


ARTICLE

Wine rankings and the Borda method

Salvador Barberà¹, Walter Bossert² and Juan D. Moreno-Ternero³ 

¹MOVE, Universitat Autònoma de Barcelona, and Barcelona School of Economics, Facultat d'Economia i Empresa, 08193. Bellaterra (Barcelona), Spain, ²CIREQ, University of Montreal. P.O. Box 6128, Station Downtown, Montreal QC H3C 3J7, Canada and ³Department of Economics, Universidad Pablo de Olavide, Carretera de Utrera, Km.1. 41013. Seville, Spain

Corresponding author: Juan D. Moreno-Ternero, email: jdmoreno@upo.es

Abstract

We propose to establish wine rankings using scores that depend on the differences between favorable and unfavorable opinions about each wine, according to the Borda rule. Unlike alternative approaches and specifications, this method is well-defined even if the panelists' quality relations are not required to exhibit demanding properties such as transitivity or acyclicity. As an illustration, we apply the method to rank wines assessed by different experts and compare the resulting ranking with that obtained according to Condorcet's method of majority voting.

Keywords: Borda method; wine rankings; wine ratings

JEL classifications: C18; D71; L15; L66

1. Introduction

The modern era of wine journalism has popularized the use of numerical rating systems for wines to the extent that they largely influence consumers' decisions and investments.¹ Robert Parker first introduced his famous 50-to-100 point rating system in the *Wine Advocate*, and *The Wine Spectator* and *The Wine Enthusiast* quickly followed suit. The British wine expert Jancis Robinson rates wines on a 0-to-20 scale. As of today, it is difficult to find wine reviews without numerical ratings. In the academic literature, Amerine and Roessler (1983) were probably the first to raise the importance of seeking a consensus among tasters and exploring procedures to do so. More recently, Gergaud, Ginsburgh, and Moreno-Ternero (2021) provided a formal and comprehensive framework to obtain a consensus, building on contributions from political science, social choice, game theory, and operations research. There are now numerous contributions that focus on wine rankings and wine ratings.²

¹See, for example, Jaeger (1981), Ali, Lecocq, and Visser (2008), Friberg and Grönqvist (2012), Fogarty and Sadler (2014), and Ashton (2016).

²See, for instance, Lindley (2006), Ashton (2012, 2013), Ginsburgh and Zang (2012), Cao (2014), Cardebat and Paroissien (2015), and Bodington (2017, 2020).

In this paper, we challenge the need for detailed numerical valuations on the part of the jurors and question the methods that can only be applied if such inputs are available. Instead, we defend the use of the Borda (1781) count as a method to aggregate the opinions of experts about different wines because it is operational in many contexts where alternative methods are likely to fail. We claim that most other procedures used in practice or proposed in the literature impose unnecessarily high demands on the inputs to be provided by experts and that this may result in poor-quality expressions of actual opinions. In contrast, the Borda count can operate based on pairwise comparisons that do not need to be transitive (or even acyclical). This allows the method to process the data generated by individuals who form their judgments through reasoning that violates some classical notions of rationality; as the behavioral literature has shown, these violations occur quite frequently. In addition, we argue that the Borda count allows us to exploit databases that incorporate data from diverse contests to reach significant rankings. We emphasize that we retain the usual assumption of completeness—that is, all pairs of wines are compared by all jurors. Thus, there is no potential for any bias caused by unranked wines.

The Borda method assigns numerical values (the Borda scores) to alternatives, which are then translated into rankings. In our view, rankings are the essential output of any contest, and numerical representations are just a method to facilitate comparisons. But to respect tradition, we note that the Borda scores may be used to represent these rankings, and, as we shall see, these scores have an intuitive interpretation because they tell us, for each wine under consideration, the difference between the number of times this wine is considered better than other wines and the number of times it is considered worse.

The numbers used in applying the Borda method have an intuitively appealing definition and interpretation. For each of the experts i and for each wine x , we count the number of wines that are ranked as worse than x by i and subtract from this value the number of wines that are considered better than x by expert i . Clearly, this difference could be positive, negative, or equal to zero. For each of the wines, we now add these differences across all experts to arrive at the overall Borda scores of the wines. The overall ranking of the wines is obtained by using these overall scores as the criterion—a wine is at least as good as another if the overall Borda score of the former is greater than or equal to the overall Borda score of the latter. Some readers may have come across formulations of the Borda method that differ from the one we just introduced. These alternatives include the possibility of using the ranks of alternatives, or the number of their wins alone, to establish the Borda scores. Although the corresponding counts are equivalent to the one in our definition when the relations of the individual jurors are strict and complete or even strict orderings, these alternative proposals fail in more general informational environments, like those we consider in this paper. We comment on their failure in more detail in Section IV once the requisite formal definitions have been provided. Hence, our description of the method represents the (only) genuine definition of the Borda method that is universally applicable without any additional restrictions on the quality relations of the experts.

Let us elaborate on the idea that the methods currently employed can be viewed as being rather demanding on the experts who act as jurors. They must submit highly

detailed information to be used in the wine judgment process. This task may often be extremely difficult to perform in a reliable fashion due to some inherent complexities, through no fault whatsoever on the part of the experts. The methods currently in use require the experts to supply not only a complete and transitive ranking of the wines to be judged but even to assign numerical values to each of the wines. These jurors may be extremely competent and experienced, but, nevertheless, some concerns seem to arise immediately. For instance, it may be the case that the difference between two wines x and y is so minuscule that they cannot reasonably be distinguished even by the most experienced taster, and a similar situation may apply to wine y when it is compared to a third wine z . Yet, the difference between x and z may be sufficient to establish a strict ranking between them. Such cases are well documented in the psychology literature as arising from thresholds of perception, and they induce natural violations of transitivity.³

Hence, we think that valid concerns already arise even if merely a complete and transitive ranking of the wines is to be produced by each juror. If, in addition, numerical values are to be assigned, matters become even more complex—and potentially unreliable, even when the best possible experts are being consulted.

