The voting impossibilities of Arrow, Gibbard & Satterthwaite, and Young

Warren D. Smith*
WDSmith@fastmail.fm

December 11, 2005

Abstract — **Arrow and Gibbard's theorems are usually considered the two most important impossibility theorems on voting systems. Also important, but less famous, are some theorems by Young et al. We state and prove them, survey their extensions, criticize them, and discuss their limitations. We also present some new and original extensions of Arrow's theorem to make it cover voting systems whose goal is to output a single winner, and also give some other new and old impossibility theorems due to Schulze and the author.**

1 Introduction

Suppose there are some finite number V of voters who must, with the aid of some voting system, choose one winner from some finite number N of possible alternatives or candidates.

K.L.Arrow proved his impossibility theorem in 1952. It won him the 1972 Nobel prize in Economics. It shows, very roughly, that if $N \geq 3$ then "no good voting system can exist."[1] Although Arrow originally required an entire book to prove it, nowadays (thanks to John Geanakoplos and others) its proof can be done on a single page.

A different kind of (and more profound) impossibility theorem was shown by A.Gibbard in 1973 and independently by M.A.Satterthwaite in 1975. This theorem concerns the idea that voters can either try to maximize their impact on the election, or they can try to honestly state their opinions in their votes. In a "strategyproof" voting system, these two acts would be the same thing. Unfortunately, Gibbard's theorem shows that if $N \geq 3$ then no useful strategyproof voting system can exist. Although this theorem originally seemed difficult and profound, nowadays it too has a similar short proof, due to Jean-Pierre Benoît.

Finally, J.H.Smith and H.P.Young showed some still more profound (but for some reason far less well known) Arrow-like theorems in 1973-1978, with another such theorem proven by H.Moulin in 1988. These apply (in different versions) *both* to voting systems whose goal is to elect a single winner, and to those whose goal is to rank-order all the candidates. They provide a complete understanding of "symmetric separable" systems.

These are the most famous impossibility theorems about voting systems, and Arrow's, Gibbard's, and Moulin's theorems nowadays may each may be proven in a single page entirely with pre-college-level mathematics.

Our goals here are

1. To state and prove them.
2. To survey various extensions of these theorems.
3. To criticize them and discuss their limitations.

Thus this chapter is almost entirely expository and unoriginal. There is, however, one exception: we shall provide some new extensions of Arrow's theorem which seem to increase its usefulness significantly.

2 Arrow's theorem and I.I.A.

An *Arrovian preference ordering* is a ranking of the N alternatives from top to bottom with ties allowed. It is *transitive* if $a \geq b$ and $b \geq c$ implies $a \geq c$. An *Arrovian voting system* (AVS) is a function mapping the V voters' Arrovian preference orderings to a single (output) societal Arrovian preference ordering. An AVS has the *unanimity* property if society ranks $a > b$ whenever every voter does. It has the *I.I.A.* (for "Independence of Irrelevant Alternatives") property if the societal ranking of a relative to b (i.e. above, below, or neither) depends only on a and b's relative rankings by each voter. Finally, an AVS is a *ab-dictatorship* by voter X, if, whatever X's relative ranking of a versus b is, society's is the same; and X is simply a *dictator* if this is true of every a, b.

Theorem 1 (Arrow). *Let $N \geq 3$. Any AVS that respects transitivity, I.I.A., and unanimity is a dictatorship.*

The most obvious deficiency of this theorem is that most people's notion (as opposed to Arrow's notion) of a "voting system" produces as output *not* an ordering of the candidates, but merely the identity of the election winner. Call such a system an "MVS" (the M standing for "most people's"). Is there an Arrow-like theorem for MVSs? The answer is yes, and it is trivial to produce via the following trick.

Suppose A wins. If A is removed from the election (and from all voter preference orderings) then suppose B wins. If both A and B are removed, suppose C wins. And so on. The point is that this defines an ordering $A > B > C \cdots$ among the candidates. Now: Arrow's theorem applies to the new voting system that outputs this ordering, consequently showing:

*Non-electronic mail to: 21 Shore Oaks Drive, Stony Brook NY 11790.
[1]Here "good" means, more precisely, "satisfying a certain list apparently-obviously desirable/necessary properties."

Theorem 2 (New version of Arrow for single winner voting systems). *Any MVS with $N \geq 3$ candidates which*

- *inputs voter preference orderings*
- *outputs single winner*
- *obeys unanimity and I.I.A. concerning the implied ordering got by successive winner-removal*

must be a dictatorship.

Perhaps because of the trivial way it was produced, this theorem leaves us unsatisfied. Somehow, requiring I.I.A. of the somewhat unnatural ordering induced by winner-removal seems possibly unnatural. Therefore – appearing for the first time here – we now produce apparently *the* natural generalization of Arrow to single-winner voting systems.

We first must make a few observations and state a few definitions. If a voting system inputs preference orderings and outputs a single winner, then that same voting system fed the *reversed* preference orderings, would output a single *loser*. So without loss of generality we shall regard our voting system as outputting both (1) a single "winner" candidate and (2) a single "loser" candidate. We shall also regard there as being (although it is not necessary for the voting system to output or know it) a *partial* ordering $\prec$ among the candidates. This partial ordering and the winner and loser shall be required to obey the following properties:

Noncontradiction: It is impossible for both $a \prec b$ and $b \prec a$ to hold. (However, it can be entirely possible that *neither* holds; that is why $\prec$ is a *partial* as opposed to a total order; indeed we are imagining there to be as few $\prec$-related candidate-pairs as possible without our voting system totally losing meaning.)

Transitivity: If $a \prec b$ and $b \prec c$ then $a \prec c$; if $a \not\prec b$ and $b \not\prec c$ then $a \not\prec c$.

Winner-domination: If a is the winner then $x \prec a$ for all $x \neq a$. If b is the loser then $b \prec x$ for all $x \neq b$.

Pair-comparison unanimity: If every vote says that $a < b$, then $a \prec b$; if every vote says that $b \geq a$, then $b \not\prec a$.

