Soc Choice Welf
DOI 10.1007/s00355-015-0914-3

ORIGINAL PAPER

Aggregation of binary evaluations: a Borda-like approach

Conal Duddy[1] · **Ashley Piggins**[1] · **William S. Zwicker**[2]

Received: 6 September 2014 / Accepted: 16 August 2015

Abstract We characterize a rule for aggregating binary evaluations—equivalently, dichotomous weak orders—similar in spirit to the Borda rule from the preference aggregation literature. The binary evaluation framework was introduced as a general approach to aggregation by Wilson (J Econ Theory 10:89–99, 1975). In this setting we characterize the "mean rule," which we derive from properties similar to those Young (J Econ Theory 9:43–52, 1974) used in his characterization of the Borda rule. Complementing our axiomatic approach is a derivation of the mean rule using vector decomposition methods that have their origins in Zwicker (Math Soc Sci 22:187–227,

Financial support from the Spanish Ministry of Economy and Competitiveness through MEC/FEDER grant ECO2013-44483-P, and the Irish Research Council co-funding from the European Commission, is gratefully acknowledged. Conal Duddy conducted this research while visiting the Centre for Philosophy of Natural and Social Science at the London School of Economics. He is grateful for their hospitality. Thanks to Felix Brandt, Franz Dietrich, Nicolas Houy, Vincent Merlin, Marcus Pivato and the participants at PET 13 for their helpful comments. This joint project began after we attended a workshop on New Developments in Judgement Aggregation and Voting Theory, held in Freudenstadt, September 2011. The workshop was organised by Clemens Puppe and Klaus Nehring. We are extremely grateful to them for organizing this workshop, and for providing financial support. Finally, we would like to thank the Managing Editor, the Associate Editor and a referee, for their helpful comments on an earlier version of this paper.

✉ Ashley Piggins
ashley.piggins@nuigalway.ie

Conal Duddy
conal.duddy@nuigalway.ie

William S. Zwicker
zwickerw@union.edu

1 J.E. Cairnes School of Business and Economics and the Whitaker Institute, National University of Ireland Galway, University Road, Galway, Ireland

2 Department of Mathematics, Union College, Schenectady, NY 12308, USA

Published online: 10 September 2015

1991). Additional normative appeal is provided by a form of tension minimization that characterizes the mean rule and suggests contexts wherein its application may be appropriate. Finally, we derive the mean rule from an approach to judgment aggregation recently proposed by Dietrich (Soc Choice Welf 42:873–911, 2014).

JEL Classification D71

1 Introduction

The Borda rule is a way to resolve the problem of "cyclic" profiles in preference aggregation. This problem is illustrated by the majority voting paradox. However, other aggregation problems have their own versions of this paradox. One that has been discussed recently is the "doctrinal paradox" from the judgment aggregation literature.[1] As with preference aggregation, when one aggregates individual judgments on related issues by applying majority rule separately to each, the result may be an incoherent collective judgment. It is natural to ask, then, if the Borda rule from the preference aggregation literature has a counterpart in other aggregation problems.[2] This question is important because some advocate the Borda rule as the "best" solution to the paradoxes that arise from preference aggregation.[3] Perhaps there are other aggregation problems that a suitable "Borda" rule can solve too? In any case, developing alternatives to majoritarian aggregation in these contexts should be of considerable interest.

How can we construct a Borda-like rule in other social choice settings? One possible strategy is to consider Young's (1974) original characterization of the Borda rule via four axioms. By translating these axioms into another setting we might identify a plausible analogue of the Borda rule for a different aggregation problem. We show that binary evaluations represents at least one aggregation context wherein Young's axioms translate into versions that characterize a particular aggregation rule. It seems reasonable to claim, therefore, that the resulting "mean rule" is the counterpart to the Borda rule for this problem.

Here is an example of the mean rule. In Table 1, there are four students, s_1, s_2, s_3 and s_4, and nine teachers.

The teachers have decided to partition the set of students into two non-empty subsets: a *higher-ability* group, denoted by H, and a *lower-ability* group, L. The two groups are then to be taught separately, according to their abilities. The teachers have been asked to provide their individual opinions as to who should go into which group; each is to submit a binary evaluation of each student, expressed in the form of a binary

[1] List and Pettit (2002) started the modern literature on this topic. List and Polak (2010) is a recent survey. The doctrinal paradox appears in Kornhauser and Sager (1993).

[2] This question has been addressed recently by Dietrich (2014). We consider Dietrich's approach in Sect. 6 of this paper.

[3] Donald Saari is the most prominent holder of this view. His arguments appear in Saari (1995, 2001a, b, 2008) and in many articles.

Table 1 Majority vs. mean rule

	s_1	s_2	s_3	s_4
Four teachers	H	H	L	L
Three teachers	L	H	H	H
Two teachers	H	L	H	H
Majority	H	H	H	H
Mean rule	H	H	L	L

partition (a bipartition).[4] According to Table 1, the first four teachers believe that s_1 and s_2 should be in H, and that s_3 and s_4 should be in L, the next three judge s_1 to be so weak as to deserve the special attention of a teacher, etc. The collective bipartition is to be determined by these individual bipartitions by aggregating them in some way. Applying majority rule to Table 1 leads to all of the students being put into H, which violates the requirement that the outcome of the aggregation exercise be a bipartition.[5] Under the mean rule, we calculate the mean number of H evaluations across all of the students. In the example, this is $5\frac{3}{4}$. The mean acts as a threshold. Those students with a total number of H evaluations greater than $5\frac{3}{4}$ are put into H, and those with less than $5\frac{3}{4}$ are put into L. Therefore, the mean rule puts s_1 and s_2 into H, and s_3 and s_4 into L.[6] We show that the mean rule is characterized by Young's axioms when they are adapted to this formal framework.

Another argument in favor of the mean rule comes from the idea of *tension*. When submitting their bipartitions, a natural concern of the teachers is to make the students in each group as similar as possible, and so minimize the total amount of tension within H and L. We take the pairwise tension between two students to be the absolute difference between their levels of ability, according to some (as yet, unspecified) ability measure; total tension is then the sum of pairwise tension over all pairs chosen from the same group. We will become more precise shortly, using "tension" informally for the moment. A teacher who considers a small number of students to be significantly more able than the rest may believe that tension is minimized by placing only those few into H. Similarly, if there is one particularly weak student, it may be optimal to have only that single student in L. Clearly, unless all of the students have exactly the same

4 Technically, such an evaluation is an *ordered bipartition* in the form of an ordered pair (H, L). That is, the unordered pair $\{H, L\}$ bipartitions the set S of students, while the ordering of the pair (H, L) identifies which piece of the bipartition is the "Higher" piece. We use the term "bipartition" in this paper to always mean an ordered bipartition and we assume that both pieces of the bipartition are non-empty. Note that the same information as (H, L) would be conveyed by providing H alone (with the understanding that $L = S \setminus H$). A binary evaluation is thus formally equivalent to an *approval voting* ballot. A binary evaluation can also be identified with a *dichotomous weak order relation*, and we take this approach in Sect. 5.

5 We can write this majority evaluation as $(+1, +1, +1, +1)$ to denote that students 1, 2, 3 and 4 respectively are all in H. In this way, evaluations can be regarded as vectors. Note that the majority method satisfies an "independence" property in that the collective decision on a student is independent of the teachers' evaluations of the other students. Significantly, the mean rule that we present in this paper does not satisfy this property.

6 In our model, if a student has a number of H evaluations equal to the mean then we take it that there are multiple solutions which are in an intuitive sense "tied".

Table 2 The ability-levels of the four students

	s_1	s_2	s_3	s_4
Ability-levels	6	7	5	5
Tension-minimization	H	H	L	L

Table 3 The ability-levels of five students

s_1	s_2	s_3	s_4	s_5
0	2	2	2	5

level of ability, placing them all into one group will never be tension-minimizing. This justifies our assumption that the teachers submit bipartitions and that the collective outcome is also a bipartition.

Suppose we take the number of teachers who assign a given student to H to be a measure of that student's level of ability. Then, as noted above, for any two students the absolute difference between these numbers serves as a measure of the tension between them, and the sum of these differences—taken over every pair of students that belong to the *same* member of the pair (H, L)—represents the total tension in (H, L). To give an example, the ability-levels of the four students in Table 1 are given in Table 2.

Based on these ability-levels, the total tension in $(\{s_1, s_3\}, \{s_2, s_4\})$ is simply the tension in $\{s_1, s_3\}$ plus the tension in $\{s_2, s_4\}$, which is $1+2 = 3$. It is easy to determine that the bipartition that minimizes total tension is $(\{s_1, s_2\}, \{s_3, s_4\})$ which leads to a total tension of $1 + 0 = 1$. This is identical to the outcome under the mean rule. We show that this is always the case when a natural monotonicity requirement is satisfied, lending further normative support to the mean rule.

To justify this requirement, consider the example of five students and five teachers depicted in Table 3. The zero below s_1 indicates that none of the teachers assigned that student to H while two teachers assigned s_2 to H, and so on.

If we assign students s_1 and s_5 to H, and the other three students to L then total tension is 5 $(=5+0)$. It is easy to verify that this is the tension-minimizing bipartition. However, this tension-minimizing bipartition is problematic from an intuitive point of view. It places the most able and the least able students together into H which we take to be undesirable for ability grouping.[7] To avoid this, we impose a natural monotonicity requirement. This says that the lowest level of ability among the students in H must be at least as great as the greatest level of ability in L. Under this constraint, the unique tension-minimizing bipartition places only student s_5 into H (for a total tension of 6). Therefore, we have $H = \{s_5\}$ and $L = \{s_1, s_2, s_3, s_4\}$ under this tension-minimizing approach.[8] As noted above, this is identical to the outcome under the mean rule. Note that if there is a student whose ability-level is equal to the mean then there are multiple solutions (all of which achieve a common lowest tension).

[7] We are grateful to Nicolas Houy for proposing the first example of this kind.

[8] In a planned sequel to this paper (*Social Dichotomy Functions*) we show that a measure of *external displacement*, related to total tension, is also optimized by "cutting" at the mean in this way, with no need to impose a monotonicity condition.

In this paper we also show how the mean rule emerges when we apply Zwicker's (1991) vector decomposition technique to this social choice problem (expressed in its dichotomous weak order form). Zwicker's technique allows us to decompose a standard preference profile into a portion that tends towards a majority cycle, and a portion that tends towards no cycle in majority preference.[9] He then shows that the Borda ranking at the original profile is identical to the (transitive) majority preference at this second, acyclic portion of the original profile. It is noteworthy that two very different approaches to social choice theory (one axiomatic, based on Young, and the other based on linear algebra) both lead to the mean rule.

Before we outline the paper, we should explain how our aggregation problem differs from the classic voting problem. In that problem, the objective is to devise voting rules that will help us identify the "best" candidate(s). The goal here, however, is not to identify the best student. If we wanted to do that, then, given our formal set-up (described in footnote 4) we would use Approval Voting. Instead, we want to identify the best binary partitioning of the students. This, we have argued, is the one that minimizes total tension. The social outcome that concerns us, therefore, is different to the classic problem of voting.[10]

The paper is organized as follows. We describe our basic model in Sect. 2. We prove in Sect. 3 that the mean rule determines the tension-minimizing, monotone bipartition. In Sect. 4 we present an axiomatic approach to this social choice problem. In Sect. 5 we derive the mean rule using the vector decomposition approach. In that section we also introduce the "Borda" mean rule, which is a generalization of the mean rule. Under the Borda mean rule, the inputs are rankings and the output is a bipartition. Finally, in Sect. 6, we consider applying the mean rule to the problem of judgment aggregation. We consider an approach to judgment aggregation due to Dietrich (2014). Dietrich develops the idea of "scoring rules" for judgment aggregation. We present a new scoring rule for judgment aggregation, and consider its relationship to the mean rule and the classic Borda rule. Section 7 concludes the paper and suggests directions for future research.

2 Binary evaluations model

Our model is based on the binary evaluations model presented in Wilson (1975) and then developed by Rubinstein and Fishburn (1986), Dokow and Holzman (2010a) and Nehring and Puppe (2010).[11] X is a finite set of cardinality k containing the objects, issues or propositions under consideration. An evaluation f is an element of $\{+1, -1\}^k$ where $f(j)$ is the evaluation on issue j. Assigning $+1$ to an issue in X means that

[9] These ideas are developed in Saari (1994).

