Follow-on Election Simulation Leads to Definitive Proposal

By Roy A. Minet (Rev. 02/19/20)

[Abstract: A 2019 election simulation study shed considerable new light on, and understanding of how various voting methods work, largely clarifying what is important and what is not. This follow-on study builds upon the first and concludes with a definitive voting method proposal in a bid to bring the 250-year-long debate to a productive (and overdue) conclusion.**]**

Background and Definitions

The serious shortcomings of Plurality voting have been recognized for well over two centuries. The ensuing 250-year debate has produced a healthy paper publishing industry, universal agreement that Plurality is awful and needs to be replaced, but no solid consensus on a best replacement. Many experts (including this author) believe Plurality is an exacerbating factor to political polarization which is increasing to alarming levels in the US. Instead of just publishing papers for another century or two, Plurality needs to be replaced soon, if not with a "perfect" method, at least by one that makes a large improvement.

During 2019, an extensive election simulation study[1] was conducted which was aimed at understanding how various voting methods work while clarifying what is important and what is not. (The manner in which elections are simulated is described in detail in the 2019 paper.) Only the most pertinent elements of that study will be summarized here. This project ties up a loose end to the prior study, but most importantly, builds upon its results with the objective of arriving at an actionable proposal for a practical voting method to actually replace and very significantly improve upon Plurality (and IRV wherever implemented).

Before attempting to design anything (certainly including a voting method), it is essential to clearly define what the design is intended to accomplish. Surprisingly, this step seems to sometimes be skipped or glossed over. The following two foundational definitions underlie all the reported simulation work:

> **The primary design objective for an election mechanism must be for it to most consistently render the best possible decisions** (with the caveat that decision-making power be kept reasonably dispersed).
>
> **The best possible decision is that result which maximizes voter satisfaction, net of dissatisfaction, when summed over all voters who voted.**

[1] See "Election Simulation Sheds New Light on Voting Methods" at http://royminet.org/voting-elections/

A large amount of effort has been expended over the years investigating peripheral issues – the "fairness" of voting methods comes immediately to mind. Such issues are subsidiary and important only to the extent that they actually do affect consistently rendering the best possible decisions.

"The Handbook of Good Voting Method Design"[2] recommends the following reasonably straightforward three-step process:

1. Acquire a good understanding of how various voting methods work with sincere voters.
2. Select a few of the best performing methods. Analyze their susceptibility to manipulation by strategic voters and see if they can be made more resistant to such attacks.
3. Consider and optimize the human interface (presentation, ballots, instructions, etc. for voters and election officials). Note that any voting method suitable for use in public elections will need to satisfy the Jones[3] rule: Everything about voting must be understandable by a reasonably bright high school student. Most voters must understand it. Election officials and polling place workers certainly must understand it; and, of course, the politicians who would pass the laws to implement it will need to understand it.

The 2019 Election Simulation Project

Here is a summary of some of the most pertinent conclusions drawn from the 2019 simulation project which are valid to the extent that real-world elections were meaningfully simulated:

- Plurality is indeed every bit as bad as everybody already knows it is.
- IRV is only a small improvement over Plurality, and then only in elections which have 4 or more candidates.
- Pairwise (with IRV fallback) is only microscopically better than plain IRV, in spite of being much more complex. Indeed, the neat concept of a Condorcet winner which has received so much attention over the years does not appear to be at all important in real elections.
- A previous paper[4] argued that MRCV is the best possible ordinal voting method. The simulation confirms this in that MRCV is considerably better than IRV or Pairwise/IRV and no better ordinal method is known.
- No ordinal (ranked choice) method can achieve better than "mediocre" performance.
- The simplest cardinal method, Score1 (which is the same as Approval) matched or slightly exceeded the performance of the best ordinal method (MRCV).
- Although Score1/Approval substantially duplicates the reasonably good performance of MRCV with sincere voters, its ability to choose the correct winner deteriorates quite rapidly with insincere voting, especially in elections which have 4 or more candidates.
- Good cardinal methods can perform much better than any ordinal method.

---

[2] Does not currently exist!

[3] Douglas W. Jones, University of Iowa, Department of Computer Science

[4] See "A Comprehensive, Conclusive Analysis of Ordinal Voting Methods" at http://royminet.org/voting-elections/

- Not at all surprisingly, there are many cases where a voter has no opinion on a particular candidate. That is, the voter either does not know enough or care enough about some candidate to even indicate an opinion on the ballot. Approximately 25% of all possible voter/candidate opinions had "no opinion" in 2-candidate elections and the percentage increases to about 45% for 7-candidate elections. It is important to simulate these and take them into consideration.
- Not too surprisingly, it is very important to give voters a clear way to register not only some level of satisfaction for a candidate, but also some level of dissatisfaction for candidates they do not like. This additional information enables voting methods which properly utilize it to more consistently identify the correct winner under widely varying circumstances.
- Without a way to clearly convey dissatisfaction for a candidate, the data gathered from voters cannot support the ability of any voting method to make a well-defined distinction between candidates voters don't like and the ones about which they just have no opinion.

The remaining conclusions of the 2019 project require a more detailed discussion as they lead directly into the work done in this project.

In 2016, a voting method called True Weight Voting[5] or TWV was proposed. It allows voters to indicate either satisfaction or dissatisfaction for as many candidates as they have opinions on a ballot scale consisting of the 2n + 1 integers from –n to +n. Negative integers indicate degrees of dissatisfaction, positive integers indicate degrees of satisfaction and zero is either "don't care" or "no opinion." If a way can be found to gather sincere data from voters, TWV has to work very well. TWV was tested in the 2019 simulation with the relatively high "resolution" of n = 100 (called TWV100). With sincere data from the simulated voters, TWV100 is spectacularly good as was expected. It is able to choose the correct winner with substantially zero error over the entire range of candidates tested (2 to 7).

The next obvious question was: how much resolution is actually required to achieve good performance? Simulations were run with n set to lower and lower values, which requires that voters project their opinions onto an increasingly coarse ballot scale. The surprise was that n could be as low as 3 without much degradation in performance. TWV2 showed a noticeable degradation and TWV1 was worse. However, TWV1 was still very good and considerably outperformed all non-TWV methods. TWV1 allows voters only three score values: -1, 0, and +1. All simulated elections had 1,000 voters. It was suspected (but not tested) that small elections (fewer voters) probably would require greater resolution to maintain low error rates.

Unless there is yet another promising voting method to be tested, it appeared that step 1 of the voting method design procedure had been completed at this point.

[5] See "Voting for Better Decisions" at http://royminet.org/voting-elections/

The 2019 project did progress a short way into step 2. The susceptibility of TWV1 to one easy-to-simulate form of strategic voting was investigated. Voters may decide not to provide any "extra" information which might help a candidate other than the voter's first choice (later harm). For TWV1, that means marking the ballot only for the most (and/or least) favored candidate even if the voter actually does have opinions about other candidates. TWV1 was very resistant to this type of insincere voting and its performance actually improved slightly in the four-candidate elections for which this was tested.

Thus, TWV1 was recommended as a replacement for Plurality (and IRV). This is still a reasonable recommendation. However, it seems possible to do even better.

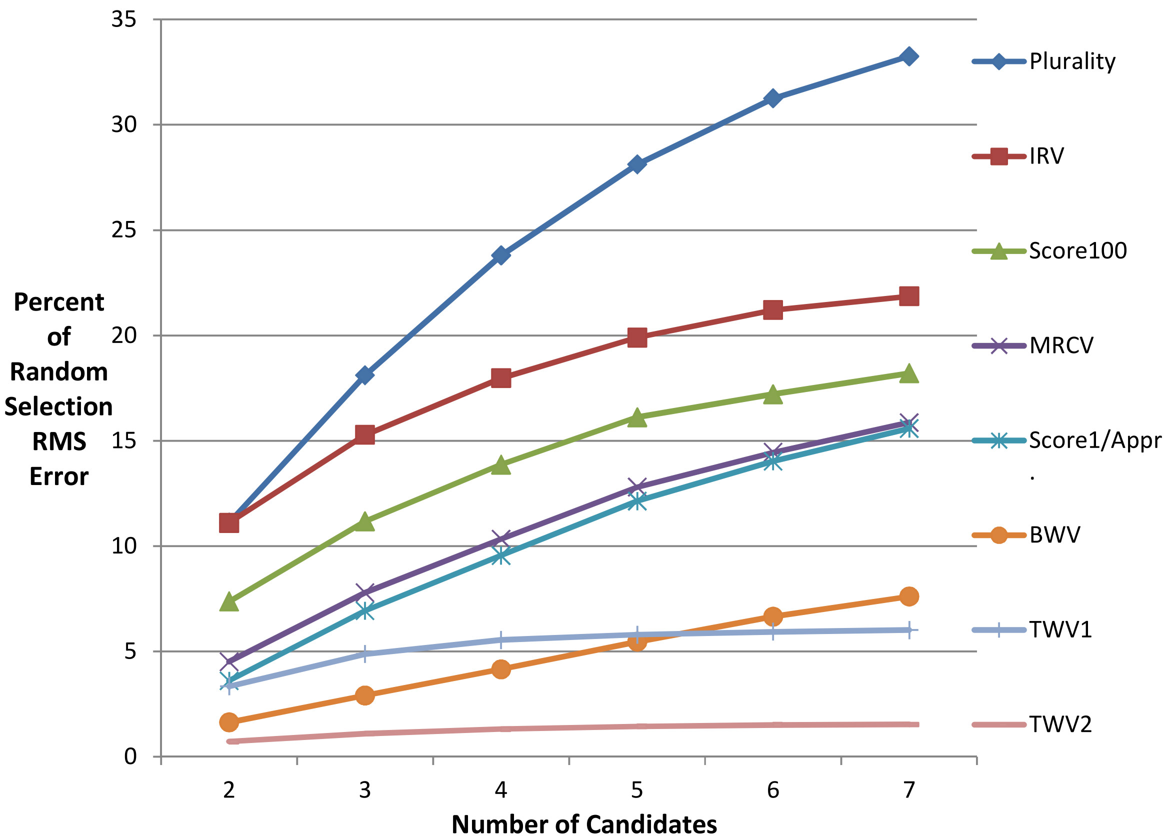
<u>This Follow-on Simulation Project</u>