I.I.A.: The question of whether $a \prec b$ or $b \prec a$ or neither, depends only on the pairwise relative comparisons of a versus b in the votes, and not on comparisons among other candidate-pairs.

Some immediate consequences of these properties include:

Unanimous top-ranking: If every vote top-ranks a, then a must be the winner.

Unanimous bottom-ranking: If every vote bottom-ranks b, then b must be the loser.

Winner≠loser: If $N \geq 2$ then the winner and loser must differ.

Undominated⟹Winner: A candidate x such that no y exists with $x \prec y$, must be the winner. (And if no y exists with $x \succ y$, then x must be the loser.)

A voter X now is an *ab-vetoer* if whatever relative ranking of a versus b he has, society's does not disagree, i.e. if X says $a < b$ then society says $a \not\succ b$; and X is simply a *vetoer* if this is true of every a, b. Note that (due to Undominated⟹Winner) a vetoer can singlehandedly determine both the winner and loser, i.e. is a *win/loss-dictator.*

Theorem 3 (Our new extension of Arrow). *Let $N \geq 3$. Any voting system which inputs Arrovian preference orderings among the N candidates (one per voter), and which outputs a single winner and a single loser candidate, and for which there implicitly exists a partial ordering $\prec$ such that the properties listed above hold, must contain: a vetoer, and hence a win/loss-dictator.*

Criticism: *Initially* Arrow's theorem 1 and our theorem 3 seem to be a severe problem, in a sense proving that no good voting system can exist. That is because "obviously" any good voting system must obey I.I.A., transitivity, and unanimity, and must not be a dictatorship, hence by Arrow's theorem cannot exist.

But in fact, Arrow's theorem is not nearly as important as it at first seems. That is because

1. In the real world, there is no reason voters must input preference orderings; they could input utility vectors.
2. In that case I.I.A. is actually plainly *not* a desirable property of a voting system, from the point of view of society-wide utility. Example: suppose the utilities of alternatives A, B are 10, 0 respectively from the viewpoint of 51% of the voters, and 0, 99 respectively from the viewpoint of the remaining 49%. For the good of society, B should win. But if the "10" were replaced by "100" then, for the good of society, A should win. The relative rankings of A and B by the voters have not changed, *but* the best societal choice has! So it is clearly wrong to demand I.I.A. as Arrow does, in the wider scenario where more general input is allowed.

These two criticisms are unanswerable in the sense that "honest utility voting" is a single-winner-output, utility-vector-input voting system which obeys all of Arrow's axioms (as well as monotonicity), except for I.I.A. which as we've seen *shouldn't* then be obeyed.

Other Arrow-like theorems: Kelly [7] surveyed work before 1978 on other Arrow-type theorems. His theorem 4-12 (due to J.H.Blau) replaces I.I.A. with the equivalent, but seemingly weaker, assumption of the independence of a *single* irrelevant alternative. An Arrow-like theorem permitting voters to input *partial* orderings (under certain assumptions) is Kelly's theorem 4-10. Fishburn [3] showed the assumption that V is finite is essential in Arrow's theorem. A different formulation of Arrow's theorem [9] allowing *non-deterministic* voting systems replaces Arrow's I.I.A. demand by a stochastic version, continues to require unanimity, and demands further a "regularity" demand that the probability of X being chosen cannot increase if more alternatives are made available, and concludes that if there are ≥ 4 alternatives, the voting system must be a "weighted random dictatorship."

Amartya Sen once claimed in print [11] that Gibbard had found a different extension of Arrow's theorem to make it cover voting systems that output *winners* rather than orderings. Sen claimed a proof would be found in his (Sen's) book [12] on pp.76-77, but I disagree that there is a proof there. Gibbard then told me he believed that Sen had in mind his (Gibbard's) theorem from [6] (despite the fact that it was published 8 years after Sen's book, and despite the fact that if this is really what Sen had in mind, then Sen misstated the theorem). Unfortunately, my examination of [6] has merely left me feeling that Gibbard's tricky notion there of what a "voting system" and/or a "winner" is, is different from everybody else's, and thus really his theorem says little or nothing about voting systems that anybody cares about. So we shall not survey [6] here, aside from warning the reader to be careful.

3 Gibbard's theorem and the quality of being strategyproof

A *Gibbardian preference ordering* is a ranking of the N alternatives from top to bottom with *no* ties allowed, i.e. a permutation of $\{1, 2, \ldots, N\}$. A *Gibbardian voting system* (GVS) is a function mapping the V voters' reported Gibbardian preference orderings to the identity of a single "winning" alternative. A GVS has the *unanimity* property if society always selects an alternative that is top-ranked by everyone. It is *strategyproof* if a best strategy for each voter always is to report his preferences truthfully.

Theorem 4 (Gibbard). *Let $N \geq 3$. The only strategyproof GVS that respects unanimity is a dictatorship.*

This is a severe limitation on non-dictatorial voting systems – voters will necessarily, in ($N \geq 3$)-candidate elections, sometimes be motivated to vote dishonestly, i.e. to misrepresent their true preference orderings. Furthermore, the proof we shall provide in §6 still works even if the voting system asks voters for utility N-vectors instead of N-permutations as votes – they will still sometimes be motivated to provide vectors with misordered entries *if* it is forbidden for those vectors to contain two *equal* entries.

There then remains a narrow escape hatch, which approval and range voting[2] exploit to partially evade Gibbard's theorem. Specifically, these two voting systems *allow* equalities (indeed, in approval voting in $\geq$ 3-candidate elections, each voter is *forced* to express a preference equality). They both *are* strategyproof in the weaker sense that each voter finds it strategically best, in a 3-candidate election, to provide a vote-vector which either obeys his true preference $>$-inequalities, or is a limiting case of them, i.e. is arbitrarily near to a vector satisfying them.

Indeed, *approval and range voting are strategyproof* in this weaker sense, for *any* number of candidates. That is, if an approval or range voter knows the exact vote totals from all the other voters, then he may always choose an approval vote vector which is both optimally strategic and is a *limit* of vectors which obey his true preference ordering $<$-relations.

