Completion of Gibbard-Satterthwaite impossibility theorem; range voting and voter honesty

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

October 23, 2006

Abstract — **Let S be a "reasonable" single-winner voting system. (The precise definition of "reasonable" will vary from theorem to theorem and is not stated in this abstract.) Then**

(a) For each $C \geq 3$: if S is based on rank-order ballots (with equalities either permitted or forbidden) then there exist C-candidate election situations with "complete information" (i.e, the voter knows everybody else's votes) in which voting honestly is not best voting strategy.

(b) If S is range voting, then in every C-candidate election situation ($C \geq 1$) with complete information, and also in every C-candidate election situation with incomplete information ($1 \leq C \leq 3$), there is a "semi-honest" vote (i.e, in which the $<$, $>$, and $=$ relations among the candidate-scores are valid for a *limit* of scores obeying the honest relations) which is strategically best.

(c) If S is based on *either* rank-order ballots (with equalities either permitted or forbidden) *or* candidate-scoring "range vote" type ballots where each candidate is rated with a real number, then for each $C \geq 4$ there exist C-candidate election situations with incomplete information in which no *semi*-honest vote is best voting strategy.

Part (a) is Gibbard & Satterthwaite's impossibility theorem. These results show a sense in which range voting is a best possible deterministic single-winner voting system. These theorems also hold for certain classes of probabilistic voting systems (in which chance plays a role in determining the winner) but not all. We conclude by introducing and beginning the study of the "Nash model" of voter honesty.

Keywords — Voter honesty, semi-honesty, strategy, strong Nash equilibria, Nash model.

1 Introduction

The Gibbard-Sattherthwaite theorem [1][2][9][10][14][17][19] said, essentially, that, in reasonable single-winner voting systems with rank-order ballots (with equalities allowed [14] or forbidden [9]), "honest voting" and "strategic voting" sadly are not the same thing.

More precisely: the only deterministic voting system based on strict-rank-order ballots in which

(a) voting honestly is always best (or co-equal best) strategy,
(b) unanimously top-ranked contenders get elected,
(c) there are at least 3 candidates and at least 3 voters

is a "dictatorship" in which some particular voter (the "dictator") gets whatever he wants regardless of the other votes. (A review is [17]; the present report now solves the top open problem posed there.)

We observe that condition (c) is best possible because:

- with only 2 candidates, voting honestly is best strategy in the usual majority-vote system;
- and with ≤ 2 voters, voting honestly is best strategy in the plurality with random (or deterministically biased in favor of the first voter) tie-breaking system. (However, in what would seem to be the *most sensible* 3-candidate 2-voter voting system based on rank-order ballots, illustrated in figure 4.1, best strategy can involve dishonest voting.)

Define "range voting" to be the following voting system:

1. Each voter gives to each candidate a real-number score in the interval $[0, 1]$.
2. The candidate with the highest average score wins. (Ties broken randomly or according to some predetermined preference order.)

"Approval voting" is the same except the allowed scores are either 1 or 0 with intermediate real numbers now forbidden.

Define a "semi-honest vote" to be one in which the scores for the candidates obey a set of $<$, $>$, and $=$ relations which arise as a *limit* of the true relations.[1] For example, if you truly feel $A > B$, then a range vote which says $A > B$ would be both honest and semi-honest, one which says $A = B$ would be dishonest but semi-honest, and $A < B$ would be neither honest nor semi-honest. For a second example, if you truly feel $A = B$, then every semi-honest vote says $A = B$.

Range voting *evades* the Gibbard-Sattherthwaite "impossibility" theorem in the following two senses which concern semi-honest voting:

I. In any 3-candidate range-voting election situation (including ones in which you have incomplete information about the

*21 Shore Oaks Drive, Stony Brook NY 11790.
[1] Brams & Fishburn [4] called this "sincere" voting in the context of "approval voting."

other votes), there exists a strategically-best range vote which happens to be semi-honest.

Proof: If in your view all candidates are equal, then you do not care what your vote is, so cast a semi-honest vote. Otherwise: It can never hurt you (i.e. never decrease the expected election utility) to give your true-favorite(s) the maximum score 1. Symmetrically, it also can never hurt you to give your truly-most-hated candidate(s) the minimum score 0. Then any vote with these properties is semi-honest. Q.E.D.

II. In any N-candidate range-voting election in which you have complete information about every other vote, there always exists a semi-honest strategically-best range vote.

Proof: Let X be the greatest expected utility (in your view) of the winner that (by choosing your vote correctly) you can achieve. Then deliver the vote in which all candidates better than or equal to X are 1, and all candidates worse than X are 0. (It never is useful to give the latter class of candidates a positive score, since that might help one tie for the win, which would decrease the utility of the tied-winner set. Symmetrically it never is useful to give the former class a below-maximum score.) Q.E.D.

Remark 1: The above proof may have left you with the impression that it always is strategically optimal to range-vote in "approval voting" [4] style in which every score is either "maxed-out" or "minned-out" to 1 or 0, with no intermediate scores. But that impression is *false* in *in*complete information scenarios. For example, consider a three-voter three-candidate situation where the other two voters (in sum) *either* cast the summed-vote $A = 1$, $B = 1.1$, $C = 1.2$ *or* $A = 0$, $B = 1.1$, $C = 1.2$ but you do not know which. Suppose your candidate-election utilities are $U_A = 10$, $U_B = 5$, $U_C = 0$. Then it is strategically best for you to cast the honest range vote $A = 1$, $B = 0.5$, $C = 0$, (or a slightly-distorted vote such as $A = 1$, $B = 0.6$, $C = 0$ would be equally strategically good); this is strictly superior to every possible approval-style vote such as $(1, 0, 0)$, $(1, 1, 0)$, or $(1, 1, 1)$.

