goals, and further may be viewed as (b) is not desirable in and of itself because voters should be using the ballot box to indicate their sincere preferences. Thus the supposed reduction in the need for strategic calculations in AV as opposed to plurality can be considered an independent argument for the superiority of AV, above and beyond any greater efficacy it may have in selecting Condorcet winners.

For any situation in which we can find strategic voting incentives under Coombs we can create a situation with strategic incentives under AV. However, the incen-

\(^{30}\) We conjecture that these data sets satisfy net single-peakedness, but, unfortunately, reanalysis is not possible (Personal communication, John Chamberlin, May 2003). The present authors (and Michel Regenwetter) are in the process of examining data from other contests using STV/AV, including data from more recent elections of the American Psychological Association.
tives for strategic voting under Coombs and AV are not really the same. As Chamberlin et al. (1984) point out, strategic voting under AV exploits its non-monotonicity property, in that a candidate may actually be helped if fewer of its supporters votes for it. But this can occur only if the votes that would have gone to that candidate are shifted in such a way as to eliminate in the early rounds a candidate with substantial later round strength who could defeat that candidate, and to leave in the competition only candidates whom that candidate can defeat in later rounds of voting. But it requires a certain level of sophistication to implement such a strategy, i.e. to persuade supporters that, ‘in order to help B win,’ they must ‘vote for D’ (Chamberlin et al., 1984: 497). In contrast, as Chamberlin et al. note (1984: 495), strategic voting under Coombs involves lowering the rank in your submitted preference ordering of the candidate(s) whom you regard as the greatest ‘threat’ to your candidate. This is a more ‘natural’ kind of preference misrepresentation, and therefore, probably more likely to actually occur in practice.

Another way to think about strategic manipulability is in terms of the size of the smallest coalition that would be needed to change the outcome in a particular election to one more to the members of that coalition’s liking (Chamberlin et al., 1984: 494–495). Using data on five real-world (American Psychological Association) elections, reconstructing full preferences from the actual data, these authors show that in only one of ten situations would AV have been manipulable. For Coombs, the mean coalition size needed for manipulability was 421.6 (st. dev. 476.4); while for plurality the mean coalition size needed was 424.1 (st. dev. 256.4).

3. Discussion

The alternative vote (in the United States, recently relabeled as the ‘instant runoff’) has been proposed as a replacement for simple first past the post (plurality) voting in situations where there is a single alternative to be chosen or where there are single seat elections to a parliament. One of its advantages is a greater likelihood than plurality of selecting the Condorcet winner (i.e. that candidate, if any, who can receive a majority in paired competition against each and every other candidate), both in situations where the Condorcet winner is candidate or party who is centrist on some ideological dimension and in situations where centrism is defined in terms of conciliatory views about potentially polarizing racial/ethnic/religious divisions. Also, the alternative vote has been argued for on the grounds that, unlike plurality, it permits voters to express their ‘true’ preferences for candidates who, in terms of first place preferences, have little chance of being the plurality winner.

31 See earlier footnote for definition.
32 This same point has been emphasized to us by Arend Lijphart (personal communication, 24 May 2002).
33 In the one simulated election in which strategic manipulation under AV was possible, to change the outcome would have required a coalition of at least 35 voters voting together.
34 For important formal work on manipulability of voting methods when preferences are single-peaked see Moulin (1980).
without, at the same time, creating an increased likelihood that the voter’s least preferred candidate will gain victory. In other words, it is argued that AV has lower incentives for strategic voting than does plurality. Moreover, it has been argued that, if in multi-candidate/multi-party competition AV has a much lower chance than plurality of having ‘sincere’ voting yielding results that are undesirable, AV consequently reduces the need to arbitrarily reduce the number of alternatives being voted on by raising barriers to entry of the sort that reinforce exclusive two-party competition and/or perpetuate the dominance of existing parties.