However, *neither* range nor approval voting are strategyproof – even in this weaker limit-sense – for 4-candidate elections with a very large number of voters if a voter does *not* know the exact vote totals from all the other voters, but instead only knows the vote totals (and perhaps also the covariance matrix) of a random subset of (say) 0.1% of the other voters.

Example #1: Let the 4 candidates be A, B, C, D. Suppose their election utilities, from your point of view, are: $U_A = 0$, $U_B = 20$, $U_C = 70$, $U_D = 25$. Assume the prior likelihoods of election of the candidates (based on vote totals from a pre-election poll – we assume the covariance matrix is proportional to the 4×4 identity matrix) are $L_A \gg L_B \gg L_C \gg L_D$. (To be completely concrete, assume our model of all the other voters is that they approve candidates A, B, C, D with independent probabilities $0.7, 0.6, 0.5, 0.1$ respectively.) Then your best strategic approval vote is $(0, 1, 1, 0)$ for (A, B, C, D) respectively, which is dishonest about your preference for the candidate-pair B, D. ▲

Example #2: There are two liberal candidates L_1, L_2 and two conservatives C_1, C_2. You are pretty sure that L_1, L_2 will get a near-equal number of the other people's votes, and ditto for C_1, C_2, but don't know whether the liberals or conservatives will be ahead. This kind of situation can be modeled with a highly ellipsoidal Gaussian probability distribution, e.g. arising from a covariance matrix of the form

$$\begin{pmatrix} 1 & 1-O(\epsilon) & \pm O(\epsilon) & \pm O(\epsilon) \\ 1-O(\epsilon) & 1 & \pm O(\epsilon) & \pm O(\epsilon) \\ \pm O(\epsilon) & \pm O(\epsilon) & 1 & 1-O(\epsilon) \\ \pm O(\epsilon) & \pm O(\epsilon) & 1-O(\epsilon) & 1 \end{pmatrix} \tag{1}$$

when $\epsilon \to 0^+$ (whereas the previous example worked even with a spherical, i.e. correlation-free, Gaussian). Then your best strategy is to vote in the style $(1, 0, 1, 0)$ for L_1, L_2, C_1, C_2 respectively, even if you prefer both L's over both C's (or both C's over both L's). This example also can be generalized to $2N$ candidates falling into N ultra-correlated pairs. ▲

These examples indicate that Gibbard was employing a wrongheaded notion of "strategyproof." In Gibbard's notion, voters are assumed to know all other votes. A more realistic world-model would involve voters with only *partial* information about the other votes, for example knowing only a small random subset of them or only a probabilistic model of them. Of course, if strategyproof voting systems are impossible even under Gibbard's exact-information scenario, then they are impossible under the more general scenario where voters might only have partial information. So in that sense, Gibbard's is the strongest possible theorem – but if we are going in the other logical direction then Gibbard's is the *weakest* possible theorem.

To repair that flaw we would like to have a theorem saying that nondictatorial strategyproof voting systems are impossible in (≥ 4)-candidate elections where voters may only have access to partial information about the other votes, even for "voting systems" permitting utility vectors as votes, and even with the weakened notion of "honesty" where vector votes which are *limits* of vectors obeying honest $<$-preference inequalities are permitted. So far, that theorem remains a conjecture, although situations like example #2 above convince me that it must be true.

Several extensions of Gibbard's theorem are known. We shall state but not prove them.

Let a *GVS with chance* be a map from the voters' Gibbardian preference orders *and* some random bits to the identity of a single winning alternative.

[2] We remind the reader that in *range voting* in an N-candidate election each vote is a real N-vector each of whose entries x obeys $0 \leq x \leq 1$; whereas in *approval voting* $x = 0$ or $x = 1$. The vote vectors are summed and the greatest entry in the sum-vector corresponds to the winner. Ties are broken randomly.

Theorem 5 (Gibbard's probabilistic extension of Gibbard [5]). *The only strategyproof GVS's with chance are*

Random dictator: *Choose a voter at random and proclaim his top-rank choice the winner.*

Random pair vote: *Choose a pair (a, b) of distinct alternatives at random, and perform an ordinary 2-alternative election using the voters' preference relations restricted to a, b only.*

Combination: *Some probabilistic combination of the above two ideas.*

Unfortunately, as Gibbard remarks, these GVS's, despite being strategyproof, seem unacceptable in practice since they "leave too much to chance."

Let a *GVS with ties* be like a GVS, except the voters now are allowed to express indifference in their preference orderings, and the outcome can be an arbitrary *subset* of "tied" co-winners. The voters are each assumed to have private utilities for each possible winner-subset that are compatible with each of them assuming that some (unspecified) lottery mechanism will be used to break the ties. (Voters are allowed to have differing theories about what those lottery mechanisms will be.) It obeys the *near-unanimity* property if, whenever all voters (except possibly one) strictly top-rank some alternative, it alone wins.

Theorem 6 (Benoît's extension of Gibbard [1]). *Let $N \geq 3$ and $V \geq 3$. There is no strategyproof GVS with ties that respects near-unanimity.*

Observe that range voting does *not* obey the near-unanimity property (although it does obey unanimity), because 999 voters ranking A above B by 0.001 are outweighed by one voter ranking B above A by 1. This in fact makes it plain that near-unanimity is *not* always a desirable property (from the standpoint of maximizing society-wide utility) for a voting system to have. Thus Benoît's extension does not apply to range voting and also, like Arrow's theorem, is somewhat wrongheaded in that one of its axioms that "obviously" should be obeyed by "good" voting systems, in reality is unjustified.

P.K.Pattanaik invented a weaker notion of "strategyproofness" involving the idea that "threats" of manipulation could be met by "counterthreats." This weaker kind of strategyproofness is also known to be impossible, see p.71 of [7].

Finally, note that ordinary majority voting in an ($N = 2$)-alternative election is strategyproof, obeys I.I.A., near-unanimity, and unanimity, and is not a dictatorship. In other words, as far as Arrow, Gibbard, and Benoît are concerned, it is perfectly wonderful. Thus these impossibility theorems all only apply when $N \geq 3$.