Remark 2: Van Hees and Dowding [11] in a manuscript titled "in praise of manipulation," defined two types of voter-dishonesty (in ranked-ballot single-winner voting systems) – namely "sincere" and "insincere" manipulation! They consider the former less damaging. According to them, a subset S of voters "manipulates sincerely" if:

1. *assuming* everybody else continues to rank their current top-ranked candidate top,
2. all members of S changing their vote to a single rank-ordering which dishonestly top-ranks some candidate A, leads to A winning the election,
3. which is an outcome all S-members prefer to whatever otherwise would have happened;
4. Furthermore, there is no other outcome that S can achieve by a unilateral switch that they all prefer.

Van Hees and Dowding then defined a voting system to be "immune to insinsere manipulation" if whenever insincere manipulation works, then sincere manipulation does also. Then then *proved* (their theorem 1) that voting systems based on rank-order ballots are immune to insinsere manipulation if and only if they are "monotonic." Finally, we note that the Borda count and many Condorcet voting systems (including Schulze beatpaths [15]) are monotonic, but "instant runoff" (IRV) is not monotonic [3]. (Range voting is also monotonic, but it is not based on rank-order ballots.)

I have not checked that van Hees & Dowding's proof is valid, but assuming it is, their arguments "in praise of manipulation" still leave me unimpressed, because it is quite clear to me that Borda Count is *not* a good voting system precisely because of its enormous vulnerability to manipulation, including the "DH3 pathology" pictured in figure 1.1.

#voters	their vote
x_1	$A > D > B > C$
x_2	$A > D > C > B$
y_1	$B > D > C > A$
y_2	$B > D > A > C$
z_1	$C > D > A > B$
z_2	$C > D > B > A$

Figure 1.1. DH3 pathology. Assume A, B, C are three excellent candidates and D is a mediocrity ("dark horse"). All voters honestly regard all three of $\{A, B, C\}$ to be far superior to D, but their opinions are split concerning the ordering within $\{A, B, C\}$. Therefore, each voter strategically ranks his favorite top, and his two top rivals artificially "last" (exaggerating to get more "discriminating power"). The result is the scenario here with $x_1 + x_2 \approx y_1 + y_2 \approx z_1 + z_2$. Then the uniquely-worst candidate D becomes the Condorcet-Winner and wins the election under *any* voting system that elects Condorcet-Winners or Smith-Set members! D also becomes the Borda winner. (However, Plurality, IRV, Approval, and Range would elect one of $\{A, B, C\}$.) This scenario is both *common* in practice and *maximally bad*. It illustrates the devastating vulnerability of Borda and systems obeying Condorcet's principle to strategic manipulation. Observe that any one of the three voter groups, if it decided to vote honestly, would lose to one of the less-honest competitor groups, so the members of any group feel they "must fight fire with fire." ▲

Effects of that nature indeed were immediately observed [13] in the only government (Kiribati) which adopted Borda voting (which soon caused them to abandon Borda in favor of the plurality system). Indeed, I think almost everyone would agree (subjectively and see [6][7]) that Borda is much more horribly vulnerable to manipulation than IRV, despite Van Hees & Dowding's theorem arguing in the opposite direction because Borda is "immune to insincere manipulation" while IRV isn't. (Also, objectively, note that IRV is immune to the DH3 pathology.) In short, just because a voting system is "immune to insinsere manipulation" in Van Hees & Dowding's sense, is in my view a very insufficient condition for the acceptability of that voting system.

So I believe that the notion of "semi-honesty" advanced here is a better way to develop things – but with roughly the same goal in mind as van Hees & Dowding.

Remark 3: It is possible to conduct "range voting with rank-order ballot votes" by first *converting* all the rank-order ballots (which are allowed to contain equalities) to range-style ballots, then conducting a range-voting election. The conversion procedure is to score the top candidate 1, the bottom candidate 0 and the remaining equality-classes equally spaced between. In that case we get a voting system based on rank-order ballots (with equalities permitted) in which semi-honest voting is strategic, etc (to be more precise, it has the same honesty-properties as range voting).

However, all the above facts left the following **question** open: Does there exist a reasonable voting system in which a semi-honest vote is always strategically best, even in election situations in which you have *incomplete* information about the other votes?

We shall prove in §2 that the answer is YES if there are 3 or fewer candidates (range and approval voting both work), but NO if there are 4 or more candidates. This result depends on a certain defintion of the word "reasonable" and it will even be valid if some kinds of probabilistic voting systems are permitted, although the theorem statement and proof both will simplify (and be more attractive) if the voting system is deterministic.

Thus, this investigation has provided a sense in which range voting is a *best possible* voting system, and better than (or at least as good as) any system based on rank-order ballots. Approval voting also is best possible in the same sense but is rather peculiar in that it actually *forbids* honest voting in generic scenarios with ≥ 3 candidates, although permitting semi-honesty (range voting, in contrast, enables honesty, but it can be strategically unwise).

Another interesting way to look at the question of how much voters are motivated to be honest is in what we call the **Nash model**. There are two ways to view this model:

1. There is *public knowledge* of every voter's honest preferences and utility values but their actual *votes* are secret (at least until after the election is over).
2. There is no public knowledge of anything about the other voters at all – there is simply a black box that inputs votes and outputs elections results. An enormous number of experiments are conducted with the black box by the voters to find voting strategies that cause the black box to output results which they consider to be best for them in the sense that each voter, by altering his voting strategy while the other voters remain with fixed strategies (which may however be randomized), cannot increase the expected utility (to him) of the black-box output.