[10] Approval Voting implicitly produces a partition into "winners" and "losers", but this will not be tension-minimizing in general. As we can see from Table 1, the outcome under approval voting is $H = \{s_2\}$ and $L = \{s_1, s_3, s_4\}$ with the winner being s_2.

[11] Using a similar model, Brams et al. (1998) consider the "paradox of multiple elections" where voters vote yes/no on a sequence of propositions on a referendum. Generalizations of the binary evaluations model can be found in Dokow and Holzman (2010b) and in the literature on combinatorial domains (chapter 8 of Xia 2011 is an introduction).

the evaluator judges that this issue should receive a "high" evaluation, a -1 means that it should receive a "low" evaluation. This is equivalent to specifying a subset of X containing all of the issues evaluated as $+1$. For example, if X contains a set of propositions then this subset is a *judgment set* containing those propositions that the evaluator believes to be true (or accepts). Up to this point, our model is as general as the original model of binary evaluations and judgment aggregation theory.

We impose a feasibility constraint on evaluations (consistent with our idea of tension minimization in Sect. 1). An evaluation is *feasible* if it does not involve $+1$ on all k issues nor -1 on all k issues. This constraint means that our model is a special case of the original binary evaluations model (and a special case of judgment aggregation) in that the constraint always takes this particular form and does not vary from problem to problem. There are important social choice contexts where this feasibility constraint applies. We mention four of them.

1. List's (2008) model of judgment aggregation on a single proposition (where this proposition is modeled by a set of possible worlds). Suppose the members of a group each form a judgment on which worlds in a given set are possible, subject to the constraint that at least one world is possible but not all are. The group then seeks to aggregate these individual judgments into a collective judgment, subject to the same constraint.
2. Kasher and Rubinstein's (1997) group identification problem. The members of a group each make a judgment on which members of that group have a given property (i.e. having a particular religious affiliation), subject to the constraint that at least one individual has that property but not all individuals do.
3. The problem of aggregating dichotomous weak preferences, where issues are interpreted as social alternatives, and where no evaluator can be indifferent between all of the alternatives.[12]
4. The problem of aggregating linear orders over three alternatives. This is the classic setting of the majority voting paradox. We show in Sect. 6 how our model applies to this problem.

Of course, feasible evaluations are formally equivalent to ordered bipartitions (see footnote 4). In our model, each individual holds a feasible evaluation, and these are aggregated to determine a feasible evaluation (or a set of feasible evaluations) for the group.

We now present some formal definitions. Let N be a countably infinite set of individuals and let S be the set of all finite subsets of N (including the empty set). Let V be the set of all feasible evaluations.

A *profile* is a function P from V to S such that $P(v)$ and $P(v')$ are disjoint for all distinct evaluations v and v'. Given an evaluation v we interpret $P(v)$ as the set of individuals with evaluation v at profile P. The requirement that $P(v)$ and $P(v')$ are disjoint means that at profile P each individual submits no more than one evaluation. The set of participating individuals at P is the union of $P(v)$ over all evaluations v in V.

[12] The literature on aggregating dichotomous preferences is large and covers a wide range of questions. See Bogomolnaia et al. (2005) and Maniquet and Mongin (2015) for a taste of these.

It may happen that the aggregated bipartition yields a tie between distinct evaluations that differ with respect to issue i. To indicate such a tie we will assign 0 to issue i in the *outcome vector*. For example, we say that the two evaluations in $\{(-1, +1, +1), (-1, -1, +1)\}$ "agree" with the vector $(-1, 0, +1)$, we identify $(-1, 0, +1)$ with this set, and we interpret this single vector as a tie between those two evaluations. Formally, we take the *set of feasible outcomes* (for the group), to be the set $\{-1, 0, +1\}^k$ less both $\{-1\}^k$ and $\{+1\}^k$, and identify each feasible outcome with the set of feasible evaluations that agree with it. To give another example, we identify $(0, 0, 0)$ with the set of all six feasible evaluations over three issues. Thus 0 serves as shorthand for a certain type of tie, rather than as a distinct, third evaluation in our model.[13]

Binary aggregation rule. A *binary aggregation rule* is a function F from the set of all profiles to the set of feasible outcomes. We denote by $F(P)_i$ the ith component of outcome $F(P)$.

Given an issue i and profile P, let $y(P)_i$ be the number of individuals who give issue i a $+1$ evaluation, and $n(P)_i$ be the number of individuals who give issue i a -1 evaluation, at profile P. Let $y(P)$ be the k-tuple $(y(P)_1, \ldots, y(P)_k)$, and let $n(P)$ be $(n(P)_1, \ldots, n(P)_k)$.

We now give a formal definition of the mean rule. For all profiles P, let $\bar{y}(P)$ denote the sum of the components of $y(P)$ divided by k. In other words, $\bar{y}(P)$ is the mean number of $+1$ evaluations received by the elements of X. Define a particular binary aggregation rule F_M as follows.

Mean rule. For all profiles P and all issues i,

$$F_M(P)_i = \begin{cases} +1 & \text{if } y(P)_i > \bar{y}(P) \\ -1 & \text{if } y(P)_i < \bar{y}(P) \\ 0 & \text{if } y(P)_i = \bar{y}(P). \end{cases}$$

[13] Nonetheless, this notational convention has two substantive implications. First, it allows certain types of ties while forbidding others. For example, the two-way tie $\{(+1, -1, +1, +1), (+1, -1, -1 - 1)\}$ is disallowed, as it does not arise as the set of all feasible evaluations that agree with some vector in $\{-1, 0, +1\}^4$. The effect is to eliminate variants of the mean rule that agree with it except that they yield some forbidden ties. Such variants *might* otherwise satisfy the four characterizing axioms of Theorem 7 …"might" because in the absence of the convention, the consistency axiom must be reformulated (which can be done in more than one way). Still, it seems worth considering whether some altered version of the Theorem 7 axioms selects the mean rule uniquely from among these variants, without any assistance from our *feasible outcome* convention; this seems plausible, as the *tension minimization* characterization does select the mean rule uniquely.

Note, as well, that the feasible outcomes $y = (+1, +1, 0)$ and $y' = (+1, +1, -1)$ give rise to the same single feasible evaluation $(+1, +1, -1)$ (because $(+1, +1, +1)$ is not feasible). Thus, two rules that are formally distinct only for this reason (one outputs y at some profile for which the other outputs y') should actually be considered the same rule. This happens only for outcomes having a single 0 with all other coordinates equal, so one could alternately address the issue simply by further restricting the definition of *feasible outcome* so as to exclude such vectors.

3 Tension-minimizing bipartitioning

In this section we demonstrate that the mean rule determines the tension-minimizing bipartition. The "tension" between any two students is simply the absolute value of the difference between their ability levels. In a classroom of students, the total amount of tension is the sum of tension over every pair of students in the classroom. More generally, let S be a finite set and let s be a function that assigns a "score" (any real number) to each element of S. Then the total amount of tension within S, denoted $\mathcal{T}(S)$, is defined by the following equation.

$$\mathcal{T}(S) = \sum_{i,j \in S} |s(i) - s(j)| .$$

If we split the students into two groups, the total amount of tension is the sum of the tension within each of the two groups. More generally, the total amount of tension for a pair of sets (S_H, S_L) is defined as follows.

$$\mathcal{T}(S_H, S_L) = \mathcal{T}(S_H) + \mathcal{T}(S_L).$$

We seek to bipartition a set S into a pair of subsets S_H and S_L in a way that minimizes $\mathcal{T}(S_H, S_L)$ under the requirement of monotonicity. Monotonicity requires that for all i in S_H and all j in S_L we have $s(i) \geq s(j)$.

When we restrict our attention to monotone solutions, we can represent this bipartitioning problem in the following, intuitive way. First, we mark the scores of the elements of S along a number line. We then need to identify a point at which to "cut" the line into two parts; separating those scores that lie above the cut point and those that lie below. In other words, by cutting, we create a bipartition of the original set S into two subsets. Note that cutting at any point that is strictly between the lowest score and the highest score must lower total tension. The reason is that for every pair of elements that straddle the cut point, the tension between them becomes zero once the cut has occurred. The only tension that remains is the tension between those elements below the cut point, and the tension between those above the cut point.

If some elements have a score equal to the number where the cut is made then (as mentioned earlier) we treat this as implying that there are multiple solutions. More precisely, this means that there is a tie among all bipartitions that can be obtained by assigning these elements randomly and independently to either part of the bipartition. All of these bipartitions will achieve the same lowest tension (which is why we say that they are "tied"). We claim that cutting at the mean of the scores is tension-minimizing when monotonicity is required.

Theorem 1 *Let μ be the mean of $s(i)$ over all $i \in S$. If (S_H, S_L) is an ordered bipartition of S such that $s(i) > \mu$ implies $i \in S_H$ and $s(i) < \mu$ implies $i \in S_L$ then, for any other monotone ordered bipartition (S'_H, S'_L) of S, we have $\mathcal{T}(S_H, S_L) \leq \mathcal{T}(S'_H, S'_L)$.*

Proof Suppose that we make the cut initially at μ to generate the ordered bipartition (S_H, S_L). Then we "move" the cut point to a number less than the mean so that one

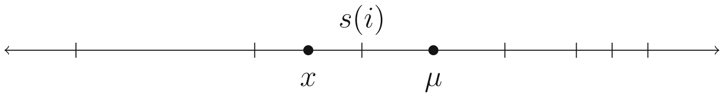


Fig. 1 Cutting tension

or more elements are transferred from the lower part to the upper part, generating a new ordered bipartition (S'_H, S'_L). Figure 1 gives an example of this; when the cut is moved from μ to x, element i is thereby transferred from the lower part to the upper part.

By doing this we reduce the amount of tension in the lower part and increase the amount of tension in the upper part. That is, we have $\mathcal{T}(S'_L) < \mathcal{T}(S_L)$ and $\mathcal{T}(S'_H) > \mathcal{T}(S_H)$.

Let M be the set of elements in S'_H that have a score less than μ. The tension *within* that set M, that is, $\mathcal{T}(M)$, is transferred from the original lower part, S_L, to the new upper part S'_H. The transferral of this tension from the lower to the upper part has no impact on the total amount of tension. But, of course, this is not the only effect of transferring those elements. The tension in the lower part is further reduced by the elimination of the tension that existed between M and $S_L \setminus M$. And we have now introduced tension between the members of M and the members of S_H. So what, then, is the net impact?

The tension between each element m of M and the set $S_L \setminus M$ is the sum of $s(m) - s(i)$ over every i in $S_L \setminus M$. Similarly, the tension between m and S_H is the sum of $s(i) - s(m)$ over every i in S_H. We claim that for every m in M the first of these two sums is less than the second. This implies that the total amount of tension increases as a result of transferring the members of M from the lower part to the upper part.

To see that this claim is true let us note the following well-known property of the mean. The sum of $\mu - s(i)$ over every i is zero. In other words, the sum of $\mu - s(i)$ over every i in S_L is equal to the sum of $s(i) - \mu$ over every i in S_H.

It follows that for any number α that is less than the mean, the sum of $\alpha - s(i)$ over every i in S_L must be *less* than the sum of $s(i) - \alpha$ over every i in S_H. If the former sum is taken over a subset of S_L, instead of all of S_L, then the difference just becomes even larger. Hence, the sum of $s(m) - s(i)$ over every i in $S_L \setminus M$ is less than the sum of $s(i) - s(m)$ over every i in S_H, which is what we claimed.

So moving the cut-point away from μ to a lower number so that one or more elements m with $s(m) < \mu$ are transferred from the lower part to the upper part results in an increase in total tension. A symmetric argument shows that moving the cut to a number greater than the mean so that one or more elements with scores greater than the mean are transferred from the upper part to the lower part also results in an increase in total tension.

