The 3-candidate left-middle-right scenario

Warren D. Smith*
warren.wds at gmail.com

July 17, 2006

Abstract — **The most common nontrivial scenario in single-winner political races is the 1-dimensional political scenario consisting of a "left," "middle" and "right" candidate in a 3-way contest. We attempt an exhaustive computer-aided analysis of how well commonly proposed voting systems handle this scenario. The clear conclusion is that (within this model) *range voting* is the best voting system among all common proposals.**

Keywords — Democracy, single winner voting systems, Bayesian regret, computer simulation.

1 Introduction

The left-middle-right scenario is important because it is the commonest and simplest nontrivial scenario in single-winner democratic elections. It also has the advantage that, because there are only three candidates, many different voting systems degenerate to the same system, allowing great simplification of the analysis.

2 Mathematical model of voters, candidates, and utilities

Imagine that "Left" is located at -1, "Right" is located at $+1$, and "Middle"' is located at some x with $-1 < x < 1$ all on a line. There are three kinds of voters.

Leftist voters: utility 100 if Left elected, utility 0 if Right elected, and utility $\ell(x)$ if Middle elected. Here $\ell(x)$ is a monotonic decreasing continuous function of x for $-1 \le x \le 1$ with $\ell(-1) = 100$ and $\ell(1) = 0$.

Rightist voters: utility 100 if Right elected, utility 0 if Left elected, and utility $r(x)$ if Middle elected. Here $r(x)$ is a monotonic increasing continuous function of x for $-1 \le x \le 1$ with $r(-1) = 0$ and $r(1) = 100$.

Centrist Voters: utility 100 if Middle elected, utility 0 if the "bad" extremist is elected, and utility g for some g with $0 < g < 100$ if the "good" extremist is elected. Which extremist is "bad" or "good" depends on the centrist voter where there are two types of centrist voters: leftist centrists and rightist centrists.[1]

We shall demand that $r(x)$ be one of the following three functions:

$$50(1 + x), \quad 25(1 + x)^2, \quad 100 - 25(1 - x)^2 \qquad (1)$$

and that $\ell(x)$ be one of the following three (mirrored) functions:

$$50(1 - x), \quad 25(1 - x)^2, \quad 100 - 25(1 + x)^2. \qquad (2)$$

These functions are respectively straight-line, concave-∪, and concave-∩.

So the entire model is described by the following six nonnegative parameters (really only five parameters since the first four are constrained to sum to the total voter population, which we may regard as fixed)

1. Number of leftist voters
2. Number of left-central voters
3. Number of right-central voters
4. Number of rightist voters
5. Value of g with $0 < g < 100$
6. Value of x with $-1 < x < 1$

plus the $3 \times 3 = 9$ possible choices of the $\ell(x)$ and $r(x)$ functions.

We restrict g to be a member of the 9-element set $\{10, 20, 30, 40, 50, 60, 70, 80, 90\}$. We restrict $3x$ to be a member of the 5-element set $\{-2, -1, 0, +1, +2\}$. We also restrict the numbers of the four different kinds of voters to be in the 22-element set $\{1, 2, 3, \ldots, 21, 22\}$ (and to sum to 23 voters in all) so that we get in total at most $5 \times 9 \times 9 \times 22^4 = 94873680$ possible scenarios. Actually, this is a large overestimate because it has ignored the sum=23 constraint on the first four parameters. In view of that constraint, the true[2] number of scenarios is exactly $623700 = 5 \times 9 \times 9 \times \binom{22}{3}$.

*21 Shore Oaks Drive, Stony Brook NY 11790.

[1] I.e, we view the voters as located at -1, near to but on both sides of 0, and $+1$. The computer will explore all possible g values. It would have been possible to have two kinds of g, say g_ℓ and g_r, and/or to make them depend on x, but we did not bother with those more complicated models. It also would have been possible to distribute the voters more smoothly along the line, e.g. as a normal distribution instead of three "sharp density peaks." Again for reasons of simplicity we did not do that. We are not aiming in the present paper for maximal realism, but rather for maximal simplicity and exact reproducibility of the results. My previous study [2] was much more ambitious, comprehensive, and seeking high realism. However, a price was paid in terms of simplicity and reproducibility.