4 Impossibility theorems by Smith, Young, Levenglick, Moulin, and Perez

J.H.Smith [14], Young [17], and Levenglick [18] proved some wonderful theorems about voting systems which input preference-orderings of the candidates. Smith and Young's discoveries overlapped and were independent. Although Arrow's focus was on voting systems that output a full ordering of all the candidates, while Gibbard's focus was on simply selecting a single winner, the Smith/Young/Levenglick theorems are available in different versions addressing either focus.

A voting system is *separable*[3] if, whenever the results of elections on two disjoint voter subsets are the same, then the result of the election on the *combined* vote set must also be the same. Warning: this yields two very different definitions of "separable" depending on whether we are considering "election results" which are an ordering of all candidates, versus just the name of a winner (or the names of several "winners").

A voting system is *symmetric* if it produces the same results, regardless of how the voters are permuted and regardless of how the candidates are permuted (except that in the latter case, the results are permuted).

A voting system is *weighted positional*[4] if it works as follows: each time a voter ranks a candidate kth we award that candidate S_k points, where S_1, S_2, S_3,..., S_N are N pre-fixed real constants. The candidate with the most points wins, or in the ordering-as-output version, we rank the candidates by their number of points. Smith and Young also considered breaking ties by means of a second weighted positional system (with different real weights), and then ties in *this* system could be broken by votes using a third weighted positional system, and so on. Young called these "*composed* weighted positional systems."[5] Alternatively, and equivalently, we could use just *one* set of weights S_1, S_2,..., S_N but these would not be real numbers, but rather a larger field containing both reals and infinitesimal quantities. Yet another equivalent view would be to regard the weights as *tuples* of reals, ordered lexicographically.

Theorem 7 (Smith & Young's separability theorem [14][17]). *A symmetric and separable voting system whose output is the name of a winner (or the names of several unordered winners) must be a composed weighted positional system (or equivalent to one). If the voting system also satisfies a further condition, which Young calls "continuity,[6] " then it must be a (plain) weighted positional system.*

A "Condorcet-Winner" is a candidate who would beat each other candidate in a head-to-head 2-candidate election, using the same preference orderings as votes (but with all the other $N - 2$ candidates erased from those orderings).[7]

[3] J.H.Smith's name. Young instead called this property "consistency," and others have called instances where it is disobeyed, the "multiple districts paradox."

[4] J.H.Smith's name; Young's name was "point scoring system."

[5] J.H.Smith called them "generalized point scoring systems."

[6] But Smith calls it "Archimedean." Roughly, the Archimedean property states that a sufficiently large set of voters with a given distribution of preferences can impose its will on any fixed-size set of voters.

[7] Condorcet invented this notion and realized that Condorcet-Winners need not exist. Condorcet also realized that Borda's weighted-positional system (with weights $S_K = N - K - 1$) can select a winner different from the Condorcet-Winner: with 6 votes $A > B > C$, 4 votes $B > C > A$, and 1 vote $C > A > B$ the Borda winner is B but the Condorcet-Winner is A. Theorem 8 shows that similar counterexamples exist for *every* composed weighted positional voting system. If a Condorcet-Winner does exist, it is necessarily unique.

Theorem 8 (Hansson's no-Condorcet theorem [18]). *If the number of candidates $N \geq 3$, there is no symmetric and separable voting system whose output is the name of a winner (or the names of several unordered winners) and which always uniquely elects a Condorcet-Winner whenever one exists.*[8]

In order to generalize theorem 8 to handle voting systems which output full orderings rather than just the name of a winner, Young and Levenglick introduced the *quasi-Condorcet property*, which we now define.

Suppose i and j and the votes are such that

1. Every voter ranks candidate i either immediately above or immediately below candidate j, and
2. In a head-to-head election between candidates i and k (got by erasing all the $N-2$ other candidates from all voter preference orderings) the two would come out exactly tied, for *every* k.

A quasi-Condorcet voting system is one whose output ordering, in such an "ultra-tied" scenario, would always rank i and j equal (or, if equalities are disallowed via a tie-breaking scheme which always outputs exactly *one* ordering) which would be equally likely to rank i immediately above or immediately below j.

Theorem 9 (Young-Levenglick [18]). *There is exactly one symmetric and separable voting system whose output is an ordering of all the candidates, and which obeys the quasi-Condorcet property: Kemeny's voting method.*[9]

J.H.Smith [14] considered single-winner voting systems based on "successive elimination." These systems proceed in rounds: each round, the candidate (or several candidates) with the fewest points under a weighted positional voting system is eliminated. Such systems are not weighted positional systems, and hence by theorem 7 are not separable. But, more severely, Smith showed they always are "non-monotonic" even in the 3-candidate case.

Theorem 10 (Smith's elimination$\Longrightarrow$nonmonotonicity theorem [14]). *Any elimination-based single-winner voting system where the round eliminations are based on a nontrivial weighted positional scoring system, is non-monotonic. That is, changing some of the votes in a candidate's favor can remove him from first place.*

Moulin [8] and Perez [10] extended Hansson's no-Condorcet theorem to show that not only must voting systems that always elect Condorcet-Winners (when they exist) disobey separability, they must in fact exhibit a particularly nasty kind of pathology called the "strong no-show paradox":[10]

Theorem 11 (Condorcet$\Longrightarrow$No-Show-Paradox [8][10]). *Any (≥ 4)-candidate voting system inputting preference orderings as votes, and outputting a set of "winners," and which uniquely elects a Condorcet-Winner whenever one exists, must exhibit "no-show paradox" situations in which*

1. *A is the unique winner, and*
2. *by adding a some identical new votes*[11] *strictly preferring A over B, we get a scenario in which A is no longer the unique winner, but B is a winner.*

(See our theorem 20 for a proof of a slightly altered statement.)