When an alternative has a score equal to the mean, we can see that it does not matter which part of the ordered bipartition it is assigned to. If an issue j with $s(j) = \mu$ is assigned, say, to S_L then the amount of tension generated is the sum of $s(j) - s(i)$ (or, equivalently, $\mu - s(i)$) over every score $s(i)$ that is less than μ. If, on the other hand, j is assigned to S_H then the amount of tension generated is the sum of $s(i) - s(j)$ (or, equivalently, $s(i) - \mu$) over every score $s(i)$ that is greater than μ. As we have noted, a well-known property of the mean is that these two sums are equal. □

Tension minimization requires the mean rule to treat the cardinalities of sets S_H and S_L in the aggregated bipartition (S_H, S_L) as endogenous to the aggregation. For example, in our student partitioning example the tension minimizing partition might well place a single student in the higher-ability group. Nonetheless, the *number* of terms in the sum forming $\mathcal{T}(S_H, S_L)$ is lower when the sets are similar in size; consequently, the mean rule "leans" towards a more equal split when the distribution of individual scores $s(i)$ lacks outliers.

4 Axiomatic approach

Given a permutation σ on $\{1, \ldots, k\}$ and a profile P, let P^σ denote the profile such that P and P^σ have the same set of participating individuals and where an individual evaluates $+1$ on issue i at profile P if and only if that individual evaluates $+1$ on issue $\sigma(i)$ at profile P^σ. We write $\mathbf{0}$ to denote the vector $\{0\}^k$.

The following are properties that a binary aggregation rule F may satisfy. As mentioned in Sect. 1, they are essentially the axioms Young (1974) used in his characterization of the Borda rule, but translated into the framework of this paper.[14]

Cancellative. For all profiles P such that $y(P) = n(P)$, $F(P) = \mathbf{0}$.

Cancellation says that at profiles in which every issue is assigned $+1$ by half of the individuals and -1 by the other half, the social outcome is 0 over all of the issues. Cancellation is a natural requirement when a profile is "balanced".

Profiles P and P' are said to be disjoint if the electorates are disjoint, i.e. $P(v) \cap P'(v') = \emptyset$ for all $v, v' \in V$. The sum of two disjoint profiles, defined by $(P+P')(v) = P(v) \cup P'(v)$, corresponds to pooling the two electorates.

Consistent. For all disjoint profiles P and P', and all issues i, (1) if $F(P)_i = +1$ and $F(P')_i \geq 0$ then $F(P + P')_i = +1$, and (2) if $F(P)_i = -1$ and $F(P')_i \leq 0$ then $F(P + P')_i = -1$.

Consistency says the following. Imagine that two disjoint groups are merged together. If an issue is assigned $+1$ by one subgroup (using binary aggregation rule F), and the other subgroup assigns either $+1$ or 0 to this issue (using the same binary aggregation rule F), then F must assign $+1$ to this issue when we apply it to the union of these two subgroups. Conversely, if an issue is assigned -1 by one subgroup (using binary aggregation rule F), and the other subgroup assigns either -1 or 0 to this issue (using the same aggregation rule F), then F must assign -1 to this issue when we apply it to the union of these two subgroups. Consistency has a flavor of the traditional "Pareto principle" from welfare economics, and is normatively attractive for that rea-

[14] Note that Young's paper characterized the Borda rule as a social choice function. A characterization of the Borda rule as a social welfare function can be found in Nitzan and Rubinstein (1981). An alternative characterization of the mean rule based on principles from approval voting can be found in Duddy and Piggins (2013).

son. In terms of our original student partitioning problem, if the arts teachers assign a student to the higher-ability group, and the science teachers either agree or regard the matter as being "tied", then consistency requires that the student be assigned to the higher-ability group when the teachers are combined.

Neutral. For all permutations σ on $\{1, \ldots, k\}$ and all profiles P, $F(P)_i = F(P^\sigma)_{\sigma(i)}$ for all issues i.

Neutrality says that the criterion for determining whether an issue is assigned $+1, -1$ or 0 by F is the same for all of the issues. Different criteria/thresholds do not apply to different issues, in other words. Neutrality is a standard social choice requirement. In terms of the student partitioning problem, neutrality ensures that all of the students are treated fairly.

Faithful. For all profiles P such that the range of P contains only the empty set and one singleton, $P(v) \neq \emptyset$ implies $F(P) = v$.

Faithful says that when society consists of one individual, the social evaluation of the issues is identical to this individual's evaluation of the issues. This is a reasonable requirement of a one person society.

We now present five lemmas before proving our characterization theorem. We then show that the axioms used in the characterization are logically independent. First, we note the following property.

Strong Anonymity. For all profiles P and P' such that $y(P) = y(P')$ and $n(P) = n(P')$, $F(P) = F(P')$.

Lemma 2 *A binary aggregation rule that is cancellative and consistent is strongly anonymous.*

Proof Let F be a binary aggregation rule that is cancellative and consistent. Take two profiles P and P' such that $y(P) = y(P')$ and $n(P) = n(P')$. By way of contradiction, assume that $F(P) \neq F(P')$. Let us assume, without loss of generality, that $F(P)$ and $F(P')$ differ over issue i.

Construct a third profile P_*, disjoint from both P and P', such that $y(P_*) = n(P)$ and $n(P_*) = y(P)$. It follows that $y(P_*) = n(P')$ and $n(P_*) = y(P')$. Cancellation implies both that $F(P + P_*) = \mathbf{0}$ and that $F(P' + P_*) = \mathbf{0}$.

Since $F(P)_i$ and $F(P')_i$ differ, they cannot both be zero. Assume without loss of generality that $F(P)_i$ is not zero. If $F(P)_i = +1$ then $F(P_*)_i$ cannot be $+1$ or zero, since, by consistency, that would imply $F(P + P_*)_i = +1$, contradicting $F(P+P_*) = \mathbf{0}$. So if $F(P)_i = +1$ then we must have $F(P_*)_i = -1$. By a symmetrical argument, if $F(P)_i = -1$ then $F(P_*)_i = +1$. So $F(P_*)_i$ must take the opposite value to $F(P)_i$. However, $F(P')_i$ must also take the opposite value to $F(P)_i$ or else be zero, since $F(P)_i$ and $F(P')_i$ are not equal. It follows, by consistency, that $F(P' + P_*)_i$ must take the opposite value to $F(P)_i$, which is either $+1$ or -1. This contradicts $F(P' + P_*) = \mathbf{0}$. □

Lemma 3 *If a binary aggregation rule F is cancellative and consistent then for all disjoint profiles P and P' such that $y(P) = n(P')$ and $n(P) = y(P')$, $F(P) + F(P') = \mathbf{0}$.*

Proof Let F be a binary aggregation rule that is cancellative and consistent. Take any profile P and let P' be a profile that is disjoint from P and such that $y(P) = n(P')$ and $n(P) = y(P')$. By way of contradiction, assume there is an issue i such that $F(P)_i + F(P')_i \neq 0$. That is, despite P' being a kind of inverse of P, the outcomes are not opposites. If $F(P)_i = +1$ and $F(P')_i \geq 0$ then consistency implies $F(P+P')_i = +1$. On the other hand, if $F(P)_i = -1$ and $F(P')_i \leq 0$ then consistency implies $F(P + P')_i = -1$. So $F(P + P')_i$ must be either $+1$ or -1. However, since $y(P + P') = n(P + P')$, cancellation implies $F(P + P') = \mathbf{0}$. This is a contradiction. □

Lemma 4 *If a binary aggregation rule F is cancellative and consistent then for all disjoint profiles P and P', and all issues i, $F(P)_i = F(P')_i = 0$ implies $F(P+P')_i = 0$.*

Proof Let F be a binary aggregation rule that is cancellative and consistent. Take an issue i and disjoint profiles P and P' such that $F(P)_i = F(P')_i = 0$. Let P_* be a profile disjoint from P and P' such that $y(P_*) = n(P)$ and $n(P_*) = y(P)$. Let P'_* denote a profile disjoint from P, P' and P_* such that $y(P'_*) = n(P')$ and $n(P'_*) = y(P')$. Lemma 3 implies $F(P) + F(P_*) = \mathbf{0}$ and $F(P') + F(P'_*) = \mathbf{0}$. Therefore, since $F(P)_i = F(P')_i = 0$, we have $F(P_*)_i = F(P'_*)_i = 0$.

If it were the case that $F(P + P')_i = +1$, then, by consistency, we would have $F(P + P' + P_*)_i = +1$ and $F(P + P' + P_* + P'_*)_i = +1$. Similarly, if $F(P + P')_i = -1$ then $F(P + P' + P_* + P'_*)_i = -1$. However, cancellation implies that $F(P + P' + P_* + P'_*) = \mathbf{0}$. So it must be that $F(P + P')_i = 0$. □

In terms of the earlier scenario involving teachers and students, the next lemma says that if every student is assigned to the higher-ability group by the same number of teachers then neutrality, consistency and cancellation jointly imply that all feasible evaluations are tied.

Lemma 5 *If a binary aggregation rule F is neutral, consistent and cancellative then for all profiles P such that the components of $y(P)$ are all equal, $F(P) = \mathbf{0}$.*

Proof Let F be a binary aggregation rule that is neutral, consistent and cancellative. Take any profile P such that the components of $y(P)$ are all equal. By way of contradiction, assume that $F(P) \neq \mathbf{0}$. Without loss of generality, assume that $F(P)_i = +1$ for some issue i. By the definition of a binary aggregation rule we know that $F(P) \neq \{+1\}^k$. So $F(P)_j \leq 0$ for some issue j. Let P_* be a profile disjoint from P such that $y(P_*) = n(P)$ and $n(P_*) = y(P)$. Let σ be a permutation on $\{1, \ldots, k\}$ that transposes i and j. Since all the components of $y(P)$ are equal, we have $y(P^\sigma) = y(P)$. So $y(P^\sigma) = n(P_*)$ and $n(P^\sigma) = y(P_*)$. Since $y(P^\sigma) = n(P_*)$ and $n(P^\sigma) = y(P_*)$ this implies that $y(P^\sigma + P_*) = n(P^\sigma + P_*)$. Cancellation implies, therefore, that $F(P^\sigma + P_*) = \mathbf{0}$.

However, since $F(P)_i = +1$, neutrality implies that $F(P^\sigma)_j = +1$. Lemma 3 implies that $F(P)_j + F(P_*)_j = 0$ and we know that $F(P)_j \leq 0$. Therefore, we can deduce that $F(P_*)_j \geq 0$. So, by consistency, $F(P^\sigma + P_*)_j = +1$. This is a contradiction. □

Lemma 6 *If a binary aggregation rule F is faithful, neutral, consistent and cancellative then for all profiles P and all issues i, if the ith component of $y(P)$ is strictly greater than all of the other components of $y(P)$ then $F(P)_i = +1$, and if the ith component of $y(P)$ is strictly less than all of the other components of $y(P)$ then $F(P)_i = -1$.*

Proof Let F be a binary aggregation rule that is faithful, neutral, consistent and cancellative. Take any profile P and issue i such that the ith component of $y(P)$ is either strictly greater than or strictly less than all of the other components of $y(P)$. Without loss of generality, assume that the former is true. Construct a profile P', disjoint from P, such that the total number of participating individuals is $y(P)_i$, and, for all elements j, the number of individuals who say $+1$ to j is equal to $y(P)_i - y(P)_j$. Note that at profile P' every individual says -1 to issue i. So if we decompose profile P' into a series of single-individual profiles then faithfulness requires that the social evaluation for issue i must be -1 at every one of those profile. If we then recompose P', joining these single-individual profiles together one-by-one, we find that consistency requires $F(P')_i = -1$. Note also that at profile $P + P'$ every element receives the same number of $+1$ evaluations, $y(P)_i$. So Lemma 5 implies that $F(P + P') = \mathbf{0}$. Since we have $F(P + P') = \mathbf{0}$ and $F(P')_i = -1$, consistency implies that $F(P)_i = +1$. □

Theorem 7 *The binary aggregation rule F_M is the only one that is faithful, consistent, cancellative and neutral.*

Proof Let F be a binary aggregation rule that is faithful, consistent, cancellative and neutral. Take any profile P and issue i. Let σ be a permutation of $\{1, \ldots, k\}$ with two orbits $\{1, \ldots, k\} \setminus \{i\}$ and $\{i\}$. We write $P^{(\sigma)}$ to denote a profile such that P and $P^{(\sigma)}$ have the same number of participating individuals and, for all issues i, the number of individuals who say $+1$ to i at profile P is equal to the number of individuals who say $+1$ to $\sigma(i)$ at profile $P^{(\sigma)}$. Importantly, P and $P^{(\sigma)}$ may be disjoint (unlike P and P^{σ}). However, if the binary aggregation rule F is both neutral and strongly anonymous then $F(P^{\sigma}) = F(P^{(\sigma)})$ since $y(P^{\sigma}) = y(P^{(\sigma)})$ and $n(P^{\sigma}) = n(P^{(\sigma)})$. We know, by Lemma 2, that F is strongly anonymous.