The Borda method has already been discussed in the requisite literature,⁴ but still in the context where experts are asked to submit complete and transitive quality relations. Thus, our proposal to employ the method on less informationally demanding datasets is new.

II. A summary of standard approaches

We begin with a brief review of some aggregation methods that appear in the earlier literature so that we can later compare their informational requirements with the milder ones that are needed for the application of the Borda rule. Suppose there is a finite set $N = \{1, \dots, n\}$ of experts (or jurors) with $n \geq 2$ members whose task it is to assess the members of a finite set X of $m \geq 2$ wines with respect to their relative quality. Typically, the demands on the experts regarding the quality assessment that they are expected to provide are rather strong. In most cases, these jurors are asked to assign a numerical value (a rating) within a predefined range (such as between 0 and 20, for example, as in the case of the so-called *Judgment of Paris*) to each wine in X . In discussing these methods, we use the notation $r_i(x)$ to indicate the numerical rating of wine $x \in X$ by expert $i \in N$.

The first method employs the arithmetic mean of all experts to determine an overall quality rating; this is sometimes referred to as the usual consensus rating. More precisely, for each wine $x \in X$, the arithmetic-mean rating $r^{AM}(x)$ is defined as

$$r^{AM}(x) = \frac{1}{n} \sum_{i=1}^n r_i(x).$$

³According to Armstrong (1939, p. 122), intransitive equal goodness can result from “the imperfect powers of discrimination of the human mind whereby inequalities become recognizable only when of sufficient magnitude.” Luce (1956) and Luce and Raiffa (1957) illustrate this point using the well-known coffee-sugar example. Fishburn (1970) provides a comprehensive survey of intransitive equal goodness.

⁴See, for instance, Huklower (2009) and Gergaud, Ginsburgh, and Moreno-Ternero (2021).

A modification of this method can be obtained if a threshold is employed. Formally, let π denote a threshold within the range of possible rating values. The threshold-dependent quality rating $r_i^{TD}(x)$ of expert $i \in N$ for a wine $x \in X$ is

$$r_i^{TD}(x) = \begin{cases} r_i(x), & \text{if } r_i(x) \geq \pi \\ 0, & \text{if } r_i(x) < \pi \end{cases}$$

and, in analogy with the arithmetic-mean rating, the threshold-dependent arithmetic-mean rating (also known as the proportional approval consensus rating) for a wine $x \in X$ is given by

$$r^{TD}(x) = \frac{1}{n} \sum_{i=1}^n r_i^{TD}(x).$$

Strictly speaking, these threshold-dependent individual and overall ratings depend on the choice of the threshold π , but we suppress this dependence for ease of exposition; we do not think that this move involves any danger of ambiguity.

An alternative method that eliminates all wines below a threshold π in the individual ratings homogenizes the ratings of the wines that remain so that they all receive the same numerical rating of one. The corresponding approval-vote quality rating $r_i^{AV}(x)$ of expert $i \in N$ for wine $x \in X$, also referred to as the approval-consensus rating, is defined by

$$r_i^{AV}(x) = \begin{cases} 1, & \text{if } r_i(x) \geq \pi \\ 0, & \text{if } r_i(x) < \pi \end{cases}.$$

The overall approval-vote rating $r^{AV}(x)$ of wine $x \in X$ is obtained as the arithmetic mean of these individual ratings so that

$$r^{AV}(x) = \frac{1}{n} \sum_{i=1}^n r_i^{AV}(x).$$

Again, we suppress the dependence on the threshold π in our notation for the sake of simplicity.

There is one feature that the approval-vote rating method has in common with the Borda method that we advocate—it does not rely on the numerical ratings provided by the experts. However, we think that the approval-vote method goes too far in that regard because it also eliminates all distinctions among the wines that survive the elimination process—all of them are treated equally, even if an expert may have a firm view that some of them are better than others. The Borda method retains this (as we think, vitally important) information regarding the relative quality of the wines under consideration.

The primary motivation behind the use of a threshold is to exclude wines that perform rather poorly in the individual ratings. As we argue later in the paper, the Borda method guarantees that wines with a relatively poor quality rating by all experts do

not influence the aggregate ranking of the remaining wines, so that the use of a (somewhat arbitrary and potentially distorting) threshold becomes superfluous.

It is straightforward to define relative variants of the latter two threshold-based methods. The corresponding relative overall ratings are obtained by dividing the original (non-relative) ratings by the sum of the requisite individual threshold-adjusted ratings instead of using the total number of experts in the denominator.

III. Individual wine assessments

The informational requirements on which the methods described in the previous section are based are very demanding; it is by no means obvious that an expert can provide such finely nuanced assessments. In fact, it may very well be the case that even a mere ordering (i.e., a complete and transitive relation) of the wines in X is difficult to elicit from a juror. The method we propose—the Borda method—takes these concerns into account in that it only relies on a very modest amount of information to be provided by the experts. Any complete relation will do; there is no need to require these relations to possess any coherence properties such as transitivity or acyclicity.

Each expert $i \in N$ is assumed to have a complete quality relation R_i defined on the set X of wines that are to be assessed. Thus, the statement xR_iy represents the view that wine $x \in X$ is at least as good as wine $y \in X$ according to expert $i \in N$. As is commonly done, we use xP_iy to denote that wine x is better than wine y according to juror i , and we write xI_iy to indicate that i considers x and y to be equally good. To be precise, the betterness relation P_i and the equal-goodness relation I_i are derived from the at-least-as-good-as relation R_i by letting xP_iy whenever xR_iy and not yR_ix are true, and xI_iy whenever it is the case that xR_iy and yR_ix . We stress that we do not need to assume that the individual quality relations of the experts are transitive; they do not even have to be acyclical.