The ability of TWV1 to "shrug off" one form of strategic voting is noted; especially the fact that its performance actually improved slightly in 4-candidate elections is a bit counterintuitive. Normally, one expects that allowing voters to provide more information should be helpful. Nevertheless, a voting method identical to TWV1 except that voters are allowed to indicate only the candidate they like best and the one they think is worst was tested. It is identified as BWV (Best/Worst Voting). Another series of 600,000 election simulations was carried out with BWV added to the menagerie of voting methods. The table below shows the statistics for these elections.

Number of Elections	100,000	100,000	100,000	100,000	100,000	100,000
Number of Voters	1,000	1,000	1,000	1,000	1,000	1,000
Number of Candidates	2	3	4	5	6	7
Majority Winners	99,838	55,553	21,558	6,450	1,605	340
Incorrect Majority Winners	7,578	2,591	568	69	11	0
Condorcet Winners	99,838	72,233	50,825	38,178	30,737	25,473
Incorrect Condorcet Winners	7,578	4,183	2,212	1,272	860	633
Zero Winners	0	3	4	2	0	0
Negative Winners	0	5	2	2	0	0
Lowest Winning Sats	1	-3	-2	-2	1	1
Average Winning Sats	43	34	29	26	24	22
Highest Winning Sats	61	61	60	57	55	53
Least Winning Voters	492	175	51	81	90	110
Least Winning Voter Percent	53.4	17.9	5.2	8.2	9.3	11.0
Zero Opinions	360,827	751,157	1,157,890	1,564,588	1,977,645	2,390,761
No Opinions	59,152,433	112,255,744	163,358,217	214,243,850	264,585,201	314,973,782
No Opinion % of All Opinions	26.1	37.4	40.8	42.8	44.1	45.0
Candidate A Wins (%)	86.44	72.52	60.69	51.96	45.11	39.58
Candidate B Wins (%)	13.56	22.39	26.86	27.98	27.90	27.41
Candidate C Wins (%)		5.09	10.22	13.81	16.06	17.21
Candidate D Wins (%)			2.22	5.17	7.66	9.57
Candidate E Wins (%)				1.08	2.78	4.42
Candidate F Wins (%)					0.49	1.50
Candidate G Wins (%)						0.32

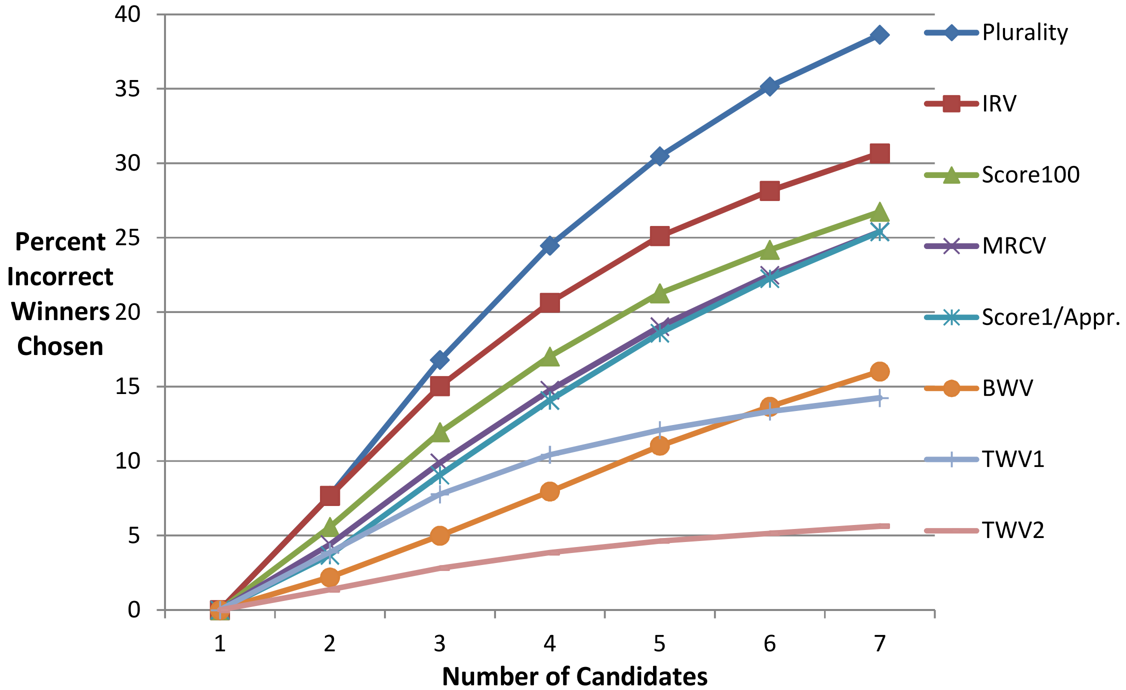
The three following sets of tables and charts show the performance of the various voting methods.

Candidates	Plurality	IRV	Score100	MRCV	Score1/Appr.	BWV	TWV1	TWV2
2	11.110	11.107	7.368	4.507	3.605	1.635	3.351	0.726
3	18.115	15.284	11.175	7.793	6.933	2.913	4.869	1.098
4	23.799	17.972	13.868	10.334	9.551	4.154	5.545	1.313
5	28.119	19.903	16.117	12.794	12.138	5.463	5.794	1.440
6	31.251	21.209	17.212	14.456	14.035	6.652	5.939	1.516
7	33.256	21.865	18.208	15.863	15.587	7.604	6.014	1.534



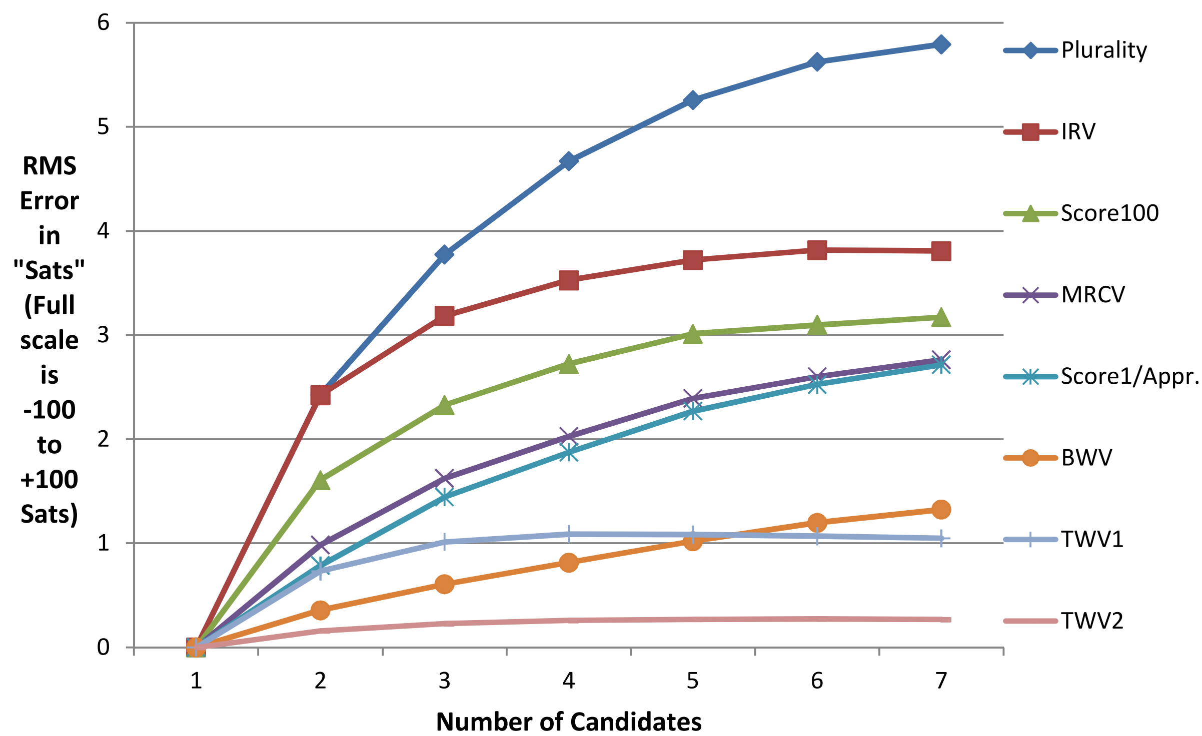
(Note: Pairwise/IRV is substantially equal to IRV)
(Note: For any number of Candidates, TWV100 is 0.0% and Random is 100%)