Compared to Arrow's impossibility theorem, these theorems seem far superior. Many of them apply to either ordering- or winner-as-output voting systems. They start from axioms (separability, symmetry) which seem more clearly desirable. The axiom whose desirability I *do* dispute is the requirement that input votes be preference orderings instead of real-vectors. These theorems perhaps indeed may be thought of as *damning indictments* of every preference-orderings-as-votes system. While the axiom that Condorcet-Winners must be elected perhaps is desirable if only preference orderings are available, it clearly is *un*desirable if real vectors are allowed as votes (we already gave a counterexample in §2).

5 Proof of both Arrow's theorem and our new extension

We prove theorem 3 (and as an immediate consequence, Arrow's original theorem 1) following Geanakoplos [4]. Both this proof and the proof we shall give in §6 of Gibbard's theorem are similar in that they begin by showing the existence of a "pivotal" voter P.

The proof: Arbitrarily focus on some outcome b. Let a "profile" mean all the votes.

Lemma 12 (TopBot$\Longrightarrow$TopBot). *In any profile in which each voter either puts b at the top, or at the bottom, of his preference ordering, society must also (even if some voters put b at the top and others at the bottom), i.e. b must either be the winner or loser.*

Proof of lemma 12: Suppose for a contradiction that there exist a and c (distinct from each other and from b) with $a \not\prec b$ and $b \not\prec c$ in the societal ordering. (By winner-domination, this must happen if b is neither the winner nor the loser.)

By I.I.A., $a \not\prec b$ and $b \not\prec c$ will still hold even after each voter moves c above a because (since b is topmost or bottommost in every vote) this would not disturb any ab or cb votes. Hence by transitivity the social preference would have $a \not\prec c$. But by pair-unanimity it would have $a \prec c$, a contradiction. Q.E.D.

Lemma 13 (Pivotal voter). *There must exist a "pivotal" voter P who, solely by changing his vote, can (in at least one profile) move b from the bottom to the top of the social ordering, i.e. from loser to winner status.*

[8] This is theorem 2 in [18] and in its present simplifed form is credited to Bengt Hansson.

[9] Kemeny's voting method is the following. Let the L_1*-distance* between two compatibly-sized matrices A, B be $\text{dist}(A,B) \stackrel{\text{def}}{=} \sum_{i,j} |A_{ij} - B_{ij}|$. With each preference ordering among the N candidates associate the $N \times N$ anti-symmetric matrix with ij entry $+1$ if i is preferred to j, -1 if j is preferred to i, and 0 if $i = j$. The *election matrix* is the average of all the matrices arising from the votes. Kemeny's voting method outputs a preference ordering whose matrix has minimum L_1-distance to the election matrix. (Warning: it is known that this minimization task is NP-hard; essentially it is a traveling salesman problem. J.Rothe and H.Spakowski have indeed recently claimed to have shown that it is $P_{\parallel}^{NP}$-complete to determine Kemeny winners.)

[10] Note: separability is actually logically unrelated to the no-show paradox, despite the fact that they seem similar.

[11] In Perez's strengthened versions, under certain conditions, only a *single* new vote is needed.

Proof of lemma 13: We shall show that P can change b from a state in which no x exists with $x \prec b$, to a state in which no x exists with $b \prec x$; by undominated$\Longrightarrow$winner this is the same as changing b from loser to winner.

Let each voter put b bottommost in his vote. By unanimity, b must be the loser and $b \prec x$ for all x. Now go though the voters in order from voter 1 to voter V having each successively move b from the bottom to the top of their vote-ranking (while leaving all other relative rankings unchanged). At some point b must change social rank (since at the end of this process, by unanimity, it obeys $b \succ x$ for all x, i.e. has top rank and is the winner); suppose this happens during the change-of-vote of voter P. By lemma 1 this one change must shift b all the way from loser to winner. Q.E.D.

Define "profile ❶" to be the profile just before P changes his vote, and "profile ❷" just after.

Lemma 14 (Pivotal$\Longrightarrow$ *ac*-vetoer). *The pivotal voter P from lemma 13 and its proof must be an ac-vetoer for any outcome pair ac not involving b.*

Proof of lemma 14: Construct profile ❸ from profile ❷ by letting P move a above b in his vote (so that $a >_P b >_P c$) and then by letting all voters besides P arbitrarily alter their relative rankings of a and c while leaving b in its extreme position. By I.I.A. the societal ordering corresponding to profile ❸ must have $a \not\prec b$ (since all ab votes are the same as in profile ❶) and $b \not\prec c$ (since all bc votes are the same as in profile ❷). So by transitivity society must put $a \not\prec c$ in profile ❸. But now by I.I.A. the social preference must have $a \not\prec c$ whenever $a >_P c$. Q.E.D.

In particular, note that P singlehandedly can control the identity of the winner, or the loser, or both, since at least one of these cannot be b.

Finale: We will now argue P must also be a vetoer over every pair ab, and hence a full vetoer. To see this, consider a third distinct outcome c which we put at the bottom in every vote in the construction in the proof of lemma 13. Now by the argument in the proof of lemma 14, there must be a pivotal voter P' who is an $\alpha\beta$-vetoer for any pair $\alpha\beta$ not involving c – such as ab. Note that P, acting alone, can reverse some societal $a \prec / \succ b$ ranking (namely at profiles I and II) since by winner-domination this societal ranking exists in both cases so that vetoing it in fact reverses it, i.e. is a genuine accomplishment. This accomplishment would be impossible if anybody besides P were an ab-vetoer. Since P' is an ab-vetoer, we conclude that $P = P'$ so that P is a full vetoer and hence a win/loss-dictator. Theorems 3 and 1 are now proven. Q.E.D.

A far even simpler proof still! The fact that an Arrovian Voting System obeying I.I.A. cannot exist for (≥ 3)-candidate elections is a *triviality* if we add the demand that the system reduce to majority vote in the 2-candidate case. That is because: suppose in some Condorcet-cyclic profile, candidate A wins. Then by omitting all the candidates besides A and B (where B is a candidate superior to A pairwise), B must win, i.e. those omitted candidates were not "irrelevant" to the A versus B battle, contradicting Arrow's I.I.A. assumption. Q.E.D.

6 Proof of Gibbard's theorem

We follow Benoît [2] to prove theorem 4.[12]

The proof: Suppose we are given a strategyproof Gibbardian voting system obeying unanimity. We shall show some voter is a dictator.