In either picture, if there are 2 voters, then best voting-strategy in the resulting 2-player Von Neumann "matrix game" is (in general) *randomized*, and if there are ≥ 3 voters, then there will be one or more "Nash equilibria" in which no player can improve his randomized strategy if the other players hold theirs fixed.

Despite its fundamental logical importance, the Nash model has been examined little or not at all in preceding voting systems literature, and we shall barely begin its investigation here.

The **top open question** about voter honesty in the Nash model is: "Does there exist a reasonable voting system in which an honest, or semi-honest, voting strategy is always strategically best (in the sense that the vote *averaged* over that voter's randomness is honest or semi-honest)?"

Observe that in the Nash model, some approval voter could, in principle, provide an "honest-mean" probabilistic combination of approval votes no matter how many candidates there were, so that the peculiarity of approval voting (i.e. that it forbids voter honesty) vanishes. And indeed when we analyze the apparently most-sensible 3-candidate 2-voter voting system based on rank-order ballots (illustrated in figure 4.1), we shall exhibit an election situation in which each strategic voter casts semi-honest-mean votes even though some component votes inside that average are dishonest.

However, it is easy to see that in 3-candidate scenarios where, e.g. your favorite candidate has no hope of winning, your best strategy is only semi-honest and not full-honest, even in the Nash model. Thus the Nash model apparently does *not* enable escaping from the iron grip of the Gibbard-Satterthwaite impossibility theorem (although semi-honesty does enable such an escape).

What if the Nash model and semi-honesty *both* are put in? Then we shall show that approval voting fails the Nash-*semi*-honesty test in 6-candidate scenarios but passes it in 3-candidate scenarios – the cases of 4 and 5 candidates remain open. More importantly, the question of whether some unknown reasonable voting system exists that is superior to approval voting in the sense it passes the Nash-semi-honesty test in 4 or 5-candidate scenarios – or even in *every* scenario – remains open.

2 Statements and proofs of main results about semi-honesty

Definition. A "voting system" inputs votes (which we here demand be either $[0, 1]$-range-style ballots, or rank-order ballots with equalities permitted) and outputs the identity of a winner.

Consider the following $(2N + 2M)$-voter 4-candidate **situation** (we have written this for $[0, 1]$-range-style ballots, but if you prefer voting systems based on rank-order ballots with equalities permitted, then please translate the votes below into the equivalent votes in that format):

#voters	their vote
N	$A = 1$, $B = C = D = 0$
N	$B = 1$, $A = C = D = 0$
M	$B = 1$, $A = C = 1/2$, $D = 0$
M	$A = C = 1$, $B = D = 0$

Figure 2.1. A $(2N+2M)$-voter 4-candidate situation, where $1 \leq M \ll N$. Normally we shall only consider the case $M = 1$, but at the end of §2 is a remark about the uses of larger M. ▲

Definition: "Reasonable" voting systems are:
(i) invariant under permutations of the candidate names (or, more weakly, we may merely demand that the truths of the following *properties* be invariant), and
(ii) in the *situation* above (if $N \gg M \geq 1$) *but* with the last $2M$ votes altered arbitrarily, it elects either A or B, but never C or D – or more weakly we may merely demand that in the $N/M \to \infty$ limit, the probability of electing C or D tends to 0. [This is a "nearly-unanimous top-2" property.]

(iii) in the *situation* above (if $N \gg M \geq 1$) it elects A with probability greater than the probability of electing B.

(iv) in the *situation* above (if $N \gg M \geq 1$) *but* with the last M votes altered arbitrarily to a form obeying $1 \geq B \geq A \geq 0$ and $0 \leq C \leq 1$, $0 \leq D \leq 1$, it elects B with probability greater than or equal to the probability of electing A. [This is a "unanimous top-2 with majority for one" property.]

(v) in the *situation* above (if $N \gg M \geq 1$) but with the last M votes altered arbitrarily to a form obeying $1 \geq A \geq B \geq 1/2 \geq C \geq D \geq 0$ or obeying $1 \geq C \geq D \geq 1/2 \geq A \geq B \geq 0$ (or *if* we are considering rank-order ballots instead of range-voting-type ballots with the last vote altered to say either $A > B > C > D$ or $C > D > A > B$ or any of these with any $>$s replaced by $=$s), it elects B with probability greater than or equal to the probability of electing A. [This is a "near-unanimous top-2 with equal-or-majority for one" property.]

(vi) in the *situation* above but with the last M votes replaced by ones of form $1 \geq A > B > C > D \geq 0$ (or a limit form), altering that vote to $1 \geq A > C > B > D \geq 0$ (or a limit form) by increasing the value of C and/or decreasing the value of B, holding all else fixed, cannot increase the probability-ratio for B winning versus A.

(vii) in the *situation* above but with the last M votes replaced by one of form $1 \geq C > D > A > B \geq 0$ (or a limit form), altering that vote to $1 \geq C > A > D > B \geq 0$ (or a limit form) by decreasing the value of D and/or increasing the value of A, holding all else fixed, cannot increase the probability-ratio for B winning versus A. [Vi and vii are weakened monotonicity properties and as we shall see in remark 3 below, they may largely be dropped.]

(viii) the *gap* between the probability ratio for A versus B winning in iii, and that probability-ratio in iv, v, is at least some positive constant k. Indeed, we may weaken the statements "greater than 1/2 and "greater than or equal to 1/2" in the preceding properties; it suffices *merely* for a gap$> k$ to exist between the A/B probability ratios in these cases that contains a number near 1 inside the gap. [Note that in the *deterministic* case all probabilities are 1 or 0 and so these gap statements reduce to trivialities.]