[2] If you have 23 pennies in a row, then there are 22 inter-penny gaps. If you choose 3 of the gaps to put "walls" in, that divides the 23 pennies into 4 – automatically nonempty – subsets. So the number of ways to do this is $20 \times 21 \times 22/6 = 1540$. This number was also confirmed by our computer enumeration.

Note that in every one of our scenarios, each voter has a *strict* preference ordering of the three candidates, i.e. never feels unsure about the relative ranking of any candidate-pair.

3 The voting systems we shall consider

RANGE – Range voting: each voter gives a "score" to each candidate from the interval $[0, 100]$. The score is the honest utility except for the following strategic exaggeration that the centrist voters employ: they each pretend their "good" extremist has utility 100 or 0 (whichever is closer to g) instead of its true utility g (but they use the honest g if g is exactly at the halfway point $g = 50$). Since this is not completely honest range voting, call it "near-honest."[3] The candidate with the greatest sum-of-scores wins.

COND – Condorcet voting: each voter honestly ranks the candidates in order from best-to-worst. Then in our left-middle-right scenario, by a famous theorem ("Duncan Black's single-peakedness theorem" [1]) a Condorcet winner automatically exists, whom these votes pairwise-prefer over both opponents, unless there is an exact tie. [Many kinds of Condorcet voting systems have been proposed, but all reduce to the same one in this scenario.]

PLUR – Plurality voting: each voter honestly names his favorite candidate. The most-named candidate wins.

P+KING – Plurality with kingmaker: each voter honestly names his favorite candidate. The last-place (i.e. least named) "loser" candidate then donates all his votes to the candidate located closest to him, and then whoever has the most votes wins. [This approximates what might happen under three-candidate "asset voting" with strategic voters. It also, in our model, approximates what happens with strategic plurlaity voters.]

RUNOFF: each voter honestly ranks the candidates in order from best-to-worst. The candidate top-ranked by the fewest voters is eliminated and then a plurality election is conducted to decide between the two remaining candidates. [Instant runoff and delayed runoff are the same thing under our assumptions.]

IRV+KING – runoff with kingmaker: Same as runoff system, except that after that is done, the second-place finisher (non-eliminated non-winner) has the option of "retroactively dropping out" which causes the plurality-winner among the remaining two candidates to win (which may cause a difference, or may not). He chooses to drop out if that alters the winner to become closer to him.

Note: under our assumptions, the computer found that the IRV+KING winner was always the *same* as the Condorcet winner. Jan Kok and James J. Faran then both observed that this was a theorem that could be proven:

Theorem 1. *The IRV+KING winner is the same (at least, ignoring ties) as the Condorcet winner in the 3-candidate, 1-dimensional case.*

Proof: (Jan Kok) By Black's singlepeakedness theorem there is always a Condorcet Winner (CW), who will either be Middle or an extremist. In order for an extremist to be a CW, he has to have > 50% of the first-choice votes, in which case it doesn't matter if the runner-up withdraws. In the other case, where Middle is the CW, if Middle wins, then it doesn't matter what the runner-up does; if Middle was eliminated first, then the runner-up will drop out to make Middle win; Middle can't be the runner-up if he is the CW! ❄

BORDA: each voter honestly ranks the candidates in order from best-to-worst. The candidate with the greatest "Borda score-sum" wins, where you score 2 for being top-ranked and 1 for being middle-ranked, by a voter.

NANSON-BALDWIN: The candidate with the least Borda score is eliminated. Then a plurality election is conducted to decide between the two remaining candidates.

Note: under our assumptions, the computer found that the NANSON-BALDWIN winner was always the *same* as the Condorcet winner. And indeed

Theorem 2. *The NANSON-BALDWIN winner is the same (at least, ignoring ties) as the Condorcet winner in the 3-candidate, 1-dimensional case.*

Proof: Except in a tie-scenario, it is not possible for the Borda loser to be the Condorcet winner. By Black's singlepeakedness theorem there is always a Condorcet Winner (CW) in our model, and hence CW must be the pairwise victor among the two non-eliminated Nanson-Baldwin candidates. ❄

Note: both of the preceding theorems were first observed empirically in the computer output, then proven. Then these two methods were removed from our computer program as redundant.

