Multi winner Approval Voting: An Apportionment Approach

Brams, Steven J. and Kilgour, D. Marc and Potthoff, Richard F.

New York University, Wilfrid Laurier University, Duke University

26 March 2017

Multiwinner Approval Voting: An Apportionment Approach

Steven J. Brams
Department of Politics
New York University
New York, NY 10012
UNITED STATES
steven.brams@nyu.edu

D. Marc Kilgour
Department of Mathematics
Wilfrid Laurier University
Waterloo, Ontario N2L 3C5
CANADA
mkilgour@wlu.ca

Richard F. Potthoff
Department of Political Science and Social Science Research Institute
Duke University
Durham, NC 27708
UNITED STATES
potthoff@duke.edu

Abstract

We extend approval voting so as to elect multiple candidates, who may be either individuals or members of a political party, in rough proportion to their approval in the electorate. We analyze two divisor methods of apportionment, first proposed by Jefferson and Webster, that iteratively depreciate the approval votes of voters who have one or more of their approved candidates already elected. We compare the usual sequential version of these methods with a nonsequential version, which is computationally complex but feasible for many elections. Whereas Webster apportionments tend to be more representative of the electorate than those of Jefferson, the latter, whose equally spaced vote thresholds for winning seats duplicate those of cumulative voting in 2-party elections, is even-handed or balanced.

1. Introduction

The properties of approval voting (AV)—whereby voters can vote for as many candidates as they like, and the candidate with the most votes wins—have been extensively studied in the case of single-winner elections (Brams and Fishburn, 1983; Brams, 2008; Laslier and Sanver, 2010). Although AV has been and continues to be used in multiwinner elections (e.g., to elect members of the Council of the Game Theory Society), this usage has been questioned, because it enables a 51% majority, if it votes as a bloc for all of its preferred candidates, to win all the seats on a committee or council, leaving the 49% minority unrepresented.

To address this problem, alternative ways of aggregating approval votes have been proposed to better represent factions in the electorate roughly proportionally to their approval. For example, Brams, Kilgour, and Sanver (2007) analyzed a "minimax procedure," which chooses the committee that minimizes the maximum Hamming distance to the ballots, weighted by their proximity to other voters' ballots. They applied this procedure to the 2003 election of the Game Theory Society Council and analyzed differences between the 12 candidates elected under AV and the 12 that would have been elected under the minimax procedure. For a generalization of this approach, see Sivarajan (2016).

Several other ways of aggregating approval ballots have been analyzed, including "satisfaction approval voting" (Brams and Kilgour, 2014), which maximizes the sum of voters' satisfaction scores, defined as the fraction of the approved candidates who are elected. Another approach, first proposed by Monroe (1995), and generalized by Potthoff and Brams (1998) using integer programming, minimizes voters' dissatisfaction, which can be based on approval votes. Brams (1990, 2008, ch. 4) analyzed "constrained approval voting," in which winners are determined by both their vote shares and the categories of voters who approve of them, with constraints put on the numbers that can be elected from different categories.

Kilgour, Brams, and Sanver (2006), Kilgour (2010), Kilgour and Marshall (2012), and Elkind, Faliszewski, Skowron, and Slinko (2017) review a variety of methods for electing committees using approval ballots. A number of analysts (Sánches-Fernández, Fernández Garcia, and Fisteus, 2016; Subiza and Peris, 2014; Brill, Laslier, and Skowron, 2016) have suggested using divisor methods of apportionment (more on these methods later) as a basis for selecting multiple winners. We adopt this approach here, addressing, among other topics, the representativeness of the candidates that different versions of these methods elect, which previous studies have not analyzed.

We focus on the use of weights from two most prominent divisor methods of apportionment (Balinski and Young, 1982/2001; Pukelsheim, 2014), one of which was independently proposed by Thomas Jefferson and Viktor d'Hondt, the other by Daniel Webster and André Saint-Laguë. We identify them by the surnames of their American discoverers (Jefferson and Webster), because their discoveries preceded those of their European discoverers.

The standard Jefferson and Webster methods are based on iterative procedures, whereby winners are determined sequentially until a body of requisite size is obtained. With approval ballots, if the candidates are individuals, then once a candidate is elected, the method is applied to the remaining unelected candidates in order to fill the desired number of seats. If the candidates are not individuals but instead political parties, then these methods determine the numbers of seats in a legislature or other body that each party receives (if any).[1]

Each of these methods has a nonsequential analogue, which is not used in practice and may produce different, even disjoint, winners from those of the sequential version. We ask whether the nonsequential winners are *more representative* than sequential winners: Do more voters approve of one or more of the nonsequential winners than the sequential winners?[2] We also ask whether the Webster winners (both sequential and nonsequential) are more representative than the Jefferson winners.

We next turn from the election of individual candidates to the election of different numbers of candidates from political parties. We show that Webster tends to choose a member of a relatively small party before choosing an additional member of a larger political party, thereby giving more voters at least one representative, whereas Jefferson has the opposite tendency.

If there are only two candidates or political parties, we assume that voters prefer, and vote for, only one. Thus, the opportunity afforded by an approval ballot—of approving of more than one candidate or party—is irrelevant in the 2-candidate or 2-party case.

For approval balloting with more than two parties or candidates, it turns out that vote thresholds for electing one or more candidates under Jefferson coincide with those

---

[1] In lieu of determining apportionments sequentially, one can search for a divisor which, when divided into the vote shares of political parties, yields the number of seats—after some kind of rounding—that each party will receive in the legislature. The Jefferson method rounds down the exact entitlements (later called "quotas") of the parties, which are not typically integers, whereas the Webster method rounds in the usual manner (rounding up the exact entitlement if its remainder is equal to or greater than 0.5, rounding down otherwise). For details, see Balinski and Young (1982/2001).

[2] A downside to the nonsequential versions of the apportionment methods is that they are computationally complex, not implementable in polynomial time (Brill, Laslier, and Skowron, 2016).

for cumulative voting, but Jefferson does not require that parties nominate only as many candidates as are commensurate with their approval totals in order to ensure proportional representation. The thresholds for Jefferson are equally spaced, which renders them even-handed or balanced compared with the unequally spaced Webster thresholds.

In elections in which voters can vote for only one candidate or party, the Jefferson and Webster apportionment methods satisfy several desirable properties (Balinski and Young, 1982, 2001), but they are not flawless. Like all divisor apportionment methods, they are vulnerable to manipulation; furthermore, they may not always give political parties the number of representatives to which they are entitled after rounding (either up or down).

Nevertheless, these methods seem the best possible way not only to elect single winners but also to elect multiple winners, using approval ballots. By expressing their support for *sets* of candidates that cross ideological or party lines, voters may ultimately diminish the gridlock one sees in voting bodies, especially in the United States.

2. The Jefferson and Webster Methods Applied to Candidates

The development and use of apportionment methods has a rich history. It is recounted in the American case by Balinski and Young (1982/2001), wherein the best-known application has been to the apportionment of members of the U.S. House of Representatives to states according on their populations. In the European case, these methods have been applied to the apportionment of seats to political parties in a parliament according to the numbers of votes they receive in an election (Pukelsheim, 2014). In both of these applications, the methods determine how many seats each state, or each party, receives, respectively, in the House of Representatives or in the parliament.

Other jurisdictions in the United States,[3] and parliaments around the world, especially in Europe, use divisor methods of apportionment, of which there are exactly five that lead to *stable apportionments*: No transfer of a seat from one state or party to another can produce less disparity in apportionment, where "disparity" is measured in five different ways (there are other ways of measuring disparity, but they do not produce stable apportionments using a divisor method). Two of the five ways of defining disparity are given by the Jefferson and Webster methods, which we describe next.[4]

Though originally devised for allocating seats to parties, based on votes, or to states, based on population, apportionment methods can also be used to allocate seats to select individual candidates based on a set of approval ballots. In this role, they iteratively depreciate the value of a voter's approvals as more and more of his or her approved candidates are elected. More specifically, the *sequential versions* of these methods, which are the standard ones, give one seat, on each round, to the candidate, i, who maximizes a deservingness function, $d(i)$.

Let β denote the set of ballots, and let $B(i) \subseteq \beta$ denote the set of ballots that include an approval vote for candidate i. On any round, let $a(b)$ denote the number of

---

[3] This includes ten states that use multimember districts in the apportionment of their legislatures (https://ballotpedia.org/State_legislative_chambers_that_use_multi-member_districts).