Remark 3: In our main theorem below, it should actually be possible to drop or weaken some of these reasonability assumptions, i.e. prove them instead of assuming them. For example if vi or vii were disobeyed sufficiently strongly that the "increase" were bounded below by some positive constant, then it clearly would pay to vote dishonestly so that the theorem would be proven immediately; so we only need to prove the theorem in the case where vi and vii are obeyed (or only disobeyed weakly)

Remark 4: It is rather embarrassing that Gibbard & Satterthwaite had a very *simple* definition of "reasonable" voting system, whereas we needed a long and complicated one. It is unclear how much our definition can be simplified. However, let us point out that in spite of its length, our definition *is* analogous to Gibbard & Satterthwaite's in the sense that all of our criteria (and all of theirs) can be expressed in the form: "in this election scenario XXX, in the limit when the number of certain kinds of voters are taken to be infinite, the following winner or winners must happen YYY."

Remark 5: An inequivalent different possible definition of "reasonable" could be to demand that the voting system always elect a "Condorcet winner" when one exists. But it is known already (some proofs, most originally arising from Kevin Venzke, are on the CRV website [18]) that dishonest and not-semi-honest voting (which indeed "betrays your favorite") always is uniquely strategically optimal in at least some 3-candidate complete information election scenario if the voting system is (1) based on rank-order ballots (with or without equalities permitted) and (2) always elects Condorcet winners.

Theorem 1 (Main result). *In any reasonable voting system based on either* $[0,1]$*-range-style ballots, or rank-order ballots with equalities permitted, there exists a 4-candidate election scenario in which you have partial information about the other votes, in which your only strategically-optimal votes are not semi-honest.*

Proof: [2] Let your true candidate election utilities be

$$U_A = X + Y + Z, \quad U_B = Y + Z, \quad U_C = Z, \quad U_D = 0 \quad (1)$$

where X, Y, and Z are positive constants. Let the other votes be either of the form in situation 1 or situation 2 (see tables 2.2 and 2.3) *but* you do not know *which*.

#voters	their vote
N	$A = 1,\ B = C = D = 0$
N	$B = 1,\ A = C = D = 0$
M	$B = 1,\ A = C = 1/2,\ D = 0$

Figure 2.2. Situation 1. ▲

#voters	their vote
N	$C = 1,\ A = B = D = 0$
N	$D = 1,\ A = B = C = 0$
M	$D = 1,\ B = C = 1/2,\ A = 0$

Figure 2.3. Situation 2. ▲

In that case, if you vote (non-semi-honestly)

$$A = C = 1, \quad B = D = 0 \quad (2)$$

[2] Brams & Fishburn [4] state as "theorem 2.3" on pge 30 "every voting system is sincere for dichotomous voters. But only approval voting is sincere for trichotomous voters, and no voting system is sincere for multichotomous voters." An "Nchotomous" voter is one whose candidate-utility values lie in an ($\leq N$)-element set. This theorem seems at first to be the same result as ours! However, actually, it is a much weaker result (whose proof, indeed is so trivial that B&F omit it) *because* the definition B&F here use for "voting system" is very restrictive: The only "voting systems" allowed are those in which each voter delivers 1 or 0 to each candidate, with the total number of 1s he delivers being from a fixed set of allowed cardinalities (e.g. for plurality, the cardinality-set is $\{1\}$, for approval it is $\{0, 1, 2, 3..., C\}$ in a C-candidate election), and with the candidate with the highest sum winning. Thus, for the purposes of B&F's theorem, neither Borda, instant runoff, nor range voting would be considered "voting systems." Still, though, this is a natural predecessor of our result.

then in situation 1, you will cause A to win with probability$> k+1/2$, and B and D to win with probability 0 (in the large-N limit); while in situation 2, you will cause C to win with probability$> k+1/2$ (for some constant $k>0$) and A and B to win with probability 0. (By reasonability i, ii, iii, viii.) However, if you cast any semi-honest vote – the possibilities are

$$A>B>C>D, \quad A=B>C>D, \tag{3}$$

$$A>B=C>D, \quad A>B>C=D, \tag{4}$$

$$A=B=C>D, \quad A>B=C=D, \tag{5}$$

$$A=B>C=D, \quad A=B=C=D \tag{6}$$

then by reasonability iv, v, *either* in situation 1 you will cause B to be elected with probability$\geq 1/2$ while C and D lose (costing you expected utility$> kX$), *or* in situation 2 cause D to be elected with probability$\geq 1/2$ while A and B lose (costing expected utility$> kZ$) but there will be no compensating benefits in the other situation, due to reasonability vi and vii, so these costs cannot be recompensed. (Note, we have oversimplified slightly in the above by saying "$k+1/2$" and "$1/2$," but the truth by reasonability viii is these are just two quantities near $1/2$ that are separated by a gap of at least k, and that is good enough for our purposes.) Q.E.D.

Theorem 1 and its proof still leave open the possibility of having certain kinds of probabilistic voting system in which semi-honest voting is always strategic. Here is an interesting new one:

Random-triplet range voting:

1. Each voter gives to each candidate a positive real-number score.
2. A triple of candidates (call them A,B,C) is selected randomly.
3. All voter-scores for all candidates other than A,B,C are ignored.
4. All range votes are rescaled by a linear transformation so the highest scorer among A,B,C now gets score 1 and the lowest 0. (Except "null" votes $A=B=C$ are discarded.)
5. Using these rescaled range votes, the election winner is the member of $\{A,B,C\}$ with the greatest average score (break ties randomly or according to some predetermined preference order).