APPROVAL: each voter "approves" all candidates with above-average utility in that voter's view. The most-approved candidate wins. This is the same thing as range voting if every voter "strategically exaggerates" his score for each candidate to be the max (100) or min (0) possible depending on whether that candidate has above-average utility or not.

[3] In other words, the extremist leftist voters say "We'll rate Left 100 and Right 0." (That so far is pretty realistic.) "Now what should we rate Middle? 0? 100? or honest? We'll do honest." The middle voters say "Middle is the best and we'll give him 100, and I'll give the worst extremist 0 so I can use the full range, i.e. not be a strategic idiot." (So far, pretty realistic.) "Now what should we give the good extremist?" Well, if they rated him honestly then RANGE voting would become in our model identical to BEST, i.e. would be perfect in this model (which would be both boring and already treated), so therefore the author made them "strategically exaggerate." This still is a moderately realistic model of what happens under the 2-party system because the 2 main candidates effectively *define* the line and everybody else is effectively in the middle whether they truly are or not (perhaps with a very high or low ratio of the number of leftist-centrist to rightist-centrist voters, but that is fine because our model covers such scenarios) and then it is strategic to exaggerate for the two majors. Exaggerating on a minor is of less strategic importance and hence is rarer. In the present paper we cannot *really* model strategic voters because there are no "pre-election polls" – but the present voter score-choosing-behavior with range voting may be viewed as a kind of weak attempt to be somewhat like strategic voters. Incidentally, it is important ot note that although in our particular setup here, fully honest range voting would be a perfect voting system (equivalent to BEST), that is *not* the case, even with fully honest voters, in general. That is because in general, the candidates would span different utility ranges for different voters.

APRNG: A half and half mix of near-honest range voting (above) and approval voting interpreted as a minned-and-maxxed-out range vote (also above). That is, each voter's score for each candidate is the sum of both of these two kinds of votes. The candidate with the greatest score-sum wins. This attempts to simulate a scenario with half near-honest and half strategic voters, which (based on experience polling real human voters [4]) seems a decent approximation of reality.

WORST: The candidate with the least summed-utility for all of society, magically wins.

BEST: The candidate with the greatest summed-utility for all of society, magically wins.

RANDOM: A pseudo-candidate with mean of the three summed (over all voters) candidate-utilities wins.

How we handle ties: The rightmost candidate always wins a tie (or in tied runoff elimination decisions, the rightmost always survives).

Roundoff errors: everything is done in exact integer arithmetic (and every integer fits in a 32-bit machine word), so there are no roundoff errors. Our computational results should therefore be exactly reproducible.

4 A few simple illustrative examples

A typical example in which RUNOFF voting performs badly is this.

#voters	their vote
36	$L > M > R$
35	$R > M > L$
29	$M > L > R$

Figure 4.1. In this 100-voter election, the Leftist L wins; the middle candidate M is eliminated in the first IRV round and then L beats R by 65-to-35. *However*, M would beat both opponents by at least 64-to-36 in a head-to-head race, so M is the Condorcet (and Borda) winner. ▲

This example is a simplification of what actually happened in the Chilean presidential election of 1970 (L=Allende, M=Tomic, R=Allesandri Rodriguez). Allende won and Tomic finished last (the election was actually held with plurality voting[4]) but probably Tomic "should" have won in the sense that he, almost unquestionably, would have defeated either opponent pairwise (because, e.g, the rightist's supporters would have seen him as the "lesser evil").

A typical example in which CONDORCET voting performs badly is this.

#voters	their vote
21	$L >> M > R$
20	$R >> M > L$
2	$M > L > R$

This is, in fact, the same example but with altered numbers. (It is a simplification of what probably would have happened in the 2000 USA presidential election between L=Gore, R=Bush, and M=Nader if it had been held with strategic Borda voters and if the straight popular vote had been employed without the "electoral college.") Although M is the Condorcet winner, I would feel much happier giving this victory to the plurality, Borda, and runoff winner L.

