wine economics

AMERICAN ASSOCIATION OF WINE ECONOMISTS

AAWE WORKING PAPER
No. 274
Economics

WINE RANKINGS AND THE BORDA METHOD

Salvador Barberà, Walter Bossert and
Juan D. Moreno-Ternero

May 2022

www.wine-economics.org

Wine rankings and the Borda method*

Salvador Barberà
MOVE, Universitat Autònoma de Barcelona
and Barcelona School of Economics
Facultat d'Economia i Empresa
08193 Bellaterra (Barcelona)
Spain
salvador.barbera@uab.cat

Walter Bossert
Centre Interuniversitaire de Recherche en
Économie Quantitative (CIREQ)
University of Montreal
P.O. Box 6128, Station Downtown
Montreal QC H3C 3J7
Canada
walter.bossert@videotron.ca

Juan D. Moreno-Ternero
Department of Economics
Universidad Pablo de Olavide
Carretera de Utrera, Km.1.
41013 Seville
Spain
jdmoreno@upo.es

This version: May 2, 2022

Abstract. We propose to use the classical Borda rule to establish wine rankings. Unlike alternative approaches, this method has a powerful attribute: it is well-defined even if the panelists' quality relations are not required to exhibit demanding properties such as completeness or transitivity. We apply the method to rank Bordeaux wines assessed by different experts, some of whom do not rate all of the wines under consideration. *Journal of Economic Literature* Classification Numbers: C18, D71, L15, L66.

Keywords: Wines; Ratings; Rankings; Opinions; Borda rule.

* 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), from the Spanish Ministry of Economy and Competitiveness and FEDER through grant PID2020-116771GB-I00, and from the Generalitat de Catalunya, through grant 2017SGR-0711. Juan D. Moreno-Ternero acknowledges financial support from 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.

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 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, 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]

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 opinion 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 on the basis of pairwise comparisons that do not need to be complete or transitive (or even acyclical). This allows the method to process the data generated by individuals who abstain from certain comparisons, or who form their judgements through reasonings that the behavioral literature has shown to be frequently used and yet contradicting classical notions of rationality. In addition, we also argue that the Borda count allows us to exploit databases that incorporate data from diverse contests to reach significant rankings.

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 in order 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.

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

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

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

assign numerical values to each of the wines. These jurors may be extremely competent and experienced but, nevertheless, some concerns seem to appear immediately. To begin with, it is perfectly plausible that some wines are very difficult to compare because the criteria used may be incommensurable—and, thus, an expert's judgment may quite possibly be incomplete. Moreover, 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 to be 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 induce natural violations of transitivity.

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,[3] but still in the context where experts are asked to submit complete and transitive quality relations. In that case, the rule can be described in several equivalent forms, and the one we propose here is the only one of them that is able to operate in more modest informational environments while retaining the desirable properties of Borda's proposal. We discuss these equivalences and the inadequacy of the other two variants when extending our scope to larger classes of admissible inputs.

2 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 on compare their informational requirements with the milder ones that are needed for the application of the Borda rule. Suppose there is a finite set $N = \{1, \ldots, n\}$ 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 usual consensus method employs the average rating of all experts to determine an overall quality rating. That is, for each wine $x \in X$, the overall usual consensus rating $r^{UC}(x)$ is defined as

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

Several modifications of this method can be obtained if a threshold is employed. The primary motivation behind the use of such a threshold is to exclude wines that perform

[3]See, for instance, Hulkower (2009) and Gergaud, Ginsburgh, and Moreno-Ternero (2021).

rather poorly in the individual ratings. Let π denote a threshold within the range of possible rating values.

The first of these methods eliminates all wines below the threshold and, moreover, homogenizes the ratings of the wines that remain so that they all receive the same numerical rating of one. Thus, the approval consensus rating $r_i^{\pi,AC}(x)$ assigned by expert $i \in N$ to wine $x \in X$ for the threshold value π is defined by

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

and the overall approval consensus rating $r^{\pi,AC}(x)$ of wine $x \in X$ is obtained as the arithmetic mean

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

The above-described method discards much of the information contained in the individual ratings supplied by the jurors—note that all wines above or at the threshold level are treated equally. To avoid this feature, the proportional approval consensus method can be employed. This alternative method also removes the wines below the threshold from consideration but it retains the original ratings above and at the threshold. The corresponding proportional approval consensus rating $r_i^{\pi,PC}(x)$ of expert i for wine x is

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

Again, the overall proportional approval consensus rating $r^{\pi,PC}(x)$ of wine $x \in X$ is given by the arithmetic mean of the individual modified ratings so that

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

It is straightforward to define relative variants of these two approval-based methods. The requisite overall relative approval consensus rating is obtained by dividing the overall (proportional) approval consensus rating by the sum of the individual (proportional) approval consensus ratings instead of using the total number of experts in the denominator.