We write $P^{(\sigma)2}$ for $(P^{(\sigma)})^{(\sigma)}$, $P^{(\sigma)3}$ for $(P^{(\sigma)2})^{(\sigma)}$ and so on. Let us construct these profiles such that they are all disjoint from one another and from P. Let P' be the profile $P + P^{(\sigma)} + P^{(\sigma)2} + \cdots + P^{(\sigma)(k-2)}$. If k is 2 then P' is P.

Since $\sigma(i) = i$, neutrality and anonymity together imply that $F(\mathbf{.})_i$ is identical over all of the profiles P, $P^{(\sigma)}$, $P^{(\sigma)2}$, ..., $P^{(\sigma)(k-2)}$. If $F(\mathbf{.})_i$ is $+1$ over all of these profiles then consistency implies that $F(P')_i = +1$. So, if $F(P)_i = +1$ then $F(P')_i = +1$. Similarly, consistency implies that if $F(P)_i = -1$ then $F(P')_i = -1$. Lemma 4 implies that if $F(P)_i = 0$ then $F(P')_i = 0$. So $F(P)_i = F(P')_i$.

Let $Y(P)$ denote the sum of the components of $y(P)$. At profile P' element i receives a number of $+1$ evaluations equal to $(k-1) \times y(P)_i$. Every other element receives a number of $+1$ evaluations equal to $Y(P) - y(P)_i$.

Case 1. Assume $y(P)_i > \bar{y}(P)$. Multiplying both sides by k yields $k \times y(P)_i > Y(P)$. Subtracting $y(P)_i$ from both sides yields $(k-1) \times y(P)_i > Y(P) - y(P)_i$. By construction the left hand side equals $y(P')_i$ and the right hand side equals $y(P')_j$ for all issues j in $\{1, \ldots, k\} \setminus \{i\}$. Therefore, $y(P')_i > y(P')_j$ for all $j \in \{1, \ldots, k\} \setminus \{i\}$. Lemma 6 implies that $F(P')_i = +1$. Hence $F(P)_i = +1$.

Case 2. Assume $y(P)_i < \bar{y}(P)$. Simply replace everywhere the relation $>$ with $<$ in the argument for Case 1 to see that $F(P)_i = -1$.

Case 3. Assume $y(P)_i = \bar{y}(P)$. Simply replace everywhere the relation $>$ with $=$ in the argument for Case 1 to see that $y(P')_i = y(P')_j$ for all $j \in \{1, \ldots, k\} \setminus \{i\}$. Lemma 5 implies then that $F(P') = \mathbf{0}$. Hence $F(P)_i = 0$.

So the four properties jointly imply the mean rule. And of course it is readily checked that the mean rule satisfies all four properties. □

We conclude this section by demonstrating the logical independence of the four properties. If there are just two elements in X then any binary aggregation rule that is faithful, cancellative and consistent must also be neutral. The four properties are logically independent if X contains more than two elements, and so we assume that this is the case for the fourth binary aggregation rule listed below.

1. Consider the rule that is the inverse of the mean rule. That is, we collectively assign +1 to an issue if the number of individuals who assign +1 to it is *less* than the mean and if that number is *greater* than the mean then we assign −1. We continue to assign a value of zero to an issue if it receives a number of +1 evaluations equal to the mean. This rule is neutral, consistent and cancellative but not faithful.
2. Consider the following representation of Approval Voting. We collectively assign +1 to an issue only if the number of individuals who assign +1 to it is maximal. We make an exception to this when *all* issues satisfy that condition. In that case we collectively assign zero to every issue, since it is not feasible to assign +1 to every issue. This rule is neutral, faithful and cancellative but not consistent. To see that it is not consistent, suppose that we have a set of three issues and two disjoint profiles P and P' with $y(P) = (3, 2, 0)$ and $y(P') = (0, 2, 3)$. Both electorates collectively assign −1 to the middle issue since the number two is not maximal in either list. Yet, we have $y(P + P') = (3, 4, 3)$ and so the joint electorate assigns +1 to the middle issue.
3. Suppose that we order the individuals in N by seniority. At each profile, the collective evaluation is identical to the evaluation of the most senior individual who is present. Note that in our variable-population model the most senior individual will not be the *same* individual at every profile. This rule is neutral, consistent and faithful but not cancellative.
4. Let us arbitrarily single out one issue in X and label it i. This issue will be treated differently from the others, so that neutrality is violated. The collective evaluation for i is determined in just the same way as it is under the mean rule. That is, at every profile, the value assigned to i by this rule and the value assigned to i by F_M are the same. Each of the other issues is compared to i. We assign +1 to an issue if it has received more +1's from the individuals than i has and we assign −1 to it if it has received fewer +1's than i has. If it has received the same number of +1's as i has then we assign to it the same value that we assigned to i (be that +1, −1 or zero). This binary aggregation rule is faithful, consistent and cancellative but not neutral.

To illustrate the violation of neutrality under rule 4, consider the following example. There are 3 alternatives in X (i, j and k respectively) and 5 individuals. The first 4 people hold the evaluation $(+1, +1, -1)$ while the fifth holds $(+1, -1, +1)$. The outcome is $(+1, -1, -1)$ under rule 4. If we now transpose i and j in each person's evaluation then the social outcome is now $(+1, +1, -1)$ under rule 4. This is a violation of neutrality since i is $+1$ at the new profile, but j was -1 at the original profile.

It is interesting to note that if we allow every element of $\{-1, 0, +1\}^k$ to be a feasible outcome, and do not exclude $\{-1\}^k$ and $\{+1\}^k$ then simple majority rule, applied to each issue individually, satisfies the four properties. Theorem 7 tells us that any adaptation of majority rule to the original problem, in which $\{-1\}^k$ and $\{+1\}^k$ are not feasible, must either yield the mean rule or lose at least one of the four properties. For example, suppose we adapt majority rule as follows: in case all issues are rated -1 by a majority, take the issue with the highest score and change its rating to $+1$ (and proceed analogously, if all issues are rated $+1$). It is not difficult to see that the resulting rule fails consistency.

5 Vector decomposition approach

Axioms are not the only way to characterize aggregation rules. There is a compelling analysis of the Borda rule that uses an *orthogonal decomposition* to separate the part of the profile information that is actually used, from the part discarded, in determining the Borda outcome. Our purpose here is to explain the similar role played by the same decomposition in the mean rule, and also in the *Borda mean rule*, which we introduce here—it is a generalization of the mean rule wherein Borda scores appear explicitly. The parallels provide independent support for our claim that the mean rule is the analogue of the Borda rule for the context of aggregating binary evaluations.

In this section, we either provide brief proof sketches or omit proofs entirely, relying instead on examples and intuitions.[15] We begin by describing the role played by the flow of pairwise net preferences and its decomposition, focusing first on *social welfare functions*—aka *SWF*s—such as Borda or Copeland, which yield a social ranking of alternatives as their outcome, and then shifting back to the mean and Borda mean rules.

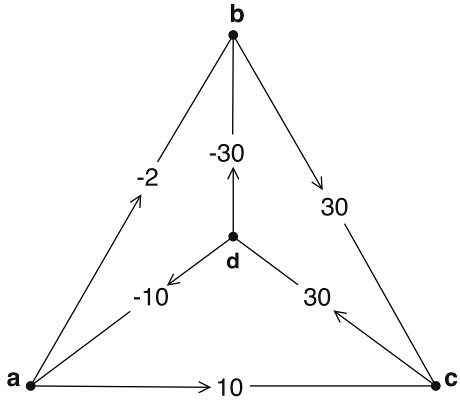
5.1 The flow of net preference

Imagine a network (directed graph), such as that in Fig. 2, consisting of labeled nodes (vertices), each pair of which are connected by a channel (edge), with each channel having both a pre-assigned direction and a real number label indicating the amount of flow (of electricity, or water, or—as will be the case for us—of net preference by the electorate) in that direction.[16] The negative label of -10 on the $d \rightarrow a$ edge of this figure indicates a flow of $+10$ in the *opposite* direction, from a to d; as we see

[15] Most of the lengthy and somewhat tedious proofs do not rely on our application to preference profiles, and can be found in standard sources such as Croom (1978) and Harary (1958).

[16] A *weighted tournament*, in the language of social choice.

Fig. 2 A flow, later identified as the net preference flow $\mathbf{v}_{P_1}$ induced by profile P_1



shortly, one cannot arrange for uniformly non-negative labels by re-orienting edges. We'll refer to such a labeled diagram as a *flow*.

We explain, below, how to decompose such a flow into mutually orthogonal components—a *cyclic* part representing a purely circulatory flow, and a *cocyclic* part representing sinks and sources. The idea originated in the study of the boundary map of algebraic topology in the one-dimensional case. When applied to electric circuit theory this decomposition serves as a mathematical basis for Kirchoff's laws; edges represent wires, nodes stand for connections, and the numerical labels indicate the amount of electric current. Applications of the same decomposition to preference aggregation go back to Zwicker (1991) and the idea was subsequently exploited in Saari (1994), and in his later articles.

For the preference aggregation case, the network's nodes are the alternatives in some finite set X, and the stuff flowing along edges is *net pairwise preference of the voters*, extracted (as given below in Eq. 1) from some profile P of weak orders. Throughout this section we identify a binary evaluation with the corresponding *dichotomous weak order* $X_H > X_L$,[17] and use $P = \{\geq_i\}_{i \in N}$ to denote a profile of unspecified weak orders. Thus, P may equally well represent a profile of linear orders, or a vector of binary evaluations.

To extract the net pairwise preference flow from the profile, we set the numerical label on the edge $x \to y$ equal to:

[17] Consider the set X of alternatives to be evaluated (identified, in Sect. 2, with the index set $\{1, 2, \ldots, k\}$). Recall that a *linear order* on X is a binary relation $\geq$ that is *reflexive, transitive, complete* (or *total*), and is also *anti-symmetric*—satisfying for all $x, y \in X$ that whenever $x \geq y$ and $y \geq x$ both hold, $x = y$. *Weak order* relations relax this last requirement, allowing both $x \geq y$ and $y \geq x$ to hold with $x \neq y$, and writing $x \sim y$ in this situation; $\sim$ is thus an equivalence relation, with the equivalence classes known as I-classes (I for "indifference," arising from the interpretation of $\geq$ as "weakly preferred to"). A weak order can be thought of as a linear order of I-classes, and a *dichotomous weak order* $X_H > X_L$ is one containing exactly two distinct I-classes, which we call X_H and X_L, with $x \geq y$ holding if and only if $x \in X_H$, or $y \in X_L$ (or both). The weak order corresponding to a binary evaluation places each alternative evaluated as $+1$ into X_H and each evaluated -1 into X_L.

Table 4 The profile P_1

No. of individuals	Preferences
14	$a > b > c > d$
10	$b > c > d > a$
6	$b > a > c > d$

$$Net_P(x > y) = [\text{the number of voters } i \text{ with } x \geq_i y] \\ -[\text{the number of voters } i \text{ with } y \geq_i x]. \tag{1}$$

Figure 2 is extracted in this way from P_1, a profile of linear orders (Table 4).

For example, P_1 tells us that while 10 voters prefer d to a, 20 prefer a to d. This is why the $d \to a$ edge of Fig. 2 is labeled with $Net_P(d > a) = 10 - 20 = -10$. When the edge labels arise in this way from a profile, we'll use the term *net preference flow* to refer to the entire diagram (of nodes and labeled, directed edges).

It's worth pointing out that a significant amount of information is already discarded in passing from a profile such as P_1 to the corresponding net preference flow. Aggregation rules that require only the pairwise information retained in the flow include several Condorcet extensions (such as Copeland, Simpson—aka maximin—and Kemeny), one positional scoring rule (Borda), and both of the mean rules considered here; these are "C2" rules, in Fishburn's (1977) classification. Rules in "C3" require some of the lost (non-pairwise) information, and cannot be calculated from the flow; examples include every positional scoring rule other than Borda, as well as all scoring run-offs (such as Hare—aka STV, Coombs, Baldwin, and Nanson).