Although these are my mere subjective preferences, I think most people would agree with them, and also they could be objectively justified if these examples were supplemented with actual candidate-utility numbers designed to make the case. The point is that if this latter situation occurred in practice with honest voters, then probably automatically some of the preferences usually would be *large* ones, indicated by ">>." If so, *then* it becomes clear L is the correct winner. Intensity of preference is a concept *range* voters are capable of expressing, while Condorcet, Borda, and Runoff voters cannot.

A typical example in which BORDA and CONDORCET disagree about the winner is

#voters	their vote
53	$A > B > C$
47	$B > C > A$

Condorcet says A should win but Borda says B should win. Who really should win? If you ask people their opinion, I doubt you will get much agreement. However, with *range voting* we get clarity. *Both* Borda and Condorcet are easily made to be clearly correct, by assigning appropriate *intensities* to these preferences. That is something that happens automatically in range voting, which therefore is capable of making a sensible decision while Borda and Condorcet flounder.

Indeed, it seems that range voting would usually handle *all three* of these examples right. That gives you some intuitive, if nonquantitative, idea of why it is that the experiments in this paper found a clear superiority of range voting over all the other methods.

5 Results

There are two ways to assess the results. First, there is "Bayesian regret," which is the expected (over all scenarios) summed (over all voters) difference between the utilities of the BEST winner and the actual winner. Voting systems with smaller regret are better.

Second, we keep track, for each ordered pair (A, B) of voting methods, of the number of scenarios in which A gives a greater-summed-utility winner than B.

Finally, we also keep track (for each method) of the number of scenarios in which a tie occurred. Voting systems that yield exact ties less often, are better. We chose an odd prime number 23 of voters both in the hope that it would make some kinds of ties unlikely, and also because it was just small enough to keep all integers fitting in 32-bit words without overflow. But we also tried the highly composite even number 24 of voters in the hope that it would make ties *more* likely. (Note, if we have a tied Runoff elimination decision,

[4] Some have disputed this in the sense that they believe the 29 final votes in our table would be accurately be $M > R > L$ so that, more dramatically, the plurality winner Allende actually would have lost a head-to-head vote against either opponent. That dispute will, however, be irrelevant to us here.

that is counted as a "tie" even if both ways of breaking the tie ultimately result in the same election winner, and similarly for Nanson, which was the way we used to implement Condorcet for this purpose.)

The entire computer run required under 1 second. The tie-count results were

```
with 23 voters:
    Method    Tie Count    %Tied
 0: RANGE          3352    0.537
 1: COND          14580    2.338
 2: PLUR          32400    5.195
 3: PKING         43335    6.948
 4: RUNOFF        56700    9.091
 5: BORDA         14580    2.338
 6: APPROVAL          0    0.000
 7: APRNG             0    0.000
 8: WORST          3878    0.622
 9: BEST            546    0.088

with 24 voters:
    Method    Tie Count    %Tied
 0: RANGE          4677    0.652
 1: COND          61155    8.526
 2: PLUR          28350    3.953
 3: PKING         83754   11.677
 4: RUNOFF       101655   14.173
 5: BORDA         12150    1.694
 6: APPROVAL          0    0.000
 7: APRNG           190    0.026
 8: WORST          8441    1.177
 9: BEST            761    0.106
```

The remaining program output is below. It features a number of surprises.

1. It was entirely expected that Range Voting would enjoy by far the smallest tie-counts among genuine methods (where WORST, BEST, and RANDOM are not "genuine"). That turned out to be true *except* that for some reason presumably highly dependent on the number-theoretic properties of our particular scenarios with 23 and 24 voters, APPROVAL turned out to be incapable of yielding a tie, and this caused APRNG to have very few ties also. (Obviously, in real life RANGE will enjoy far fewer ties than APPROVAL.)