Theorem 2. *In random-triplet range voting, voting semi-honestly is always best strategy.*

Proof sketch. It suffices to prove (for any candidate pair A,B) you will never want to vote $B>A$ if you truly feel $A>B$ or $A=B$. Suppose for a contradiction such misordering was strategically helpful. Then it must help in some 3-candidate range-voting (with rescaling) election. However, as a lemma, we can prove that in a 3-candidate range-voting (with rescaling) election, if you cast a vote that is not semi-honest, then

1. if the mis-ordered pair is adjacent in your vote (such as $B>A>C$ when honestly $A>B>C$ or $A=B>C$, or $B>A=C$ when honestly $A=B=C$ of $A>B=C$) then you can always replace the vote by one with that pair equal ($A=B>C$ or $A=B=C$) without hurt;
2. and if it is non-adjacent ($B>C>A$ when honestly $A>C>B$ or $A>C=B$ or $A=C>B$) then swapping the pair values (swapping A,B) never hurts your expected utility [we state but actually shall not use this middle fact], and
3. Also setting $A=B$ in the non-adjacent situation while preserving whichever inequality about C was true (note: both cannot be) to get $A=B>C$ or $C>A=B$ cannot hurt you, or finally if truly $A>C>B$ (so *all* inequalities were fully-dishonest) then the null vote $A=B=C$ cannot hurt you.

(Throughout by "hurt" we mean more precisely, "decrease expected utility of election result" where each candidate has some election-utility for each voter.)

So any non-semi-honest N-candidate vote can always be altered to make it nearer to semi-honesty by replacing all dishonest inequalities by equalities, without hurt. And after a finite number of such alterations, your vote must reach a semi-honest state. Q.E.D.

3 Conclusions about semi-honesty in voting

This all has shown senses in which

1. range voting is superior to every rank-order ballot system, and
2. no reasonable deterministic (or we can even permit certain kinds of probabilistic) voting system can do better than range voting in terms of inspiring voter (semi)honesty.

However, there are certain probabilistic voting systems (i.e. in which chance plays a role in selecting the winner) which are superior both to range voting and to every deterministic system, in the sense that honest (or semi-honest) voting is *always* strategically optimal in those systems.

Random-triplet range voting is one such probabilistic voting system. The other two I know of (found by Gibbard [9]) are "do whatever a random voter wants" and "use rank-order ballots to conduct a majority vote on a random candidate pair." I conjecture there are no others besides these three (and probabilistic mixtures of them).

A different kind of vote is "rank-order with threshold" votes, e.g. $A>B>$ threshold $>C\geq D$ where "threshold" is an artificial extra candidate who can never win the election. It is also possible to prove a version of theorem 1 with this sort of vote permitted; you need to make a more complicated definition of "reasonability" and our proof-examples can be converted to threshold-style votes by converting each vote to a *set* of votes with the threshold located in every possible position. We intentionally do not give a detailed exposition (which would more than double the length of this paper).

Also, although theorem 1 (and our "reasonableness" definition) focused on 4-candidate elections, they may easily be redone for C-candidate elections for any $C\geq 4$. Simply adjoin $C-4$ additional "no hope" candidates whom every other voter scores 0. These extra candidates will not affect the logic.

Compared to Gibbard & Satterthwaite's original theorems, our theorems are considerably stronger since they completely settle the question of what happens when we extend "honesty" to permit "semi-honesty," and also extends everything to allow incomplete-information scenarios. But unfortunately, our theorems may be criticized as being *weaker* in the sense that they need to make considerably more "reasonability" assumptions about the voting system. In particular, what most people regard as the or least-reasonable reasonability assumption, are the demands in iii, iv, v, vi that the election result be swung by *one* vote in situations of perfect balance aside from the swinging-votes.

This criticism may be blunted as follows. Consider allowing M in the tabulated election situations with now $1 \le M \ll N$, including large M. (Previously we had only considered the case $M = 1$.) In that case it is M (or $2M$) votes that are doing the swinging, and we can refrain from this demand except for *large* M. In that case, the theorems and proof still work except that the "strategic voting" is to be performed by a *bloc* of M voters rather than just one. But then we observe that if the members of this bloc switch voting policies *one by one*, then the election first is swung by just *one* bloc-member, and then we get a situation in which just that one voter is strategically impelled to vote dishonestly. So we can get a similar theorem out despite considerably weaker assumptions in.

4 Voting as a Von Neumann matrix game (if there are 2 voters)

$V_1 \backslash V_2$	$A>B>C$	$A>C>B$	$B>A>C$	$B>C>A$	$C>A>B$	$C>B>A$
$A>B>C$	A wins	A wins	AB tie	B wins	A wins	ABC tie
$A>C>B$	A wins	A wins	A wins	ABC tie	AC tie	C wins
$B>A>C$	AB tie	A wins	B wins	B wins	ABC tie	B wins
$B>C>A$	B wins	ABC tie	B wins	B wins	C wins	BC tie
$C>A>B$	A wins	AC tie	ABC tie	C wins	C wins	C wins
$C>B>A$	ABC tie	C wins	B wins	BC tie	C wins	C wins

Figure 4.1. A 2-voter 3-candidate Condorcet voting system – which would seem to be the most sensible 2-voter 3-candidate voting system based on rank-order ballots – as a 2-player 6×6 matrix game in the style of Von Neumann and Morgenstern [8][20]. The candidates are A, B, and C. Each event such as "A wins" has a numerical expected utility ("payoff") for each of the two voters (players); in general these payoffs differ for each player. ▲

In the matrix game of figure 4.1, assume "A wins" has utility U_A to player 1 (and W_A to player 2), and all ties are broken randomly so that, e.g. "AB tie" has utility $(W_A + W_B)/2$ to player 2.