3 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 (that is, 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. In particular, any binary relation will do; there is no need to require these relations to possess any richness properties such as completeness or any coherence properties such as acyclicity.

Each expert $i \in N$ is assumed to have a 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, 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 possess any properties such as completeness or transitivity (or even acyclicity).

The quality relations $R_1, \ldots, R_n$ can be collected in a quality profile $\mathbf{R} = (R_1, \ldots, R_n)$. 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 four experts $N = \{1, 2, 3, 4\}$ and a set of three wines $X = \{x, y, z\}$. A quality profile $\mathbf{R} = (R_1, R_2, R_3, R_4)$ is composed of the individual quality relations given by

$$
\begin{gathered}
yP_1z, zP_1x, yP_1x, \\
xP_2y, yP_2z, zP_2x, \\
zI_3x, xI_3y, zP_3y, \\
yP_4x.
\end{gathered}
$$

In this example, expert 1 submits a complete ordering for 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. It turns out that this is a situation that is perfectly acceptable as an input for the application of the method we shall present below. As for expert 3, there is an instance of intransitive equal goodness in the quality ranking delivered; (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. Finally, expert 4 expresses only a single opinion—he or she considers wine y to be better than wine x but is silent regarding comparisons between z and x and between z and y. We note that, in our flexible setting that allows for incomplete and intransitive individual quality relations, it does not matter whether the absence of a betterness relation between two wines is interpreted as equal goodness or as non-comparability; the resulting rankings according to the method presented below are the same for both of these options. More generally, instances of equal goodness do not influence the wine ranking to be established by means of the Borda method we shall present; what matters is the number of times a wine is considered better than another.

The method we propose is not new by an means, although the arguments we offer here are novel, as they refer to its extension to additional data sets. But the idea that each single instance of betterness has value in itself and should be counted as one out 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.[4] 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.

4 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 (that is, 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 $\mathbf{R} = (R_1, \ldots, R_n)$. 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 above, the individual Borda score of wine x according to expert i is given by

$$b_i(x) = |\{z \in X \mid xP_iz\}| - |\{z \in X \mid zP_ix\}|,$$

where the notation $|S|$ is used to indicate the number of elements in the finite set S. Note that this difference is unaffected if the quality relation R_i is used in place of its associated betterness relation P_i. This is the case because whenever we have an instance of equal

[4] See Barberà, Bossert, and Suzumura (2021) for a detailed discussion.

goodness between wines x and y, it follows by definition that xR_iy and yR_ix cancel each other out when calculating the above difference. This argument applies to any two wines x and y, no matter whether x and y are distinct or identical. Either or both of the two sets in the definition may be empty. Furthermore, note that these individual Borda scores are well-defined even if the individual quality relation does not possess any regularity properties such as completeness or 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 $b(x)$ of a wine x is obtained by adding the individual scores over all experts so that

$$\begin{aligned} b(x) = \sum_{i=1}^{n} b_i(x) &= \sum_{i=1}^{n} (|\{z \in X \mid xP_iz\}| - |\{z \in X \mid zP_ix\}|) \\ &= \sum_{i=1}^{n} |\{z \in X \mid xP_iz\}| - \sum_{i=1}^{n} |\{z \in X \mid zP_ix\}|. \end{aligned} \tag{1}$$