The quality relations R_1, \dots, R_n can be collected in a quality profile $\mathbf{R} = (R_1, \dots, R_n)$. Thus, a quality profile consists of n individual quality relations—one relation for each of the experts.

An example may be instructive at this point. Suppose that there are a set of three experts $N = \{1, 2, 3\}$ and a set of three wines $X = \{x, y, z\}$. A quality profile $\mathbf{R} = (R_1, R_2, R_3)$ is composed of the individual quality relations given by

$$\begin{aligned} &yP_1z, zP_1x, yP_1x, \\ &xP_2y, yP_2z, zP_2x, \\ &zI_3x, xI_3y, zP_3y. \end{aligned}$$

In this example, expert 1 submits an ordering of the three wines, considering y the best wine, followed by z as the second-best, and wine x as the worst. Expert 2's quality relation, on the other hand, cannot be expressed in terms of an ordering because it is cyclical: for this expert, wine x is better than wine y , wine y is better than wine z , and wine z is better than x , leading to a cycle. We do not mean to suggest that such cyclical quality relations are likely to occur, but we include them in the example to emphasize the point that the Borda method is perfectly capable of accommodating such

relations. There is an instance of intransitive equal goodness in the quality relation of expert 3: (s)he considers wines z and x to be equally good, and the same judgment applies to wines x and y . In violation of transitivity, however, wine z is considered better than wine y . The relation assigned to expert 3 appears to be quite plausible. As alluded to in the introduction, a quality relation of this nature can easily result from a threshold of perception that prevents the expert from distinguishing wines z and x as well as wines x and y but allows for the unambiguous judgment that z is better than x .

IV. From quality assessments to wine rankings

We now address the issue of aggregating the experts' quality relations into a ranking of the wines in X . The principle underlying Borda's method has considerable intuitive appeal. For each of the wines x and for each expert i , we count (a) the number of times wine x is judged to be better than one of the other wines and (b) the number of times wine x is judged to be worse than another wine. The difference between the first of these numbers and the second is the individual Borda score of wine x according to expert i . Now, for each wine x , these scores are added over all experts, and the resulting sum is the overall Borda score of wine x . Finally, the overall wine ranking is established by comparing any two wines according to their respective Borda scores. Thus, the criterion to assess the relative overall quality of the wines is the difference between the number of times a wine beats another one in a pairwise contest and the number of times this wine is beaten by another in such a contest. A fundamental and attractive feature of this method is that there is no need whatsoever to invoke any properties of the experts' quality relations—only pairwise comparisons matter, and these are well-defined for any binary relation.

In our context, a wine-ranking method assigns an ordering (i.e., a complete and transitive relation) on the set of candidate wines X to each quality profile composed of the experts' individual quality relations. We emphasize that the objective is to rank all possible wines. Anything short of that would be rather unsatisfactory from the viewpoint of a consumer; a cycle or a large degree of non-comparability are attributes that may render the entire ranking exercise close to meaningless. We use R to label the wine ranking that is associated with the profile $\mathbf{R} = (R_1, \dots, R_n)$. The ranking generated by the Borda method is denoted by R^B .

Consider any expert $i \in N$ and suppose that his or her quality ranking on the set of wines X is given by R_i . Furthermore, let $x \in X$ be any of the wines to be assessed. In line with the informal description provided, the individual Borda score of wine x according to the quality relation R_i of expert i is given by

$$b(x; R_i) = |\{z \in X | xP_i z\}| - |\{z \in X | zP_i x\}|,$$

where the notation $|S|$ is used to indicate the number of elements in the finite set S .

Note that the above difference is unaffected if the quality relation R_i is used in place of the associated betterness relation P_i . This is the case because whenever we have an instance of equal goodness between wines x and y , it follows by definition that $xR_i y$ and $yR_i x$ cancel each other out when calculating the above-defined

difference. This argument applies to any two wines, x and y , no matter whether x and y are distinct or identical. As explained in more detail in the following section, the observation that the Borda method possesses this cancellation property is a major reason why there is no need for any thresholds to eliminate the influence of wines that perform poorly according to the quality relations of all experts.

Either or both of the two sets in the definition of the individual Borda scores may be empty. Furthermore, note that these scores are well-defined even if the individual quality relation does not possess any coherence properties such as transitivity; again, this is an important and desirable feature of the Borda method that is not shared by most alternative ranking procedures.

The overall Borda score $b(x; \mathbf{R})$ of a wine x is obtained by adding the individual scores of all experts so that

$$\begin{aligned} b(x; \mathbf{R}) &= \sum_{i=1}^n b(x; R_i) = \sum_{i=1}^n (|\{z \in X \mid xP_i z\}| - |\{z \in X \mid zP_i x\}|) \\ &= \sum_{i=1}^n |\{z \in X \mid xP_i z\}| - \sum_{i=1}^n |\{z \in X \mid zP_i x\}|. \end{aligned} \quad (1)$$

Finally, the Borda wine ranking R^B is obtained by declaring a wine x to be at least as good as a wine y if and only if the overall Borda score of x is greater than or equal to the score of y , that is,

$$xR^B y \Leftrightarrow b(x; \mathbf{R}) \geq b(y; \mathbf{R}).$$

The method we propose is by no means new, although the arguments we offer here are novel, as they refer to its extension to additional data sets. But the idea that every single instance of betterness has value and should be counted as one of many opinions, all of which have the same value, is stressed in an essay by Morales (1797). Morales was a stout defender of Borda's (1781) voting rule, on which the Borda method is based. Not everyone was as supportive of Borda as Morales; for example, voting rules founded on the majority principle were advocated by Condorcet (1785). Daunou (1803), a strong critic of Borda, proposed an alternative voting rule based on a lexicographic combination of the Condorcet criteria and the plurality rule.⁵ Following Morales, we also use the term opinion when referring to a single instance of betterness in an expert's assessment of the wines under consideration.