Suppose player 1 has utilities $U_A = 9$, $U_B = 3$, $U_C = 0$. An incomplete account of best play with this player 1 is

1. If player 2's favorite is A then both honestly vote "$A > \dots$" and A is elected.
2. If player 2's favorite is C by a large enough margin, then V_1 is $A > B > C$ and V_2 is $C > B > A$ (but $C > A > B$ would be unstrategic even if honest) and we get a 3-way ABC tie (which is optimal play for both players in the sense tht either player would get a worse result by retrospectively changing their vote with the other player holding fixed). Note that in this scenario player 2 can be uniquely strategically forced to vote dishonestly.
3. If player 2's honest assessment is $W_B = 9$, $W_C = 3$, $W_A = 0$ then a cyclic chain of best moves and countermoves is
$$1: A>B>C \implies 2: B>C>A \implies 1: A>C>B$$
$$\implies 2: B>C>A \implies 1: A>C>B \implies \dots \quad (7)$$
where each move-countermove pair respectively leads to these results (where "ABC tie" repeats forever): B wins, ABC tie, ABC tie, ABC tie...
4. If player 2's honest assessment is $W_B = 9$, $W_C = 6$, $W_A = 0$ then a cyclic chain of best moves and countermoves is
$$1: A>B>C \implies 2: B>C>A \implies 1: A>C>B$$
$$\implies 2: C>B>A \implies 1: A>B>C \implies \dots \quad (8)$$
where each move-countermove pair respectively leads to these 4 results in cyclic order: B wins, ABC tie, C wins, ABC tie.

As the last of these makes clear, it is strategically *foolish* for either player to reveal his vote to the other, since that knowledge allows the certain selection of the best countermove. As is well known, best strategy instead is *randomized*. In the last situation, player 1's best strategy is to flip a coin and vote $A > B > C$ with probability 1/2 and $A > C > B$ with probability 1/2 (which in some "averaged" sense is the semi-honest vote $A > B = C$), while player 2's best strategy is to vote $B > C > A$ and $C > B > A$ with probability 1/2 each (which in some averaged sense is the semi-honest vote $B = C > A$). With these strategies, the election result probabilities in table 4.2 arise, and neither player can improve his or her expected utility by altering their probability-vectors.

result	probability	payoff_1	payoff_2	summed payoff
B wins	1/4	4	9	13
C wins	1/4	0	6	6
ABC tie	1/2	4	5	9

Figure 4.2. Election results with best randomized voting strategy for player 1 (who has utilities $U_A = 9$, $U_B = 3$, $U_C = 0$) and player 2 (who has utilities $W_A = 0$, $W_B = 9$, $W_C = 6$). ▲

The best result for the 2-player society as a whole is obtained if both vote honestly and B wins (summed-payoff=13) *but* each player's use of partially-dishonest voting strategy leads to a worse result for society as a whole (but a better result for each player alone given that the other player is voting strategically).

Thus for these two particular players, this voting system is equivalent to the 2×2 game in figure 4.3 (in which each player either votes "honestly" or "dishonestly")

player 1\ player 2	honest ($B > C > A$)	dishonest ($C > B > A$)
honest($A > B > C$)	3, 9	4, 5
dishonest($A > C > B$)	4, 5	0, 6

Figure 4.3. Each player in the preceding scenario can chose to vote either "honestly" or "dishonestly," and these are the payoffs (for player 1 and player 2 respectively) that result. Best strategy for both players is to flip a coin. ▲

To conclude: This example has made it clear that in 2-voter, 3-candidate scenarios using the "most sensible" such voting system based on rank-order ballots with equalities forbidden (i.e. the system in figure 4.1), dishonest voting can be uniquely strategically best in both complete-information scenarios and in the Nash model. However, it has also shown that averaging votes over randomness in the Nash model can sometimes make strategic voters be "more honest."

5 Nash equilibria if there are more than 2 voters

If there are more than 2 players, the Von Neumann theory of games breaks down and there is no longer, in general, any clear meaning of the term "best strategy." However, there is, according to a famous theorem of John Nash [12], one or more "Nash equilibria" in which each voter individually cannot improve his strategy, given that all the others keep their strategies fixed. And again, in general, these voter strategies each will be randomized, i.e. described as probability distributions over all the possible votes that player could cast.

Perhaps the first thing we should say about Nash equilibria is that some of them can be exceedingly stupid. For example, consider a 2-candidate V-voter ($V \geq 3$) simple-majority-vote election where every voter prefers $A > B$. If every voter stupidly votes "$B > A$" then B wins. This situation is a (stupid) Nash equilibrium because no voter, acting alone, can change the election result. And it of course involves dishonest votes. However, dishonest voting obviously is not strategically *forced* in the sense that any voter can change his vote to become honest without suffering a worsened result, and indeed if most of them do so, they get a better election result[3] Nash examples are more interesting when dishonesty is strategically forced. An idea which gets rid of a great deal of stupidity in one stroke is to define a "*strong*" Nash equilibrium to be one in which each player (voter), by changing his strategy actually suffers a *worse* (in expectation) result – as opposed to "weak" equilibria where the result merely "does not get better."