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,

$$xR^By \Leftrightarrow b(x) \geq b(y).$$

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 incomplete and intransitive 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

$$|\{z \in X \mid zP_ix\}| + |\{z \in X \mid xP_iz\}| = m - 1,$$

is valid for all wines x if R_i is a strict complete relation defined on the set of m wines X. In the absence of completeness, however, this is not the case—removing the second part of the difference in (**??**) leads to a different criterion to rank the wines. In the context of quality relations that are not necessarily complete, this second option leads to highly undesirable consequences. For example, suppose that there are four experts $N = \{1, 2, 3, 4\}$ and six wines $X = \{x, y, z, u, v, w\}$. Furthermore, consider the quality profile $\mathbf{R} = (R_1, R_2, R_3, R_4)$ given by

$$\begin{gathered} xP_1z, xP_1w, \\ yP_2z, yP_2w, \\ uP_3v, \\ vP_4u; \end{gathered}$$

thus, expert 1 favors wine x over wines z and w, expert 2 considers wine y to be better than wines z and w, wine u is better than wine v according to expert 3, and expert 4 expresses a single opinion that is the reverse of that submitted by expert 3.

According to the alternative ranking that is obtained by counting the number of wins only, it follows that x and y are equally good because the number of wines beaten by x and by y are the same—x beats two wines in expert 1's quality relation and y beats two wines in expert 2's quality relation. Now consider another quality profile $\mathbf{R}' = (R'_1, R'_2, R'_3, R'_4)$ defined by

$$\begin{gathered} xP'_1z, xP'_1w, \\ yP'_2z, yP'_2w, \\ uP'_3v, zP'_3y, wP'_3y, \\ vP'_4u, zP'_4y, wP'_4y. \end{gathered}$$

Experts 1 and 2 have the same quality relations as in the previous quality profile but now both experts 3 and 4 favor wines z and w over wine y, in addition to their original assessments of wines u and v. If we count the number of wins only but do not modify this score by subtracting the number of losses suffered by a wine, x and y are again equally good because their respective number of wins is the same as in the quality profile $\mathbf{R}$. This seems difficult to accept because now wine y is beaten by z and by w in the quality relations of experts 3 and 4, whereas x does not suffer any such losses. The use of the criterion expressed in (**??**) avoids this troublesome conclusion.

A second alternative definition of the Borda rule consists of considering it as a special case of a scoring method.[5] Assuming that the individual quality relation of each 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. 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 $m-1$ position, and its weight is $0-(m-1) = -(m-1)$. Thus, according to the Borda method, the positional weights $w_1, \ldots, w_m$ assigned to the m positions from top to bottom are given by

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

Adding over all experts yields the overall Borda scores as defined in (**??**). A generalization of this method consists of the class of scoring methods. These are obtained by assigning arbitrary weights $w_1, \ldots, w_n$ to the positions, where the restriction $w_1 \geq \ldots \geq w_n$ with at least one strict inequality is usually imposed to ensure that better positions cannot receive

[5] See Young (1975).

lower weights than worse positions. Because the notion of a position cannot even be defined if the experts' quality relations fail to be orderings, it is clear that this alternative method cannot be applied in our informationally austere framework. Thus, with the exception of the Borda rule, general scoring methods are not suitable for the purposes of this paper and, therefore, we cannot treat it as a special case of these more general rules.

5 Properties of the Borda method

5.1 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 relation defined on the set of wines to be judged, without having to require this relation to possess any properties that are often required. The diversity in the criteria that the jurors are likely to invoke when forming their assessments can easily lead to instances of incommensurability and, as a consequence, several of the wines under consideration may turn out to be extremely difficult to compare with a high degree of confidence. In addition, violations of transitivity are bound to arise if wines are difficult to distinguish with respect to some of their attributes. Thus, the observation that the experts are able to 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 seem from their definitions, all of the consensus methods illustrated in Section 2 rely not only on individual orderings but even on numerical ratings provided by the experts.

5.2 Merging 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.[6] The property states that if the overall quality relations obtained for 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.

[6]See also Fine and Fine (1974a,b).

5.3 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 cancelled 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 (**??**). Moreover, if one expert consider 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 the method.[7]

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 opinions xP_iz and yP_iz cancel each other out when it comes to the relative ranking of x and y—the overall Borda score of both of these wines 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 the expert i and expert j are the same person or two different individuals. To go one step further, if there are two opinions xP_iz and yP_iw with w and z being distinct, the overall ranking of x and y remains the same if the two opinions are removed, 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 above, 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 of them so that this elimination process may turn out to be superfluous.

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, and no matter 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.