To illustrate the definition of the Borda method, let us return to the example defined in the previous section. It follows that

$$\begin{aligned} b(x; R_1) &= 0 - 2 = -2, & b(y; R_1) &= 2 - 0 = 2, & b(z; R_1) &= 1 - 1 = 0; \\ b(x; R_2) &= 1 - 1 = 0, & b(y; R_2) &= 1 - 1 = 0, & b(z; R_2) &= 1 - 1 = 0; \\ b(x; R_3) &= 0 - 0 = 0, & b(y; R_3) &= 0 - 1 = -1, & b(z; R_3) &= 1 - 0 = 1. \end{aligned}$$

⁵See Barberà, Bossert, and Suzumura (2021) for a detailed discussion.

Adding over all experts, we obtain the overall Borda scores

$$b(x; \mathbf{R}) = -2 + 0 + 0 = -2, b(y; \mathbf{R}) = 2 + 0 - 1 = 1, b(z; \mathbf{R}) = 0 + 0 + 1 = 1$$

and, therefore, the overall Borda ranking is given by yI^Bz, yP^Bx, zP^Bx . Thus, wines y and z are tied for the top, and wine x is at the bottom. This example confirms that the Borda method always generates an overall ordering even if the individual quality relations fail to be acyclical.

A few words of explanation may be in order to clarify why two commonly used alternative definitions of the Borda method are not suitable in our setting, owing to the possibility of equal goodness or violations of transitivity in the individual quality relations.

If the experts' quality relations are strict and complete (so that one of any two distinct wines must be better than the other), the definition of the Borda method can be simplified. The conjunction of these two properties guarantees that the ranking of any two wines is unchanged if the second term in the difference that defines the overall Borda score is omitted. This is the case because the equality

$$|\{z \in X | zP_i x\}| + |\{z \in X | xP_i z\}| = m - 1$$

is valid for all wines x if R_i is a strict complete relation defined on the set of m wines in X . If individual quality relations may include instances of equal goodness, however, this is not the case—removing the second part of the difference in Equation (1) leads to a different criterion to rank the wines. In the context of quality relations that are not necessarily strict, this second option leads to highly undesirable consequences. For example, suppose that there are three experts $N = \{1, 2, 3\}$ and three wines $X = \{x, y, z\}$. Furthermore, consider the quality profile $\mathbf{R} = (R_1, R_2, R_3)$ given by

$$\begin{aligned} & xI_1y, xP_1z, yP_1z, \\ & xI_2y, xI_2z, zP_2y, \\ & xI_3y, xI_3z, zP_3y. \end{aligned}$$

According to the Borda method, the overall Borda score of wine x is given by $1 - 0 = 1$ because x beats wine z according to expert 1's quality relation, and the overall score of wine y is $1 - 2 = -1$ because y beats z according to expert 1 and is beaten by z according to experts 2 and 3. Thus, x is better than y according to the Borda method. However, if we were to count the number of wins only and disregard the number of losses when calculating the requisite scores, x and y would be ranked as equally good because each of them registers one win. This seems difficult to accept because wine y is beaten by z in the quality relations of experts 2 and 3, whereas x does not suffer any such losses. The use of the criterion expressed in Equation (1) avoids troublesome conclusions of this nature.

A second alternative definition of the Borda rule consists of considering it a special case of a scoring method.⁶ Assuming that the individual quality relation of each

⁶See Young (1975).

expert is a strict ordering, each wine is given a position: the top wine is in position 1, the next-to-best is in position 2, and so on until we reach the bottom wine, whose position corresponds to m —the total number of wines under consideration. In this environment that is restricted to strict individual orderings, the Borda method can be thought of as assigning a weight to each position, where these weights are given as follows: Position 1 receives a weight of $(m - 1) - 0 = m - 1$ because this top position is better than the $m - 1$ remaining ones and worse than none of the others. The weight of position 2 is $(m - 2) - 1 = m - 3$ because position 2 is superior to $m - 2$ of the remaining positions and inferior to one position—the top position. This process can be continued until we reach the bottom position m , which is superior to none of the others and inferior to the remaining $m - 1$ positions, and its weight is $0 - (m - 1) = -(m - 1)$. Thus, according to the Borda method, the positional weights w_1, \dots, w_m assigned to the m positions from top to bottom are given by

$$w_1 = m - 1, w_2 = m - 3, \dots, w_{m-1} = -(m - 3), w_m = -(m - 1).$$

Adding the scores of all experts yields the overall Borda scores as defined in Equation (1). A generalization of this method consists of the class of scoring methods. These are obtained by assigning arbitrary weights w_1, \dots, w_m to the positions, where the restriction $w_1 \geq \dots \geq w_m$ with at least one strict inequality is usually imposed to ensure that better positions cannot receive lower weights than worse positions. Because the notion of a position cannot even be defined if the experts' quality relations fail to be orderings, this alternative method cannot be applied in our informationally austere framework. Thus, except for the Borda rule, general scoring methods are not suitable for the purposes of this paper, and, therefore, we cannot treat Borda's method as a special case of these more general rules. Moreover, we note that the Borda rule is perfectly well-defined even if the individual relations are complete and transitive but not necessarily strict—the definition of the criterion does not have to be amended in any way. For other scoring rules, this is not the case because modifications are needed to make sure that weights can be assigned in instances of equal goodness; this is often accomplished by employing arithmetic means to determine the positional weights involved in a tie.