Candidates	Plurality	IRV	Score100	MRCV	Score1/Appr.	BWV	TWV1	TWV2
1	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
2	7.670	7.663	5.572	4.406	3.685	2.196	3.884	1.357
3	16.794	15.030	11.935	9.883	9.063	4.992	7.774	2.815
4	24.481	20.648	17.030	14.756	14.084	7.960	10.431	3.845
5	30.477	25.115	21.272	19.033	18.594	11.034	12.095	4.634
6	35.162	28.159	24.190	22.491	22.260	13.658	13.335	5.161
7	38.640	30.672	26.746	25.422	25.394	16.022	14.244	5.635



(Note: Pairwise/IRV is substantially identical to IRV)
(Note: TWV100 is 0.000% for any number of Candidates)

Candidates	Random	Plurality	IRV	Score100	MRCV	Score1/Appr.	BWV	TWV1	TWV2
1	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
2	21.813	2.424	2.423	1.607	0.983	0.786	0.357	0.731	0.158
3	20.823	3.772	3.183	2.327	1.623	1.444	0.607	1.014	0.229
4	19.619	4.669	3.526	2.721	2.027	1.874	0.815	1.088	0.258
5	18.690	5.256	3.720	3.012	2.391	2.269	1.021	1.083	0.269
6	17.993	5.623	3.816	3.097	2.601	2.525	1.197	1.069	0.273
7	17.413	5.791	3.807	3.171	2.762	2.714	1.324	1.047	0.267



(Note that Pairwise/IRV is substantially identical to IRV)
(Note: Twv100 is 0.0 RMS error for any number of Candidates)

Note the interesting tradeoff that results from the reduction in information that BWV collects from voters. BWV's ability to accurately pick correct winners is markedly improved for elections which have a small number of candidates, but it is less able than TWV1 to identify the correct winner in elections with a large number of candidates. However, BWV is to be strongly preferred as few real-world elections have more than 3 or 4 significant candidates and BWV is not worse than TWV1 until there are 6 or more candidates.

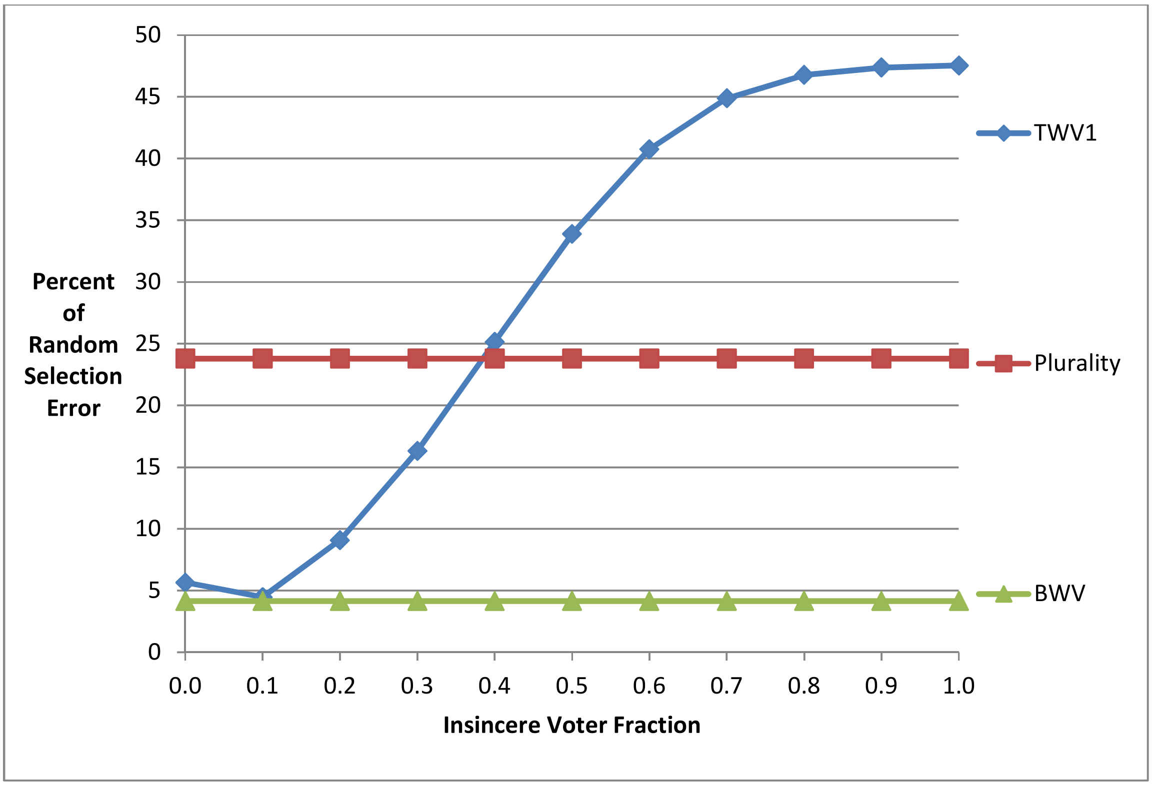
At this point, we might at least tentatively consider step 1 of the design process complete (again).

Step 2 and Strategic Voting

Reduced vulnerability to strategic attack also is an important advantage as insincere voting can ruin performance. Any voting method which allows voters to score candidates on a scale is susceptible to a very obvious form of manipulation. Voters will realize fairly quickly that they can maximize the impact of their ballot by marking the highest allowed value for candidate(s) they like and the lowest allowed value for all others. BWV (and TWV1) essentially eliminate that opportunity by virtue of their minimalist scale. Fortunately, the statistics of large numbers of voters (1,000 or more) enables good performance with such low resolution voting methods.

Any voting method which allows voters to vote for and against multiple candidates is vulnerable to another fairly obvious attack. To maximize the impact of their ballot, voters can vote for the candidate they most favor and against all other candidates whether they actually dislike them or not. This form of strategic attack is easy and straightforward to simulate, so TWV1's performance under this type of insincere voting was evaluated. A run of 100,000 four-candidate elections was made for each 10% increment of strategic voters. The table and chart below show the TWV1 results (Plurality and BWV performance levels are shown for reference).

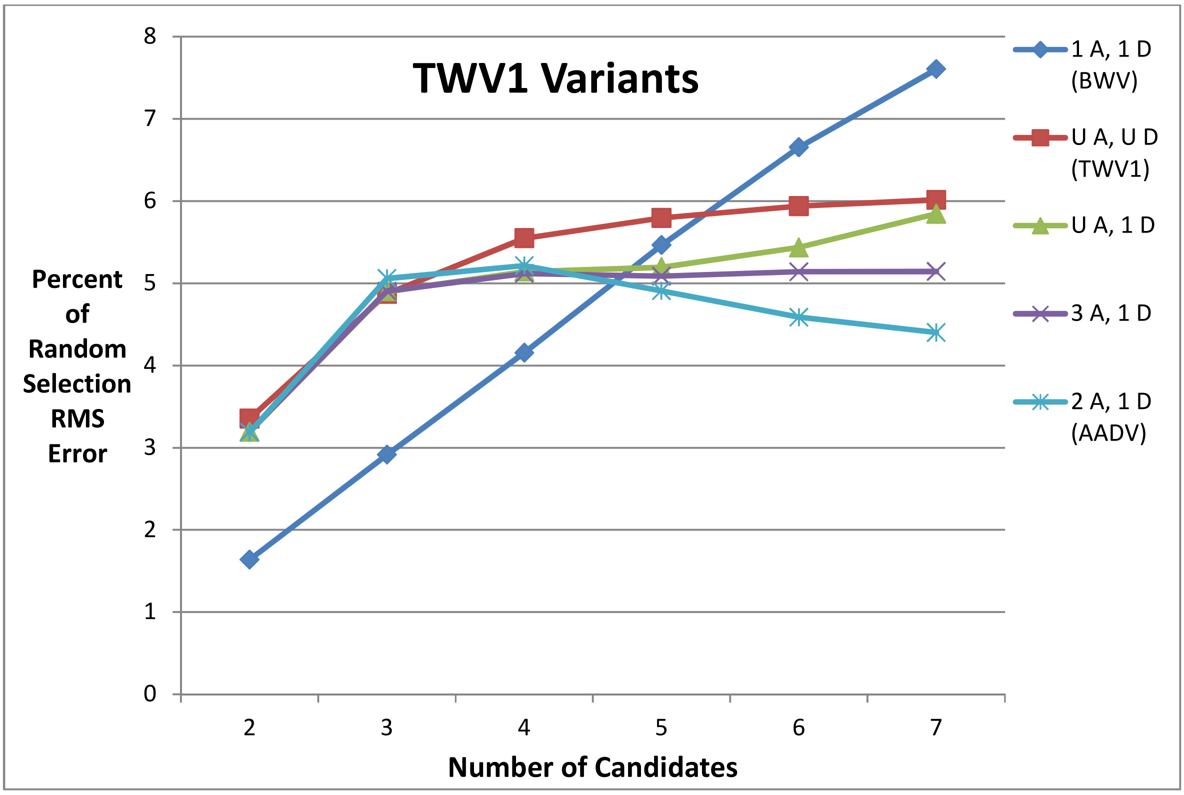
Insincere Fraction	TWV1	BWV	Plurality
0.0	5.649	4.154	23.799
0.1	4.469	4.154	23.799
0.2	9.079	4.154	23.799
0.3	16.309	4.154	23.799
0.4	25.128	4.154	23.799
0.5	33.878	4.154	23.799
0.6	40.753	4.154	23.799
0.7	44.875	4.154	23.799
0.8	46.776	4.154	23.799
0.9	47.380	4.154	23.799
1.0	47.536	4.154	23.799