[7] See also Barberà and Bossert (2021).

5.4 Additional favorable opinions

Suppose that a quality profile is augmented by an additional favorable opinion for a wine x against another wine z, all else unchanged. This could be the consequence of an expert changing his or her initial assessment of non-comparability 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 removed (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 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 the overall quality relation is assumed to be transitive and, moreover, all wines are equally good if no expert expresses any opinion. Thus, the only method that respects these intuitively appealing conditions is the Borda method.[8]

6 An application to Bordeaux 2020 future wines

We conclude with an application of our method to a real-life scenario—the recent tasting of some 2020 Red Bordeaux gathered on *Bordoverview*.[9] This tasting involves ratings for some Bordeaux 2020 future wines produced by nine international experts: the *Wine Advocate* (WA), Neal Martin (NM), Jancis Robinson (JR), Tim Atkin (TA), Michel Bettane and Thierry Desseauve (BD), Jeff Leve (JL), *Decanter Magazine* (DM), *Revue du Vin de France* (RV), and Chris Kissack (CK).[10]

Table 1 exhibits a selection of 12 wines with the ratings assigned by the above experts. Because the Borda method merely requires a 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 90–92, we took this to mean that the wine in question is considered better than a wine with a rating of 90 and worse than a wine with a rating of 92. In one case, there appears a rating of 97 for one wine and a rating of 96–98 for

[8]See Barberà and Bossert (2021) for details.

[9]See `https://www.bordoverview.com`.

[10]See the appendix for some details on the nine experts.

another by the same expert, and we treated these two wines as being considered equally good. Analogously, a wine with the rating 16+, for instance, was treated as better than a wine with a rating of 16, for instance. 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. A blank entry means that the corresponding wine was not rated by the corresponding expert. As the table illustrates, some experts did not rate all of the wines in the sample; in fact, only Michel Bettane and Thierry Desseauve provide a rating for each one. The fact that ratings are not complete for all experts would prevent all the standard aggregating methods in Section 2 from providing an overall ranking of all the wines in the sample. In other words, the lack of completeness would require to truncate the sample of wines or experts in order to use standard aggregating methods. As already pointed out, the Borda method that we advocate does not require the experts' induced relations to be complete so that we can work directly with these individual relations.

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 each of the experts. For instance, Clos Lunelles is assigned an individual Borda score of five by expert Jeff Leve because, according to Table 1, Clos Lunelles beats seven other wines and is beaten by two other wines so that the score is given by $7 - 2 = 5$. Zero entries result either from non-comparability (as is the case, for example, for Clos Puy Arnaud and the expert Tim Atkin) or from recording an equal number of pairwise wins and losses (see, for example, the individual Borda score of $2 - 2 = 0$ given to Guiraud by *Decanter Magazine*). The latter possibility illustrates the cancellation property of the Borda method alluded to in Subsection 5.3.

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 12 wines according to the Borda method. Cos d'Estournel comes in first with an overall Borda score of 64, followed by Valandraud with a score of 47 and so on, until we reach Joanin-Bécot in the next-to-last position with a score of -24 and, finally, Cap de Faugères at the bottom with a score of -44. 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.

Thus, thanks to the flexibility and the power of the Borda method, we can provide a complete ranking for the full sample of 12 wines, even though only one of the nine experts rates them all. This could not be achieved with any other standard aggregation method.

Wines	WA	NM	JR	TA	BD	JL	DM	RV	CK
Cos d'Estournel	96–98	95–97	18.5	96	95	97–99	97	98–100	95–97
Valandraud	95–97	95–97	16.5+	93	94–95	97–99	96	96–98	95–97
Clos Lunelles	92–94	89–91	16+	89	93–94	94–96			91–93
Coutet		92–94	17+	93	93	94–96	94	97	93–95
Clos Puy Arnaud		90–92			92–93	88–90			91–93
Guiraud		91–93	16.5+	92	92–93	93–95	94	93–94	90–92
Joanin-Bécot	90–92	87–89	16		91–92	92–94			
Virginie de Valandraud		91–93		90	91	91–93	91		91–93
D'Aiguilhe		90–92		89	90–91	93–95			91–93
Monbousquet	91–93+	90–92	16+	90	90–91	93–95	90	91–92	92–94
Mont-Pérat					90				
Cap de Faugères		87–89		86	87–88	91–93			83–85

Table 1: Ratings from nine experts for 2020 Red Bordeaux wines.