Condorcet's (1785) method of majority decision shares the flexibility of the Borda rule.⁷ It is also applicable if the individual experts' relations fail to be orderings. According to the Condorcet rule R^C , wine x is at least as good as wine y if and only if the number of experts who rank x higher than y is greater than or equal to the number of experts who rank y higher than x , that is,

$$xR^C y \Leftrightarrow |\{i \in N | xP_i y\}| \geq |\{i \in N | yP_i x\}|.$$

In analogy to the Borda rule, the betterness relation P_i can be replaced with the goodness relation R_i without changing the comparisons according to this method—again, instances of equal goodness cancel each other out. Because only pairwise comparisons appear in this definition, the individual relations do not have to possess any

⁷See May (1952, 1953) for a characterization of the majority rule.

coherence properties such as transitivity. As is well known, the Condorcet method suffers from the serious drawback that it may generate cyclical overall relations, thus making the rule of little use when it comes to providing guidance to consumers. For this reason, we do not advocate Condorcet's method of majority decision; we introduce it merely for comparative purposes in the applied section.

V. Properties of the Borda method

A. Informational parsimony

One of the major characteristics of the Borda method applied to wine rankings is its remarkable adaptability to environments in which very little information is available. As we have been pointing out repeatedly, all that is required of the experts is that they submit a complete relation defined on the set of wines to be judged without having to require this relation to possess any properties that are often required. Because violations of transitivity are bound to arise if some wines are difficult to distinguish with respect to some of their attributes, this is a very important feature. Thus, the observation that the experts can submit quality relations without having to conform to any restrictive properties provides a forceful argument in favor of the Borda method.

We stress that most—if not all—competing methods to establish wine rankings or wine ratings do not share this ability to accommodate such informationally austere environments. As can be seen from their definitions, all the methods illustrated in Section II rely not only on individual orderings but even on numerical ratings provided by the experts.

B. Merging expert panels

Merging data coming from different panels, with varying experts or alternative subsets of wines under consideration, is likely to influence the choice of a suitable aggregation method. A prominent example of a requirement that emerges in this context is what Smith (1973) refers to as separability.⁸ The property states that if the overall quality relations obtained by two disjoint groups of experts agree on the relative ranking of two wines, this relative ranking is preserved if the two groups are merged into one. The Borda method satisfies this requirement, along with several other methods such as the scoring methods. An analogous property is used by Young (1974), Richelson (1978), and Ching (1996), among others, in the context of aggregation methods that generate choices rather than overall quality relations. While the Borda method certainly is not the only one satisfying this (quite natural) requirement, the observation that it is compliant with this condition provides yet another argument in favor of its use.

C. Independence of dominated wines

Consider a situation in which there are some wines that are beaten by all others according to all experts. More precisely, suppose that there is a non-empty strict subset Y of X such that, according to the quality relation R_i of every expert i on the panel

⁸See also Fine and Fine (1974a, 1974b).

N , every wine y in Y is worse than every wine x in the complement $X \setminus Y$ of Y —that is, we have xP_iy for all wines $x \in X \setminus Y$, for all wines $y \in Y$, and for all jurors $i \in N$. Following the argument that is often provided in favor of using thresholds (see Section II), it is desirable to require that the (inferior) wines in Y do not influence the ranking of the remaining (superior) wines in the complement $X \setminus Y$ of Y . Thresholds of this nature are not required to accomplish this objective if the Borda method is employed. If the relations among the wines in Y change according to all experts, but all of these wines remain worse than the remaining wines in all individual assessments, nothing happens to the collective ranking of the superior wines (those in the complement $X \setminus Y$), provided that the relations between the wines in $X \setminus Y$ are unchanged. This follows immediately from the definition of the method: the Borda scores of the wines in $X \setminus Y$ do not depend on the relative rankings among the wines in Y ; all that matters is that all the wines in Y are considered worse than all of the wines in $X \setminus Y$ according to all experts. Analogously, removing the wines in Y altogether does not change the ranking of the remaining wines according to the Borda method; this is a consequence of the observation that such a move reduces the Borda score of each of the wines in $X \setminus Y$ by the same value—the number of wines in Y . Thus, all comparisons of the superior wines remain unchanged once the wines in Y are eliminated from the contest. Note that a (hypothetical) elimination of this nature is very different from the elimination of wines based on thresholds because it is not based on some (arguably, arbitrarily chosen) threshold—rather, it is an intrinsic feature of the method itself. It is also worth noting that externally imposed thresholds cannot help but introduce biases because a wine that receives a low rating from some experts may receive a high rating from others. Reducing all individual ratings below the threshold to zero would seem to stack the deck against such a wine for, as we believe, no convincing reason.

D. All opinions count equally

The Borda method also possesses an interesting cancellation property. We already mentioned that any individual assessment of equal goodness does not affect the overall wine ranking. This is the case because an instance of one wine being judged to be at least as good as another is canceled out by an instance of the latter being at least as good as the former; this is an immediate consequence of the criterion expressed in terms of the difference in Equation (1). Moreover, if one expert considers a wine x to be superior to another wine y and a second expert submits the reverse judgment, the two again cancel each other out once the sum of the differences between wins and losses is calculated. This cancellation property of the Borda method is already noted by Young (1974) and used in his characterization of Borda's voting rule.⁹

There is yet another form of cancellation property that applies to pairwise comparisons according to the Borda method. Consider three distinct wines, x , y , and z , and suppose we focus on the overall ranking of x and y . If there is an expert i who considers x superior to z and an expert j who considers y superior to z , the statements $xP_i z$ and $yP_j z$ cancel each other out when it comes to the relative ranking of x and

⁹See also Barberà and Bossert (2023).

y : if both opinions are replaced by equal goodness, the overall Borda score of each wine is reduced by one so that the criterion arrives at the same comparison. Of course, the ranking of the third wine z relative to others may be affected by this move because it now experiences two fewer losses as compared to the initial situation. Note that the conclusion holds no matter whether experts i and j are the same person or two different individuals. To go one step further, if there are two opinions, $xP_i z$ and $yP_j w$, with z and w being distinct, the overall ranking of x and y remains the same if the two betterness statements are replaced with equal goodness, even if z and w are different. In addition, an analogous conclusion applies in the opposite direction—that is, if the number of pairwise losses of two wines x and y is reduced by one. This is a consequence of using the difference between the number of wins and the number of losses as the criterion to rank two wines according to the Borda method.