For example, by referring to Fig. 2 one can quickly identify the outcome for our profile P_1, under any Condorcet consistent SWF, as follows: we observe only positive labels on edges directed *out* of b and only negative labels on edges directed *in* to b. All edges incident to b thus indicate pairwise majorities in b's favor, so b is a Condorcet alternative for P_1. Extending this reasoning, we see that pairwise majority rule yields the transitive ranking $b > a > c > d$.

The standard vector of Borda scoring weights for 4 alternatives is (3, 2, 1, 0), and for m alternatives is $(m - 1, m - 2, \dots, 1, 0)$.[18] The *symmetric* vector of Borda scoring weights for 4 alternatives is $(3, 1, -1, -3)$; and for m alternatives is $(m - 1, m - 3, \dots, -(m - 1))$.[19]

Definition 8 Given a profile P and an alternative x the symmetric Borda score x_P^β is the sum of individual point awards to x (as determined by the above symmetric scoring weights), made by the voters of P.

18 Each *linear* order ballot awards $m - 1$ points to its top-ranked alternative, $m - 2$ to its second-ranked, etc. One separately sums the points awarded to each alternative x by all ballots, and the resulting *Borda scores* determine the outcome ranking of alternatives. We discuss how to score indifferences (arising in weak order ballots) after Lemma 9.

19 *Any* vector of evenly spaced weights $(w, w - d, w - 2d, w - 3d, \dots)$ for $d > 0$ yields the same Borda SWF as the standard vector.

In practice we often drop the P subscript, writing x^β. A side benefit of the symmetric version is that the total point award is 0, so that the average Borda score is 0, independent of the profile. The principal benefit, however, is that x^β may be directly calculated from the net preference flow, as *net flow out of* x:

Lemma 9 *For any profile P of weak orders and alternative x,*

$$x^\beta = \sum_{z \in X} Net_P(x > z).$$

For example, we can calculate $b^\beta_{P_1}$ from Fig. 2 by summing the labels on the three b-incident edges, after reversing the sign for the two edges directed *in* to b: $b^\beta_{P_1} = 2 + 30 + 30 = 62$. The remaining scores are $a^\beta = 18$, $c^\beta = -10$, and $d^\beta = -70$. Thus for profile P_1 the *Borda outcome ranking* happens to agree with the Condorcet outcome: $b > a > c > d$.

Although our example P_1 consists of linear order ballots, it's worth emphasizing that Lemma 9 applies more broadly to weak order ballots, providing one defines x^β appropriately:

Averaging method To assign (symmetric)[20] Borda scoring weights based on a weak order ballot:

- Choose an arbitrary linear extension of the ballot.
- Award, to each alternative in an I-class of the weak order ballot (see fn.17), the *average* (symmetric) scoring weight that would be awarded to the members of that class by the extension.

Within the class C2, the Borda rule is distinguished from Condorcet extensions in that it discards additional information retained by the flow of net preference—the information contained in the cyclic component. In fact, we will see that this feature essentially characterizes the Borda rule among SWFs, and characterizes the mean rule and Borda mean rule among aggregators of their respective types. Justifying these claims requires a more exact understanding of the decomposition.

5.2 The decomposition into cycles and cocycles

A *basic cycle* is a directed unit flow around a simple closed loop. Figure 3 shows two basic cycles. Our convention is that when the flow along an edge is zero (as it is for each edge off the cycle), the edge is dotted and we omit the numerical label. Note also that the edge label is -1 when the direction of an edge on the loop is opposed to the direction in which the loop is taken. A *cycle* is any linear combination of basic cycles.[21] Thus a flow **v** is a cycle if and only if we can decompose **v** as a sum of scalar

[20] There are arguments for using this averaging method with *any* vector of Borda weights (see fn. 19) but for Lemma 9 we require the symmetric version.

[21] Suppose we fix the m nodes, the $j = \frac{(m)(m-1)}{2}$ edges, and the edge directions of a network. Then any single labeling of edges (with real numbers) corresponds to a vector of length j, and the space of all such labeling is thus identified with the Euclidean vector space $\Re^j$. The sum $\mathbf{v}+\mathbf{w}$ (of two edge labelings) assigns

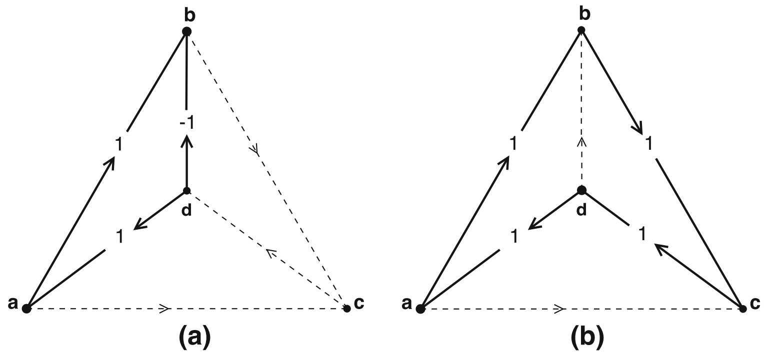


Fig. 3 Two basic cycles

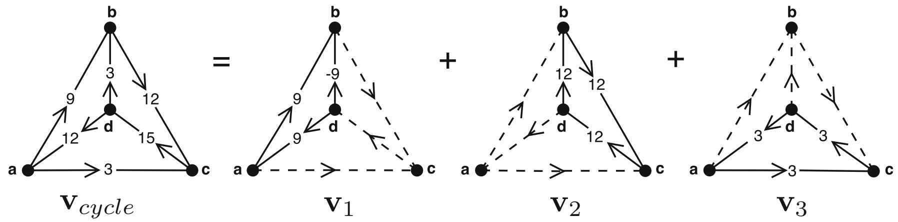


Fig. 4 Expressing $\mathbf{v}_{cycle}$ as a linear combination of basic cycles

multiples of basic cycles.[22] To an electrical engineer, a cycle is a "superposition of loop currents"—see Fig. 4 for an example.

For any alternative x the *basic cocycle for* x is a flow that labels each $x \rightarrow y$ edge with $+1$, each $x \leftarrow y$ edge with -1, and each edge not incident to x with 0. Figure 5 shows the basic cocycle for alternative b. A *cocycle* is any linear combination of basic cocycles.[23] Thus a flow $\mathbf{v}$ is a cocycle if and only if we can decompose $\mathbf{v}$ as a sum of scalar multiples of basic cocycles.[24] To an electrical engineer, a cocycle is a "superposition of sinks and sources."

Theorem 10 *Let* $\mathbf{v}$ *be any flow. Then* $\mathbf{v}$ *has a unique decomposition*

$$\mathbf{v} = \mathbf{v}_{cycle} + \mathbf{v}_{cocycle}$$

Footnote 21 continued
to each edge the sum of the labels separately assigned to that edge by $\mathbf{v}$ and $\mathbf{w}$, while scalar multiplication of $\mathbf{v}$ by s simply multiplies each of $\mathbf{v}$'s edge labels by s. The set of cycles, then, is the linear span of the basic cycles, and forms the *cycle subspace* $\mathbf{V}_{cycle}$—a linear subspace of $\Re^j$.

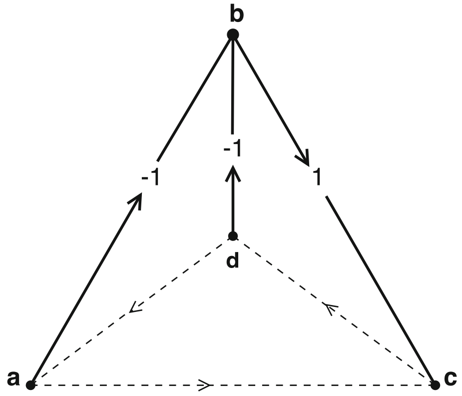
The inner product $\mathbf{v} \cdot \mathbf{w}$ of two edge labelings is found by multiplying all pairs of corresponding edge labels, and then adding these products, and provides a notion of orthogonality.

22 This decomposition is not unique; the basic cycles are not linearly independent.

23 The cocycle subspace $\mathbf{V}_{cocycle}$ is the linear span of the basic cocycles in $\Re^j$; see fn. 21.

24 Once again the decomposition is not unique.

Fig. 5 A basic cocycle



as the sum of a cycle and a cocycle. Moreover, $\mathbf{v}_{cycle} \perp \mathbf{v}_{cocycle}$.[25]

We'll wave our hands at the proof shortly. First, let's look at the decomposition for our running example, $\mathbf{v}_{P_1}$ (see Fig. 6).

Proof sketch for Theorem 10 It is easy to see why the inner product of a basic cycle with a basic cocycle must be zero. (Try the examples from Figs. 3 and 5, and consider as well what happens for a basic cycle that does not pass through node b.) As inner product is a bilinear operator (it distributes over addition and respects scalar multiplication), it follows that every cycle is orthogonal to every cocycle. Theorem 10 then follows immediately from basic linear algebra, after showing that the dimensions of the cycle and cocycle subspaces (see fn. 21, 23) add to *at least* the dimension $j = \frac{(m)(m-1)}{2}$ of the entire space.[26] □

5.3 Borda scores and the cocyclic component

Observation 11 *Borda scores have the following relationship to the decomposition of the net preference flow:*

1. *Borda scores depend only on the* $\mathbf{v}_{cocycle}$ *information. Thus the Borda SWF uses only the profile information in the cocyclic component of the net preference flow.*

[25] Thus subspaces $\mathbf{V}_{cycle}$ and $\mathbf{V}_{cocycle}$ are *orthogonal complements* in $\Re^j$ and for each $\mathbf{v} \in \Re^j$, the components $\mathbf{v}_{cycle}$, $\mathbf{v}_{coycle}$ may be obtained via *orthogonal projection* of $\mathbf{v}$ onto the corresponding subspaces.

[26] The last point can be established by induction on m: one assumes existence, for m alternatives, of a set S_{cycle} of $\frac{(m-1)(m-2)}{2}$ linearly independent length-3 cycles and a set $S_{cocycle}$ of $m-1$ linearly independent cocycles. Note that these cardinalities add to $\frac{(m)(m-1)}{2}$. One then shows that with the addition of one new alternative one can construct at least one more cocycle, which maintains linear independence of $S_{cocycle}$ when added in, expanding $|S_{cocycle}|$ to m, and can construct at least $m-1$ more length-3 cycles, which maintain linear independence of S_{cycle} when added in, expanding $|S_{cycle}|$ to $\frac{(m-1)(m-2)}{2} + (m-1) = \frac{(m)(m-1)}{2}$, as desired.

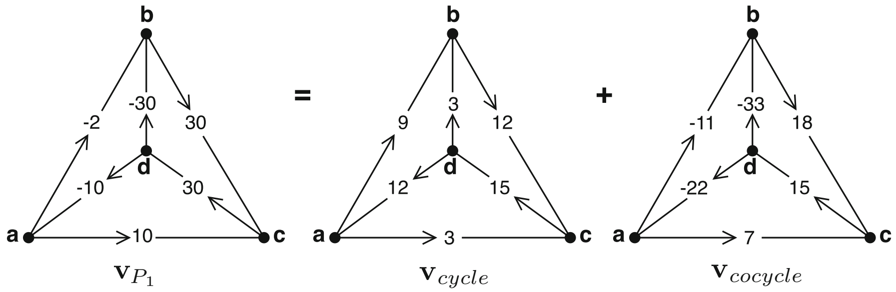


Fig. 6 The decomposition $\mathbf{v}_{P_1} = \mathbf{v}_{cycle} + \mathbf{v}_{cocycle}$

2. *In fact, the vector of Borda scores and $\mathbf{v}_{cocycle}$ are essentially the same thing. Note in Fig. 6 that the edge label on the $b \to c$ edge of $\mathbf{v}_{cocycle}$ is a scaled version $\frac{b^\beta - c^\beta}{4}$ of the difference in their symmetric Borda scores. More generally, $\frac{x^\beta - z^\beta}{m}$ yields the edge labels for the cocyclic component for m alternatives. Conversely, from these edge labels, we can recover the symmetric Borda scores of the alternatives.*[27]
3. *Suppose we treat the cocyclic component $\mathbf{v}_{cocycle}$ as if it were actually the entire flow of pairwise net preference (ignoring that $\mathbf{v}_{cycle}$ has been discarded), and use it to extract a version of the pairwise majority relation R (by setting $x R z \Leftrightarrow$ the $x \to z$ edge has a nonnegative label in $\mathbf{v}_{cocycle}$, or the $z \to x$ edge has a nonpositive label). Then the result is always a transitive relation, and is identical to the order induced by Borda scores. Thus, using the Borda rule is tantamount to applying pairwise majority rule after suppressing $\mathbf{v}_{cycle}$.*
4. *The components $\mathbf{v}_{cycle}$ and $\mathbf{v}_{cocycle}$ have opposing implications for transitivity of the majority preference relation. If $\mathbf{v}_{cocycle}$ is large enough (relative to $\mathbf{v}_{cycle}$) then $\mathbf{v}_{cocycle}$ dominates, so that the Borda and Condorcet outcomes are identical (as was the case for our sample profile P_1).*[28] *If $\mathbf{v}_{cocycle}$ is too small, so that $\mathbf{v}_{cycle}$ dominates, then the majority preference relation is intransitive—a Condorcet cycle exists. In an intermediate situation, the majority preference relation that would result from $\mathbf{v}_{cocycle}$ alone changes when one adds $\mathbf{v}_{cycle}$ back in, but without being thrown into a cycle.*

Note that a basic cycle contributes zero to the symmetric Borda score of each alternative, whence Observation 11.1 follows immediately. Before turning to the implications of the decomposition for the mean rule, we illustrate Observation 11.4 with an example, by adding multiples of a 4-voter Condorcet cycle to our initial profile P_1 (Table 5).