Note that a small percentage of strategic voters added to the mix actually reduced errors somewhat. However, performance then turns around and rapidly deteriorates, slightly exceeding Plurality at 40% and maxing out at approximately double Plurality's error. BWV substantially eliminates this strategic attack opportunity by allowing a vote for at most one candidate and against at most one candidate. We have "conditioned ourselves" to believe that more information from voters is always a good thing. That may be true if useful and *sincere* additional data could be collected, but BWV seems to be clearly the better choice, even though it restricts the data voters may provide.

Because of the above result and also because of the somewhat "dramatic" change in the shape of the TWV1 curve to a nearly linear one when approvals and disapprovals are limited to just one each, it was decided to examine the family of TWV1 variants more closely and completely. Runs of 100,000 four-candidate elections were made to test TWV1 versus TWV1 when restricted to 1 disapproval (so as to prevent the above demonstrated strategic attack) while approvals were restricted to 1, 2, 3 and unlimited. Those results are shown by the below data table and chart.

Candidates	TWV1 Unlimited A Unlimited D	TWV1 Unlimited A 1 D	TWV1 3 A 1 D	TWV1(AADV) 2 A 1 D	TWV1(BWV) 1 A 1 D
2	3.351	3.188	3.191	3.185	1.635
3	4.869	4.897	4.903	5.057	2.913
4	5.545	5.140	5.118	5.215	4.154
5	5.794	5.194	5.086	4.907	5.463
6	5.939	5.434	5.138	4.585	6.652
7	6.014	5.844	5.142	4.400	7.604



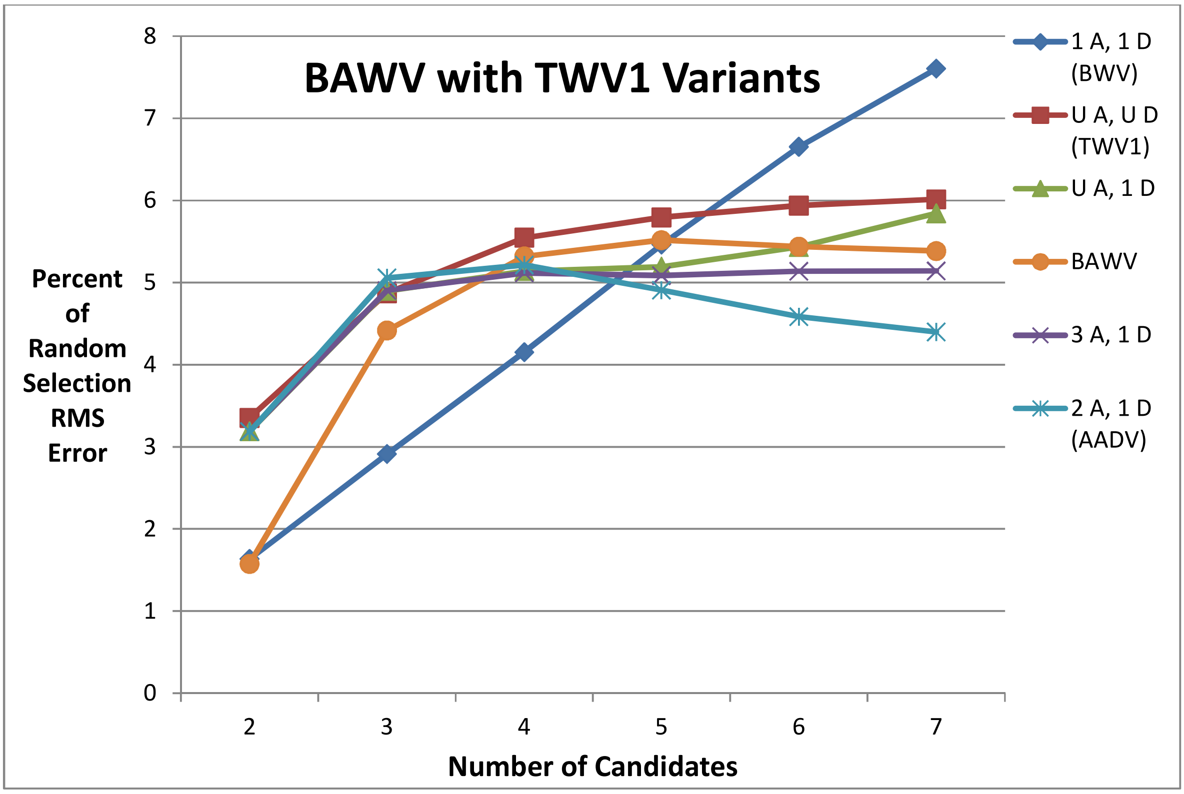
All methods in the TWV1 family have somewhat similar and very good performance with sincere data. BWV reduces opportunities for strategic attack as much as possible. However, BWV certainly is still very much susceptible to "vote for the lesser evil" problems.

Unfortunately, the *opportunity* for VLE strategic voting cannot ever be eliminated. As long as elections are open, free and fair, voters will always be able to vote for the lesser evil. Thus, the approach to reducing this kind of strategic vulnerability must be to eliminate the *motivation* of the voter to vote insincerely. This type of strategic voting is considerably more complex to simulate since it depends upon voters' analyses of the probability that various candidates have of winning as formed by the amalgam of information (polls, etc.) that they may have seen. The validity of any such assumptions can always be questioned, so this project has, thus far, steered clear of such simulations.

It should first be noted that BWV would make it much more difficult for strongly divisive candidates to win. Therefore, the nomination of such candidates will be discouraged in favor of nominating candidates that have broad appeal since their path to victory is much easier. Surely, the best possible way to reduce the motivation to vote for the lesser evil is to have fewer evil candidates! However, it would also be good, if possible, to modify the voting method to reduce or eliminate this motivation.

A way to eliminate vote-for-the-lesser-evil motivation would be to give voters the ability to designate an alternate best choice. Voters must understand that designating such an alternate can in no way either help or hurt any candidate. They must also have complete confidence that, if the candidate they have designated as best does not win, then their alternate will be counted exactly as though it had originally been marked as their best choice. This should dispatch any motivation to insincerely mark the "lesser evil" as best choice. Similarly, there can be no (later harm) concern about designating an alternate (which could, of course, be the "lesser evil"). Call this BAW Voting (Best/Alternate/Worst Voting). A BAWV method was coded and its performance measured. The results for four-candidate elections are shown below with the TWV1 variants.

Candidates	**TWV1 Unlimited A Unlimited D**	**TWV1 Unlimited A 1 D**	**TWV1 3 A 1 D**	**TWV1(AADV) 2 A 1 D**	**TWV1(BWV) 1 A 1 D**	**BAWV**
2	3.351	3.188	3.191	3.185	1.635	1.574
3	4.869	4.897	4.903	5.057	2.913	4.417
4	5.545	5.140	5.118	5.215	4.154	5.319
5	5.794	5.194	5.086	4.907	5.463	5.518
6	5.939	5.434	5.138	4.585	6.652	5.437
7	6.014	5.844	5.142	4.400	7.604	5.385



As can be seen, BAWV works extremely well with sincere voter data. Actually, its performance curve seems to be slightly preferable to any of the other variants. BAWV is a huge improvement over Plurality or IRV, and certainly is far better than Approval. Also, since it is maximally immune to strategic manipulation, it should be possible to obtain sincere information from voters and actually achieve that excellent performance. Clearly, BAWV is an extremely strong contender for the best possible voting method.

However, BAWV necessarily adds some complexity. Things still are simple enough for voters. They just have one additional option (Alternate) for each candidate. But tallying the votes and determining the winner becomes considerably more complicated. In order to honor the contract with voters regarding their alternate candidate selection, BAWV must be an iterative eliminations (sometimes called a "last man standing") method. Very roughly, that means tallying the votes as would be done for BWV, but instead of designating a winner, eliminate the weakest candidate, then promote any alternates into eliminated best choice spots and repeat the process until only one candidate remains. The BAWV tally process surely would be automated. With the technology available today, there is absolutely no reason that can't be done in an entirely secure, reliable, transparent and auditable way (see Future Work).

