A Comprehensive, Conclusive Analysis Of Ordinal Voting Methods

By Roy A. Minet

© 2018 - 2019 by Roy A. Minet. All rights reserved.

1. Introduction

Although it certainly could be argued that there are better ways to make good decisions, we employ the mechanism of voting when it is important to diffuse decision-making power among large numbers of people and thereby prevent its concentration within a small faction or an individual. Clearly, the primary objective of voting must be to make the best possible decision as it affects all those who vote. We have no choice but to assume (regardless of veracity) that those voting are at least somewhat interested, informed, rational and therefore able to contribute toward making a good decision. The best possible decision must then be defined as the outcome which results in the highest satisfaction, net of dissatisfaction, summed over all those who voted.

The simplest possible case is a binary decision, a referendum. Suppose that 51% of voters vote "No" and 49% vote "Yes" on some referendum. By traditional "majority rule" the measure is defeated.

However, further suppose that the 51% who voted "No" don't really have a very strong opinion regarding this particular issue. On the other hand, the 49% who voted "Yes" believe it to be a super-good thing and favor it very strongly. In this case, the majority rule decision to defeat the referendum would not likely be the best choice by the previously stated definition.

Now suppose that a machine (a "satometer") is available which is trusted by voters and which can accurately measure each voter's satisfaction regarding the issue at hand on some absolute scale of "sats" (dissatisfaction would be negative "sats"). If the net total sats for all voters is positive, the referendum result is "yes" and if total sats is negative, the result is "no" (a zero-sat tie is resolved by majority rule, subtracting the number of voters who registered negative sats from those who registered positive sats). This hypothetical referendum illustrates the ideal best possible decision that we aspire to most closely approach in real-world voting. It is based upon complete, accurate, comparable and sincere information from each voter on the given issue.

Considering now the general case where voters are to select n of m options, we set up our voting booth satometer to grab a quick reading on each of the m options from each voter. In essence, it's a separate referendum on each option. This is indeed a large and complete data set. The winners will be the n options with the highest net sat totals (or in extremely sad cases, the positive-most or least negative totals). A bitter lesson swiftly learned is not to nominate candidates with large negatives. Buh bye Donald J. Trump and Hillary R. Clinton; evidence indicates that both almost undoubtedly would have had negative net sat totals in the 2016 election. The sun breaks forth from behind the turbulent dark clouds of polarization and the lands grow happier and more peaceful.

Lacking a satometer, obtaining such information-rich and *sincere* cardinal data from voters is extremely problematic. At least one method has been proposed (True Weight Voting) which promises to obtain and utilize valid cardinal data. Possibly, it could closely approach the ideal behavior described above. However, it is untested and enough of a "radical" change that it is unlikely to find acceptance in the near term. Therefore, the remaining discussion will focus on how best to utilize more familiar and easier-to-collect ordinal data, but will always strive to most closely approach the above ideal decision model.

Any voting method can be considered as two steps: 1) gathering data from voters, and 2) processing that data in some way to yield the result. These steps are not entirely independent. Obviously, the data collected can affect the processing and vice versa. Perhaps just slightly less obvious is the fact that voters' understanding of the processing can affect the sincerity of the data they provide. However, if it can be shown that, for a certain combination of the two steps, the best possible data is being gathered and that this data is being processed in the best possible way to yield the result, then we can claim to have defined the best possible ordinal voting method.

2. Gathering Data From Voters

It is abundantly obvious that, when restricted to ordinal data, there is, unfortunately, way much less of it. We have only a pale shadow of the satometer's rich and complete data from which we must infer and estimate as best we can what the missing cardinal data most likely looks like and what decision it would render. Here are some important considerations affecting this step:

- Since the data which can be obtained is so limited, it is critical that all useful ordinal data be collected and utilized
- Voters must understand what information they are to provide and, to a reasonable extent, how it will be used to determine the result
- Voters must be capable of providing valid data and be willing to provide it
- Voters must be motivated to provide sincere (and not insincere or "strategic") information

It will not be possible to obtain a complete ranking of all options from all (or even most) voters. If write-in choices are freely allowed, it is not even theoretically possible as voters have no way of knowing what or how many different write-ins other voters may have supplied, not to mention knowing enough about them to intelligently rank them. Also, many voters will only have a first choice, or possibly a first and a second choice in mind. They may not have sufficient knowledge of other options to be able to rank them on any rational or meaningful basis. It is counterproductive to force voters to provide rankings as the data will be just unhelpful noise which can't improve and likely would degrade decision quality.