Lemma 15 (Old or new winner). *Suppose some profile*[13] *causes outcome a to be the winner. Modify the profile by raising some outcome x in voter i's ranking (holding all else fixed). Then: either a or x is the new winner.*

Proof of lemma 15: Suppose for a contradiction that when x rises some other winner c is chosen. Then if i prefers a over c he would not report the change, whereas if he preferred c to a he would have falsely reported the change earlier. Q.E.D.

Lemma 16 (Bottom feeders are not winners). *Consider an arbitary profile in which all voters bottom-rank b. Then b cannot be the winner.*

Proof of lemma 16: If b were a winner then by strategyproofness it would still have won if the voters one at a time raised a to the top of their votes (since otherwise some voter would have). But that would contradict unanimity. Q.E.D.

Now start with the profile in lemma 16 and starting with voter 1 and continuing through each other voter in order, have that voter raise b from the bottom to the top of their vote (leaving all else fixed). Let P be the pivotal voter whose change causes b to be elected. "Profile ❶" is the profile before P changes his vote, and "profile ❷" after.

Consider profile ❷ with winner b. Outcome b must still win if any other voter $i > P$ changes his ranking, otherwise i would misrepresent. Also b must still win if any voter $i \leq P$ changes his vote-ranking with b still ranked top (otherwise i would not honestly report his ranking).

Lemma 17. *If voters $1, 2, \ldots, P$ top-rank b, then b must win.*

Proof of lemma 17: Consider profile ❶ (with b not the winner). Outcome b must still not win if any voter $i < P$ changes his ranking, or else i would have done just that. We can make this same argument considering these voters acting one at a time. Q.E.D.

Similarly b must still not win if any voter $i \geq P$ changes his vote-ranking (but still ranking b bottom), or else i would not honestly report his ranking, and we can make this same argument considering these voters acting one at a time, thus similarly proving:

Lemma 18. *If voters $P, P+1, \ldots, V$ bottom-rank b, then b cannot win.*

With these lemmas in hand, we now are ready to show that the pivotal voter P is, in fact, a dictator.

Let profile ❸ mean any profile of the form

1	2	...	$P-1$	P	$P+1$	...	V
?	?	...	?	k	?	...	?
?	?		⋮	⋮	⋮		⋮
b	b	b	b	b	b	b	b

[12] A different short proof, instead based on induction on V, has been given by Arunava Sen [13].

[13] Remember, all profiles consist entirely of *strict* preferences in theorem 4.

First raise k to the top for all voters; then by unanimity k wins. Now raise b to the top for voters 1 through $P-1$ one at a time to get profile ❹:

1	2	…	$P-1$	P	$P+1$	…	V
b	b	…	b	k	k	…	k
k	k	…	k	?	?	…	?
⋮	⋮		⋮	⋮	⋮		⋮
?	?	?	?	b	b	b	b

By lemma 18 outcome b does not win, so by lemma 15 k does. Finally raise b to the second position for voter P. Now k still does not win or else P would not report this change, i.e. k wins in profile ❺:

1	2	…	$P-1$	P	$P+1$	…	V
b	b	…	b	k	k	…	k
k	k	…	k	b	?	…	?
⋮	⋮		⋮	⋮	⋮		⋮
?	?	?	?	?	b	b	b

Now reconsider profile ❸ and suppose g wins ($g \neq k$). Raise b to the top for voters $1, 2, \ldots, P-1$ one at a time. By lemma 18 we know b does not win so by lemma 15 the winner is still g. Now raise b in voter P's vote to the second position to get profile ❻:

1	2	…	$P-1$	P	$P+1$	…	V
b	b	…	b	k	?	…	?
?	?	…	?	b	?	…	?
⋮	⋮		⋮	⋮	⋮		⋮
?	?	?	?	?	b	b	b

Here, if b is not the winner, then by lemma 15 g must still win. From lemma 17 we know b will still win when it is raised to the top of P's vote. Hence P should now falsely report this preference since he prefers b to $g \neq k$, contradicting strategyproofness. Hence the winner in profile ❻ must be b.

Now raise k to the second position for voters $1, 2, \ldots, P-1$ and to the top for voters $P+1, \ldots, V$. Then b still wins or else the first group of voters would not truthfully report this change, whereas the second group of voters would have misreported it. But this modification of profile ❻ is the same as profile ❺ with winner k! That contradiction proves g ($g \neq k$) cannot have been the winner in profile ❸. In short,

Lemma 19 (Pivot voter selects non-b winner). *In profile ❸, whomever P top-ranks (as long as it is not b) must win.*

We have nearly shown that P is a dictator; we shall now go the rest of the way to showing that. Consider an arbitrary profile in which P ranks some outcome k ($k \neq b$) on top. First modify this profile by dropping b to the bottom for all voters. By lemma 19 the winner is k. Now restore b to its initial position for all voters (one at a time). By lemma 15, either k or b must now win. Now consider profile ❼:

1	2	…	$P-1$	P	$P+1$	…	V
b	b	…	b	b	a	…	a
⋮	⋮		⋮	⋮	⋮		⋮
c	c	…	c	c	c	…	c

where $c \neq b$ and $c \neq k$. Similarly to the argument in lemmas 16-18 have c jump to top rank in the votes one at a time until we find the pivotal voter M for outcome c. Similarly to lemma 19 M's top choice in profile ❻ must win. On the other hand from 15 we also know that b wins in profile ❼. Hence $M \leq P$. But a symmetric argument (beginning with M then finding P) shows $P \leq M$. Hence $P = M$ and voter P is pivotal with respect to c as well as b. So not only do we know either k or b wins in our initial arbitrary profile, we also know it must be k or c. Since $c \neq b$ that forces it to be k.

Finally, if $k \neq b$ a similar argument shows P is pivotal for a as well as c and that b wins. We conclude that P is a dictator: P's top choice always wins. That proves theorem 4. Q.E.D.