It is easy to see that $\mathbf{v}_Q = 2\mathbf{v}_{3b}$; that is, to get $\mathbf{v}_Q$, just double all the edge labels in the basic cycle of Fig. 3b. Now consider the profile $P + rQ$, obtained by adding r

[27] With non-symmetric vectors of scoring weights, one cannot fully recover the scores from their scaled differences (corresponding to $\mathbf{v}_{cocycle}$ edge labels), as the vector of scores also encodes the number n of voters, while from $\mathbf{v}_{cocycle}$ it is not possible to recover n.

[28] We are using "large enough" as a loose metaphor here—the precise version (Zwicker 1991) does not use a single numerical measure of size.

Table 5 The profile Q

No. of individuals	Preferences
1	$a > b > c > d$
1	$b > c > d > a$
1	$c > d > a > b$
1	$d > a > b > c$

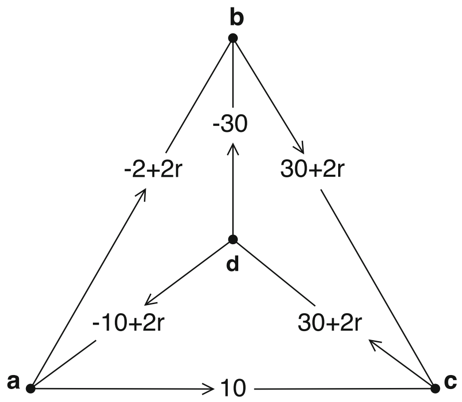


Fig. 7 The net preference flow for $\mathbf{v}_{P_1+rQ}$

disjoint groups of voters (each with 4 voters, and each having Q as its profile) to the 30-voter electorate for P.

What is the effect of gradually increasing r on the Borda outcome ... or on the pairwise-majority relation? As the cocyclic component of Q is $\mathbf{0}$, Q contributes 0 to the symmetric Borda score of each alternative, and the effect on the Borda outcome is nil; regardless of the r value, the Borda outcome is $b > a > c > d$. The situation for the pairwise majority outcome is quite different, however (see Fig. 7).

With $r = 0$, the two outcomes are the same. With $r = 2$ the sign of the edge label on $a \to b$ switches from $-$ to $+$, and the pairwise majority relation (Condorcet outcome) changes from $b > a > c > d$ to $a > b > c > d$, reflecting the intermediate situation described in Observation 11.4; any Condorcet extension (Black's rule, for example) now declares a to be the winner. If we let r grow to the point where $r \geq 6$ then the sign of the edge label on $d \to a$ also becomes $+$, and the pairwise majority relation becomes intransitive; at this point Black's rule switches back to declaring b as the winner—a failure of *reinforcement* (aka "consistency").

5.4 The Borda mean rule, the mean rule, and the decomposition

We now show that the mean rule is like the Borda count SWF, in that it uses only the profile information in the cocyclic component of the net preference flow. Our first argument shows that the Borda mean rule—a natural generalization of the mean rule, introduced below—uses only the information in the symmetric Borda scores. A second, independent argument notes that when a profile of preference rankings consists entirely of dichotomous weak orders, the resulting net preference flow satisfies $\mathbf{v}_{cycle} = \mathbf{0}$.[29]

Definition 12 The Borda mean rule takes as input a profile P of weak orders on X. Alternatives x having positive symmetric Borda score $x_P^\beta > 0$ are placed in X_H, those with negative score $x_P^\beta < 0$ are placed in X_L, and the output of the rule is the corresponding dichotomous weak order $X_H > X_L$. Alternatives with $x_P^\beta = 0$ are treated as described in Sect. 3 (just before Theorem 1); they yield a tie among all dichotomous weak orders that can be obtained by assigning these elements randomly and independently to X_H or to X_L.[30]

Proposition 13 *When restricted to profiles of dichotomous weak orders, the Borda mean rule is identical to the mean rule.*

Proof Recall that

- $y(P)_x$ is the number of individuals j who evaluate x as $+1$—equivalently, the number who place x in the higher I-class X_H of their dichotomous ranking $\geq_j$
- $\overline{y}(P) = \frac{\sum_{z \in X} y(P)_z}{k}$ (where k is the number of elements in X)
- The mean rule places x in X_H when $y(P)_x > \overline{y}(P)$.

Thus, it suffices to establish

$$x_P^\beta > 0 \Leftrightarrow y(P)_x > \overline{y}(P). \tag{2}$$

For a profile P of dichotomous weak orders,

$$Net_P(x > z) = y(P)_x - y(P)_z. \tag{3}$$

Hence

$$\begin{aligned} x_P^\beta &= \sum_{z \in X} Net_P(x > z) = \sum_{z \in X} \left[y(P)_x - y(P)_z \right] \\ &= \sum_{z \in X} [y(P)_x] - \sum_{z \in X} \left[y(P)_z \right] = k \left[y(P)_x - \overline{y}(P) \right], \end{aligned}$$

from which condition (2) follows immediately. □

Now that we know that both the mean rule and Borda mean rule are determined by the symmetric Borda scores, which in turn depend only on $\mathbf{v}_{cocycle}$ (Observation 11), we immediately obtain:

[29] In a sense, then, it is thus doubly true that the mean rule depends only on $\mathbf{v}_{cocycle}$. *First*, the Borda mean rule discards $\mathbf{v}_{cycle}$, using only $\mathbf{v}_{cocycle}$. *Second*, for the mean rule $\mathbf{v}_{cycle}$ need not even be discarded. We say "in a sense" because the second argument is subtly incomplete—it fails to rule out the possibility that the mean rule relies on profile information not present in the net preference flow. In fact, the vector $\{y(P)_x\}_{x \in X}$ used as input for the mean rule incorporates some information about the total number of voters wiped out (by taking differences, in Eq. 3 below) when it is converted to net preferences.

[30] To keep things short, we've suppressed all mention of ties in the remainder of this section. The modifications needed to take them into account are straightforward.

Corollary 14 *Both the mean rule and the Borda mean rule use only the profile information in the cocyclic component of the net preference flow.*

Our promised second argument consists of the following:

Proposition 15 *Let P be any profile of dichotomous weak orders, and $\mathbf{v}$ be the net preference flow induced by P. Then $\mathbf{v}_{cycle} = \mathbf{0}$.*

Proof First we show that $\mathbf{v} \cdot \mathbf{c} = 0$ for every basic cycle $\mathbf{c}$. Let

$$\mathbf{c} = x_1 \to x_2 \to x_3 \to \cdots \to x_t \to x_1$$

be a basic cycle, for which we assume, without loss of generality, that edges on $\mathbf{c}$ agree in direction with the cycle's orientation and we identify x_{t+1} with x_1. Thus $\mathbf{c}$ assigns weight 1 to each edge $x_r \to x_{r+1}$ on $\mathbf{c}$ and assigns weight 0 to each edge off $\mathbf{c}$, while Eq. (3) tells us that $\mathbf{v}$ assigns weight $\left(y(P)_{x_r} - y(P)_{x_{r+1}}\right)$ to each $x_r \to x_{r+1}$ edge. Thus the value of $\mathbf{v} \cdot \mathbf{c}$ is

$$\left(y(P)_{x_1} - y(P)_{x_2}\right)(1) + \left(y(P)_{x_2} - y(P)_{x_3}\right)(1) + \cdots + \left(y(P)_{x_t} - y(P)_{x_1}\right)(1) = 0.$$

It follows that $\mathbf{v}$ is orthogonal to the subspace $\mathbf{V}_{cycle}$ spanned by the basic cycles, whence $\mathbf{v}$ belongs to its orthogonal complement $\mathbf{V}_{cocycle}$ (see fn. 25), with $\mathbf{v} = \mathbf{v}_{cocycle}$ and $\mathbf{v}_{cycle} = \mathbf{0}$. □

The same idea may be expressed via the following *quantitative transitivity* condition:

$$Net_P(x > z) + Net_P(z > w) = Net_P(x > w), \tag{4}$$

which is strictly stronger than transitivity of P's pairwise majority relation, and equivalent to $\mathbf{v}^P_{cycle} = \mathbf{0}$. It is satisfied by any profile of dichotomous weak orders, because

$$(y(P)_x - y(P)_z) + (y(P)_z - y(P)_w) = (y(P)_x - y(P)_w)\,.$$

At this point, we can summarize our vector theory argument as follows: because they use the same restricted profile information—$\mathbf{v}_{cocycle}$, or equivalently the symmetric Borda scores—the mean rule is the analogue of the Borda SWF. The argument would be stronger if we could justify the analogy:

The Borda SWF: SWFs :: The mean rule: ____s

by showing that the way the Borda rule exploits the common information in the social welfare function context is analogous to the way that the mean rule (or Borda mean rule) exploits it in the ____ context.[31] That would require filling the blanks—identifying a larger context exemplified by the mean and Borda mean rules. This matter is one we will take up in a planned sequel, "Social Dichotomy Functions."

[31] Perhaps by showing each exploits that information in the *only* reasonable way for its context.

5.5 Applying vector methods more broadly to Judgment Aggregation

Observation 11.3 can be restated as follows: *By projecting the agents' preferences onto the subspace* $\mathbf{V}_{cocycle}$ *we wipe out all underlying tendency towards intransitivity, obtaining a modified version of net preference flow that satisfies a strong, quantitative form of transitivity. The outcome is then extracted from this "strongly transitive" version by treating it as if it were the original.*

A central problem in the field of judgment aggregation arises when each of several agents submits a judgment on a sequence of interrelated binary issues. It can happen that the result of applying majority rule separately on an issue-by-issue basis gives rise to an infeasible collective judgment on those issues—one inconsistent with the predetermined interrelationships—even though the sequence of judgments submitted by any single agent is feasible.

Might it be possible to proceed in parallel with Observation 11.3, after substituting some other form of *feasibility* for *transitivity*? Given a judgment aggregation context, one would identify some subspace $\mathbf{V}_{feasible}$ of strongly feasible judgments, each of which has zero tendency towards infeasibility, then project individual judgments onto that subspace, to obtain a modified version of the issue-by-issue net majorities, which would safely lead to a feasible collective judgment.

Preliminary investigation suggests that the effectiveness of such an approach varies depending on the particular judgment aggregation context. In some cases there seem to be no acceptable choices for $\mathbf{V}_{feasible}$. Suppressing all tendency to infeasibility comes at an unacceptably high cost for these cases, limiting (to an unreasonable degree) the range of possible collective judgments. It would be quite interesting to characterize those judgment aggregation contexts for which such a vector decomposition method does yield useful forms of aggregation, but that seems to require a finer understanding of "useful form of aggregation" than we now possess.

6 Judgment aggregation

There has been a great deal of interest in the topic of judgment aggregation since List and Pettit's (2002) discussion of the now-famous doctrinal paradox. In this section we consider ways in which the mean rule may be relevant to this topic. We begin by defining some terminology from the literature on judgment aggregation.