[4] The Webster method is used in four Scandinavian countries, whereas the Jefferson method is used in eight other countries. None of the other three divisor methods is currently used, except for Hill for the U.S. House of Representatives. The nondivisor Hamilton method, also called "largest remainders," is used in nine countries (Blais and Massicotte, 2002; Cox, 1997). For a review of apportionment methods, see Edelman (2006a), who proposed a nondivisor method that is mathematized in Edelman (2006b). In section 4, we will say more about why we favor the Jefferson and Webster methods for allocating seats to political parties in a parliament.

candidates mentioned on ballot b who are already elected. The deservingness functions of the sequential versions of Jefferson (J) and Webster (W) are[5]

$$d_J(i) = \sum_{b \in B(i)} \frac{1}{a(b)+1}$$

$$d_W(i) = \sum_{b \in B(i)} \frac{1}{a(b)+\frac{1}{2}}.$$

Simply put, on any round, each approval ballot supporting unelected candidate i is depreciated by an amount that reflects the number of already elected candidates mentioned on the ballot.

On the first round, no candidate has yet received a seat, so $a(b) = 0$; for every ballot, the Jefferson fraction equals 1 and the Webster fraction equals 2. Thus, the first candidate elected, according to both methods, will be the candidate who obtains the maximum number of approvals, or the AV winner. The following example shows that the two methods may produce different winners, beginning in the second round.

Example 1. 2 of 4 candidates $\{A, B, C, D\}$ to be elected. The numbers of voters who approve of different subsets of candidates are

2: *A* 5: *AB* 3: *AC* 2: *BC* 4: *D*.

A, *B*, *C*, and *D* receive, respectively, 10, 7, 5, and 4 approvals, so *A* is the candidate elected first. For Jefferson on the second round, *B*'s ballots (the 5 supporting *AB*, and the

[5] The deservingness functions of the three other divisor methods are defined similarly. The denominators of the summands are $[a(b)(a(b)+1)]^{1/2}$ for Hill or "equal proportions," $2a(b)[a(b)+1]/[2a(b)+1]$ for Dean or "harmonic mean," and simply $a(b)$ for Adams or "smallest divisors" (Balinski and Young, 1982/2001; Pukelsheim, 2014).

2 supporting *BC*) are counted as follows: $a(b) = 1/2$ for each of the 5 *AB* ballots, and $a(b) = 1$ for each of the 2 *BC* ballots. Thus, on the second round, *B*'s deservingness score is

$$B: 5 \times (1/2) + 2 \times (1) = 4\ 1/2.$$

Similarly, on the second round the deservingness scores of *C* and *D* are

$$C: 3 \times (1/2) + 2 \times (1) = 3\ 1/2; \quad D: 4 \times (1) = 4,$$

so the second-round winner is *B*.

For Webster on the second round, *B*'s ballots (the 5 supporting *AB*, and the 2 supporting *BC*) are counted as follows: $a(b) = 2/3$ for each of the 5 *AB* ballots, and $a(b) = 2$ for each of the 2 *BC* ballots. Thus, on the second round, *B*'s deservingness score is

$$B: 5 \times (2/3) + 2 \times (2) = 7\ 1/3.$$

Similarly, on the second round the deservingness scores of *C* and *D* are

$$C: 3 \times (2/3) + 2 \times (2) = 6; \quad D: 4 \times (2) = 8,$$

so the second-round winner is *D*. To summarize Example 1, Jefferson elects *AB* and Webster elects *AD*.

Notice the difference in the sequences of fractions that are added to determine deservingness scores under Jefferson and Webster. As $a(b)$ increases, for Jefferson $1/[a(b) + 1]$ decreases according to the sequence

$$1, 1/2, 1/3, 1/4, 1/5, \ldots,$$

whereas for Webster $1/[a(b) + 1/2]$ decreases according to the sequence

$$2, 2/3, 2/5, 2/7, 2/9, \ldots,$$

or, equivalently,

$$1, 1/3, 1/5, 1/7, 1/9, \ldots.$$

Later we generalize these sequences to the *h*-sequence, defined by

$$1, \frac{h}{h+1}, \frac{h}{h+2}, \frac{h}{h+3}, \ldots$$

where $h \geq 0$. Note that setting $h = 1$ produces the Jefferson sequence and $h = ½$ produces the Webster sequence.

For both methods, the contributions of voters to deservingness scores are depreciated more and more as candidates whom they approve are elected. But as can be seen by comparing the corresponding fractions in the Jefferson sequence and the Webster sequence (which is normalized to start at 1), candidates who approve of the AV winner—and subsequent candidates who may be elected on later rounds—are less depreciated under the Jefferson method than under the Webster method. This means that the Jefferson method more than the Webster method tends to favor candidates (e.g., *B*) whose voters have approved of a candidate already elected (e.g., *A*) than candidates (e.g., *D*) whose voters have not yet had an approved candidate elected.

The same sequences can be used as the basis for a nonsequential method of committee election. In a nonsequential method, each possible committee is assigned a score measuring the total satisfaction that it would deliver to voters; any committee with a maximum score wins. Assuming there are n candidates and a committee of size $m < n$

is to be elected, there are $\binom{n}{m}$ possible committees to be compared. Denote the set of all possible committees by Ω.

To construct a nonsequential rule from any *h*-sequence, we measure the satisfaction of having one candidate elected as 1, the satisfaction of having two candidates elected as $1 + \frac{h}{h+1}$, the satisfaction from three candidates as $1 + \frac{h}{h+1} + \frac{h}{h+2}$, etc. The Jefferson (*J*) and Webster (*W*) nonsequential scores for a committee $C \in \Omega$ are obtained by setting $h = 1$ and $h = ½$, respectively:

$$s_J(C) = v_1(C) + \left(1 + \frac{1}{2}\right)v_2(C) + \left(1 + \frac{1}{2} + \frac{1}{3}\right)v_3(C) + \left(1 + \frac{1}{2} + \frac{1}{3} + \frac{1}{4}\right)v_4(C) + ...$$

$$s_W(C) = v_1(C) + \left(1 + \frac{1}{3}\right)v_2(C) + \left(1 + \frac{1}{3} + \frac{1}{5}\right)v_3(C) + \left(1 + \frac{1}{3} + \frac{1}{5} + \frac{1}{7}\right)v_4(C) + ...,$$

where $v_k(C)$ is the number of voters who approve of exactly *k* members of *C*.[6] Formally,

$$v_k(C) = |\{j \in V : |B_j \cap C| = k\}|.$$

[6] Notice in the parenthetic expressions that 1, and the fractions that follow are decreasing (e.g., from 1 to 1/2 under Jefferson, from 1 to 1/3 under Webster, for voters who have more than one candidate approved). Thus, as with the sequential versions of each method, getting a second approved candidate elected does not come close to doubling a voter's satisfaction over electing the first. It is important to point out, however, that Jefferson and Webster never proposed the weighting sequences, based on the deservingness functions, given by the decreasing fractions. (In addition, because a voter can reside in only one state, he or she can be counted only for that state, whereas with approval voting, a voter can vote for more than one candidate or party.) Instead of weighting sequences, they proposed divisors that determine, after rounding, the numbers of seats that states are entitled to in the U.S. House of Representatives. For more on the history of using weighted sequences in apportionment, which can be traced back to the late 19th century, see Brill, Laslier, and Skowron (2016). What we call the sequential and nonsequential versions of the Jefferson method, in particular, are referred to in the literature as sequential proportional approval voting (SPAV) and proportional approval voting (PAV). Aziz, Brill, Conitzer, Elkind, Freeman, and Walsh (2017) show that PAV but not SPAV satisfies "justified representation" and "extended justified representation," but SPAV may be more "representative" than PAV, as we show in section 4.

where V is the set of all voters, and B_j is the ballot (set of approved candidates) of voter j. Of course, the C's that maximize $s_J(C)$ and $s_W(C)$ are the ones chosen by each method.

We illustrate the Jefferson method in Example 1:

2: A 5: AB 3: AC 2: BC 4: D.