7 Proofs of Smith-Young-Levenglick-Hansson-Moulin-Perez theorems

We shall not prove all of these theorems; we shall instead only prove the simplest among them and sketch most of the remaining proofs.

Table 7.1 shows that no weighted positional scoring system can elect a Condorcet-Winner whenever one exists.

#voters	their vote
5	$A > B > C$
2	$B > A > C$
4	$B > C > A$
2	$C > A > B$

Figure 7.1. Nasty 13-voter example. A is the unique Condorcet-Winner. However, the total numbers of (top-ranked, mid-ranked, bottom-ranked) votes garnered are respectively $A : (5, 4, 4)$, $B : (6, 5, 2)$, and $C : (2, 4, 7)$ so that in any weighted-positional score-sum system, no matter what weights $S_1 \geq S_2 \geq S_3$ are employed (provided not all S_k are equal), B would be the unique winner. ▲

Hansson's theorem 8 has a particularly simple standalone **proof.** Suppose for a contradiction that a separable Condorcet (≥ 3)-candidate voting system exists. Let ψ be a V-vote-profile in which no Condorcet-Winner exists and let the election winner-set for ψ (or for 2ψ) contain candidate A. Since A is not a Condorcet-Winner there exists another candidate B who would be preferred to A if all other candidates were ignored, say by V_{BA} votes. (And suppose we choose B to maximize V_{BA}.) Define a new profile ϕ on $2V + V_{BA}$ votes such that $V + V_{BA}$ votes have preference order $A > B > \ldots$ while V votes have preference order $B > A > \ldots$. Evidently A is a Condorcet-Winner for ϕ. Hence by separability $\psi + \phi$ and $2\psi + \phi$ both have winner A. But by construction $2\psi + \phi$ has Condorcet-Winner B, a contradiction. Q.E.D.

We **omit the proof** of Smith and Young's theorem 7 that a symmetric and separable voting system whose output is the name of a winner (or the names of several unordered winners) must be a composed weighted positional system. Essentially, its idea is to investigate convex sets in the $N!$-dimensional space of rational numbers. The set of permissible counts of each of the $N!$ possible types of votes such that candidate A wins, is evidently such a set if we have a symmetric and separable voting system. Then it is realized that these sets really are only N^2-dimensional and must arise from a basis that is the $N \times N$ permutation matrices. Then it is further realized that they must only be N-dimensional polytopal cones and must be defined by inequalities among a single linear real-valued "scoring function."

J.H.Smith's proof that any elimination system whose rounds are based on a weighted-positional scoring system, is nonmonotonic, is in figure 7.2.

#voters	their vote
8	$A > B > C$
8	$C > A > B$
8	$B > C > A$
3	$A > C > B$
3	$B > A > C \to A > B > C$
3	$C > B > A \to C > A > B$
2	$A > B > C$
1	$A > C > B$
1	$B > A > C$

Figure 7.2. J.H.Smith's 37-voter nonmonotonicity example [14]. In any elimination system whose rounds are based on a nontrivial weighted-positional scoring system, C will be eliminated in the first round and then A will win 22 to 15. But after 6 voters of two kinds make the changes in A's favor indicated by the arrows, then B gets eliminated in the first round whereupon C wins 19 to 18. (Evidently only one of the two kinds of vote need to change to stop A from winning, either in this example, or in the situation after the first kind change.) ▲

Although Smith [14] did not mention it, we remark that further, if the round-eliminations are based on any system that eliminates Condorcet Losers then this same counterexample works. (This strengthens his theorem.)

Finally, we present a marvelous 1-page **proof** by Markus Schulze,[14] of (a somewhat altered version of) the Moulin-Perez theorem 11 that voting systems that uniquely elect Condorcet-Winners must exhibit no-show paradoxes.

Theorem 20 (Our form of Moulin-Perez theorem). *It is impossible for any (possibly nondeterministic) single-winner election method (with preference orderings as votes) with* (≥ 4) *candidates to satisfy both*

Condorcet: *if there is a Condorcet-Winner, he must be elected uniquely and with certainty,*

(Strong) participation: *adding votes with A ranked top cannot decrease the probability A wins, and adding votes with B ranked below all voters who win with positive probability, cannot increase the probability that B wins.*

(Note: further, the extension to our proof will produce "maximally dramatic" no-show paradoxes requiring only a single no-show voter.)

Proof (Schulze). Suppose such a method existed. Then starting with the 15-voter scenario in table 7.3, we shall in six further steps derive a contradiction.

#voters	their vote
3	$A > D > B > C$
3	$A > D > C > B$
4	$B > C > A > D$
5	$D > B > C > A$

[14] And included with his permission.

[15] Or by induction, any number of copies of that voter.

Figure 7.3. Situation 1. ▲

Situation 2: Suppose B was elected with positive probability in situation 1. When we add 6 $B > D > A > C$ voters B must be elected with positive probability (by participation) and D must be elected with certainty according to Condorcet. Therefore, B couldn't be elected with positive probability in situation 1.

Situation 3: Suppose C was elected with positive probability in situation 1. When we add 8 $C > B > A > D$ voters C must be elected with positive probability (by participation) and B must be elected with certainty according to Condorcet. Therefore, C couldn't be elected with positive probability in situation 1.

Situation 4: Suppose D was elected with positive probability in situation 1. When we add 4 $D > A > B > C$ voters D must be elected with positive probability (by participation) and A must be elected with certainty according to Condorcet. Therefore, D couldn't be elected with positive probability in situation 1.

Situation 5: We conclude from situations 2-4 that A must be elected with certainty in situation 1. When we add 4 $C > A > B > D$ voters, B and D must be elected each with zero probability (by participation).

Situation 6: Suppose A was elected with positive probability in situation 5. When we add 6 $A > C > B > D$ voters A must be elected with positive probability (by participation) and C must be elected with certainty according to Condorcet. Therefore, A couldn't be elected with positive probability in situation 5.

Situation 7: Suppose C was elected with positive probability in situation 5. When we add 4 $C > B > A > D$ voters C must be elected with positive probability (by participation) and B must be elected with certainty according to Condorcet. Therefore, C couldn't be elected with positive probability in situation 5. Q.E.D.