Clearly, voters must be allowed to voluntarily supply their first choice as the most important single datum. Equally clearly, a single choice is not enough, as it would limit us to brain-dead Plurality. So, voters must also be allowed to voluntarily supply a second choice, if they have one. For "good measure," voters will be allowed to voluntarily supply a third choice, but that will be the limit. Few will likely have a third choice, but it will be valid data to be considered if they are motivated enough to provide it. A vanishingly small number of voters are likely to have a fourth choice. As will be evident in the next section, the impact upon the decision that fourth and lower choices could have is miniscule, so we need not mourn their absence.

Although not by mathematical rigor, straightforward logic applied to the characteristics of real-world voters and elections supports reasonable confidence that an optimum data set to collect has been specified. The remaining task as it relates to collecting data is to make sure its processing motivates voters to supply sincere and not strategic information.

3. Voter *Dis*satisfaction

With a satometer, it was possible to accurately measure both satisfaction and dissatisfaction on the same scale. Use of both enabled greatly improved decisions, "encouraged" the nomination of more widely acceptable candidates and promised to, over time, generally reduce polarization. (Anyone for True Weight Voting?) While collecting ordinal data, it might be possible to give voters a way to, say, indicate dissatisfaction with one candidate. MRCV does not do so for good reasons. First, it is not strictly ordinal data. Second, there is no way to reasonably weight the magnitude of dissatisfaction relative to positive satisfaction choices. Third, it is a more confusing option for voters than simply indicating choices 1, 2 & 3. Fourth, it is an invitation for voters to attempt to use this option strategically, thus possibly degrading data quality.

4. Processing The Data

Consider a voter's first choice. A satometer would have measured some absolute number of sats. Since that number is not available, represent this as S_1. Worse, the information is not available for any voter, so we are forced to assume that the first choices of all voters are the same, namely S_1 sats (although certainly they are different).

Similarly, we don't know the strength of the voter's support for the option ranked second, so we label that, S_2. Furthermore, we can calculate the value of S_2 for any given voter from the value of S_1 (for the same voter) by multiplying S_1 by some factor we will call a relative weighting factor. That is, $S_2 = wS_1$, where w is the weighting factor. What can be said about the value of w?

At one extreme, the voter may have liked the top two choices equally and may have flipped a coin to decide which option to rank first. In this case, $S_2 = S_1$ and, therefore, $w = 1$. We can be sure that w can't have a value greater than 1; if it did, the voter would have ranked the second option first.

At the opposite extreme, the voter may not have liked the second choice option much at all and may have seriously considered leaving that option out of the ranking altogether. Since, in fact, the voter's decision was to include the option, it must enjoy some positive amount of support (however small) from that voter. Therefore, *w* must have a value greater than zero, so $w > 0$.

Thus, $0 < w <= 1$. Again, since we have no better information, we are forced to assume that *w* is the same for all voters and all rankings, although we know that it surely is not. However, the situation may not be as bleak as that sounds. We can still make reasonable decisions when statistically large numbers of voters are involved if we know the average value of *w* for all voters. It might be possible to design some tests to measure the average *w* for lots of voters in various elections and races (indeed, that might be a fun exercise worth the effort). Statistically, it is highly unlikely that the average *w* would either be close to 1 or close to 0. In fact, the best assumption that can be made is that the average value of *w* is 0.5. That is, averaged over lots of voters for many races in many elections, voters' support for the options they rank second is half as strong as their support for the option they rank first.

By repeating exactly the logic we have just been through, we can conclude that the weight of a third choice should be assumed to be half of the weight of a second choice (and, therefore, a quarter of the weight of a first choice). In fact, this can be extended to as many choices as one would like (i.e., a fourth choice is half the weight of a third choice or an eighth of a first choice, etc.) Even if valid data could be obtained for such lower choices, their influence on the decision declines exponentially so we need not be concerned by the absence of this data.

By using proper weightings for the choices, a "scoring system" can be constructed that will rank all options in the order of their support totaled over all voters. Most people find it easier to deal with integers rather than decimal fractions, so 4 "points" are awarded for a first choice, 2 "points" for each second choice and 1 "point" for each third choice, which accomplishes the proper weightings. The option receiving the highest point total will be the choice that results in the greatest voter satisfaction. There can be no better way to utilize ordinal data.

The scoring system could very well be used to directly identify the winning option. To do so would have the advantage of making the method precinct summable, and would indeed be a strong voting method. However, it is even better not do this. Instead, the scoring system is used to very reliably identify the weakest option, which then is eliminated in an iterative, last-man-standing system (similar to Baldwin or IRV – Instant Runoff Voting). This will lead to exactly the same final result often, but not always. The method was proposed in 2007 and is called MRCV (Minet Ranked-Choice Voting).