For the subset AB, $v_2(AB) = 5$ voters approve of both A and B and $v_1(AB) = 7$ voters approve of exactly one of A and B, with 5 approving of A but not B and 2 approving of B but not A. Thus,

$$s_J(AB) = 5 \times (1 + 1/2) + 7 \times (1) = 14\ 1/2; \quad s_W(AB) = 5 \times (1 + 1/3) + 7 \times (1) = 13\ 2/3.$$

For each of the six possible committees, we can calculate the v_1 and v_2 counts and, from them, the committee's scores according to Jefferson (J) [in the format $v_2 \times (3/2) + v_1 \times (1)$] and Webster ($W$) [in the format $v_2 \times (4/3) + v_1 \times (1)$], as shown below:

J: $s_J(AB) = 5 \times (3/2) + 7 \times (1) = \underline{14\ 1/2}$; $s_J(AC) = 3 \times (3/2) + 9 \times (1) = 13\ 1/2$;

$s_J(AD) = 0 \times (3/2) + 14 \times (1) = 14$; $s_J(BC) = 2 \times (3/2) + 8 \times (1) = 11$;

$s_J(BD) = 0 \times (3/2) + 11 \times (1) = 11$; $s_J(CD) = 0 \times (3/2) + 9 \times (1) = 9$.

W: $s_W(AB) = 5 \times (4/3) + 7 \times (1) = 13\ 2/3$; $s_W(AC) = 3 \times (4/3) + 9 \times (1) = 13$;

$s_W(AD) = 0 \times (4/3) + 14 \times (1) = \underline{14}$; $s_W(BC) = 2 \times (4/3) + 8 \times (1) = 10\ 2/3$;

$s_W(BD) = 0 \times (4/3) + 11 \times (1) = 11$; $s_W(CD) = 0 \times (4/3) + 9 \times (1) = 9$.

As the underscored maxima indicate, the nonsequential versions of Jefferson and Webster choose the same committees (AB for Jefferson, AD for Webster) as the sequential versions. But this is not always the case.

Proposition 1. *The sequential and nonsequential versions of Jefferson or of Webster may elect different committees, which may have no overlap.*

Proof. We start with a simple example in which there is partial overlap.

Example 2. 2 of 3 candidates $\{A, B, C\}$ to be elected. The numbers of voters who approve of different subsets of candidates are

$$7: AB \qquad 6: AC \qquad 4: B \qquad 4: C.$$

A, B, and *C* receive, respectively, 13, 11, and 10 approvals, so *A* is the candidate elected first under sequential Jefferson. On the second round, the deservingness scores for *B* and *C* are

$$B: 7 \times (1/2) + 4 \times (1) = \underline{7\ 1/2}; \quad C: 6 \times (1/2) + 4 \times (1) = 7,$$

so *AB* wins. Under nonsequential Jefferson, the satisfaction scores of the three pairs of candidates are

$$AB: 7 \times (3/2) + 10 \times (1) = 20\ 1/2; \quad AC: 6 \times (3/2) + 11 \times (1) = 20;$$
$$BC: 0 \times (3/2) + 21 \times (1) = \underline{21},$$

so *BC* wins (and *A*, the AV winner, is excluded). Under Webster, the sequential and nonsequential winners are the same as under Jefferson.

The sequential and nonsequential winners are disjoint in the following example:

Example 3.[7] 2 of 4 candidates $\{A, B, C, D\}$ to be elected. The numbers of voters who approve of different subsets of candidates are

---

[7] Both this example and Example 4 in section 3 were found using an integer program.

$$1: AB \qquad 4: AC \qquad 4: AD \qquad 1: B \qquad 1: BC \qquad 4: BD \qquad 3: C.$$

A, B, *C*, and *D* receive, respectively, 9, 7, 8, and 8 approvals. Thus, *A* is the candidate elected first under sequential Jefferson. On the second round, the candidates' deservingness scores are

$$B: 1 \times (1/2) + 6 \times (1) = \underline{6\ 1/2;} \qquad C: \ 4 \times (1/2) + 4 \times (1) = 6;$$
$$D: 4 \times (1/2) + 4 \times (1) = 6,$$

so *AB* is the winning pair. Under nonsequential Jefferson, the satisfaction scores of the six pairs of candidates are

$$AB: 1 \times (3/2) + 14 \times (1) = 15\ 1/2; \qquad AC: 4 \times (3/2) + 9 \times (1) = 15;$$

$$AD: 4 \times (3/2) + 9 \times (1) = 15; \qquad BC: 1 \times (3/2) + 13 \times (1) = 14\ 1/2;$$

$$BD: 4 \times (3/2) + 7 \times (1) = 13; \qquad CD: 0 \times (3/2) + 16 \times (1) \ = \underline{16},$$

so *CD* is the winning pair, which does not overlap *AB*. Under Webster, the sequential and nonsequential winners are the same as under Jefferson. ■

The fact that the sequential apportionment methods start by choosing the AV winner is the reason why they fail, in both Examples 2 and 3, to find the committee that maximizes the satisfaction of all voters, which excludes the AV winner. If they had started with one member of the maximizing committee, they would have found the other.

As noted in section 1, the nonsequential versions of Jefferson and Webster are computationally complex. However, if the committee (or council) size is small, and the number of candidates is not much larger, then the calculation of satisfaction scores for all committees is certainly feasible with modern computers.

3. Representativeness of a Voting Body

While it is desirable that as many voters as possible be represented by at least one candidate they approve of on the committee, it is also desirable that voters who approve of the same or similar subsets of candidates get them elected in numbers roughly proportional to the numbers of voters who approve of them. Different election procedures, including sequential and nonsequential Jefferson and Webster, may clash on these criteria.

Examples 2 and 3 illustrate this clash in the case of both the Jefferson and Webster apportionment methods. In Example 2, both versions of these methods elect *B*, but the sequential version first elects *A* (the AV winner), and only then *B*, making *AB* the winning pair, whereas the nonsequential version elects *BC*. *BC* gives all 21 voters one approved member of the committee, whereas *AC* gives 6 voters two approved members and 11 voters one approved member (in total, 17 voters have at least one approved member), but 4 voters approve of no member.

The clash between these criteria is also evident in Example 3, wherein the sequential versions of Jefferson and Webster choose *AB* and the simultaneous versions choose *CD*. *CD* provides 16 of the 18 voters with at least one approved committee member, whereas *AB* provides only 15 voters with at least one approved member.

Recall that the nonsequential version of each procedure compares all possible committees on the basis of the satisfaction that they give to voters, with satisfaction decreasing as additional approved candidates are elected. The sequential version starts by electing the AV winner (*A* in both Examples 2 and 3), who is not even a member of the nonsequential version's winning pair in either case.

We define the *representativeness* of a committee to be the number of voters who approve of at least one member of that committee. This makes *BC* more representative than *AB* or *AC* in Example 2, and *CD* more representative than *AB* in Example 3. In these examples, as we showed, nonsequential Jefferson and Webster produce more representative committees than their sequential counterparts. But this is not always the case.

Representativeness is not a new concept, although the idea that electoral methods might have different tendencies toward representativeness is. Representativeness is the REP-1 scoring procedure proposed by Kilgour and Marshall (2012), which is also known as the Chamberlin-Courant (1983) procedure. In Generalized Approval Voting, the score of a subset is the sum over all voters of a measure of the worth of the subset to the voter that depends only on the number of candidates in the subset that the voter supports (Kilgour and Marshall, 2012). In fact, representativeness, nonsequential Jefferson scores, and nonsequential Webster scores are all generalized approval scores.

The Generalized Approval score is

$$S(C) = r_1\, v_1(C) + r_2\, v_2(C) + r_3\, v_3(C) + \ldots$$

for any committee, $C \in \Omega$, where $r_1, r_2, r_3, \ldots$ is a *rep sequence* that characterizes the procedure. Thus, the score of subset C, $S(C)$, is a sum of contributions from the voters: 0 for voters who did not support any candidate in C; r_1 for each voter who supported one candidate in C; r_2 for each voter who supported two candidates in C; etc. In particular, an

h-sequence corresponds to a rep sequence defined by $r_1 = 1$ and, for $J = 2, 3, 4, \ldots$,

$r_J = 1 + \sum_{j=2}^{J} \frac{h}{h+j-1}$. Therefore, nonsequential Jefferson and Webster are Generalized Approval procedures, based on, respectively, the rep sequences

Jefferson: $r_1 = 1, r_2 = 1 + 1/2, r_3 = 1 + 1/2 + 1/3, \ldots$

Webster: $r_1 = 1, r_2 = 1 + 1/3, r_3 = 1 + 1/3 + 1/5, \ldots$.

Because $R(\mathrm{C}) = v_1(C) + v_2(C) + v_3(C) + \ldots$ gives the number of voters who approve of at least one candidate in C, representativeness is measured by the score under the rep sequence corresponding to $h = 0$,

Representativeness: $r_1 = 1, r_2 = 1, r_3 = 1, \ldots$,

Proposition 2. *If the sequential or nonsequential versions of Jefferson or Webster elect different committees, either the sequential or nonsequential committee may be more representative*.