Because of such comparisons between plurality and AV, with respect to each of which AV is hypothesized to dominate plurality, as we have emphasized earlier, AV is currently being advocated by various electoral reformers in England and the US as a replacement for plurality—both to make less likely the election of candidates with limited overall support and to mitigate ethnic conflict in divided societies. These reformers have begun to score some successes; e.g. AV and has already replaced the use of plurality-based elections in the national parliaments of two countries (Fiji and Papua New Guinea), and it has been adopted for use in a major US city (San Francisco).

Reviewing our four criteria, we see that Coombs and AV have similar attractiveness with respect to our four criteria. AV and Coombs are identical in simplicity, and identical in guaranteeing avoidance of Condorcet losers, and if AV seems superior to Coombs in its diminished likelihood of eliciting strategic voting behavior, under the criterion of expected Condorcet efficiency, at least if we believe unidimensionality to be plausible and (near) single-peaked preferences to be far more likely than (near) single-troughed ones, then we find Coombs to be preferred to AV.35

35 In demonstrating the superior Condorcet efficiency of the Coombs rule to both AV and plurality, we have assumed (net) single-peaked preferences over a single dimension. Some readers might object to this assumption (as did one of the anonymous reviewers of this paper). In particular, if we replace single-peakedness with its inverse, single-troughedness, then the superiority of Coombs over AV reverses itself, i.e., if voters have single-troughed preferences over a single dimension, then AV always selects the Condorcet winner (i.e. the alternative supported by the median voter), but Coombs need not. (The proof is essentially just the inverse of the proof for Proposition 2, given in the Appendix.) Single-troughed preferences occur when there exists a continuum along which alternatives can be ordered such that the utilities of each voter can be plotted as a single-troughed graph, i.e., a curve which changes its slope at most once, from down to up (Black, 1958; see also Sen, 1970 for an alternative definition of single-troughedness in terms of the NB (not best) condition on a triple of alternatives). Such preferences arise in situations where voters are unwilling to adopt the centrist alternative. Arguably this was the case for some US ‘hawks’ during the latter phases of the Vietnam War: while their first choice was to pursue the defeat of North Vietnam with no holds barred, they preferred unilateral withdrawal to the continuation of what they saw as an indecisive and ineffective policy of muddling through. However, it is unlikely that Vietnam era ‘doves’ had single-troughed preferences with ‘bomb them back to the Stone Age’ as their second choice. Thus, empirically, we believe single-troughedness to be a far rarer case than single-peakedness. Similarly, in the context of ethnic-based voting, although we can imagine arguments that letting the other side’s extremists set policies would aid in motivating the members of one’s own side, we think it unlikely that extremists of a given ethnicity would actually prefer rule by extremists of the other ethnicity to rule by moderates of that same ethnicity. Thus, since, empirically, we view single-peakedness as a much more plausible assumption than single-troughedness, we would still take our comparisons of Propositions 1 and 2 to imply the superiority of Coombs over AV with respect to the likelihood of picking a Condorcet winner.
Furthermore, Coombs has exactly the same claim as AV to the title (and putative advantages) of being an ‘instant’ runoff.\textsuperscript{36}

Note, however, that we have been very careful to make no claim that, because Coombs and AV both seem superior to plurality, we ‘ought’ to use either AV or the Coombs rule. Indeed, if we use Condorcet efficiency as our criteria then clearly, a Condorcet extension method might seem the most desirable of all, albeit such methods tend to be hard to explain in simple terms; or if, like Donald Saari (1995), we do not find the Condorcet criterion normatively compelling,\textsuperscript{37} then we might prefer a method such as the Borda count for which other strong arguments can be made; or we might follow Steven Brams and his co-authors in opting for approval voting, which combines substantial simplicity with a strong likelihood of choosing a Condorcet winner when one exists. Finally, if we value simplicity above all, then a rule which requires voters to do no more than place an x in front of the first choice would be the best rule.