Nevertheless, BAWV's tally complexity could be a stumbling block for some. Thus, it would be wise to determine a "second best" fallback recommendation that is simple. BWV is extremely simple in all respects; no problem at all with the Jones rule. But as pointed out, BWV is still quite susceptible to VLE pressures.

By the time they are filling out a ballot in a voting booth, most voters are considering (for a single-winner race) either 1, 2 or 3 candidates, usually not more. A small percentage of voters simply vote for one candidate, their favorite. The vast majority have one candidate they want to win and one (the horrible competition or enemy) that they strongly do not want to win. In certain elections, a significant number of this vast majority may favor a candidate to win that they perceive as weaker than a "lesser evil" candidate. The 3-candidate case is the most challenging one that voters usually face. The most useful information to help a voting method choose the correct winner obviously is the voter's first choice candidate. The second most important is the candidate the voter most strongly does NOT want to win. The third useful data item is the "lesser evil" candidate, whenever there is one. A voting method should gather and utilize those three data items in that priority order.

The question now is whether or not it would be an improvement to allow voters to input any additional data. Some voters may indeed occasionally have additional sincere information that might improve the ability to choose the correct winner. However, additional input options always increase the opportunity for all voters to think (correctly or incorrectly) that they may be able to amplify the impact of their ballot through some manipulative strategy. Simulation results show that additional sincere information can only very slightly improve decision accuracy, while on the other hand, insincere, strategic or garbage data usually degrades decision making quickly and seriously. For two-candidate elections, there simply is no additional useful information; it is obvious that just the most important two data items can possibly be helpful. For three-candidate simulated elections, only 8% of voters have any additional sincere information to contribute and for four-candidate elections it was 22%. The voting method will be the same for any number of candidates and it seems clear that it should err on the side of minimizing the risk of damaging strategic or garbage data. Contrary to the mantra that more information is always better, more than the top three data items almost surely is worse.

Those considerations lead rather directly to the selection of AADV from the TWV1 family as the second best method to recommend. Voters may approve of up to two candidates (but disapprove of only one). This does not remove the motivation to vote for the lesser evil, but it does allow the voter to also approve the (perceived to be weaker) best candidate without later harm risk. This is the same favorable characteristic touted for Approval voting. Per the discussion in the previous paragraph, there is no value in allowing more than two approvals. AADV has performance comparable to BAWV with sincere data and has minized opportunities for strategic manipulation. Unlike BAWV, it does not eliminate VLE motivation, but should result in more votes cast for good, but weaker candidates. AADV can be directly scored, thus eliminating BAWV's tallying complexity.

AADV and BAWV are fleshed out in more detail in the next section.   At this point, step 2 of the design process might at least tentatively be considered complete.

Step 3 and the Human Interface

The best general way to present AADV or BAWV to voters and politicians is to explain that a separate yes/no referendum will be conducted for each and every candidate in a race. Voters have the option to vote "yes" in the referendum of the candidate they think is the best choice and/or to vote "no" in the referendum of the candidate they think is the worst choice. The candidate that wins its referendum by the largest (positive) majority is the winner of the race.

> **AADV Instructions to Voters:** Mark an "**X**" in the "**Approved**" box for any one or two candidate(s) (if any) that you really like and believe would be the best one(s) to win this race. Mark an "**X**" in the "**Disapproved**" box for any one candidate (if any) that you strongly believe would be the worst choice and which you would not want to win this race. If you do not know enough about a candidate or do not have a strong opinion one way or the other, leave both boxes unmarked for all such candidates. Do not mark more than one box for any single candidate.
>
> **AADV Instructions to Election Officials:** Disqualify any ballots which have more than one box marked for the same candidate. Disqualify any ballots which have more than two candidates marked "Approved." Disqualify any ballots which have more than one candidate marked "Disapproved." Total the "Approved" votes for each candidate; call this total "A." Total the "Disapproved" votes for each candidate; call this total "D." Add "A" and "D" for each candidate; call this sum "V." Eliminate any candidate whose "V" is less than one plus one percent (rounded to the nearest number of voters) of the largest "V" that any single candidate received. Subtract "D" from "A" for each remaining candidate; call this difference "N." Eliminate any candidate which has a zero or negative N. The remaining candidate that has the largest "N" is the winner.

Note that the logical possibility does exist for both AADV and for BAWV that there could be no winner for a race (no remaining candidate with a positive "N"). Of course, a rule could certainly be put into place to crown the "best" negative "N" candidate for this highly unlikely and extremely sad case. It would not seem wise to elect a candidate that more people dislike than like. A far better alternative would be to hold another election. No candidate that received a negative "N" could be re-nominated. This is similar to some voting procedures which always have a "NOTA" (Non-Of-The Above) "candidate" on the ballot. It should be considered a fairly serious defect of substantially all other voting methods that they are unable to sensibly handle the situation where a majority of voters dislike a candidate (or even all candidates). Voters need to be able to explicitly vote against candidates they don't like.

BAWV Instructions to Voters: Mark an "**X**" in the "**Best**" box next to any one candidate (if any) that you believe would be the best candidate to win this race. Mark an "**X**" in the "**Worst**" box next to any one candidate (if any) that you believe would be the worst candidate to win this race. Mark an "**X**" in the "**Alternate**" box next to any one candidate (if any) that you would like to have counted as your "Best" choice in the event that the candidate you have marked as "**Best**" is eliminated. Do not mark more than one box for any single candidate. If you do not know enough about a candidate or do not have a strong opinion one way or the other, leave all three boxes unmarked for all such candidates. Note that an "**Alternate**" designation will be completely ignored unless and until your "**Best**" candidate is eliminated, at which time your "**Alternate**" will immediately and thereafter be counted as your choice for "**Best**" candidate.

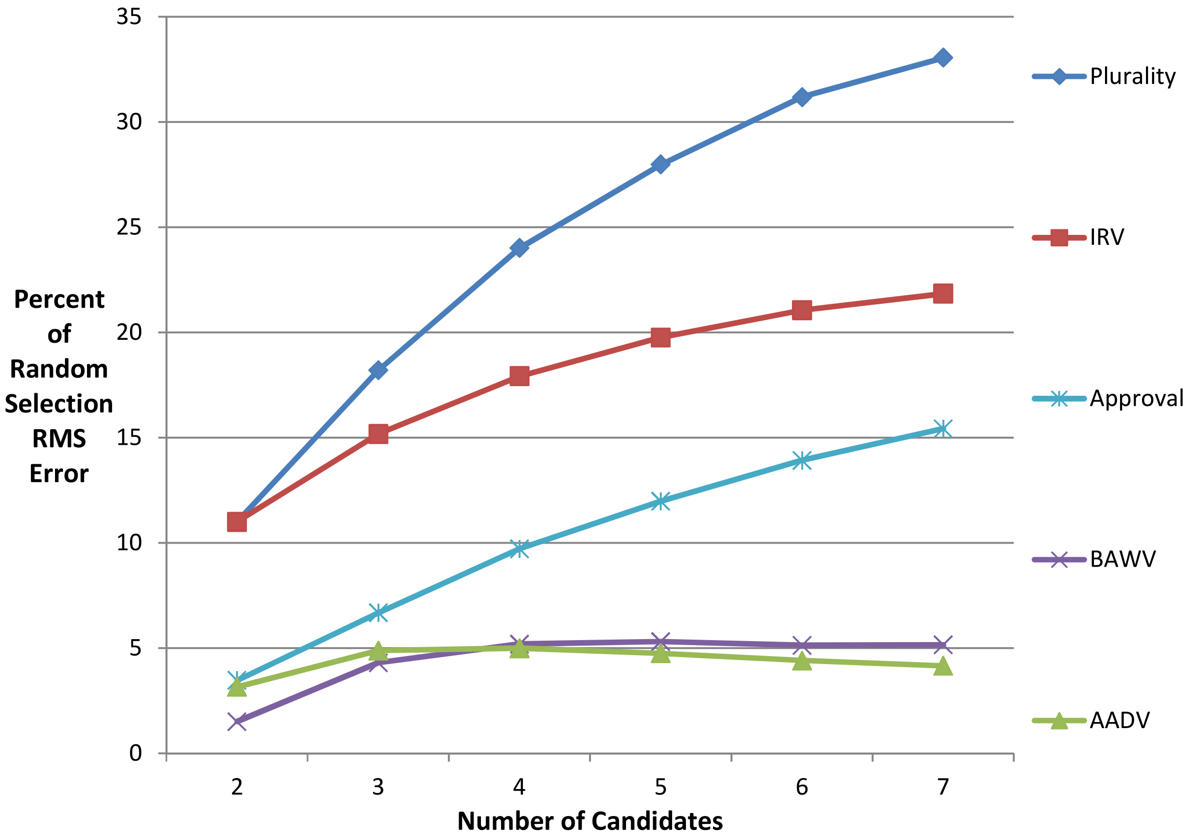
BAWV Instructions to Election Officials:

1. Disqualify any ballots which have more than one box marked for the same candidate. Disqualify any ballots which have more than one candidate marked "Best." Disqualify any ballots which have more than one candidate marked "Alternate." Disqualify any ballots which have more than one candidate marked "Worst."
2. Total the "Best" votes for each candidate; call this total "B." Total the "Alternate" votes for each candidate; call this number "A." Total the "Worst" votes for each candidate; call this total "W." Add "B" plus "A" plus "W" for each candidate; call this sum "V." Compute one plus one percent (rounded to the nearest number of voters) of the largest "V" that any single candidate received; call this number "MV."
3. Eliminate any candidate that has a "V" less than "MV." Upon eliminating each candidate, unmark the "Best" "box" for any ballots which have selected the candidate being eliminated as "Best." For any ballot on which the "Best" box is being unmarked and for which an "Alternate" is marked, change the "Alternate" candidate to "Best."
4. Total the "Best" votes for each candidate; call this total "B." Subtract "W" from "B" for each remaining candidate; call this difference "N."
5. If only one candidate remains and its "N" is positive, declare that candidate the winner. If either no candidates remain or one remains, but with zero or negative "N," there is no winner.
6. Eliminate the candidate which has the lowest positive (or most negative) "N." (If there should be a tie for lowest "N," then eliminate the tied candidate which has the lowest "V.") Unmark the "Best" "box" for any ballots which have selected the candidate being eliminated as "Best." For any ballot on which the "Best" box is being unmarked and for which an "Alternate" is marked, change the "Alternate" candidate to "Best." Go back to step 4.

Of course, invalid ballots can be prevented for either AADV or BAWV by intelligent, user-friendly electronic voter supervision, thus obviating the need to check for and disqualify them during the tally process.

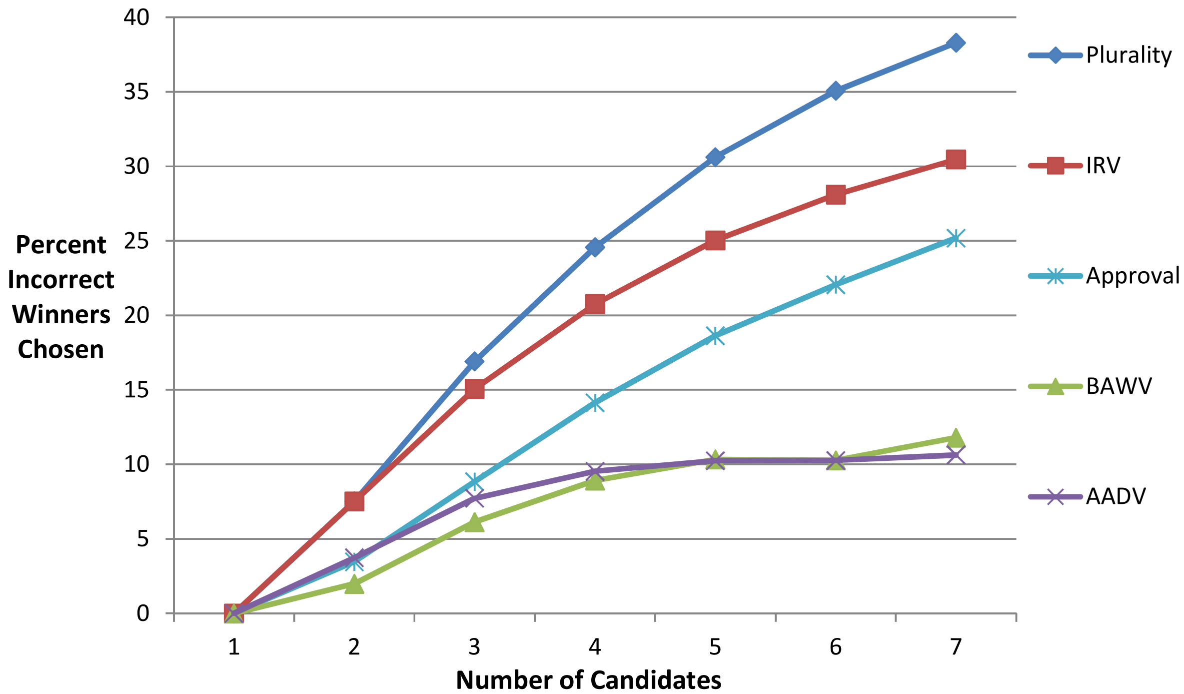
As a final summary, the following three sets of data tables and charts show the recommended BAWV and AADV methods' performance along with some commonly discussed methods for reference. Precision was increased by increasing the number of voters to 10,000 for each run of 100,000 elections.

Candidates	Plurality	IRV	Approval	AADV	BAWV
2	11.009	11.005	3.474	3.164	1.510
3	18.205	15.182	6.682	4.882	4.308
4	24.016	17.923	9.724	5.001	5.195
5	27.984	19.764	11.981	4.761	5.311
6	31.184	21.061	13.925	4.413	5.136
7	33.058	21.842	15.434	4.164	5.155



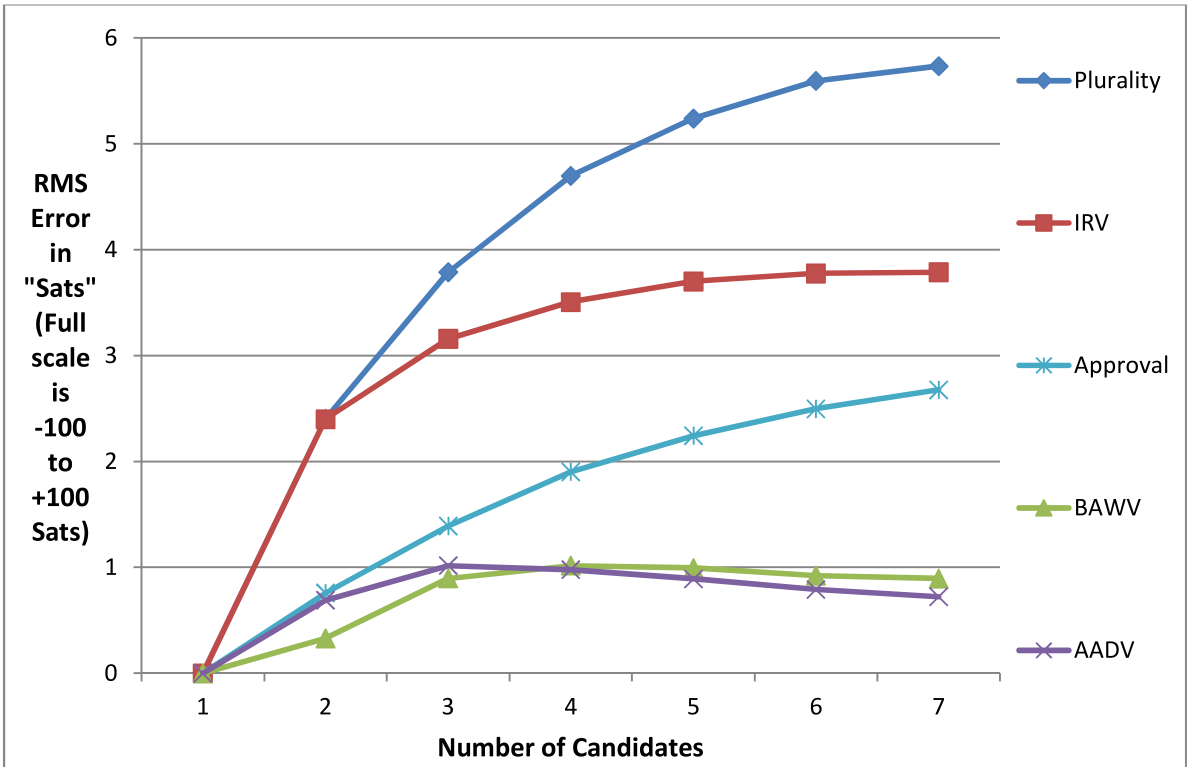
(Note: Pairwise/IRV is substantially equal to IRV)
(Note: For any number of Candidates, TWV100 is 0.0% and Random is 100%)

Candidates	Plurality	IRV	Approval	AADV	BAWV
1	0.000	0.000	0.000	0.000	0.000
2	7.524	7.522	3.476	3.739	1.984
3	16.927	15.075	8.851	7.734	6.131
4	24.572	20.764	14.139	9.544	8.924
5	30.625	25.047	18.638	10.249	10.344
6	35.077	28.109	22.057	10.268	10.268
7	38.297	30.472	25.186	10.643	11.796



(Note: Pairwise/IRV is substantially identical to IRV)
(Note: TWV100 is 0.000% for any number of Candidates)