Proof. Examples 2 and 3 demonstrated that the nonsequential committees elected by Jefferson and Webster are more representative than the sequential committees. Example 4 shows the opposite.

Example 4. 2 of 4 candidates $\{A, B, C, D\}$ to be elected. The numbers of voters who approve of different subsets of candidates are

21: AC 21: AD 21: B 9: BD 9: C 10: CD.

A, B, *C*, and *D* receive, respectively, 42, 30, 40, and 40 approvals. Thus, *A* is the candidate elected first under sequential Jefferson. On the second round, the deservingness scores are

$$B\text{: } 30 \times (1) = \underline{30;} \quad C\text{: } 19 \times (1) + 21 \times (1/2) = 29\ 1/2;$$

$$D\text{: } 19 \times (1) + 21 \times (1/2) = 29\ 1/2,$$

so *AB* is the winning pair. Under nonsequential Jefferson, the satisfaction scores of the six pairs of candidates are

$$AB\text{: } 0 \times (3/2) + 72 \times (1) = 72; \quad AC\text{: } 21 \times (3/2) + 40 \times (1) = 71\ 1/2;$$

$$AD\text{: } 21 \times (3/2) + 40 \times (1) = 71\ 1/2; \quad BC\text{: } 0 \times (3/2) + 70 \times (1) = 70;$$

$$BD\text{: } 9 \times (3/2) + 51 \times (1) = 65\ 1/2; \quad CD\text{: } 10 \times (3/2) + 60 \times (1) = \underline{75},$$

so *CD* is the winning pair, which does not overlap *AB*. Under Webster, the sequential and nonsequential winners are the same as under Jefferson.

Observe that *AB* represents $v_1(AB) + v_2(AB) = 0 + 72 = 72$ of the 91 voters, whereas *CD* represents $v_1(CD) + v_2(CD) = 10 + 60 = 70$, so *AB*, the winner under sequential Jefferson, is more representative than *CD*, the winner under nonsequential Jefferson, exactly opposite to Example 3. Thus, neither the sequential nor the nonsequential versions of Jefferson and Webster invariably produce more representative committees. ■

While the nonsequential version of each apportionment method produced the most representative 2-candidate committees in Examples 2 and 3, it was the sequential version in Example 4 that did so. But when either the sequential or nonsequential version of Jefferson or Webster gives a more representative committee, is that version the one that should be chosen?

Not necessarily. One important principle is that the method of vote aggregation should be known in advance. From the calculations we have made so far, it seems that the nonsequential outcome is more likely to be more representative than the sequential outcome when the two differ, so we recommend the nonsequential method if it is feasible. With it, one can rest assured that (by definition) one finds the committee that maximizes voter satisfaction; by contrast, the sequential committee must include the AV winner, a restriction that sometimes, but not always, reduces representativeness.[8]

It is important to point out that neither the sequential nor the nonsequential committees may be the most representative possible, because both Jefferson and Webster take into account the level of support of the candidates. This may prevent either method from choosing the set of candidates that is most representative, as our next example demonstrates.

Example 5. 2 of 3 candidates $\{A, B, C\}$ to be elected. The number of voters who approve of different subsets of candidates are

$$4: AB; \quad 1: C.$$

The most representative committees are AC and BC, representing all 5 voters, but both the sequential and nonsequential versions of both Jefferson and Webster choose AB, so the C voter is unrepresented. Clearly, the 4 voters supporting AB prevent C, with support

[8] We have not attempted to compute how often, on average, this would occur, but a computer simulation would shed light on this question.

from only one voter, from winning a seat under either version of both Jefferson and Webster.

What if there were 3 *AB* voters instead of 4, holding constant the 1 *C* voter? Whereas both sequential and nonsequential Jefferson still give *AB* an edge over *AC* and *BC*, both sequential and nonsequential Webster produce a 3-way tie among *AB*, *AC*, and *BC*. Thus, not only do Jefferson and Webster give different outcomes, but two of the three outcomes given by sequential and nonsequential Webster (*AC* and *BC*) are more representative than the unique outcome (*AB*) given by sequential and nonsequential Jefferson.

Example 1 also illustrates that outcomes produced by Webster may be more representative than those produced by Jefferson: Sequential Jefferson elects *AB*, representing 12 of the 16 voters, whereas sequential Webster elects *AD*, representing 14 out of 16. Nonsequential versions of each method yield the same outcomes, suggesting that Webster gives at least as representative, and sometimes more representative, outcomes than Jefferson for both the sequential and nonsequential versions of each method. But this is not always true.

Proposition 3. *For committees of size 2 elected by the nonsequential versions of Jefferson and Webster, the Webster committee is equally or more representative. The same is true for the sequential versions if one candidate is the unique approval winner. But for larger committees for both the nonsequential and the sequential version, either the Jefferson or the Webster committee may be more representative.*

Proof. For a committee of size 2, the nonsequential score corresponding to an *h*–sequence is

$$S_h(C) = v_1(C) + \left(1 + \frac{h}{h+1}\right) v_2(C) = R(C) + \frac{h}{h+1} v_2(C).$$

For any $h^{**} > h^* > 0$, suppose that $S_{h^{**}}(C)$ is maximized by $C = C^{**}$ and $S_{h^*}(C)$ is maximized by $C = C^*$. Then $S_{h^{**}}(C^{**}) \geq S_{h^{**}}(C^*)$ and $S_{h^*}(C^*) \geq S_{h^*}(C^{**})$ imply, respectively, that

$$\frac{h^{**}+1}{h^{**}}[R(C^*) - R(C^{**})] \leq v_2(C^{**}) - v_2(C^*)$$

and

$$v_2(C^{**}) - v_2(C^*) \leq \frac{h^*+1}{h^*}[R(C^*) - R(C^{**})].$$

It follows that $\left(\frac{1}{h^*} - \frac{1}{h^{**}}\right)[R(C^*) - R(C^{**})] \geq 0$, which establishes that $R(C^*) \geq R(C^{**})$, since $h^{**} > h^*$. In particular, the nonsequential Webster committee ($h^* = ½$) must be at least as representative as the Jefferson committee ($h^* = 1$).

In the sequential case, the two procedures based on h^{**} and h^* (where $h^{**} > h^* > 0$) produce the same first-round winner, A, because (by assumption) there is no tie for approval-vote winner. We focus on three candidates, A, B, and C, and suppose that A is the approval winner and that, in the second stage, B is more deserving than C according to h^{**}, but C is more deserving than B according to h^*. We then show that AC must be a more representative committee than AB. In particular, it may be the case that the winning committee based on h^{**} is AB, whereas the one based on h^* is AC.

For any subset $S \subseteq \{A, B, C\}$, let $n(S)$ denote the number of voters who voted exactly for S plus, perhaps, some candidates other than A, B, and C. Thus, $n(A)$ is the total number of voters who voted for A but not B or C. Similarly, $n(ABC)$ is the number of voters who voted for all of A, B, and C, and $n(\emptyset)$ is the number of voters who voted for none of A, B, and C.

Our suppositions imply that

$$\frac{h^{**}}{h^{**}+1}[n(AB)+n(ABC)]+[n(B)+n(BC)]$$

$$\geq \frac{h^{**}}{h^{**}+1}[n(AC)+n(ABC)]+[n(C)+n(BC)]$$

and that

$$\frac{h^{*}}{h^{*}+1}[n(AB)+n(ABC)]+[n(B)+n(BC)]$$

$$\leq \frac{h^{*}}{h^{*}+1}[n(AC)+n(ABC)]+[n(C)+n(BC)].$$

It follows that

$$\frac{h^{**}+1}{h^{**}}[n(C)-n(B)] \leq n(AB)-n(AC) \leq \frac{h^{*}+1}{h^{*}}[n(C)-n(B)].$$

Thus, $\left(\frac{1}{h^*}-\frac{1}{h^{**}}\right)[n(C)-n(B)] \geq 0$, which, because $h^{**} > h^*$, establishes that $n(C) - n(B) \geq 0$.