Wines	WA	NM	JR	TA	BD	JL	DM	RV	CK
Cos d'Estournel	4	9	6	8	11	9	5	4	8
Valandraud	2	9	1	5	9	9	3	1	8
Clos Lunelles	0	–6	–3	–5	7	5	0	0	–2
Coutet	0	6	4	5	5	5	0	1	5
Clos Puy Arnaud	0	–2	0	0	2	–10	0	0	–2
Guiraud	0	3	1	2	2	0	0	–2	–7
Joanin-Bécot	–4	–9	–6	0	–1	–4	0	0	0
Virginie de Valandraud	0	3	0	–1	–3	–7	–3	0	–2
D'Aiguilhe	0	–2	0	–5	–6	0	0	0	–2
Monbousquet	–2	–2	–3	–1	–6	0	–5	–4	3
Mont-Pérat	0	0	0	0	–9	0	0	0	0
Cap de Faugères	0	–9	0	–8	–11	–7	0	0	–9

Table 2: Individual Borda scores.

Wines	Scores	Ranks
Cos d'Estournel	64	1
Valandraud	47	2
Clos Lunelles	–4	5
Coutet	31	3
Clos Puy Arnaud	–12	7
Guiraud	–1	4
Joanin-Bécot	–24	11
Virginie de Valandraud	–13	8
D'Aiguilhe	–15	9
Monbousquet	–20	10
Mont-Pérat	–9	6
Cap de Faugères	–44	12

Table 3: Overall Borda scores and Borda ranks.

Appendix. 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.robertparker.com`.

- Neal Martin is a English wine critic, formerly contributing to the *Wine Journal* on `https://www.robertparker.com`, and since the 2017 vintage to *Vinous.*

- 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`.

- Tim Atkin is a UK-based Master of Wine and wine journalist with an international following, publishing on his website `http://timatkin.com`. After training in modern languages at the University of Durham, Atkin soon moved into a career of wine writing. An additional site is `https://www.wine-searcher.com/critics-34-tim+atkin`.

- Michel Bettane and Thierry Desseauve are French wine critics publishing in *TAST.*

- Jeff Leve is an American wine critic. His reviews and ratings appear on the website `https://www.thewinecellarinsider.com`.

- *Decanter* was established in 2004 by English wine critic Steven Spurrier, who was at the origin of the famous *Judgment of Paris.* In addition to publishing wine ratings on `https://www.wine-searcher.com/critics-44-decanter+world+wine+awards`, he also awards trophies and medals. Jane Anson is *Decanter*'s wine correspondent in Bordeaux.

- The *Revue du Vin de France* is a monthly French wine magazine which started in 1927. It specializes in French wines, and is highly regarded by the nation's wine industry. The wines are tasted by Olivier Poels, Hélène Durand, and Philippe Maurange.

- Chris Kissack is an English wine critic, publishing in *Winedoctor.*

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.

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. (2021). Aggregating opinions. Barcelona GSE Working Paper No. 1262.

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 McLean, I., and Urken, A.B., eds. (1995). *Classics of Social Choice.* Ann Arbor: University of Michigan Press, Chapter 5.

Cao, J. (2014). Quantifying randomness versus consensus in wine quality ratings. *Journal of Wine Economics,* 9(2), 202–213.

Cardebat, J., and Paroissien, 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 McLean, I., and Urken, A.B., eds. (1995). *Classics of Social Choice.* Ann Arbor: University of Michigan Press, Chapter 6.

Daunou, P.C.F. (1803). Mémoire sur les élections au scrutin. Paris: Baudouin, imprimeur de l'Institut National. Translated and reprinted in McLean, I., and Urken, A.B., eds. (1995). *Classics of Social Choice.* Ann Arbor: University of Michigan Press, Chapter 11.

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.

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.

Morales, J.I. (1797). Memoria matemática sobre el cálculo de la opinion en las elecciones. Madrid: Imprenta Real. Translated and reprinted in McLean, I., and Urken, A.B., eds. (1995). *Classics of Social Choice.* Ann Arbor: University of Michigan Press, Chapter 10.

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.