Candidates	Random	Plurality	IRV	Approval	AADV	BAWV
1	0.000	0.000	0.000	0.000	0.000	0.000
2	21.818	2.402	2.401	0.758	0.690	0.330
3	20.806	3.788	3.159	1.390	1.016	0.896
4	19.566	4.699	3.507	1.903	0.978	1.016
5	18.730	5.241	3.702	2.244	0.892	0.995
6	17.943	5.595	3.779	2.499	0.792	0.922
7	17.348	5.735	3.789	2.678	0.722	0.894



(Note that Pairwise/IRV is substantially identical to IRV)
(Note: Twv100 is 0.0 RMS error for any number of Candidates)

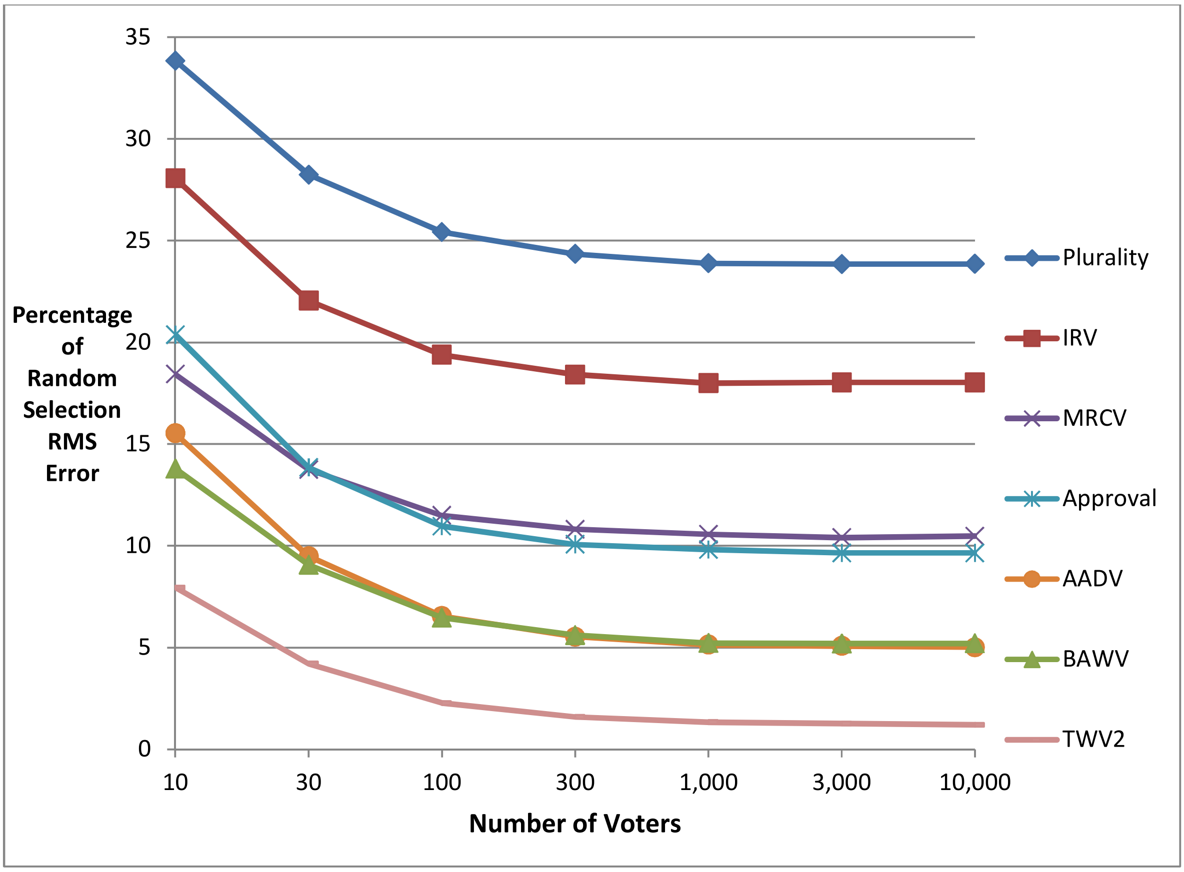
Considerations Related to the Size of Elections

Adequate precision for voting method error was considered to be a solid three decimal places. It was quickly and experimentally determined that runs of 100,000 election with 1,000 voters in each election resulted in adequate precision/reproducibility. Performance of voting methods is expected to be worse in small elections (fewer voters), especially those methods using low resolution and/or which gather smaller amounts of data from voters. A series of 4-candidate election simulations was done to investigate this dimension of the problem. Statistical noise can be expected to approximate $\sqrt{ev}$ / ev (where e is the number of elections and v is the number of voters in each election). So, ev > 1,000,000 should be maintained. To maintain precision with smaller numbers of voters, up to 1,000,000 elections were simulated for smaller numbers of voters. The following table shows the statistics for these elections.

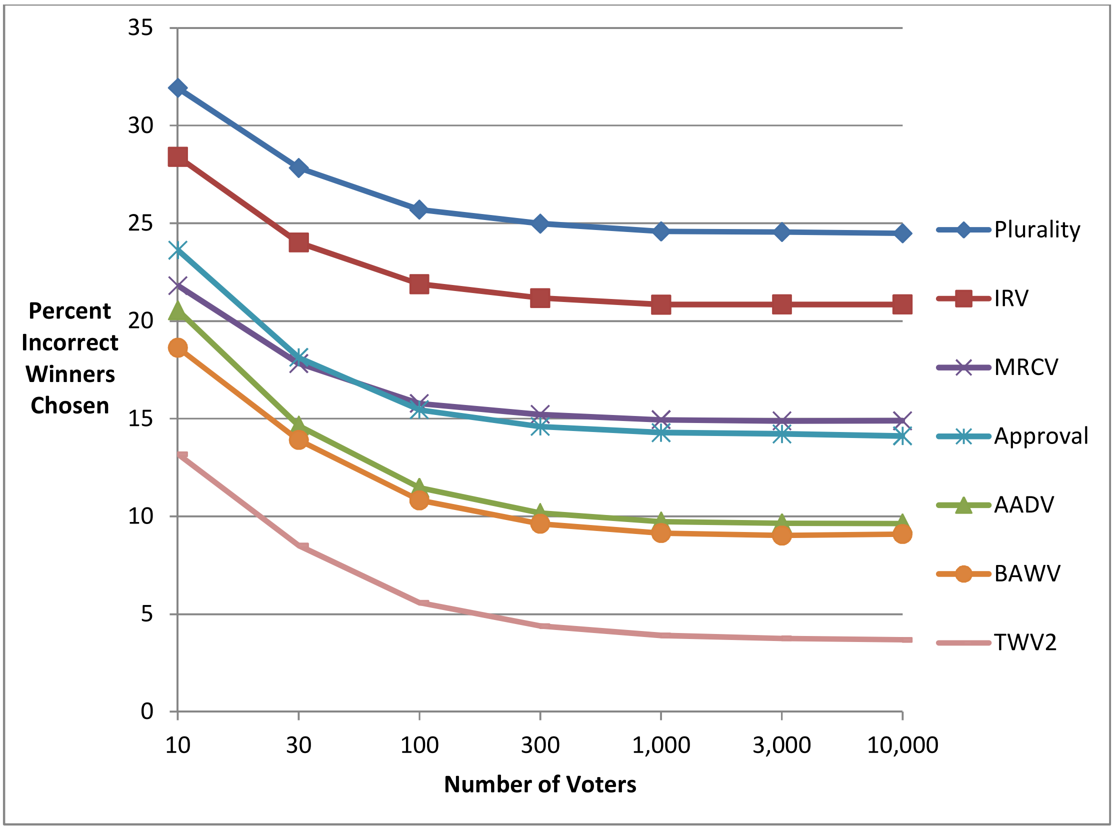
Number of Elections	1,000,000	500,000	500,000	500,000	200,000	200,000	100,000
Number of Voters	10	30	100	300	1,000	3,000	10,000
Number of Candidates	4	4	4	4	4	4	4
Majority Winners	189,558	102,510	106,240	107,723	43,275	43,204	21,773
Incorrect Majority Winners	7,456	3,229	2,950	2,803	1,071	1,081	524
Condorcet Winners	341,342	218,892	243,437	251,701	102,204	102,125	51,264
Incorrect Condorcet Winners	15,286	10,105	10,739	10,881	4,323	4,370	2,181
Zero Winners	86	24	13	12	3	3	6
Negative Winners	56	12	12	8	3	3	4
Lowest Winning Sats	-7	-4	-4	-3	-2	-4	-3
Average Winning Sats	31	30	29	29	29	29	29
Highest Winning Sats	81	73	63	61	60	59	58
Least Winning Voters	0	0	0	11	67	81	12
Least Winning Voter %	0.0	0.0	0.0	3.7	6.9	2.7	0.1
Zero Opinions	115,211	173,418	580,183	1,736,663	2,319,457	6,936,925	11,579,091
No Opinions	16,363,822	24,503,420	81,700,574	245,145,412	326,868,063	980,572,376	1,633,095,910
No Opinion % of All Opinions	40.9	40.8	40.9	40.9	40.9	40.9	40.8
Candidate A Wins (%)	56.61	59.17	60.19	60.52	60.59	60.62	60.65
Candidate B Wins (%)	27.87	27.29	27.04	26.92	26.92	26.89	26.89
Candidate C Wins (%)	12.02	10.89	10.41	10.30	10.23	10.22	10.18
Candidate D Wins (%)	3.51	2.66	2.35	2.26	2.26	2.27	2.28

The following three sets of data tables and charts shows performance of methods are affected by election size (number of voters).