The representativeness of the two subsets, AB and AC, is given by

$$R(AB) = n(ABC) + n(AB) + n(AC) + n(BC) + n(A) + n(B)$$

$$R(AC) = n(ABC) + n(AB) + n(AC) + n(BC) + n(A) + n(C) .$$

Subtraction of these equations yields

$$R(AC) - R(AB) = n(C) - n(B),$$

which is non-negative from the previous paragraph. It follows that the sequential committee based on h^* can be no less representative than the sequential committee based on h^{**}. In particular, the sequential Webster committee ($h^* = ½$) must be at least as representative as the Jefferson committee ($h^* = 1$).

For larger committees, Example 6 below shows that it is possible for a Jefferson committee to be more representative than a Webster committee under both the sequential and non-sequential procedures. ■

The following example shows that Proposition 3 does not extend to committees of size greater than 2.

Example 6. 3 of 4 candidates $\{A, B, C, D\}$ to be elected. The numbers of voters who approve of different subsets of candidates are

1: A 12: C 186: AD 186: BD 540: ABC.

For each of the four possible committees, we sum, going from left to right, the products of the number of voters in each subset and the sum of their rep sequences. For example, for the 540 ABC voters, if the committee elected under Jefferson comprises all three of their approved candidates, we multiply 540 by the sum of their rep sequence, 1 + 1/2 + 1/3 = 11/6. For nonsequential Jefferson and Webster, we have the following approval scores:

Jefferson:

$$ABC: 1(1) + 12(1) + 186(1) + 186(1) + 540(11/6) = \underline{1375}$$

$$ABD: 1(1) + 12(0) + 186(3/2) + 186(3/2) + 540(3/2) = 1369$$

$$ACD: 1(1) + 12(1) + 186(3/2) + 186(1) + 540(3/2) = 1288$$

$$BCD: 1(0) + 12(1) + 186(1) + 186(3/2) + 540(3/2) = 1287.$$

Webster:

$$ABC\text{: } 1(1) + 12(1) + 186(1) + 186(1) + 540(23/15) = 1213$$

$$ABD\text{: } 1(1) + 12(0) + 186(4/3) + 186(4/3) + 540(4/3) = \underline{1217}$$

$$ACD\text{: } 1(1) + 12(1) + 186(4/3) + 186(1) + 540(4/3) = 1167$$

$$BCD\text{: } 1(0) + 12(1) + 186(1) + 186(4/3) + 540(4/3) = 1166.$$

Thus, for Jefferson, *ABC*, which represents all 925 voters, is the committee elected, whereas for Webster, *ABD* is the committee elected, which represents only 913 voters (all except the 12 who voted for *C* only). Unlike Example 1, it is Jefferson, not Webster, that gives the more representative outcome. We show next that this result also holds for the sequential versions of each method.

A, B, *C*, and *D* receive, respectively, 727, 726, 552, and 372 approvals. Thus, under sequential Jefferson, *A* is the candidate elected first. On the second round, the candidates' deservingness scores are

$$B\text{: } 1(0) + 12(0) + 186(0) + 186(1) + 540(1/2) = \underline{456}$$

$$C\text{: } 1(0) + 12(1) + 186(0) + 186(0) + 540(1/2) = 282$$

$$D\text{: } 1(0) + 12(0) + 186(1/2) + 186(1) + 540(0) = 279.$$

Second-round Webster scores are

$$B\text{: } 1(0) + 12(0) + 186(0) + 186(1) + 540(1/3) = \underline{366}$$

$$C\text{: } 1(0) + 12(1) + 186(0) + 186(0) + 540(1/3) = 192$$

$$D\text{: } 1(0) + 12(0) + 186(1/3) + 186(1) + 540(0) = 248.$$

Consequently, both Jefferson and Webster elect *B* on the second round. Third-round Jefferson scores are

C: 1(0) + 12(1) + 186(0) + 186(0) + 540(1/3) = $\underline{192}$

D: 1(0) + 12(0) + 186(1/2) +186(1/2) + 540(0) = 186.

Third-round Webster scores are

C: 1(0) + 12(1) + 186(0) + 186(0) + 540(1/5) = 120

D: 1(0) + 12(0) + 186(1/3) +186(1/3) + 540(0) = $\underline{124}$.

Again, Jefferson elects *ABC* and Webster *ABD*, duplicating the nonsequential committees and proving that Jefferson may produce a more representative committee than Webster.

■

To illustrate the proof of Proposition 3 as it pertains to committees of size 2 for the nonsequential method, we use Example 1. Figure 1 shows the six possible committees in two dimensions—the horizontal dimension is $R = v_1 + v_2$, and the vertical dimension is v_2. In Example 1, $v_1(AB) = 7$ and $v_2(AB) = 5$, so $R(AB) = 12$, and *AB* is plotted at (12, 5).

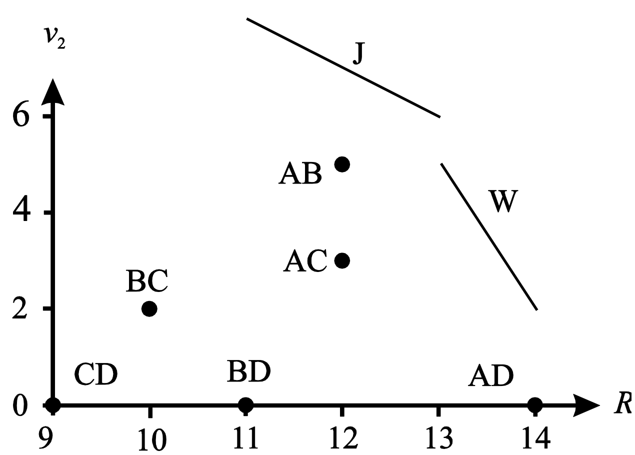

Figure 1. Properties (R, v_2) of All Possible Committees in Example 1

To visualize the maximization, observe that all six points lie on one side of the line *J*, which has slope −2. Imagine moving the line *J* parallel to itself so that it just touches one of the six points representing the committees, keeping the other five points on the same side. It is clear that the committee that comes first with respect to line *J* is *AB*.

For the Webster maximization, the process is the same, but the initial line, labelled *W*, has slope −3. It is clear again that the committee that the (extended) W line touches first is AD, which is the most representative, i.e., has the highest value of *R*.

The states in the Balinski-Young (1982/2001) model are akin to candidates in our model. In the Balinski-Young model, the states receive seats based on their populations—as if all their residents voted for their state. This determines how many seats the state receives in the U.S. House of Representatives.

By contrast, in our model, a candidate can receive votes from any voter, and a voter can vote for more than one candidate. These votes determine whether a candidate wins a seat on a committee, but not how many, because each candidate can win at most one seat. In section 4, however, we analyze the problem of apportioning different numbers of seats to parties.

4. The Jefferson and Webster Methods Applied to Parties

The standard applications of Jefferson and Webster are to apportion representatives to states, according to their populations, or seats in a legislature to political parties, according to the votes they receive. Just as people can be counted as residents of only one state to determine its population, voters can currently vote for only one party.

Now assume that voters are not restricted to voting for one party but can approve of as many parties as they like. Unlike individuals who can receive only one seat on a committee, parties can receive multiple seats in a legislature.

To calculate the numbers of seats that parties receive, we assume that each party nominates as many candidates, s, as will be elected to the legislature. Thus, party I nominates candidates $i_1, i_2, \ldots, i_s$; if an apportionment method allocates $k \leq s$ seats to I, they go to candidates $i_1, i_2, \ldots, i_k$. We assume that a voter who votes for a party approves of all its candidates.

The following example illustrates how the Jefferson method would allocate seats to parties when voters are not restricted to voting for one party but can vote for more than one:

Example 7. 2 of 6 candidates, $\{a_1, a_2, b_1, b_2, c_1, c_2\}$ from parties $\{A, B, C\}$ to be elected. The numbers of voters who approve of different parties are

7: *AB* 5: *AC* 2: *B* 3: *C*,

which translates into votes for the following sets of candidates:

7: $a_1a_2b_1b_2$ 5: $a_1a_2c_1c_2$ 2: b_1b_2 3: c_1c_2.