Now let us examine 3-candidate approval-voting elections in the context of Nash equilibria. In 3-candidate AV elections, it plainly is strategically best to vote 1 (approved) for your favorite, 0 (disapproved) for your most-hated, and then voter k ($1 \leq k \leq V$ in a V-voter election) will approve of the remaining candidate with some probability p_k.

honest preference	p
$A \gg B > C$	p
$C \gg A > B$	q
$B > C \gg A$	r

result	probability
A wins	$\overline{p}\,q\,\overline{r}$
B wins	$p\,\overline{q}\,\overline{r}$
C wins	$\overline{p}\,\overline{q}\,r$
ABC tie	$\overline{p}\,\overline{q}\,\overline{r} + p\,q\,r$
AB tie	$p\,q\,\overline{r}$
AC tie	$\overline{p}\,q\,r$
BC tie	$p\,\overline{q}\,r$

Figure 5.1. 3-vote 3-candidate cyclic approval voting election scenario. If the $A > B > C$ voter approves B with probability p while the other voters approve their middle candidates with probabilities q and r, then the election outcome probabilities are as shown in the second table. (We abbreviate $\overline{p} = 1 - p$, $\overline{q} = 1 - q$, etc.) ▲

If the "$\gg$" are as shown, then this is again a "cyclic 2-player game" (since in all cases the middle voter then always desires to make $q = 0$) and the best strategy for the first and last voters again is coin-flipping: $p = r = 1/2$. This causes

B to win, C to win, ABC tie, and BC tie,

each with probability 1/4. There again is a cyclic chain of best moves and countermoves is

$$r = 1 \implies p = 1 \implies r = 0 \implies p = 0 \implies r = 1 \implies \ldots. \tag{9}$$

Note that if we *average* over each player's coin-flip, their optimal votes for (A, B, C) are then $(1, 1/2, 0)$, $(0, 0, 1)$, and $(0, 1, 1/2)$ respectively, and so B and C each get the equally-greatest average vote-totals. These equalities between B and C with strategic voting *contradict* the obvious assessment that "morally" (or as would have happened with honest approval-voting) C should win.

On the other hand, if the *second* preference relation is $\gg$ and the first is $>$ in each vote (as was only shown in the table for the last vote, but now we consider doing that to all three) then the unique Nash equilibrium involves fully deterministic votes: all three players choose $p = q = r = 1$ and a 3-way ABC tie results. Note that the coefficient of p in the expression for the first voter's expected election utility, is proportional to $(1 - 1/3)\overline{q}\,\overline{r} + qr/3 + q\overline{r}/2 + \overline{q}r/2 = (4 - r - q)/6$ which is *positive* throughout the square $0 \leq q, r \leq 1$ so that the first voter will always desire to maximize $p = 1$ (and similarly for the other voters and q and r by symmetry).

From these two examples we see that, when we consider 3-candidate generic election situations with unique Nash equilibria with approval voting, the strategically best votes for

[3] A similar "stupid" Nash equilibrium (but the stupidity there is less obvious due to additional complexity) is given page 140 of [4] and, although Brams & Fishburn seem to think it implies something profound, I deny that.

(A, B, C) for voters with honest preferences $A > B > C$ are (averaging over randomness) sometimes of form $(1, 1/2, 0)$ and other times of the form $(1, 0, 0)$ or $(1, 1, 0)$. This demonstrates that inspiring *semi*honesty is the best we can hope for from a voting system in 3-candidate situations, even when considering strong Nash equilibria for randomized voting strategies and averaging votes over the randomness.

With our last breath of oxygen, we can demonstrate that (even when considering strong Nash equilibria for randomized voting strategies and averaging votes over the randomness), some strategic approval-voters will generate dishonest and not even semi-honest votes in 6-candidate election scenarios.

To do so, we "clone" the candidates A, B, and C in the example of figure 5.1 to get candidates A_1, A_2, B_1, B_2, C_1, C_2, where each of the voters regards the two clones as equal. However, we now add an additional fourth voter whose utility values for the candidates are

$$U(A_1) = 1608,\ \ U(B_1) = 1600,\ \ U(B_2) = 808, \quad (10)$$
$$U(C_1) = 800,\ \ U(C_2) = 8,\ \ U(A_2) = 0.$$

We claim this new voter will not affect the best strategy for the old three voters, and that his best strategy is the deterministic and *fully-dishonest* approval-vote

$$A_1 = 1,\ \ B_1 = 1,\ \ B_2 = 0,\ \ C_1 = 1,\ \ C_2 = 0,\ \ A_2 = 0 \quad (11)$$

which dishonestly gives B_2 a smaller vote than C_1.

Because this vote gives one 1 and one 0 to each of the three (A, B, and C) camps, it in no way alters the situation from the point of view of the three original voters. Indeed note the new voter's preferences are "orthogonal" to the old three voters' preferences (i.e. he cares about distinctions irrelevant to them, and vice versa).

The new voter realizes that (in his absence) with probability $1/4$ there is a B_1B_2 tie for winner (and then his approval of B_1 but not B_2 breaks the tie for expected utility gain 198), with probability $1/4$ there is a C_1C_2 tie for winner (and then his approval of C_1 but not C_2 breaks the tie for expected utility gain 198), with probability $1/4$ there is a 4-way $B_1B_2C_1C_2$ tie for winner (and then his approvals reduce that to a B_1C_1 tie, gaining expected utility 99), and finally with probability $1/4$ there is a 6-way tie for winner (and then his approvals reduce that to a $A_1B_1C_1$ tie, gaining expected utility 133). The net utility gain by casting this dishonest vote is thus 628.