We do not seek to arbitrate among these competing considerations here.\textsuperscript{38} Our aim in this paper, as we have emphasized from the start, is a quite limited one—namely to make the argument that electoral reformers who have been advocating the AV/instant runoff need to take a serious look at its close relative, the Coombs rule.

Acknowledgements

This research was partially supported by National Science Foundation grant #SBR 97-30578 (to Grofman and Anthony Marley), Program in Methodology, Measurement and Statistics. We are indebted to Clover Behrend-Gethard for library assistance.

\textsuperscript{36} However, Robert Richie of the Center for Voting and Democracy (personal communication, October 2002) has made the important point that, sometimes the Condorcet winner may be normatively undesirable in circumstances where say, we have single-peaked preferences, and the median alternative is one who is perceived of as ‘wissy-waisy’ and is someone with little first round support. All other things being equal, wissy-waisy candidates with little first place support are probably also rarely ranked last. Thus, when such candidates exist, it is possible they might be even more likely to be chosen under Coombs than under AV, but that seems to be very much a question for empirical investigation.

\textsuperscript{37} As noted above, skepticism about the abstract virtues of Condorcet winners has also been expressed by some electoral reformers, such as Robert Richie, who question the desirability of choosing Condorcet winners who lack substantial ‘real’ political support and whose chief claim to support may simply be that no voter sees them as the least desirable candidate in the race.

\textsuperscript{38} We should also note that, when (net) single-peakedness does not hold, e.g. if there are multiple dimensions of choice, then it becomes much more complicated to specify the relative merits of AV, Coombs and plurality, since, in two dimensions, unlike in one dimension, the Coombs rule, and not just AV, can eliminate the Condorcet winner—even when that alternative is also the plurality winner. Also, in this paper we have only looked at choice of a single alternative; when we are considering the elections to a legislature or committee other considerations (e.g. the proportionality of the seats-votes relationship) come to the fore. For example, Arend Lijphart, has done empirical work suggesting that ‘we should not expect AV to be any more proportional than plurality’ (personal communication, 24 May 2002: see Lijphart, 1997).
Appendix. Theorem proofs

Proposition 1. Under the assumptions of single-peaked preferences and no party holding a majority of first place preferences, when we have four parties or fewer, a candidate of the median party is more likely (or at least no less likely) to win when voting is conducted under the alternative vote than when voting is conducted under first-past-the-post.

Proof. The proof of this proposition for two or three alternatives is trivial, so we focus on the case of four alternatives.

First, we may show that, under the simplifying assumptions of four parties and single-peaked preferences, the median party (the majority winner in pairwise contest) must be a moderate party. Our proof will be by contradiction. Imagine the leftmost party is the median party. If so the leftmost party must defeat the party to its immediate right. But all the voters to the right of the midpoint between the leftmost and the next leftmost party prefer the next leftmost party to the leftmost party. But, by hypothesis the leftmost party (since it is supported by the median voter) receives a majority in paired contest against each and every other alternative, thus it must have a majority of first place votes, since otherwise it would lose to the next leftmost alternative. However, we have posited that no party has a majority of first place preferences. This contradiction demonstrates that, for single-peaked preferences among four alternatives, when there is no party with a majority of first-place preferences, no extreme party can be the median party.

With this result in hand, it is now trivial to see that, under the above assumptions, the median party will be the non-extremist party on the side of the space that includes a majority of the voter ideal points closest to it.

Now there are four cases to consider.

Case I: If the median party has the highest number of first place preferences it will win for sure under plurality, but it will also win for sure under the alternative vote.