Each of the two candidates of *A, B*, and *C* receives, respectively, 12, 9, and 8 approvals. Thus, candidate a_1 is the first candidate elected under sequential Jefferson. On the second round, deservingness scores must be compared for a_2 (since a_1 has already been elected from party *A*), b_1 (from party *B*), and c_1 (from party *C*). We put the summations in the format of Example 6 but exclude from them subsets of voters who contribute 0 to a candidate's approval score:

$$a_2\text{: } 7(1/2) + 5(1/2) = \underline{6}; \qquad b_1\text{: } 7(1/2) + 2(1) = 5\ 1/2; \qquad c_1\text{: } 5(1/2) + 3(1) = 5\ 1/2,$$

so a_2 is the second candidate elected, making the winning pair a_1a_2. Under nonsequential Jefferson, the satisfaction scores of the six possible winning pairs of candidates are

$$a_1a_2\text{: } 7(3/2) + 5(3/2) = \underline{18}; \qquad a_1b_1\text{: } 7(3/2) + 5(1) + 2(1) = 17\ 1/2;$$
$$a_1c_1\text{: } 7(1) + 5(3/2) + 3(1) = 17\ 1/2; \qquad b_1b_2\text{: } 7(3/2) + 2(3/2) = 13\ 1/2;$$
$$b_1c_1\text{: } 7(1) + 5(1) + 2(1) + 3(1) = 17; \qquad c_1c_2\text{: } 5(3/2) + 3(3/2) = 12,$$

so again a_1a_2 is the winning pair. Observe that 7 + 5 = 12 of the 17 voters are represented by this pair.

By contrast, sequential Webster, after choosing a_1, chooses c_1 on the second round, because the deservingness scores are

$$a_2\text{: } 7(1/3) + 5(1/3) = 4; \qquad b_1\text{: } 7(1/3) + 2(1) = 4\ 1/3; \qquad c_1\text{: } 5(1/3) + 3(1) = \underline{4\ 2/3},$$

so a_1c_1 is the winning pair and represents all 17 voters. Under nonsequential Webster, the satisfaction scores of the six pairs of candidates are

$$a_1a_2\text{: } 7(4/3) + 5(4/3) = 16; \qquad a_1b_1\text{: } 7(4/3) + 5(1) + 2(1) = 16\ 1/3;$$
$$a_1c_1\text{: } 7(1) + 5(4/3) + 3(1) = 16\ 2/3; \qquad b_1b_2\text{: } 7(4/3) + 2(4/3) = 12;$$
$$b_1c_1\text{: } 7(1) + 5(1) + 2(1) + 3(1) = \underline{17}; \qquad c_1c_2\text{: } 5(4/3) + 3(4/3) = 10\ 2/3,$$

so b_1c_1 is the winning pair, which again represents all 17 voters.

In applying apportionment methods to parties, we have assumed that more than one candidate can be elected from a party. In fact, as Example 7 illustrated for Jefferson, all the winners may be from the same party.

The apportionment methods are vulnerable to manipulation. To illustrate, consider the outcome, a_1a_2, under sequential and nonsequential Jefferson in Example 7. Assume that polls just before the election show that party A is a shoo-in to win one seat (a_1) and possibly two (a_1a_2). If you are one of the 5 AC voters and would prefer a committee of a_1c_1 to a_1a_2, you might well consider voting for just C to boost the chances of c_1 being the second winner, making the outcome a_1c_1.

More specifically, if you switch from AC to C, you increase the number of C voters from 3 to 4 and decrease the number of AC voters from 5 to 4. Then the outcome under both sequential and nonsequential Jefferson changes from a_1a_2 to, respectively, a_1c_1 and a tie between a_1c_1 and b_1c_1, thus producing a more diverse committee.[9] Put another way, your sincere preference for a committee comprising members of parties A and C—or at least a more diverse committee than a_1a_2—is abetted by voting for just C, demonstrating that sincerity is not a Nash equilibrium for Jefferson in Example 6.

That strategic voting may be optimal is, of course, not surprising, because virtually all voting systems are vulnerable to manipulation. What complicates matters in the case of the apportionment methods is that the determination of winners, and therefore optimal strategies to produce a preferred outcome, is anything but straightforward. This makes it difficult to use information from polls or other sources to determine optimal strategic choices, especially for nonsequential versions of the apportionment methods.

In elections for city councils and other small elected bodies in the United States, there are often only two parties (e.g., Democratic and Republican) or, in nonpartisan

---

[9] Assume one AC voter switches. Then the deservingness value of sequential Jefferson for c_1, after a_1 is elected, is 6, which is maximal (since the score for a_2 drops to 5 1/2) and makes a_1c_1 the outcome. Under nonsequential Jefferson, the maximal satisfaction value is 17 for two outcomes, a_1c_1 and b_1c_1, both of which are more diverse than a_1a_2.

elections, two factions, one liberal (e.g., change oriented) and one conservative (status quo oriented). Call the parties A and B, and assume that each voter votes for only one party. Let the fraction of voters who support A be f, so the fraction of B supporters is $1 - f$.

If there are s seats to be allocated, the question that the apportionment methods answer is how many seats are to be received by each party. Let $k = 1, 2, \ldots, s$. The apportionment method determines *thresholds* $t(s, k)$ such that party A receives k seats if

$$t(s, k - 1) < f < t(s, k).$$[10]

Recall from section 2 that the weights used in the Jefferson satisfaction function for electing 1, 2, 3, 4, … approved candidates are

$$1, 1/2, 1/3, 1/4, \ldots$$

and those used in the Webster satisfaction function are

$$1, 1/3, 1/5, 1/7, \ldots.$$

These sequences are equivalent to

$$1/(h+0), 1/(h+1), 1/(h+2), 1/(h+3), \ldots,$$

where $h = 1$ for Jefferson and $h = 1/2$ for Webster (Proposition 4 holds for other values of h besides 1 and 1/2).

Proposition 4. *Assume in a 2-party election that there are s seats to be filled, that each party has s candidates, and that every voter approves of every candidate of one*

[10] Because $f = t(s, k)$ for some k may be possible, a tie-breaking procedure, which we do not specify, may be required.

party (but no candidates of the other party). Fix $h > 0$. *In an apportionment method based on the weights*

$1/(h+0)$, $1/(h+1)$, $1/(h+2)$, $1/(h+3)$, ….

the thresholds for $k = 0, 1, 2, \ldots, s-1$, *are given by*

$$t(h,s,k) = \frac{h+k}{2h+s-1}$$

Proof. We prove the proposition first for the sequential and then for the nonsequential method. We assume that fN voters support party A and $(1-f)N$ support party B. Throughout, we disregard ties.

Sequential method. We use mathematical induction on s. First, the formula obviously holds for $s = 1$, since $t(h, 1, 0) = 1/2$ for all h and the majority party wins the one seat that is to be filled.

Now we assume that the formula holds for s and show that it holds for $(s + 1)$. Our assumptions imply that the applicable value of k ($k = 1, 2, \ldots, s-1$) satisfies

$$t(h,s,k-1) < f < t(h,s,k)\,,$$

or

$$\frac{h+k-1}{2h+s-1} < f < \frac{h+k}{2h+s-1}. \quad (1)$$

Of course, $k = 0$ corresponds to $f < t(h, s, 0)$ and $k = s$ to $f > t(h, s, s-1)$.

If party A has been awarded k out of the first s seats and $(s + 1)$ seats are to be awarded in total, let k' be the number of seats out of $(s + 1)$ to be awarded to A. The deservingness of the next A candidate is greater than the deservingness of the next B candidate, so that the final seat will go to A ($k' = k + 1$), if

$$\frac{fN}{h+k} > \frac{(1-f)N}{h+s-k}\,, \text{ or } f > \frac{h+k}{2h+s}\,;$$

but the final seat will go to B ($k' = k$) if

$$f < \frac{h+k}{2h+s}.$$

To complete the proof, we must show that

$$t(h, s+1, k'-1) < f < t(h, s+1, k'). \qquad (2)$$

If $f > \frac{h+k}{2h+s}$, then $k' = k + 1$, which makes (2) equivalent to

$$t(h, s+1, k) < f < t(h, s+1, k+1)\,, \text{ or } \frac{h+k}{2h+s} < f < \frac{h+k+1}{2h+s}.$$

If $f < \frac{h+k}{2h+s}$, then $k' = k$, which makes (2) equivalent to

$$t(h, s+1, k-1) < f < t(h, s+1, k)\,, \text{ or } \frac{h+k-1}{2h+s} < f < \frac{h+k}{2h+s}.$$

But these results follow from (1). For instance, because

$$\frac{h+k+1}{2h+s} - \frac{h+k}{2h+s-1} = \frac{h-k+s-1}{(2h+s)(2h+s-1)}$$

and $k \le s - 1$, (1) implies that $f < \frac{h+k+1}{2h+s}$, completing the proof if $f > \frac{h+k}{2h+s}$. The proof is similar if $f < \frac{h+k}{2h+s}$.