As mentioned previously, some ranking methods currently in use apply a threshold in order to exclude some wines at some stage in the ranking process. The cancellation properties just described imply that the relative overall ranking of two wines does not depend on the presence of wines that are beaten by both. This observation reiterates that an elimination process based on some more or less arbitrary threshold is superfluous if the Borda method is employed.

These cancellation properties reflect a feature of the Borda method that was already noted and emphasized by Morales (1797). He writes that, according to the Borda method, every opinion counts equally—no matter who holds it or how other wines may be judged by the experts. This is another distinguishing attribute of the Borda method. In contrast, if individual quality relations are strict orderings and a scoring rule other than that of Borda is employed, this is no longer the case: the weights assigned to the positions may be such that one opinion does not cancel out another if they are associated with different positions.

E. Additional favorable opinions

Suppose that a quality profile is augmented by an additional favorable opinion for wine x against wine z , all else remaining unchanged. This could be the consequence of an expert changing his or her initial assessment of equal goodness to a decisive choice of x as the better wine. If x is already judged to be at least as good as another wine y (which need not be equal to z) according to the overall quality ranking, it is plausible to require that this additional support for x cannot cause this wine to drop in the ranking—in fact, it would be natural to demand that x now be better than y . The dual of this responsiveness property is just as appealing: if wine x is initially at least as good as wine y according to the overall ranking and a favorable opinion for y against another wine z is changed to equal goodness (because an expert is no longer confident in his or her judgment regarding the betterness of y), this loss of support for y should not allow wine y to climb above wine x in the ranking as a result of the removed positive opinion.

The Borda method satisfies the strict forms of these requirements, whereas others, such as the method of majority decision, do not. This is the case because the majority method establishes the comparison between two wines x and y exclusively on the

Table 1. Ratings from five experts for the seven most expensive 2021 red Bordeaux wines

Wines	WA	JR	JL	JA	CK	Price (€)
Lafite-Rothschild	95–97	17.5++	96–98	97	92–94	684
Margaux	95–97+	17.5	96–98	95	93–95	599
Mouton-Rothschild	95–96	17.5+	95–97	95	93–95	599
Haut-Brion	96–98	18	95–97	95	93–95	593
Cheval Blanc	95–97	17++	96–98	96	93–95	565
Angélus	94–96+	17+	94–96	95	92–94	373
Palmer	95–97	17+	94–96	96	93–95	339

basis of the experts' opinions on these two wines. If x and y are equally good to begin with and a favorable opinion for x against a third wine z (that differs from y) is added, nothing changes in the quality assessments of x and y ; therefore, x cannot become better than y in response to the additional support for x .

In a general setting, a cancellation property and a responsiveness requirement similar to those outlined in this section can be used to provide a characterization of the Borda method, provided that all wines are equally good if no expert expresses any (strict) opinion. Thus, the only method that respects these intuitively appealing conditions is the Borda method.¹⁰

VI. An application to Bordeaux 2021 wines

We conclude with an application of our method to a real-life scenario—the recent tasting of some 2021 red Bordeaux wines gathered on *Bordoverview*.¹¹ This tasting involves ratings for some Bordeaux 2021 future wines produced by five international experts: the *Wine Advocate* (WA), Jancis Robinson (JR), Jeff Leve (JL), Jane Anson (JA), and Chris Kissack (CK).¹²

Table 1 exhibits a selection of seven wines (the most expensive ones), with the ratings assigned by these experts. Because the Borda method merely requires a complete relation provided by each of the experts as the basic input, the only information to be extracted is ordinal in nature. If, for instance, an expert submitted a range such as 95–97, we took this to mean that the wine in question is considered better than a wine with a rating of 95 and worse than a wine with a rating of 97. In one case, there is a rating of 97 for one wine and a rating of 96–98 for another by the same expert, and we treated these two wines as being considered equally good. Analogously, a wine with a rating of 17.5+ was treated as better than a wine with a rating of 17.5. We note that the scales employed by some reviewers differ from those used by others, but this is not a problem because the Borda method only requires ordinal information.

Table 2 presents the individual Borda scores that are obtained by calculating the number of wins minus the number of losses in pairwise comparisons according to

¹⁰See Barberà and Bossert (2023) for details.

¹¹See <https://www.bordoverview.com/>.

¹²See the appendix for some details on the five experts.

Table 2. Individual Borda scores

Wines	WA	JR	JL	JA	CK
Lafite-Rothschild	0	4	4	6	−5
Margaux	4	0	4	−3	2
Mouton-Rothschild	−4	2	−1	−3	2
Haut-Brion	6	6	−1	−3	2
Cheval Blanc	0	−2	4	3	2
Angélus	−6	−5	−5	−3	−5
Palmer	0	−5	−5	3	2

each of the experts. For instance, Lafite-Rothschild is assigned an individual Borda score of 4 by expert Jancis Robinson because, according to Table 1, Lafite-Rothschild beats five other wines and is beaten by only one other wine, so the score is given by $5 - 1 = 4$.

Table 3 contains the results. The overall Borda scores, obtained by adding the individual Borda scores, are listed first. The final column presents the ranking of the seven wines according to the Borda method. Haut-Brion comes in first with an overall Borda score of 10, followed by Lafite-Rothschild with a score of 9. Then, there is a tie between Margaux and Cheval Blanc (both with a score of 7). Mouton-Rothschild comes in fifth with a score of −4, Palmer is in the next-to-last position with a score of −5 and, finally, Angélus is at the bottom with a score of −24. As alluded to earlier, the Borda scores can serve as wine ratings if desired. We emphasize, however, that these ratings are nothing more than a numerical representation of the Borda ranking.