To see why that must be so, consider what happens as ballots are transferred. The party with the lowest number of first place votes is eliminated. If that party is on the same side of the line as the median party, its votes will be transferred to the median party, assuring that the median party wins. (If these votes were not sufficient to give a majority of first place votes to the moderate party we have identified as the first preference of the median voter, this would contradict our assumption that this party was the median party). If the party that is eliminated is the extreme party on the opposite side of the space, its votes will transfer the moderate party on its side. Now we have three parties in the contest. On the next elimination round, the party that will be eliminated is either the extreme party on the same side as the median party or the moderate party on the opposite side. (As per our assumptions, at this stage of the balloting the median party has more many first preferences at this stage than does the radical party of its side of the space and thus cannot be the party that will be eliminated.) But, in either case, the median party will be the only party to gain votes via transfers. If the party that is
eliminated is the moderate party on the opposite side of the space, we will have a pairwise contest between the median party and the extreme party on the same side of the space, which the median party will win. If the party that is eliminated is the extreme party on the same side of the space as the median party, the median party will have garnered a majority of the votes and will immediately be declared the victor.

Case II: If the median party has the second highest number of first place preferences it will lose for sure under in plurality, but it will have a chance to win under the alternative vote. Indeed, it will win for sure unless the transfers of the party with the lowest number of first place preferences to the party with the third highest number of first place preferences makes that party the party with the first or second highest number of first-place votes.

Under this latter scenario, we would have a three-way contest in which the median party was the party with the third highest number of first place preferences, and thus the party to be eliminated. But, unless the transfers of the party with the lowest number of first place preferences to the party with the third highest number of first place preferences put that party into the position of having the first or second highest number of first-place votes, then the median party remains the party with the second highest number of votes. But, then it is not eliminated on the second elimination round. But, if it enters the third and final round of balloting it must win, since by definition, it defeats any alternative in paired contest.

Case III: If the median party has the third highest number of first place preferences, it will lose for sure under in plurality, but it still will have a chance to win under the alternative vote.

Indeed, if the party with the lowest number of first place votes is the extreme party on the same side as the median party, then, when its votes are transferred, the median party will have a majority of first place preferences. On the other hand, if the party with the lowest number of first place votes is the extreme party on the opposite side of the space from the median party, then the median party will lose for sure, since now it remains in third place vis-a-vis the number of first-place preferences, since all the vote transfers go to the non-extreme party on the opposite side of the space. However, if the party with the lowest number of first place votes is the moderate party on the opposite side of the space from the median party, it might still be possible for the median party to win if it gets enough vote transfers from this party to move into second place vis-a-vis the number of first-place preferences. If that happens, the median party makes it into the final round and thus is assured of victory.

Case IV: If the median party has the fourth highest number of first place preferences, it will lose for sure under in plurality and it will lose for sure under the alternative vote as well.\(^{39}\) q.e.d.

\(^{39}\) The accuracy of this claim is trivially obvious.
Proposition 2. If voters have single-peaked preferences, and vote sincerely, then the Coombs rule always selects the Condorcet winner (i.e., the alternative supported by the median voter).

Proof. We only need to prove that the Condorcet winner cannot be eliminated in the successive eliminations of the Coombs process. This is done by considering the only two possibilities: either there are some alternatives to both sides of the alternative favored by the median voter (i.e., to both sides of the Condorcet winner), or all other alternatives are one side of the alternative preferred by the median voter.

Case I: If there are alternatives on both sides of the alternative supported by the median voter, then that alternative is no voter’s last choice. The only last choices are the extremes at the two ends of the dimension. So, the Condorcet winner cannot be the alternative with the most last place votes, and thus will not be eliminated.

Case II: If all other alternatives are to one side of the alternative preferred by the median voter (i.e. the median voter is an extreme alternative) then all of the voters on the side with no alternatives have the extreme alternative on the other side as their last choice. But that must be a majority of voters, and thus it is that extreme alternative that will be eliminated, rather than the alternative preferred by the median voter.

But, then the Condorcet winner is never eliminated at any round of the Coombs procedure and therefore will either be declared a winner at some early round or make it to the final round in which there are only two alternatives left. But if there are only two alternatives left, one of whom is the alternative preferred by the median voter, since that alternative is the Condorcet winner (and thus defeats each and every other alternative in paired competition), it will necessarily be chosen. q.e.d.

References