Nonsequential method. Let $Y(h, s, k)$ denote the total representativeness score if the s seats were awarded to k candidates from party A and $(s - k)$ from party B. Because there are N voters,

$$Y(h, s, k)$$

$$= fN\left(\frac{1}{h+0} + \frac{1}{h+1} + \cdots + \frac{1}{h+k-1}\right) + (1-f)N\left(\frac{1}{h+0} + \cdots + \frac{1}{h+s-k-1}\right).$$

Let $y(h, s, k) = Y(h, s, k)/N$. Then, for $k = 0, 1, 2, ..., s - 1$, the difference

$$d(h, s, k) = y(h, s, k+1) - y(h, s, k) = \frac{f}{h+k} - \frac{1-f}{h+s-k-1}$$

is decreasing in k, so the total score is maximized for the smallest k for which $d(h, s, k) < 0$, that is, for the value of k for which $d(h, s, k) < 0 < d(h, s, k-1)$. But because

$$d(h,s,k) = \frac{(2h+s-1)}{(h+k)(h+s-1-k)}\left(f - \frac{h+k}{2h+s-1}\right),$$

this double inequality is equivalent to

$$f - \frac{h+k}{2h+s-1} < 0 < f - \frac{h+k-1}{2h+s-1},$$

or to

$$t(h,s,k-1) < f < t(h,s,k)\,.$$

This proves that, for given f, the value of k that maximizes representativeness satisfies the specified thresholds. ■

The thresholds for our two apportionment methods are the following:

$$\text{Jefferson: } t(1,s,k) = \frac{k+1}{s+1}; \qquad \text{Webster: } t(1/2,s,k) = \frac{2k+1}{2s}\,.$$

If $s = 5$ and k varies from 0 to 4, these thresholds for winning 1 to 5 seats are

Jefferson: 1/6, 1/3, 1/2, 2/3, 5/6; Webster: 1/10, 3/10, 1/2, 7/10, 9/10.

Thus, to win one seat, a party needs to win at least 1/6 of the vote under Jefferson and 1/10 under Webster; to win all five seats requires 5/6 of the vote under Jefferson and 9/10 under Webster.

Evidently, the thresholds are equally spaced under Jefferson but not under Webster.[11] At the extremes, Webster requires a relatively small fraction (1/10) to win one

[11] The thresholds under Jefferson are the same as those for cumulative voting, whereby voters can spread a fixed number of votes—usually equal to the number of seats to be filled on a council—to one or more candidates (Brams, 1975/2004, ch. 3). The drawback of cumulative voting is that parties must determine

seat, and a relatively large fraction (9/10) to win all five seats. (Note, however, that Webster is equally spaced between the extremes—between 1 and 5 seats in the example.)

Define the thresholds of a divisor apportionment method to be *balanced* in the 2-party case if they render the attainment of an additional seat by a party independent of the number of seats it presently holds. By this definition, which is only applicable to the 2-party case, Jefferson intervals are balanced, whereas the Webster intervals, which make attaining the first seat "easy" and the last seat "hard," are not.[12]

Define the *quota* q_i of party i as the fraction f_i of the vote it receives times the number of seats, s, to be apportioned: $q_i = f_i s$. For example, if there are 5 seats on a council, the quota of a party that receives 32% of the vote is 0.32 × 5 = 1.6. That is, this party is "entitled" to exactly 1.6 seats.

A party's number of seats of seats must be an integer, so it is said to *receive quota* if its number of seats equals its quota, rounded up or down. From the thresholds we gave above for a 5-seat council, Jefferson would give this party one seat (because 0.32 is less than 1/3), but Webster would give it two seats (because 0.32 is greater than 3/10). As this example illustrates in the 2-party case, Webster favors the smaller party, Jefferson the larger party (68% gives it a quota of 3.4, so it would obtain four seats under Jefferson but only three seats under Webster).

---

how many candidates to run to ensure that their supporters do not spread their votes too thinly across too many candidates, thereby preventing the party's proportional representation on the council. As we noted earlier, the apportionment methods allow the parties to nominate a full slate of candidates, because they give decreasing weight to the election of additional approved candidates. This builds proportional representation into the apportionment method (i.e., without the necessity of strategizing about how many candidates to run), though what is considered "proportional" depends on the method (Jefferson or Webster) used.

[12] For other justifications of Jefferson, see Brill, Laslier, and Skowron (2016) and references therein. Brill, Freeman, „Janson, and Lackner (2017) analyze "lead-balancing" approaches that have similar justifications.

An apportionment method *satisfies quota* if every party always receives quota. If $s = 1$, it is clear that any method of allocating seats satisfies quota. If $s \geq 2$, satisfaction of quota (disregarding ties) means

$$\frac{k}{s} < t(h, s, k) < \frac{k+1}{s},$$

or $$\frac{k}{s} < \frac{h+k}{2h+s-1} < \frac{k+1}{s} \qquad (k = 0, 1, 2, \dots, s-1)\,.$$

Proposition 5 proves that, for the same 2-party case to which Proposition 4 pertains, the thresholds $t(s, h, k)$ satisfy quota provided that a simple condition on h holds when $s > 2$. Note that ties are again disregarded in the proof.

Proposition 5.[13] *If there are two parties, $s \geq 2$, and the context is the same as in Proposition 4, then quota is satisfied for all positive values of h if $s = 2$, or, if $s > 2$, for any positive value of $h < \frac{s-1}{s-2}$.*

Proof. Because of the relation $t(h, s, k) + t(h, s, s - k - 1) = 1$, we need consider only values of k satisfying $k \leq s/2$. First, it is immediate that

$$t(h, s, k) - \frac{k}{s} = \frac{k + h(s-2k)}{s(2h+s-1)} > 0$$

since k and $s - 2k$ are non-negative and cannot both be zero. Also,

$$\frac{k+1}{s} - t(h, s, k) = \frac{k(2h-1) + (s-1) - h(s-2)}{s(2h+s-1)}$$

and we can complete the proof by showing that the numerator of the fraction on the right is positive. Now if $h \geq ½$, then $2h - 1 \geq 0$, so the numerator cannot be less than

$$(s-1) - h(s-2) > 0 \text{ if } h < \frac{s-1}{s-2}.$$

[13] Special cases to which this proposition applies are for the weights of Jefferson ($h = 1$), Webster ($h = 1/2$), and Adams (h approaches 0).

Thus, the fraction is positive under this condition. To see that the same conclusion holds if $h < ½$, note that the numerator is decreasing in k, so its minimum, attained at $k = s/2$, equals

$$\frac{s}{2}(2h-1)+(s-1)-h(s-2)=2h+\frac{(s-2)}{2}>0. \blacksquare$$

Observe that, at the extreme value $h = \frac{s-1}{s-2}$, we have $t\left(\frac{s-1}{s-2}, s, 0\right) = \frac{1}{s}$ and $t\left(\frac{s-1}{s-2}, s, s-1\right) = \frac{s-1}{s}$, so both of these thresholds fall exactly on the quota boundaries. It follows that the bound on h cannot be improved.

If there are more than two parties, Proposition 5 is no longer true (Balinski and Young, 1982/2001). Nondivisor methods of apportionment (e.g., that of Hamilton) always satisfy quota, but they are subject to certain nonmonotonicity problems—for example, the Alabama paradox, whereby the apportionment of a party may decrease when the number of seats in a legislature increases (Balinski and Young, 1982/2001). It is possible, however, to marry Hamilton with Jefferson or Webster—or any of the other three divisor methods—and satisfy quota and avoid most paradoxes (Potthoff, 2014).

Balinski and Young advocate Webster in the apportionment of representatives to states, because it is least biased, showing no systematic tendency to favor some states, and it almost always satisfies quota. But in the apportionment of seats to parties in a legislature, they advocate Jefferson, because it discourages small parties.

Under Jefferson, a small party may win no seats, even when it would win one under Webster. This gives small parties a greater incentive to merge under Jefferson in order better to ensure that they win some seats. With its greater tendency to inhibit the fractionalization of parliaments into many small parties, Jefferson also facilitates the formation of a governing coalition comprising a few large parties (e.g., center-left or center-right) that together hold a majority of seats.

We believe that both Jefferson and Webster are likely to foster more cooperation among political parties if voters can approve of more than one party. In effect, voters would be able to support coalitions of parties that they prefer in a governing coalition rather than being restricted to singling out one party for exclusive support. To be sure, some voters will prefer to throw all their support to one party if they consider it the only party with which they have an ideological affinity, but other voters are likely to find more than one party—perhaps for different reasons—compatible with their views.