An *agenda* is a finite set of propositions that is closed under negation: for every proposition in the agenda, the negation of that proposition is also contained in the agenda. We denote the negation of a proposition p by $\neg p$. An example of an agenda is the set $\{p, \neg p, q, \neg q, p \wedge q, \neg(p \wedge q)\}$. For brevity, we also write this set as $\{p, q, p \wedge q\}^{\pm}$, listing just one proposition from each proposition/negation pair.

A *judgment set* is a subset of the agenda. A *minimally complete* judgment set contains one and only one of every proposition/negation pair in the agenda. For short, we will simply write *complete* instead of minimally complete. A *rational* judgment set is complete and also respects any logical interconnections between the propositions in the agenda. Using our earlier agenda, $\{p, q, p \wedge q\}$ is a rational judgment set whereas

$\{p, q, \neg(p \wedge q)\}$ is not. An incomplete judgment set is *consistent* if and only if it is a subset of a rational judgment set.

We seek to aggregate individual rational judgment sets into one or more collective judgments sets that are also rational. Can the mean rule be used for this purpose?

We know that the mean rule is suitable for at least some agendas. This is because we can describe our original problem of bipartitioning a group of students as a problem of judgment aggregation. We have an agenda $\{p_1, p_2, \ldots, p_k\}^{\pm}$, where p_i is the proposition "student i belongs in the higher-ability class". Given the requirement that every student is assigned either to the higher-ability class or the lower-ability class, $\neg p_i$ means "student i belongs in the lower-ability class". In this example, every complete judgment set is rational except for $\{p_1, p_2, \ldots, p_k\}$ and its inverse, $\{\neg p_1, \neg p_2, \ldots, \neg p_k\}$.

Applying the mean rule here is simple. We take the set $\{p_1, p_2, \ldots, p_k\}$, let us call this the *reference set*, and count the number of individuals who accept each of the propositions in that set. The omission of $\pm$ here is deliberate; this reference set is just one half of the agenda. Then we calculate the mean of those k numbers. For each proposition p in the reference set, if p is accepted by a number greater than the mean then we collectively accept p, and if p is accepted by a number less than the mean then we accept $\neg p$. Note that we arrive at the same outcome if we instead use $\{\neg p_1, \neg p_2, \ldots, \neg p_k\}$ as the reference set, but it is important that we do not calculate the mean over all $2k$ propositions in the agenda, which would always be equal to half of the number of individuals who have submitted a judgment set.

The mean rule will always generate a rational outcome when (i) there are just two complete judgment sets that are irrational and (ii) those two judgment sets are disjoint. We simply choose either one of those two irrational, complete judgment sets as the reference set and apply the mean rule to the propositions in that reference set to determine the outcome.

In Sect. 6.1 we show that the problem of partitioning a group of students into higher-ability and lower-ability classes is not the only problem that takes this form. The problem of aggregating linear orders over three alternatives does too. Moreover, we show that applying the mean rule to that problem reveals yet another connection between the mean rule and the Borda rule, in keeping with the theme of this paper.

The mean rule is not always suitable for agendas that take a different form. For instance, suppose that the agenda is $\{p, q, r\}^{\pm}$ and that the only complete judgment set that is irrational is $\{p, q, r\}$. In this case if we use $\{p, q, r\}$ as the reference set and apply the mean rule then it is impossible for $\{\neg p, \neg q, \neg r\}$ to be the best outcome according to the mean rule even though that judgment set is rational in this case. It seems undesirable that a rational outcome should be precluded in this way.

Suppose instead that there are two complete judgment sets that are irrational but that they are not disjoint. Let us say that $\{p, q, \neg r\}$ and $\{p, \neg q, r\}$ are those two irrational judgment sets. If we use $\{p, q, \neg r\}$ as the reference set then it is possible that the outcome under the mean rule will be the irrational set $\{p, \neg q, r\}$, and vice versa. This would happen if, for example, there were just two individuals in the group and one submits the judgment set $\{p, q, r\}$ and the other submits $\{p, \neg q, \neg r\}$.

In Sect. 6.2 we respond to the limited applicability of the mean rule by proposing a "general purpose" method of judgment aggregation that is an extension of both the mean rule and the Borda rule. This proposal is based on an approach to judgment aggregation introduced by Dietrich (2014).

6.1 Preference agenda with three alternatives

In this section we consider the problem of aggregating linear orders over three alternatives. Let us label the three alternatives a, b and c. Let ab denote the proposition "a is better than b". Since we are restricting our attention to linear orders, the negation $\neg ab$ is equivalent to ba. The agenda for this problem is $\{ab, bc, ac\}^{\pm}$. This is an example of a "preference agenda" (Dietrich and List 2007). Every complete judgment set is rational except for two: $\{ab, bc, ca\}$ and $\{ba, cb, ac\}$. These two judgment sets are disjoint, so we can apply the mean rule to this problem.

Consider the following example. There are five individuals in the group. Three individuals submit the judgment set $\{ab, bc, ac\}$ and two submit the set $\{ba, bc, ca\}$. To apply the mean rule we must use either $\{ab, bc, ca\}$ or $\{ba, cb, ac\}$ as the reference set. Let us use the set $\{ab, bc, ca\}$ and, thusly, count the number of individuals who accept each of the three propositions in that set. These numbers are three, five and two, respectively. So, according to the mean rule, we collectively accept bc and we reject ab and ca. In other words, the collective judgment set is $\{ba, bc, ac\}$. Note that this outcome is anti-majoritarian; a majority of the individuals believe that a is better than b, and the majoritarian outcome, $\{ab, bc, ac\}$, is perfectly rational.

Significantly, the Borda rule generates exactly the same anti-majoritarian outcome at this profile as the mean rule. Since b has a Borda score of seven, a six and c two, the Borda rule generates the ranking $b > a > c$, which corresponds to $\{ba, bc, ac\}$, the outcome under the mean rule.

However, the relationship between the two rules is deeper than simply agreeing at this profile. It turns out that when there are three alternatives, the Borda rule and the mean rule will always agree.

Here is the proof. Let $B(a)$ and $B(b)$ denote the Borda scores of a and b respectively. Furthermore, let $n(ab)$ be the number of individuals who accept ab and let μ be the mean of $n(ab)$, $n(bc)$ and $n(ca)$. So, under the Borda rule we accept ab if and only if $B(a) > B(b)$, while under the mean rule we accept ab if and only if $n(ab) > \mu$. It is convenient for the following argument to say that under the mean rule we accept ab if and only if $3n(ab) > 3\mu$.

If an individual accepts, say, ab and bc then that means that she ranks a just one place above b. Hence her ranking causes $B(a) - B(b)$ to increase by one. Her ranking also causes $3n(ab)$ to increase by three and, since she accepts a total of two propositions from the reference set, it causes 3μ to increase by two. So $3n(a, b) - 3\mu$ increases by one. Suppose instead that she accepts bc and ca. This means that $B(a) - B(b)$ decreases by two. In this case she rejects ab so $3n(ab)$ is unchanged, and, since she accepts two propositions from the reference set, 3μ increases by two. So $3n(ab) - 3\mu$ decreases by two.

It is easy to check that every ranking that is submitted causes the same change in $B(a) - B(b)$ as it does in $3n(ab) - 3\mu$. Since, loosely speaking, both of the numbers $B(a) - B(b)$ and $3n(ab) - 3\mu$ start at zero before the rankings are submitted, we can see that $B(a) > B(b)$ if and only if $3n(ab) > 3\mu$. Therefore, we accept ab under the Borda rule if and only if we accept it under the mean rule.

It is important to note, however, that this equivalence between the Borda rule and the mean rule does not hold when there are more than three alternatives. When there are four alternatives we will find that there are more than two complete judgment sets that are irrational. The mean rule then ceases to be applicable. Nevertheless, in this section we have seen that we can make a direct comparison between the mean rule and the Borda rule, albeit one that is limited to the case of three alternatives.

6.2 Other agendas and scoring functions

One weakness of the mean rule is that there are many judgment aggregation problems to which it is not applicable. Can this deficiency be overcome? In fact, suppose that we want to go even further than this and achieve the following in judgment aggregation: (i) be applicable to many agendas of interest, (ii) produce the same collective judgment as the mean rule when applied on its domain, and (iii) produce the same collective judgment as the Borda rule when applied to the preference agenda (with *any* number of alternatives). We show in this section that these objectives can be realized by building on an approach to judgment aggregation introduced by Dietrich (2014).

Let a *scoring function* be a function $s_J(p)$ that assigns a number (or *score*) to issue p given judgment set J. A scoring function can be used as the basis for a judgment aggregation rule in the following way. We first determine a total score for each proposition p by summing the scores for p over every individual's judgment set. We can then rank all of the possible outcomes, that is, all of the rational judgment sets, by the sum-total of scores over the propositions that they each contain. The one with the greatest sum of scores is the final outcome.

Dietrich proposes a number of scoring functions. These include "reversal scoring", "disjoint-entailment scoring" and "irreducible-entailment scoring". A feature of these three scoring functions that makes them particularly interesting is that they coincide with the Borda rule when applied on its domain (the preference agendas). Unfortunately, they do not also coincide with the mean rule over its domain (where there are just two complete judgment sets that are irrational and they are disjoint). Interestingly, however, we can construct a relatively simple scoring function that does agree with both the Borda rule and the mean rule over their respective domains.

Under Dietrich's reversal scoring, given a judgment set J and a proposition p in J, we determine a score for p as follows. Starting from J, we count the minimal number of reversals of propositions (replacing a proposition with its negation) required to obtain a logically consistent judgment set that contains $\neg p$. This number is the score for p. As Dietrich suggests, one may think of this score as a measure of the *strength* of the individual's judgment. Under this approach, we count not just the reversal of p itself, but also the reversals of other propositions required to "allow" p to be reversed.

Now suppose that instead of making reversals in $J - \{p\}$ that are *consistent* with the reversal of p, we make reversals in $J - \{p\}$ that *entail* the reversal of p. This variation on Dietrich's idea yields a scoring function that is consistent with both the Borda rule

Table 6 Infeasibility scoring

	Score of							
	p_1	$\neg p_1$	p_2	$\neg p_2$	p_3	$\neg p_3$	p_4	$\neg p_4$
Indiv. 1 ($p_1 \neg p_2 \neg p_3 \neg p_4$)	3	0	0	1	0	1	0	1
Indiv. 2 ($\neg p_1 p_2 \neg p_3 \neg p_4$)	0	1	3	0	0	1	0	1
Indiv. 3 ($\neg p_1 \neg p_2 p_3 \neg p_4$)	0	1	0	1	3	0	0	1
Group	3	2	3	2	3	2	0	3

and the mean rule. This distinction between our scoring function and Dietrich's will become clearer later when we present an example.

Let us explain our proposal more precisely. Given any rational judgment set J and a proposition p, if p is in J then the score for p is equal to the smallest number of judgment reversals needed to obtain from J a judgment set J^* such that

1. J^* is irrational, and
2. $J^* \setminus \{p\}$ is consistent.

By "judgment reversal" we mean the replacing of a proposition by its negation. If p is not in J then the score for p is zero. We call this scoring function *infeasibility scoring*. To explore the relationship between infeasibility scoring and the mean rule, consider Table 6.

The agenda in this example is $\{p_1, p_2, p_3, p_4\}^{\pm}$ and there are just two complete judgment sets that are irrational: $\{p_1, p_2, p_3, p_4\}$ and $\{\neg p_1, \neg p_2, \neg p_3, \neg p_4\}$.

To explain this table of scores let us take the example of proposition p_1 receiving a score of three for judgment set $\{p_1, \neg p_2, \neg p_3, \neg p_4\}$. Starting from that judgment set, the only judgment set that we can obtain by making judgment reversals and that satisfies conditions 1 and 2 in the definition of infeasibility scoring is the set $\{p_1, p_2, p_3, p_4\}$. This set satisfies condition 1 because it is irrational, and it satisfies condition 2 because removing the proposition that we are scoring, p_1, leaves the set $\{p_2, p_3, p_4\}$ which is a subset of a rational judgment set and therefore is consistent. To reach $\{p_1, p_2, p_3, p_4\}$ from $\{p_1, \neg p_2, \neg p_3, \neg p_4\}$ requires three judgment reversals. Hence the score of three for p_1. Under reversal scoring, reversing the judgment on p_1 and $\neg p_2$ at the original judgment set brings us to $\{\neg p_1, p_2, \neg p_3, \neg p_4\}$ which is a rational judgment set. The score for p_1 according to Dietrich's reversal scoring is, therefore, two.