2. Our previous much more extensive Bayesian Regret study [2] had shown very robustly that RANGE was the best among all commonly proposed voting methods, *regardless* of whether the voters were "honest" or "strategic", how "ignorant" they were, how many voters or candidates there were, and which of wide variety of randomized utility generators (e.g. based on "issue spaces" of different dimensions) were employed.

So it was not surprising that in the present study also, RANGE was found to be clearly the best method. But the following was a big surprise. We expected that since APPROVAL (which is essentially the same as RANGE using maximally-strategic voters) is a lot worse than RANGE with near-honest voters, that APRNG – i.e. a 50-50 mixture of near-honest and strategic voters – would exhibit performance intermediate between RANGE and APPROVAL. That did not happen. Remarkably, APRNG actually enjoys *better* performance than either! – as measured either by Bayesian Regret or by "better scenario counts." (And both RANGE and APRNG are better than everything else by these measures.)

These mixture results seem a "Lucky Gift from God" causing RANGE to behave better in practice than one might have any right to expect. (It perhaps is an artifact of the specific model we've employed in this paper, but even if so, should still have noticeable impact in the real world[5]) And it is not the only such. Three other such range-voting-favoring "Lucky Gifts from God" have been previously identified [4] that both cause RANGE's superiority over competing voting methods to be even larger than what mathematics alone can say:

(a) Range voting experimentally causes human voters to satiate their inner desires to be "honest" rather than "strategic," to a much greater degree than plurality voting.

(b) This causes the "nursery effect," which is that, experimentally, candidates from small "infant" third parties receive much higher range vote totals (relative to the election winners) than they would under approval voting (which in turn gives them far higher totals then under plurality voting). Indeed, this effect may be so large that third parties will all die of "infant mortality" even with approval voting, but be able to "grow to adulthood" with range voting.

(c) Human range voters actually appear to have smaller per-candidate error rate than plurality voters, leading to *fewer* "spoiled" votes than with plurality voting, in spite of plurality's superior simplicity.

3. Indeed APRNG picks a winner superior to or as good as the average-summed-utility pseudo-winner (RANDOM) over 99.95% of the time, as opposed, e.g. to PLURality which does so only 93.25% of the time[6] APRNG picks as good a winner as BEST 95.50% of the time while PLUR only does so 76.03% of the time[7] APRNG picks as bad a winner as WORST only 0.015% of the time while PLUR does so 1.17% of the time[8]

4. Every voting method here is better than every other in at least some scenarios, with the exception that WORST is of course never better than anything. Thus proving you can pick a scenario of our type to make any voting method look bad...

5. In particular, for every voting method: RANDOM is better than it in at least some scenarios! (However, RANDOM is never better than a real voting method here if averaged over *all* scenarios.)

6. Both plain plurality PLUR and APPROVAL are inferior to PKING which in turn is not as good as COND and RUNOFF. This presumably indicates that "asset voting" (introduced in [3] as a multiwinner voting method) is superior as a single-winner voting method (in this model) to plurality

[5] I also tried a 50-50 mixture of *fully* honest range voting (i.e. BEST) with APPROVAL (the data for it is not shown) and it performed even better than APRNG.

[6] These numbers are computed from the table of "better scenario counts" at the end: $1 - 42130/623700 = 0.93245$, $1 - 296/623700 = 0.99953$.

[7] Again from the table of "better scenario counts" at the end: $1 - 28062/623700 = 0.9550$, $1 - 149525/623700 = 0.7603$.

[8] Again from that table: $1 - 623606/623700 = 0.000151$, $1 - 616431/623700 = 0.011654$.

and approval voting, but is probably *not* as good as the best single-winner methods. This finding to some extent justifies the feeling of unease expressed by many at the fact that Asset Voting is an "unconventional" voting system in which the candidates also play a role in deciding who wins – not the voters alone.

7. PLUR is the worst genuine voting method here.

8. Runoff+King (i.e. the "instant runoff" voting method but where we permit the runner-up to "retroactively drop out" of the race) seems – at least in our model – to be clearly an improvement over the ordinary instant-runoff single-transferable vote scheme.