Note that both Lafite-Rothschild and Margaux are tied according to the Condorcet criterion. Each is ranked higher than the other by two experts, whereas one juror considers both wines equally good. The same happens to Haut-Brion and Cheval Blanc. However, both Lafite-Rothschild and Margaux win (in a pairwise contest) against Cheval Blanc whereas they both lose against Haut-Brion (in a pairwise contest). Thus, there are two cycles in the aggregate relation. Lafite-Rothschild beats Cheval

Table 3. Overall Borda scores and Borda ranks

Wines	Scores	Ranks
Lafite-Rothschild	9	2
Margaux	7	3.5
Mouton-Rothschild	−4	5
Haut-Brion	10	1
Cheval Blanc	7	3.5
Angélus	−24	7
Palmer	−5	6

Blanc; Cheval Blanc and Haut-Brion are tied; and Haut-Brion beats Lafite-Rothschild. The same observation applies if Lafite-Rothschild is replaced with Margaux. In other words, the Condorcet criterion is unable to provide an ordering for the full sample of seven wines, in contrast to the Borda method.

VII. Concluding remarks

The main purpose of this paper is to make a case in favor of using the Borda method in the assessment of wines by a panel of jurors. While this suggestion is not new, a novel aspect of our proposal is the observation that this method is sufficiently flexible to accommodate inputs from the experts that are complete binary relations—properties such as transitivity or merely acyclicity are not required. In principle, completeness can be dispensed with as well; this is the case because the Borda scores are entirely determined based on better-than relations. We retain the completeness assumption to conform to the view that Borda treats all jurors equally, thus allowing all of them to rank all possible pairs of wines. Because only betterness matters, one may be tempted to make the case that the assessments of experts who submit more betterness relations than others have a higher weight in determining the overall ranking. But if this argument is accepted (and we think it should not be), it immediately applies to the Borda method in the traditional framework in which all individual relations are orderings: again, only betterness matters—all instances of equal goodness cancel each other out. Thus, everyone who accepts the Borda method if individual relations are orderings cannot use the equal-treatment argument to reject the Borda method if individual relations are permitted to exhibit instances of non-comparability—the argument is valid either in both environments or in neither of them.

Another issue is the possible view that numerical ratings may be of use when it comes to the comparison of wines that are up for assessment by an expert panel at different time periods. This view rests on the assumption that these ratings are numerically and intertemporally significant: a specific quality rating means the same for every juror, for every wine, and for every time period in which an assessment is performed. To put it another way, a consumer must be confident that a wine that is given a specific rating in one year is, according to the experts being consulted, of the same quality as a wine that receives the same rating in another year. To us, this seems like a rather heroic assumption. As argued earlier, even the assumption that the experts have quality orderings is too demanding in many circumstances; for instance, the possibility of intransitive indifference generated by thresholds of perception is difficult, if not impossible, to rule out. In view of observations of this nature, the hypothesis that every expert can submit a reliable ordering (let alone assign ratings that are numerically significant) may be too much to ask for.

The illustrative example presented in Section VI does not allow us to demonstrate the remarkable flexibility of the Borda method because it utilizes a data set that is generated from individual wine ratings submitted by the experts; therefore, the underlying relations cannot but be orderings. This appears to be inevitable because non-transitive data is not available at this time. We hope, however, that this paper will contribute toward the objective of conducting assessments that permit the experts

to submit not necessarily acyclical (thus, not necessarily quasi-transitive or transitive) relations if they so desire.

Acknowledgments. The paper was presented at Sophia University. We thank an anonymous referee for thoughtful and constructive comments and suggestions. Salvador Barberà acknowledges financial support from the Spanish Agencia Estatal de Investigación (AEI), through the Severo Ochoa Programme for Centres of Excellence in R&D (Barcelona School of Economics CEX2019-000915-S), and from the Spanish Ministry of Economy and Competitiveness and FEDER through grant PID2020-116771GB-I00, and from the Generalitat de Catalunya, through grant 2021SGR00194. Juan D. Moreno-Ternero acknowledges financial support from the Spanish Agencia Estatal de Investigación (AEI) through grant PID2020-115011GB-I00, funded by MCIN/AEI/10.13039/501100011033, and from the Junta de Andalucía, through grant P18-FR-2933. Funding for open access publishing: Universidad Pablo de Olavide.