On the other hand, by casting any semi-honest approval vote, the new voter would gain less. For example, if only A_1 was approved, the gain would be 133. If A_1B_1 then the gain would be $198 + 199 + 132 = 529$. If $A_1B_1B_2$ then the gain would be 331. So the new voter is best off (in the absence of any strategy-change from the original three voters) with this dishonest vote. However, if this is the new vote, then the original three voters reason that *they* are best off staying with their votes. Hence:

Theorem 3 (Nash model dishonesty in approval voting). *We have constructed a 4-voter 6-candidate approval-voting election in which there is a strong Nash equilibrium involving a deterministic fully-dishonest (i.e. not even semi-honest) fourth vote.*

It remains an **open question** whether approval (or range) voting always inspires semi-honesty in the mean in the "Nash model" for 4- and 5-candidate approval-voting elections.

Finally, we note that although we have used Nash equilibria and Von Neumann game theory [20], all of our main examples have been constructed so that they still work in Brams' alternative "theory of moves" picture [5] instead.

6 Conclusion

The Gibbard-Satterthwaite theorem famously settled the question of which voting systems (and when) had the property that strategically-best voting was the same as honest voting. The answer was very negative.

However, the GS theorem did not address either "semi-honest" voting or "incomplete information" scenarios. Those questions are now almost completely settled. It also did not address the third "Nash model" (introduced here) of strategic voters. It remains an open question whether reasonable voting systems exist that inspire semi-honesty in their voters in the Nash model – but I conjecture the answer is NO if there are at least 6 candidates.

All these investigations are compatible, as far as they go, with the vague claim that "range voting is the best single winner voting system." It is an open question whether range voting is *uniquely* best in this paper's senses.

7 Summary Table

With complete information: strategic range votes are wlog approval-style

candidates	rank-order (= permitted or forbidden)	range & approval
1-2	honest=strategic	honest=strategic
3 or more	honesty can be nonstrategic	some semi-honest vote is strategic

With *in*complete information: strategic range votes are not necessarily approval-style

candidates	rank-order (= permitted or forbidden)	range & approval
1-2	honest=strategic	honest=strategic
3	honesty can be nonstrategic	some semi-honest vote is strategic
4 or more	honesty can be nonstrategic	semi-honesty can be nonstrategic

Nash model (strong Nash equilibrium):

candidates	rank-order (= permitted or forbidden)	range & approval
1-2	honest=strategic	honest=strategic
3-5	?	?
6 or more	?	having semi-honest-mean can be strategically bad

References

[1] Jean-Pierre Benoît: Strategic manipulation in voting games when lotteries and ties are permitted, J. Economic Theory 102,2 (2002) 421-436.

[2] Jean-Pierre Benoît: The Gibbard-Satterthwaite theorem: a simple proof, Economics Letters 69,3 (2002) 319-322.

[3] Steven J. Brams and Peter C. Fishburn: Some logical defects of the single transferable vote, Chapter 14, pp. 147-151, in Choosing an Electoral System: Issues and Alternatives (Arend Lijphart and Bernard Grofman, eds.) Praeger, New York 1984.

[4] S.J.Brams & P.C.Fishburn: Approval Voting, American Political Science Review 72,3 (September 1978) 831-847; also book with same title, Birkhäuser Boston 1983.

[5] Steven J. Brams: Theory of Moves, Cambridge Univ. Press 1994.

[6] J.R.Chamberlin: An investigation into the effective manipulability of four voting systems, Behavioral Science 30 (1985) 195-203.

[7] John R. Chamberlin, Jerry L. Cohen, Clyde H. Coombs: Social Choice Observed: Five Presidential Elections of the American Psychological Association, J. of Politics 46,2 (1984) 479-502.

[8] G.B. Dantzig: Linear Programming and Extensions, Princeton University Press, Princeton, NJ, 1963.

[9] Alan Gibbard: Manipulation of Schemes That Mix Voting with Chance, Econometrica 41 (1977) 587-600.

[10] Allan Gibbard: Strategic Behavior, and Best Outcomes, pp.153-168 in H.W. Gottinger and W. Leinfellner (eds.): Decision Theory and Social Ethics: Issues in Social Choice (Dordrecht, Holland: D. Reidel, 1978).

[11] Martin van Hees & Keith Dowding: In Praise of Manipulation, manuscript available from Dowding's web page at the London School of Economics `http://personal.lse.ac.uk/DOWDING/Papers.htm`.

[12] John Nash: Non-cooperative games, Annals of Mathematics 54 (1951) 286-295.

[13] Benjamin Reilly: Social Choice in the South Seas: Electoral Innovation and the Borda Count in the Pacific Island Countries, Int'l Political Science Review 23,4 (2002) 355-372.

[14] Mark A. Satterthwaite: Strategyproofness and Arrow's conditions, Existence and correspondences for voting procedures and social welfare functions, J.Economic Theory 10 (1975) 187-216.

[15] Markus Schulze: A New Monotonic and Clone-Independent Single-Winner Election Method, Voting Matters 17 (Oct. 2003) 9-19.

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

[17] Warren D. Smith: The voting impossibilities of Arrow and of Gibbard & Satterthwaite, #79 at `http://math.temple.edu/~wds/homepage/works.html`.

[18] A great deal of information about voting systems and range voting especially may be found at `http://RangeVoting.org` and in particular the "DH3 pathology" is discussed on the subpage `http://RangeVoting.org/DH3.html`.

[19] Alan D. Taylor: Social choice and the mathematics of manipulation, Cambridge Univ. Press, New York, 2005.

[20] John von Neumann & Oskar Morgenstern: Theory of games and economic behavior, Princeton Univ. Press 1953 (3rd ed).