Overall, this study reinforces the idea that RANGE voting is clearly the best of all the common proposals of single-winner methods as measured by either Bayesian regret, "better scenario count," or probability of yielding a tie. Furthermore, a great multitude of less-commonly proposed single-winner methods degenerate to the ones we have used here in the 3-candidate 1-dimensional scenarios here, and thus our study also supports the notion that RANGE is superior to them too.

Obviously, the unrealistic fact that we only consider 23-voter elections, is of no great import. That is since the results would not have changed if 2300-voter elections coming in 100-voter blocs were used. The important thing is that we have covered every possible 23-voter scenario exactly once, thus getting good coverage of "the whole space."

This study has not included[9] notions of "strategic voting." It mostly was not possible to put in strategic voters because our scenarios did not include "pre-election polls." However, if we *had* put in strategic voters, then our conclusion that RANGE is the best method, would presumably have only become "more true" since the other methods (e.g. Borda) would then presumably behave comparably or worse to the way they behaved with honest voters. And indeed the previous study [2] *did* examine strategic voting based on "pre-election polls" and did find that Condorcet, Borda, and Runoff all exhibited significantly worse performance when the voters were strategic than when they were honest (indeed, these methods then always performed worse than APPROVAL).

	Method	UtilSum	AvgUtil	AvgRegret
0:	RANGE	4087593924	80.911	2.033
1:	COND	3639878724	72.049	10.896
2:	PLUR	2792416224	55.274	27.670
3:	PKING	3304502274	65.410	17.534
4:	RUNOFF	3402042474	67.341	15.603
5:	BORDA	3813745224	75.490	7.454
6:	APPROVAL	3279023724	64.906	18.038
7:	APRNG	4092408024	81.006	1.938
8:	WORST	702356218	13.903	69.042
9:	BEST	4190316444	82.944	0.000
10:	RANDOM	2186015816	43.271	39.674

```
number of voters = 23     total number of scenarios = 623700
BetterCt[][]:
```

	0	1	2	3	4	5	6	7	8	9	10
	RANGE	COND	PLUR	PKING	RUNOFF	BORDA	APPROVAL	APRNG	WORST	BEST	RANDOM
0:	0	74096	140855	97481	95160	61600	89724	348	623603	0	623354
1:	23164	0	83007	27911	26539	19214	74446	22604	622834	0	619786
2:	19896	13239	0	22143	11271	32453	87685	19689	616431	0	581325
3:	22189	3246	58393	0	1197	22460	77692	21701	607990	0	597901
4:	22301	4606	60895	14571	0	23820	79052	21805	621036	0	603912
5:	23146	31692	114699	59603	58231	0	55232	22322	622674	0	619190
6:	23288	58942	141949	86853	85481	27250	0	22096	622426	0	602400
7:	1204	74380	141498	97833	95520	61620	89376	0	623606	0	623386
8:	0	0	0	0	0	0	0	0	0	0	0
9:	28918	79832	149525	103473	101767	67354	95336	28062	623700	0	623700
10:	328	3808	42130	25711	19664	4429	21094	296	623700	0	0

6 Acknowledgements

I thank Jan Kok and James J. Faran for helpful correspondence. The computer program (written in C) used to obtain this paper's results is downloadable from http://math.temple.edu/∼wds/homepage/works.html #95. A great deal of information about range voting vis-a-vis other voting systems is available at the Center for Range Voting website http://www.RangeVoting.org.

References

[1] Duncan Black: The theory of committees and elections, Cambridge Univ. Press 1958.

[2] Warren D. Smith: Range Voting, #56 at http://math.temple.edu/∼wds/homepage/works.html.

[3] Warren D. Smith: "Asset voting" scheme for multiwinner elections, #77 at http://math.temple.edu/∼wds/homepage/works.html.

[4] W.D.Smith, J.N.Quintal, D.S.Greene: What if the 2004 US presidential election had been held using Range or Approval voting? #82 at http://math.temple.edu/∼wds/homepage/works.html.

[9] Or more precisely: only included in a few cases in a rather weak way, as discussed in §3.