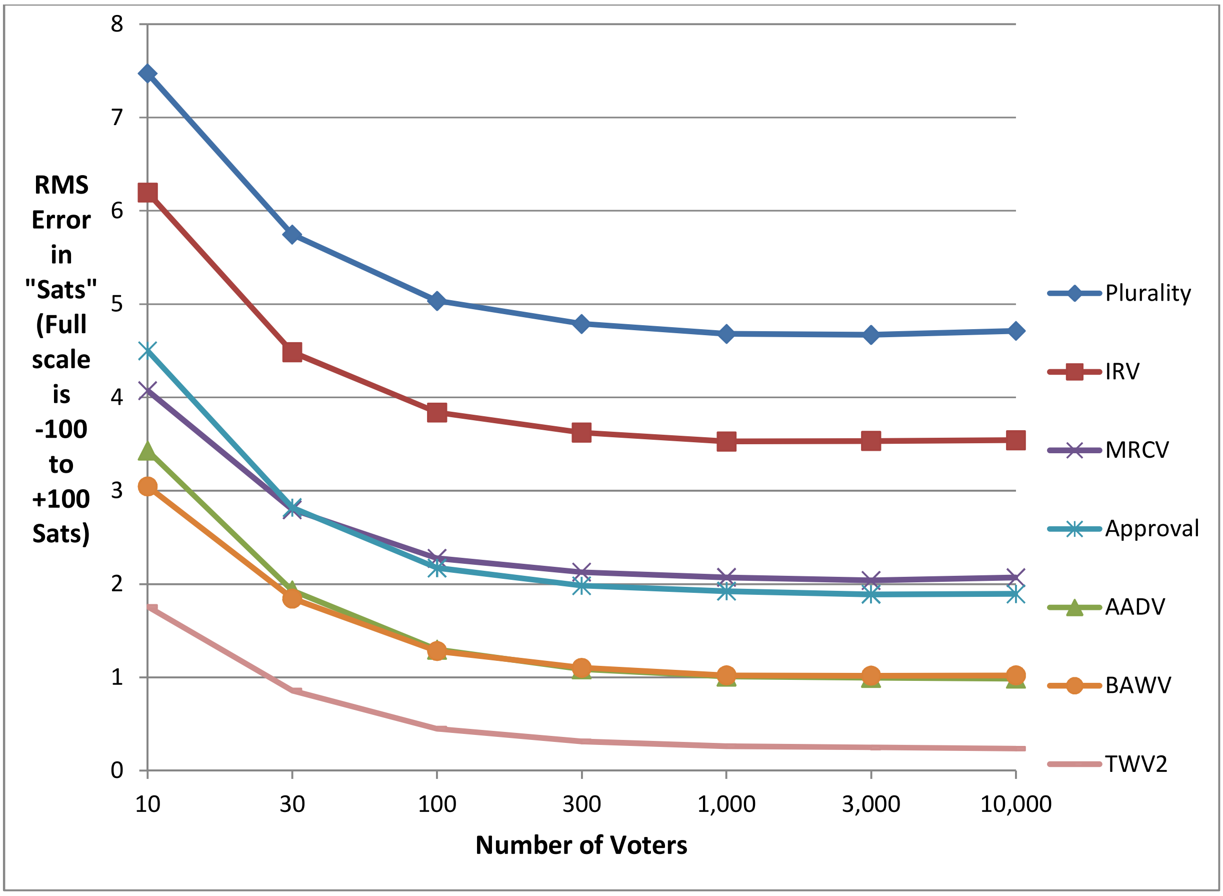
Voters	Plurality	IRV	MRCV	Approval	AADV	BAWV	TWV2
10	33.840	28.060	18.440	20.375	15.528	13.798	7.953
30	28.238	22.050	13.746	13.851	9.489	9.066	4.211
100	25.423	19.385	11.493	10.967	6.549	6.469	2.270
300	24.345	18.419	10.813	10.070	5.529	5.606	1.589
1,000	23.885	17.998	10.569	9.818	5.144	5.215	1.335
3,000	23.847	18.037	10.404	9.649	5.081	5.198	1.264
10,000	23.855	18.036	10.475	9.650	5.013	5.205	1.202



Voters	Plurality	IRV	MRCV	Approval	AADV	BAWV	TWV2
10	31.922	28.403	21.809	23.621	20.549	18.633	13.184
30	27.834	24.017	17.825	18.130	14.641	13.908	8.517
100	25.699	21.885	15.774	15.455	11.466	10.824	5.581
300	24.980	21.172	15.216	14.596	10.170	9.624	4.388
1,000	24.581	20.832	14.945	14.284	9.736	9.145	3.906
3,000	24.552	20.843	14.880	14.222	9.649	9.025	3.752
10,000	24.482	20.846	14.889	14.115	9.640	9.093	3.686



Voters	Random	Plurality	IRV	MRCV	Approval	AADV	BAWV	TWV2
10	22.086	7.474	6.198	4.073	4.500	3.430	3.047	1.757
30	20.343	5.745	4.486	2.796	2.818	1.930	1.844	0.857
100	19.805	5.035	3.839	2.276	2.172	1.297	1.281	0.450
300	19.668	4.788	3.623	2.127	1.981	1.087	1.103	0.313
1,000	19.599	4.681	3.528	2.071	1.924	1.008	1.022	0.262
3,000	19.589	4.671	3.533	2.038	1.890	0.995	1.018	0.248
10,000	19.640	4.713	3.542	2.069	1.895	0.984	1.022	0.236



The above election size simulation data ties up some lose ends by:

1. Confirming the anticipated increasing errors for all voting methods in small elections;
2. Verifying that 1,000 voters was an adequately large choice for the working number of voters used for the bulk of this work which was intended to be applicable to large public elections;
3. Showing that BAWV (or AADV) still hangs in there and is still the method to use for small elections.

It is noted that at least part of the reason for deteriorating performance in small elections likely is related to the increasing number of elections which have ties.

Conclusions

1. Error plots for the best possible ordinal voting method (MRCV) and what appears to be the best cardinal method based solely on voter satisfaction data (Score1/Approval) lie nearly on top of one another. This likely is not just a coincidence. Instead, these two methods probably are at or very near to the limit to performance (roughly 2/5 of Plurality's error) that can be achieved by any method which does not allow voters to explicitly indicate dissatisfaction information and use it to offset satisfaction. On the other hand, methods which do enable voters to input satisfaction data and dissatisfaction data, then tally ballots in such a way that satisfaction and dissatisfaction offset each other are able to achieve errors approaching zero for sincere data. The remaining challenge is to gather sincere data from voters.
2. BAWV should be implemented for public elections as soon as possible. If the tally complexity of BAWV becomes a sticking point, the second best method, AADV, is the fallback as it has a simple tally procedure. However, BAWV is strongly recommended as it could be smoothly and acceptably implemented with the technology available today. Effectively reducing or eliminating the awful and pernicious pressure to vote for the lesser evil should be well worth the increased complexity of tallying the vote.

Future Work

1. A comprehensive software package, called Election Manager,[6] was written in 2012. It is capable of managing all phases of elections: jurisdiction setup, election setup, supervision of voting in the precincts and tallying the vote for a multi-jurisdiction election. Either Plurality or MRCV can be selected as the voting method for an election. It is intended to update Election Manager and implement BAWV and AADV as selectable voting methods. The target schedule should enable testing of BAWV and AADV with actual voters at a polling place during the 2020 general election.
2. BAWV and AADV are cardinal voting methods with a resolution of 1 which have been optimized to minimize strategic manipulation opportunities and motivations. Their performance is excellent, significantly better than other methods and radically better than Plurality (or IRV). However, that performance is limited by the low resolution. Doubling resolution to 2 could further reduce error by an additional factor of about 4 (note the theoretical performance of TWV2 with sincere data). However, there may very well be no way to increase the resolution without either opening the door to unacceptable strategic voting manipulation and/or unacceptably increasing complexity. This might something to work on as long as it does not delay implementation of BAWV (or at least AADV) for public elections.
3. During this project, several cases were documented where strategic voting actually improved the performance of some voting methods; voters' insincere ballots were able to

[6] See http://royminet.org/election-manager-software/

compensate for some of a method's shortcomings and/or limitations. In some cases, a small percentage of voters voting strategically can improve performance somewhat, but further increases in that percentage then cause performance to degrade rapidly. Yet in other cases, insincere voting is undesirable and just degrades results, period. It would be of some interest to have a more complete understanding of these phenomena. One would think it should be possible to better predict the impact of strategic voting rather than having to just empirically measure its effects. Perhaps this is one non-urgent topic that can provide grist for the paper publishing mill for the next century or so.