5. Conclusions

Extending approval voting to the election of multiple winners can create a tyranny of the majority. A majority faction or party can win all the seats on a committee or in a legislature, or at least a disproportionate number of them, giving little or no voice to the views of minorities.

By depreciating the approval votes of voters who have one or more of their approved candidates elected, apportionment methods enable different individuals or groups to gain representation, and the resulting voting body to reflect a wider range of viewpoints. We focused on two well-known divisor methods of apportionment, first proposed by Jefferson and Webster, for depreciating approval votes in determining the winners in a multiwinner approval election.

We distinguished sequential and nonsequential versions of each method. The sequential versions are widely used today, both in apportioning representatives to states and apportioning seats to political parties in a legislature. Either version may elect a more representative voting body, whose members are approved of by more voters, but the nonsequential version is more likely to maximize representativeness when the two methods differ.

Nonsequential versions of the Jefferson and Webster methods are computationally complex, which is perhaps why neither is presently used in any jurisdictions. But modern computers mitigate this problem, rendering them feasible for many elections. The fact that sequential and nonsequential versions of each method can produce nonoverlapping sets of winners shows that their impact on who is elected may be decidedly nontrivial. But the fact that the different sets of winners produced by each version tend to have very similar deservingness and satisfaction scores makes the choice of one or the other less consequential.

Although the apportionment methods are vulnerable to manipulation, determining optimal manipulative strategies appears hard. The Jefferson method, which has the same vote thresholds as cumulative voting for winning seats on a council, eliminates the need for a party to strategize about the number of candidates to run to ensure its proportional representation, as also does the Webster method (but with different thresholds).

In 2-party competition, the vote thresholds for winning are evenly spaced by Jefferson but not by Webster, making the former balanced. On the other hand, if there is no restriction on the number of parties and the methods produce different winners, more voters will tend to approve of the Webster winners than the Jefferson winners, making them more representative.

It seems fitting that in electing multiple winners, voters, using approval ballots, should be able to approve of multiple candidates or parties. This enhanced ability of voters to express themselves seems likely to foster more cooperation across ideological and party lines, attenuating the oft-observed gridlock in many elected voting bodies today.[14]

[14] Partisan gerrymanders, for seats both in state legislatures and in the U.S. House of Representatives, may not survive current U.S. Supreme Court challenges to their constitutionality (Grofman, 2017). As a remedy, multiwinner elections would seem attractive as a way to combat gerrymandering, although at the federal level their implementation would require repeal of the 1967 ban on multimember congressional districts.

References

Aziz, Haris, Markus Brill, Vincent Conitzer, Edith Elkind, Rupert Freeman, and Toby Walsh (2017). "Justified Representation in Approval-Based Committee Voting." *Social Choice and Welfare* (forthcoming). http://link.springer.com/article/10.1007/s00355-016-1019-3

Balinski, Michel L., and H. Peyton Young (1982/2001). *Fair Representation: Meeting the Ideal of One-Man, One-Vote*. New Haven, CT/Washington, DC: Yale University Press/Brookings Institution.

Blais, André, and Louis Massicotte (2002). "Electoral Systems." In Lawrence LeDuc, Richard S. Niemi, and Pippa Norris (eds.), *Comparing Democracies 2: New Challenges in the Study of Elections and Voting*. London: Sage, pp. 40-69.

Brams, Steven J. (1975/2004). *Game Theory and Politics*. New York/Mineola, NY: Free Press/Dover.

Brams, Steven J. (1990). "Constrained Approval Voting: A Voting System to Elect a Governing Board." *Interfaces* 20, no. 5 (September-October): 65-79.

Brams, Steven J. (2008). *Mathematics and Democracy: Designing Better and Voting and Fair-Division Procedures*. Princeton, NJ: Princeton University Press.

Brams, Steven J., and Peter C. Fishburn (1983/2007). *Approval Voting*. New York: Springer.

Brams, Steven J., and D. Marc Kilgour (2014). "Satisfaction Approval Voting." In Rudy Fara, Dennis Leech, and Maurice Salles (eds.), *Voting Power and Procedures: Essays in Honor of Dan Felsenthal and Moshé Machover*. Cham, Switzerland: Springer, pp. 323-346.

Brams, Steven J., D. Marc Kilgour, and M. Remzi Sanver (2007). "The Minimax Procedure for Electing Committees," *Public Choice* 132, nos. 33-34 (September): 401-420.

Brill, Markus, Rupert Freeman, Svante Janson, and Martin Lackner (2017). "Phragmén's Voting Methods and Justified Representation." Preprint. https://users.cs.duke.edu/~rupert/phragmen_aaai.pdf

Brill, Markus, Jean-François Laslier, and Piotr Skowron (2016). "Multiwinner Approval Rules as Apportionment Methods." Preprint. https://arxiv.org/pdf/1611.08691v1.pdf

Chamberlin, John R., and Paul H. Courant (1983). "Representative Deliberations and Representative Decisions: Proportional Representation and the Borda Rule." *American Political Science Review* 77, no. 3 (September): 718-733.

Cox, Gary W. (1997). *Making Voices Count: Strategic Coordination in the World's Electoral Systems*. Cambridge, UK: Cambridge University Press.

Edelman, Paul H. (2006a). "Getting the Math Right: Why California Has Too Many Seats in the House of Representatives." *Vanderbilt Law Review* 59, no. 2 (March): 296-346.

Edelman, Paul H. (2006b). "Minimum Total Deviation Apportionments." In Bruno Simeone and Friedrich Pukelsheim (eds.), *Mathematics and Democracy: Recent Advances in Voting Systems and Social Choice*. Berlin: Springer, pp. 55-64.

Elkind, Edith, Piotr Faliszewski, Piotr Skowron, and Arkadii Slinko (2017). "Properties of Multiwinner Voting Rules." *Social Choice and Welfare* 48, no. 3 (March): 599-632.

Grofman, Bernard (2017). "The Supreme Court Will Examine Partisan Gerrymandering in 2017. That Could Change the Voting Map." *Washington Post*, January 31, 2017. https://www.washingtonpost.com/news/monkey-cage/wp/2017/01/31/the-supreme-court-will-examine-partisan-gerrymandering-in-2017-that-could-change-the-voting-map/?utm_term=.9375596246eb

Kilgour, D. Marc, "Approval Balloting for Multi-Winner Elections." In Jean-François Laslier and M. Remzi Sanver (eds.), *Handbook on Approval Voting*. Berlin: Springer, pp. 105-124.

Kilgour, D. Marc, Steven J. Brams, and M. Remzi Sanver (2006). "How to Elect a Representative Committee Using Approval Balloting." In Bruno Simeone and Friedrich Pukelsheim (eds.), *Mathematics and Democracy: Recent Advances in Voting Systems and Collective Choice*. Berlin: Springer, pp. 893-95.

Kilgour, D. Marc, and Erica Marshall (2012). "Approval Balloting for Fixed-Size Committees." In Dan S. Felsenthal and Moshé Machover (eds.), *Electoral Systems: Studies in Social Welfare*. Berlin: Springer, pp. 305-326.

Laslier, Jean-François, and M. Remzi Sanver (eds.) (2010). *Handbook on Approval Voting*. Berlin: Springer.

Monroe, Burt L. (1995). "Fully Proportional Representation." *American Political Science Review* 89, no. 4 (December): 925-940.

Potthoff, Richard F. (2014). "An Underrated 1911 Relic Can Modify Divisor Methods to Prevent Quota Violation in Proportional Representation and U.S. House Apportionment." *Representation* 50, no. 2: 193-215.

Potthoff, Richard F., and Steven J. Brams (1998). "Proportional Representation:

Broadening the Options." *Journal of Theoretical Politics* 10, no. 2 (April): 147-178.

Pukelsheim, Friedrich (2014). *Proportional Representation: Apportionment Methods and Their Applications*. Cham, Switzerland: Springer.

Sánchez-Fernández, Luis, Noberto Fernández Garcia, and Jesús A. Fisteus (2016). "Fully Open Extensions of the D'Hondt Method." Preprint. https://arxiv.org/pdf/1609.05370v1.pdf

Sivarajan, Shankar N. (2016). "A Generalization of the Minisum and Minimax Voting Methods." Preprint. https://arxiv.org/pdf/1611.01364v1.pdf

Subiza, Begoña, and Josep E. Peris (2014). "A Consensual Committee Using Approval Balloting." Preprint. https://web.ua.es/es/dmcte/documentos/qmetwp1405.pdf