There are several reasons for the iterative eliminations approach:

- It is very important for all voters to clearly understand that, if their first choice is eliminated, their second choice is "promoted," and will then count in every way as though it had been their first choice. It is this which exorcises the irresistible pressure to vote insincerely for the "lesser-of-two-evils."

- It would be a hard sell to convince people to completely abandon majority rule, so the iterative system eliminates options until one has a majority of the first choices (or, more rarely, is the only one remaining). In fact, the scoring system is not even used if an option initially has an outright majority of first choices. It is easier for voters to see and feel comfortable that MRCV supports majority rule.
- If the scoring system "errs" for any reason, it could pick the wrong winner. However, a small error in eliminating the worst option is not a problem, as long as the error is not so egregious as to eliminate the option that should win (a huge problem with IRV). The method normally still converges on the same correct winner. Thus, a successive elimination procedure can be seen as "safer."
- If a Condorcet winner exists, it usually will have the highest point total, but not always. However, a Condorcet winner cannot have the lowest point total and so cannot be eliminated. Therefore, the eliminations procedure guarantees election of a Condorcet winner, if one exists. (Also see section 9 and Appendix.)
- Meticulously eliminating the weakest option one at a time makes MRCV as resistant as possible to irrelevant alternatives (IIA). It also optimizes handling of clone situations.

5. Concise Definition of MRCV Procedure

Each elector may rank his or her first, second and third choices. Any choice may be a write-in. A choice may not be repeated. Electors are not forced to rank three choices, but are limited to a maximum of three. The outcome will be determined by the following (sometimes) iterative procedure:

1. **Determine Winner** – Total the first choices. If any option has a majority of the first choices, it is the winner. If only one option remains, it is the winner. If there is no winner, proceed to step 2.
2. **Eliminate Weakest Option** – Assign weighting points to the choices on each ballot. A first choice receives 4 points, a second choice receives 2 points and a third choice receives 1 point. Total the points for each option across all ballots. Eliminate the option having the lowest point total. If there is a tie, eliminate the tied option having the smallest total of first choice votes. If still a tie, eliminate the tied option that has the smallest total of second choice votes. If a tie persists, eliminate one of the tied options at random.
3. **Promote Choices** – If a ballot's first choice has been eliminated, promote the second choice (if any) to first. For any ballot where the second choice has been promoted or eliminated, promote the third choice (if any) to second. Any eliminated or promoted choice that has no lower choices will cease to exist. Return to step 1.

6. The Jones Rule

The (Douglas) Jones rule states that anything having to do with elections must be understandable by a bright high school student. At least a majority of voters need to be able to understand the voting system in order to have full confidence in it. This is particularly so for judges of elections, poll watchers and all those responsible for running, supervising or auditing elections. Legislators should not enact anything they can't understand.

Simplicity likely is the main reason Plurality is still widely used. Unfortunately, some increase in complexity is absolutely necessary for an improved system. Fortunately, the above MRCV procedure is simple enough that it can be well understood by most when explained. It probably squeaks in under the Jones rule.

Hybrid systems have been proposed which usually begin with pairwise comparison, but switch to some other eliminations method(s) if no Condorcet winner exists. It is possible that the performance of MRCV may be approached, but such complex methods surely do not satisfy the Jones rule.

7. Implementation

MRCV certainly could be done using rustic means such as hand-marked paper ballots. However, use of modern technology is strongly recommended, as long as it is carefully done (discussed extensively elsewhere). A good implementation can handle precinct non-summability smoothly and confer other advantages as well. Very briefly, the flow is:

1. Voters vote electronically. The equipment produces a voter-verifiable paper ballot.
2. Each voter reviews and verifies the paper ballot, then drops it into a traditional ballot box.
3. When the polls close, a check-summed XML file of all ballots (randomized) is produced.
4. The XML file is printed and posted at each polling place, and is also made publicly available.
5. The XML files from all polling places are processed at a central location to yield official results.
6. The paper ballots are used to verify each polling place's XML file. The central processing can be verified by anybody anywhere.

A comprehensive software implementation of MRCV is available called Election Manager. EM is capable of running entire elections. Reliable final results are available minutes after the last poll closes. All results can be verified (completely or statistically). EM may be used royalty-free for non-commercial purposes.

8. Other Interesting Observations

The value chosen for w has a profound effect on the behavior of ordinal voting. It is crucial to use the optimum value of 0.5.

As w is decreased from 0.5, choices after the first have less and less influence over the outcome. When decreased all the way to $w = 0$, they have no influence and only first choices count. At $w = 0$, behavior is the same as Plurality and the IRV elimination method.