References

- Ali, H., Lecocq, S., and Visser, M. (2008). The impact of gurus: Parker grades and en primeur wine prices. *Economic Journal*, 118(529), F158–F173. (Reprinted in 2010 in the *Journal of Wine Economics*, 5(1), 22–39.)
- Amerine, M., and Roessler, E. (1983). *Wines: Their Sensory Evaluation*. New York: W.H. Freeman.
- Armstrong, W. E. (1939). The determinateness of the utility function. *Economic Journal*, 49(195), 453–467.
- Ashton, R. H. (2012). Reliability and consensus of experienced wine judges: Expertise, within, and between? *Journal of Wine Economics*, 7(1), 70–87.
- Ashton, R. H. (2013). Is there consensus among wine quality ratings of prominent critics? An empirical analysis of red Bordeaux, 2004–2010. *Journal of Wine Economics*, 8(2), 225–234.
- Ashton, R. H. (2016). The value of expert opinion in the pricing of Bordeaux wine futures. *Journal of Wine Economics*, 11(2), 261–288.
- Barberà, S., and Bossert, W. (2023). Opinion aggregation: Borda and Condorcet revisited. *Journal of Economic Theory*, 210, 105654.
- Barberà, S., Bossert, W., and Suzumura, K. (2021). Daunou's voting rule and the lexicographic assignment of priorities. *Social Choice and Welfare*, 56(2), 259–289.
- Bodington, J. (2017). The distribution of ratings assigned to blind replicates. *Journal of Wine Economics*, 12(4), 363–369.
- Bodington, J. (2020). Rate the raters: A note on wine judge consistency. *Journal of Wine Economics*, 15(4), 363–369.
- Borda, J.-C. de. (1781). Mémoire sur les élections au scrutin. Mémoires de l'Académie Royale des Sciences année 1781, 657–665. Translated and reprinted in I. McLean and A. B. Urken (eds.), *Classics of Social Choice*, Chapter 5. Ann Arbor: University of Michigan Press, 1995.
- Cao, J. (2014). Quantifying randomness versus consensus in wine quality ratings. *Journal of Wine Economics*, 9(2), 202–213.
- Cardebat, J., and Paoissien, E. (2015). Standardizing expert wine scores: An application for Bordeaux en primeur. *Journal of Wine Economics*, 10(3), 329–348.
- Ching, S. (1996). A simple characterization of plurality rule. *Journal of Economic Theory*, 71(1), 298–302.
- Condorcet, M. J. A. N. de. (1785). Essai sur l'application de l'analyse à la probabilité des décisions rendues à la pluralité des voix. Paris: Imprimerie Royale. Translated and reprinted in part in I. McLean and A. B. Urken (eds.), *Classics of Social Choice*, Chapter 6. Ann Arbor: University of Michigan Press, 1995.
- Daunou, P. C. F. (1803). Mémoire sur les élections au scrutin. Paris: Baudouin, imprimeur de l'Institut National. Translated and reprinted in I. McLean and A. B. Urken (eds.), *Classics of Social Choice*, Chapter 11. Ann Arbor: University of Michigan Press, 1995.
- Fine, B., and Fine, K. (1974a). Social choice and individual ranking I. *Review of Economic Studies*, 41(3), 303–322.
- Fine, B., and Fine, K. (1974b). Social choice and individual ranking II. *Review of Economic Studies*, 41(4), 459–475.
- Fishburn, P. C. (1970). Intransitive indifference in preference theory: A survey. *Operations Research*, 18(2), 207–228.
- Fogarty, J., and Sadler, R. (2014). To save or savor: A review of approaches for measuring wine as an investment. *Journal of Wine Economics*, 9(3), 225–248.

- Friberg, R., and Grönqvist, E. (2012). Do expert reviews affect the demand for wine? *American Economic Journal: Applied Economics*, 4(1), 193–211.
- Gergaud, O., Ginsburgh, V., and Moreno-Ternero, J. (2021). Wine ratings: Seeking a consensus among tasters via normalization, approval, and aggregation. *Journal of Wine Economics*, 16(3), 321–342.
- Ginsburgh, V., and Zang, I. (2012). Shapley ranking of wines. *Journal of Wine Economics*, 7(2), 169–180.
- Hulkower, N. (2009). The judgment of Paris according to Borda. *Journal of Wine Research*, 20(3), 171–182.
- Jaeger, E. (1981). To save or savor: The rate of return to storing wine. *Journal of Political Economy*, 89(3), 584–592.
- Lindley, D. V. (2006). Analysis of a wine tasting. *Journal of Wine Economics*, 1(1), 33–41.
- Luce, R. D. (1956). Semiorders and a theory of utility discrimination. *Econometrica*, 24(2), 178–191.
- Luce, R. D., and Raiffa, H. (1957). *Games and Decisions*. New York: Wiley.
- May, K. O. (1952). A set of independent necessary and sufficient conditions for simple majority decision. *Econometrica*, 20(4), 680–684.
- May, K. O. (1953). A note on the complete independence of the conditions for simple majority decision. *Econometrica*, 21(1), 172–173.
- Morales, J. I. (1979). Memoria matemática sobre el cálculo de la opinion en las elecciones. Madrid: Imprenta Real. Translated and reprinted in I. McLean and A. B. Urken (eds.) *Classics of Social Choice*, Chapter 10. Ann Arbor: University of Michigan Press, 1995.
- Richelson, J. T. (1978). A characterization result for the plurality rule. *Journal of Economic Theory*, 19(2), 548–550.
- Smith, J. H. (1973). Aggregation of preferences with variable electorate. *Econometrica*, 41(6), 1027–1041.
- Young, H. P. (1974). An axiomatization of Borda's rule. *Journal of Economic Theory*, 9(1), 43–52.
- Young, H. P. (1975). Social choice scoring functions. *SIAM Journal on Applied Mathematics*, 28(4), 824–838.

Appendix 1. Experts from Bordoverview

- The *Wine Advocate* was created in 1978 by celebrated expert Robert Parker. Lisa Perrotti-Brown is the American wine critic currently providing the ratings (starting with the Bordeaux 2017 vintage), publishing on <https://www.robert-parker.com>.
- Jancis Robinson is a British wine writer and critic who rose to fame in the mid-1980s after becoming the first Master of Wine outside the wine trade. She studied mathematics and philosophy at the University of Oxford. Jancis Robinson writes a weekly column for the *Financial Times* and publishes ratings on the web, where wines are also rated by Julia Harding. The websites are <https://www.jancisrobinson.com/> and <https://www.wine-searcher.com/critics-1-jancis+robinson>.
- Jeff Leve is an American wine critic. His reviews and ratings appear on the website <https://www.thewinecellarinsider.com>.
- Jane Anson is *Decanter's* wine correspondent in Bordeaux. *Decanter* was established in 2004 by English wine critic Steven Spurrier, who was at the origin of the famous *Judgment of Paris*.
- Chris Kissack is an English wine critic, publishing in *Winedoctor*.