The rational judgment set with the greatest sum of scores is the set $\{p_1, p_2, p_3, \neg p_4\}$, with a sum-total of twelve, and so it is the outcome of the aggregation process. The mean rule generates the same outcome at this profile. This is a consequence of the following theorem.

Proposition 16 *Over the domain of the mean rule, each proposition p is accepted under the mean rule if and only if p has a greater score than $\neg p$ under infeasibility scoring.*

Proof Take an agenda in the domain of the mean rule. Let R and $\bar{R}$ be the two complete judgment sets that are irrational and are disjoint. Take any proposition p in R.

For any rational judgment set J that contains p, the infeasibility score for p is equal to the number of propositions that are in R but not in J. This is the number of judgment reversals needed to obtain R, and it is only by obtaining R that we satisfy conditions 1 and 2 in the definition of infeasibility scoring. So the score for p, in other words, is the cardinality of $R \setminus J$.

Now suppose that we represent judgment set J as a dichotomous weak order on R. A dichotomous weak order, that is, with upper equivalence class $J \cap R$ and lower equivalence class $R \setminus J$. Then the symmetric Borda score for p is the cardinality of $R \setminus J$ (the number of elements of R ranked below p). So we see that the infeasibility score for p is equal to the symmetric Borda score for p.

For any judgment set J' that contains $\neg p$, the infeasibility score for $\neg p$ is the cardinality of $J' \cap R$. This is because we must obtain $\bar{R}$ to satisfy conditions 1 and 2 in the definition of infeasibility scoring, so all of the judgments in $J' \cap R$ must be reversed.

When we represent J' as a dichotomous weak order on R, with upper equivalence class $J' \cap R$ and lower equivalence class $R \setminus J'$, we find that the symmetric Borda score for p is zero *minus* the cardinality of $J' \cap R$ (since p is now in the lower equivalence class).

So if the infeasibility score for $\neg p$ is x then the symmetric Borda score for p is $-x$. Note that there is no Borda score for $\neg p$ because we have defined the relation over R.

We have seen that the total infeasibility score for p will be greater than the total infeasibility score for $\neg p$ if and only if the symmetric Borda score for p is positive. We know, from the proof of Proposition 13, that the symmetric Borda score for p is positive if and only if p is accepted under the mean rule. □

We now consider how infeasibility scoring works over the domain of the Borda rule. Suppose there are four alternatives a, b, c and d, and so we have the agenda $\{ab, ac, ad, bc, bd, cd\}^{\pm}$, with $\neg ab$ being equivalent to ba as discussed in Sect. 6.1.

The judgment set that corresponds to the ordering $a > b > c > d$ is $\{ab, ac, ad, bc, bd, cd\}$. What does infeasibility scoring assign to, say, ad for this judgment set? With three judgment reversals we can reach the judgment set $\{ba, ca, ad, bc, bd, dc\}$. This judgment set is irrational (with $b > d > c > a > d$), but if we remove ad then we have a consistent subset (consistent with the ordering $b > d > c > a$). So infeasibility scoring yields a score of three for ad here. Notice that this is just the conventional Borda score of a minus the Borda score of d (derived from the original ordering).

This is an example of a more general equivalence. Infeasibility scoring matches Borda scoring in the sense that, under infeasibility scoring, for any two alternatives a and b the score for proposition ab is equal to the Borda score of a minus the Borda score of b.

Proposition 17 *Infeasibility scoring matches Borda scoring in the case of the preference agenda.*

Proof Take any two alternatives a and b. Let us first consider the case of an individual who accepts ab and does not rank any other alternatives between a and b. Clearly, the Borda score of a minus the Borda score of b in this case is one.

Suppose that alternative c is ranked just below b so we have ac and bc. If we reverse the judgment ac then we have ca and bc. These two judgments are incompatible with ab as they imply ba. Since just one reversal is required to obtain a binary relation in which the reversal of ab would restore transitivity, the score assigned by infeasibility scoring to ab is one. We could also have supposed that c is ranked just above a so that we have ca and cb. In that case, reversing the judgment cb gives us the same outcome; ca and bc.

So infeasibility scoring and Borda scoring are equivalent when a and b are adjacent in the ordering, with both giving a score of one to ab. But we must confirm that they are equivalent when there are some alternatives ranked between a and b. Suppose that there are t-many alternatives ranked between a and b. Then the Borda score of a minus the Borda score of b is $t + 1$. We need to prove that infeasibility scoring assigns $t + 1$ to ab.

Let d be one of the alternatives between a and b. That is, ad and db. If both of these judgments remain unchanged then transitivity requires ab. So reversing the judgment ab will not result in a transitive relation. Therefore, we must reverse either ad or db. Let us reverse the proposition ad. And let us also reverse ae for every other proposition e that is ranked between a and b. In total this will require t reversals. Once a has been demoted t times in this way we will have a transitive relation in which a is just above b. We know, by the argument above, that exactly one more reversal is needed to obtain a relation such that (i) the relation is intransitive and (ii) reversing the judgment ab would restore transitivity. So the total number of reversals required to satisfy those two conditions in the definition of infeasibility scoring is $t + 1$. Hence, infeasibility scoring assigns $t + 1$ to ab. □

We should emphasize that infeasibility scoring is not always possible. For example, consider an agenda in which each proposition/negation pair is logically independent of the others. Then we cannot obtain an irrational judgment set from a rational one by making judgment reversals. However, infeasibility scoring can be applied to many agendas in which the propositions are interconnected. For instance, it can be applied to the agenda of the doctrinal paradox.

7 Conclusion

We have introduced the mean rule as an approach to aggregating binary evaluations. We have shown that the rule (i) determines the tension-minimizing monotone partition, (ii) emerges when we translate Young's Borda axioms into the framework of binary evaluations, and (iii) can be derived using the vector decomposition approach. Points (i) and (ii) provide the main normative argument in favor of the rule, while points (ii) and (iii) support our claim that the mean rule is the analogue to the Borda rule in the framework of binary evaluations. We concluded the paper by showing how both the mean rule and the Borda rule can be derived from an approach to judgment aggregation.

We would caution, however, against applying the mean rule to other aggregation settings simply because it can be applied in those settings. Clearly, the mean rule will not always be an attractive aggregation rule. We give one example. Imagine that twenty individuals apply for a job (labeled 1 to 20), and that there are four members of the appointment panel (labeled A, B, C, D). Each member of the appointment panel must look at each application and decide whether or not the candidate should be invited to interview. Suppose that A says "yes" ($+1$) to applicants 1 to 5 and "no" (-1) to the others. Similarly B says "yes" to applicants 6 to 10 only, C to applicants 11 to 15 only, and D to applicants 16 to 19 only. In other words, A, B and C want to invite 5 applicants each, while panel member D wants to invite 4.

The mean in this situation is $19/20$ and so, under the mean rule, applicants 1 to 19 are invited to interview. In other words, virtually all of the applicants are invited to interview despite the fact that the panelists each think that only 5 or 4 should be. Moreover, if the fourth panelist says "yes" to two more candidates (applicants 20 and 1) then (with no change in the evaluations of the other panelists) the mean rises to $21/20$ and so only applicant 1 is invited to interview under the mean rule. This collapse in the number of interviewees from 19 to 1 occurs despite the fact that this panelist believes that more candidates should be invited for interview, not fewer.

The reason that the mean rule is less attractive in this example is that the panelists care about the *sizes* of the two parts of the partition.[32] However, the mean rule minimizes tension with size playing a subsidiary role (see comments at the end of Sect. 3). Therefore, while this example fits our framework, it is not an aggregation problem to which the mean rule naturally applies (and the panel would not wish to use it). Perhaps a more appropriate method in this case would be the following. First, determine the mean number of candidates voted for by the members of the panel and round this number to the nearest integer t. Then invite for interview the t applicants with the greatest total number of "yes" evaluations (using some device to break ties if necessary). Clearly, this is very different to the mean rule.[33]

Another approach would be to apply a "restricted" form of the mean rule to this problem. In this restricted version, we constrain the set of feasible outcomes. We could, for instance, require that all feasible outcomes assign $+1$ to exactly 5 candidates. We would then choose whichever of these feasible outcomes minimizes tension (as in our original student partitioning problem). Here total tension would correspond to the sum of the tension in the set of interviewed applicants and the tension in the set of non-interviewed applicants. This paper has, however, dealt with the unrestricted version of the mean rule, and we leave consideration of this restricted version to future research.

Finally, in future research, we plan to explore further the properties of what we call "social dichotomy functions". The mean rule is the first example of such a function. More generally, these functions may have dichotomies as inputs (or some other form that provides information about "separation"), and the output is a social dichotomy (or a set of such dichotomies). Other social dichotomy functions exist in addition to the mean rule.

[32] In this example, the set of applicants is partitioned into a "invited to interview" set and a "not invited to interview set".

[33] Another approach would be to determine t before the votes are cast, and then require each panelist to vote for exactly t candidates. The t candidates with the greatest total number of votes are chosen for interview. See Peters et al. (2012).

References

Bogomolnaia A, Moulin H, Stong R (2005) Collective choice under dichotomous preferences. J Econ Theory 122:165–184

Brams S, Kilgour DM, Zwicker WS (1998) The paradox of multiple elections. Soc Choice Welf 15:211–236

Croom F (1978) Basic concepts of algebraic topology. Springer, New York

Dietrich F (2014) Scoring rules for judgment aggregation. Soc Choice Welf 42:873–911

Dietrich F, List C (2007) Arrow's theorem in judgment aggregation. Soc Choice Welf 29:19–33

Dokow E, Holzman R (2010a) Aggregation of binary evaluations. J Econ Theory 145:495–511

Dokow E, Holzman R (2010b) Aggregation of non-binary evaluations. Adv Appl Math 45:487–504

Duddy C, Piggins A (2013) Collective approval. Math Soc Sci 65:190–194

Fishburn PC (1977) Condorcet social choice functions. SIAM J Appl Math 33(3):469–489

Harary F (1958) Graph theory and electric networks. IRE Trans Circuit Theory CT-6:95-109

Kasher A, Rubinstein A (1997) On the question "who is a J?": a social choice approach. Logique et Analyse 160:385–395

Kornhauser LA, Sager LG (1993) The one and the many: adjudication in collegial courts. Calif Law Rev 81:1–59

List C (2008) Which worlds are possible? Judgment aggregation problem. J Philos Logic 37:57–65

List C, Pettit P (2002) Aggregating sets of judgments: an impossibility result. Econ Philos 18:89–110

List C, Polak B (2010) Introduction to judgment aggregation. J Econ Theory 145:441–466

Maniquet F, Mongin P (2015) Approval voting and Arrow's impossibility theorem. Soc Choice Welf 44:519–532

Nehring K, Puppe C (2010) Abstract Arrowian aggregation. J Econ Theory 145:467–494

Nitzan S, Rubinstein A (1981) A further characterization of the Borda ranking method. Public Choice 36:153–158

Peters H, Roy S, Storcken T (2012) On the manipulability of approval voting and related scoring rules. Soc Choice Welf 39:399–429

Rubinstein A, Fishburn PC (1986) Algebraic aggregation theory. J Econ Theory 38:63–77

Saari DG (1994) Geometry of voting. Springer, Berlin

Saari DG (1995) Basic geometry of voting. Springer, Berlin

Saari DG (2001a) Decisions and elections: explaining the unexpected. Cambridge University Press, Cambridge

Saari DG (2001b) Chaotic elections: a mathematician looks at voting. American Mathematical Society, Providence

Saari DG (2008) Disposing dictators, demystifying voting paradoxes: social choice analysis. Cambridge University Press, Cambridge

Wilson R (1975) On the theory of aggregation. J Econ Theory 10:89–99

Xia L (2011) Computational voting theory: game-theoretic and combinatorial aspects, Ph.D Dissertation. Computer Science Department, Duke University, Durham. http://www.cs.rpi.edu/~xial/

Young HP (1974) An axiomatization of Borda's rule. J Econ Theory 9:43–52

Zwicker WS (1991) The voters' paradox, spin, and the Borda count. Math Soc Sci 22:187–227