Extension: In the scenarios in our proof with k identical no-show voters, we can – by adding those voters one at a time until the first one that changes the winner – produce a "no-show paradox" scenario with only a *single* no-show voter. Q.E.D.

Remark: Theorem 20's requirement that there be ≥ 4 candidates is essential because Condorcet's own "least reversal" voting system satisfies the conditions of the theorem in the *3*-candidate case. That is because Perez ([10] p.613) trivially showed by computing the change in the pairwise-preference-count matrix that the Simpson-Kramer minmax voting system is immune to single voter[15] "strong" no-show paradoxes of both the "positive" type where that voter, by abstaining from casting his honest vote top-ranking A, prevents A's victory, and the "negative" type where that voter, by casting his honest vote bottom-ranking B, causes B to win – but no other voting system known to him is immune to these both. Now Condorcet's system (as well as Tideman "ranked pairs," Heitzig "River," Schulze "beatpath" etc.) is equivalent to Simpson-Kramer in the 3-candidate case. (Also, although Simpson-Kramer *is* vulnerable to "weak" no-show paradoxes, it is easy to see that they cannot happen in the 3-candidate case; hence all no-show paradoxes are impossible.)

8 Condorcet implies "favorite betrayal"

Theorem 21 (Condorcet implies "favorite betrayal"). *Consider voting systems in which votes are rank-orderings of the candidates (optionally permitting equality-rankings) and in which*

1. *Condorcet winners are always elected when they exist*
2. *In a 3-candidate election without a Condorcet winner, the candidate suffering only one pairwise defeat, and among these the candidate suffering the weakest such defeat, is always elected.*

We claim that in any such voting system, voting in a manner dishonestly ranking one's favorite below top, can be strategically uniquely optimal.

Proof: The proof if ranking-equalities is permitted is in table 8.1;[16] and if they are not permitted see the election in table 8.2. Q.E.D.

#voters	their vote
2	$A > B > C$
3	$C > A > B$
4	$C = B > A$
2	$A > B > C$

Figure 8.1. Defeats are $A > B$ by 7:4, $B > C$ by 4:3, and $C > A$ by 7:4. There is an $A > B > C > A$ Condorcet cycle in which $B > C$ is the weakest defeat ("weakest" measured by either "winning votes" or by "margins"), so that C is elected.

Notice that the two $A > B > C$ voters shown on the bottom line of the table can turn the "lesser evil" B into the Condorcet Winner by "betraying" their favorite "third party" candidate A and voting $B > A > C$ or $B > A = C$ or $B > C > A$.

However, changing their vote instead to $A = B > C$ or $A > B = C$ or $A = C > B$ or $A > C > B$ or $A > B > C$ or $B = C > A$ or $A = B = C$ (or $C > B > A$ or $C > A > B$ or $C > A = B$) does not suffice: then C still uniquely wins in all cases. Hence favorite-betrayal of A was strategically necessary. ▲

#voters	their vote
8	$B > C > A$
6	$C > A > B$
5	$A > B > C$

Figure 8.2. Favorite Betrayal, or how dishonest exaggeration can pay. In this 19-voter example there is a Condorcet cycle, and the winner is B under any of a large number of voting systems.

But if the 6 $C > A > B$ voters insincerely switch to $A > C > B$ ("betraying their favorite" C) then A becomes the winner under a large number of (the same) voting systems and – more to the point for our purposes – becomes the Condorcet-winner. In the view of these voters, this is a better election result, i.e. the betrayal worked, and one may verify voting without betraying C doesn't work. ▲

9 Final note

It is perhaps worth re-iterating that most of these impossibility theorems do not apply, or only apply in weakened form to, voting systems in which votes are *not* preference orderings, but instead are real vectors; and to the extent that is true, these impossibility theorems may be though of as *indictments* of preference-ranking-based voting systems and as reasons to prefer vector votes.

10 Acknowledgements

I thank Rob LeGrand for example #2 of §3, Markus Schulze for his proof of the Moulin theorem, and Alex Small for some useful comments.

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] Peter C. Fishburn: Arrow's Impossibility Theorem: Concise Proof and Infinite Voters, J. Economic Theory 2,1 (March 1970) 103-106.

[4] John Geanakoplos: Three brief proofs of Arrow's impossibility theorem, Cowles foundation Discussion paper #1123RRR (June 2001), Cowles foundation for reserach in economics, Yale University Box 208282, New Haven CT 06520-8281.

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

[6] 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).

[7] Jerome S. Kelly: Arrow Impossibility Theorems, Academic Press 1978.

[8] Hervé Moulin: Condorcet's Principle Implies the No Show Paradox, J. Economic Theory 45 (1988) 53-64.

[9] Prasanta K. Pattanaik & Bezalel Peleg: Distribution of Power Under Stochastic Social Choice Rules, Econometrica 54,4 (1986) 909-921.

[10] Joaquín Pérez: The Strong No Show Paradoxes are a common flaw in Condorcet voting correspondences, Social Choice & Welfare 18 (2001) 601-616.

[11] Amartya Sen: Social choice theory: a re-examination, Econometrica 45,1 (1977) 53-89.

[12] Amartya K. Sen: Collective Choice and Social Welfare, Holden-Day 1970.

[16] Kevin Venzke first proved this result, as far as I know, in an electorama web post; his proof was different from this one and in some ways better, others worse.

[13] Arunava Sen: Another direct proof of the Gibbard-Satterthwaite theorem, Economics Letters 70,3 (Mar. 2001) 381-385.

[14] John H. Smith: Aggregation of preferences with variable electorate, Econometrica 41,6 (1973) 1027-1041.

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

[16] Alan D. Taylor: The manipulability of voting systems, Amer. Math'l. Monthly 109,4 (2002) 321-337.

[17] H.P. Young: Social choice scoring functions, SIAM J. Appl. Math. 28,4 (1975) 824-838.

[18] H.P.Young & A. Levenglick: A consistent extension of Condorcet's election principle, SIAM J. Appl. Math. 35,2 (1978) 285-300.