On the other hand, if w is increased, choices after the first exert a larger and larger influence upon the outcome. If increased all the way to $w = 1$, all choices count equally. Voters then lose the ability to specify which choice is their first. This is Approval Voting, which is inferior from the start since it fails to collect the most important first choice information (or in order to indicate a first choice, the voter must sacrifice the ability to vote for more than one candidate, which is what Approval Voting is supposed to be all about).

Thus, MRCV might be thought of as the optimum midpoint between Approval Voting on one side and Plurality/IRV on the other.

9. But What Of Condorcet Winners, Favorite Betrayal, Participation, Paradoxes, IIA, Etc.?

Many contributors have produced a huge body of work over the past 250 years. There have been some interesting results, guidelines, theorems and voting paradoxes. However, this colossal amount of time and effort has failed to produce a solid consensus recommending any specific voting method. Various factions advocate different methods. Meanwhile, too many people in the world still suffer with Plurality. Why?

It has been shown that, in the perfect mathematical world and under some circumstances, every possible voting method will yield decisions which are "unfair" by some quite reasonable criteria. Every possible method also is in some way and to some degree susceptible to manipulation by strategic voting. (See results by Arrow, Gibbard, Satterthwaite, etc.) A certain "paralysis" prevails with debates continuing as to which methods suffer from which defects and which defects are worse than which other defects. There is, however, some concern regarding how important said defects may or may not be in real-life elections. In order to do the math, some assumptions must virtually always be made.

It is not possible in real-world elections to have a complete ranking of all options by all voters. Even if complete rankings could in some way be forced, the data would then have a lower signal-to-noise ratio. The real world is a messy place. As has been explained, the best data that can be obtained from voters is very ragged, consisting of just first choices from some voters, first and second choices from others and perhaps three choices from a few. Therefore, any mathematics based upon the assumption that a complete ranking is available cannot be applied to real-world voting methods because at least one of the assumptions it is based upon is not valid.

A second choice carries only half the weight of a first choice, and a third only half of a second choice. Interchanging a second and third choice quite properly has much less impact on deciding the winner than swapping a first and a second choice. Therefore, any mathematics based upon the assumption that all choices carry equal weight cannot be applied to real-world voting methods because at least one of the assumptions it is based upon is not valid.

Consider the concept of a "Condorcet winner." A voting method which does not always elect a Condorcet winner when one exists is widely distained and unceremoniously cast upon the rejects pile. The existence of a Condorcet winner is determined by pairwise comparisons where a voter ranking of A over B in the 3rd and 4th places (or even the 98th and 99th places) carries the same A-versus-B weight as a voter ranking of A over B in the 1st and 2nd places! This is not reality.

Whenever an option receives a majority of first choices, it is decisively a Condorcet winner and clearly must be respected. But any reasonable voting method handles this simple case correctly, even Plurality. The challenge for voting methods is to also render the correct result when no option has received a majority of first choices. In hypothetical cases with up to four options and complete ranking data, MRCV elects a Condorcet winner if one exists. It is possible to contrive examples with more than four options where MRCV may not elect a Condorcet winner, but one must question the relevance of a Condorcet winner in such examples (or, indeed, the relevance of that entire exercise).

Finally, it should be emphasized that "fairness" must not be the overriding criterion for a voting method. As stated in the introduction, the primary objective must be simply to render the best possible decision – that result which maximizes voter satisfaction. One would not expect these two desirable characteristics be strongly at odds with each other. However, wherever a conflict may exist, it must be resolved in favor of maximizing voter satisfaction, not "fairness."

Certainly, after 250 years, a different approach is called for. Sincerely thank the mathematicians; they've had their chance. Physicists and engineers must now devise a straightforward solution which optimizes the real-world data collection and data processing mechanism of voting. A solution is needed which can be put into service as soon as possible to replace Plurality (and IRV wherever that method has been implemented).

It is argued that MRCV is that best possible ordinal voting method because it collects the most valid and useful data from voters, then does the best job possible of utilizing the data to yield the decision which maximizes voter satisfaction. That's what is important. <u>No claim is made that MRCV is perfect. The point is that whatever foibles, defects or weaknesses MRCV may have, they are minimized to the greatest degree possible.</u> <u>In order to refute the claim that MRCV is the best possible ordinal method, it is necessary to specifically define some other ordinal method which can be shown to gather better data and/or process the data in such a way that consistently produces decisions which result in higher voter satisfaction than MRCV.</u>

Of course, if MRCV truly is the best real-world voting method, one should expect it to also work very well in the theoretical mathematical world. And it does. The appendix which follows revisits a substantial collection of voting problems and paradoxes to compare the performance of MRCV with other methods. The reader is urged to try MRCV in any kind of election, real or contrived. It always renders reasonable choices and does not produce an awful result under any circumstances.