Mathematics and Social Choice Theory

Topic 4 – Voting methods with more than 2 alternatives

4.1 Social choice procedures

4.2 Analysis of voting methods

4.3 Arrow's Impossibility Theorem

4.4 Cumulative voting and proportional representation

4.5 Fair majority voting - eliminate Gerrymandering

4.1 Social choice procedures

- A group of voters are collectively trying to choose among several alternatives, with the social choice (the "winner") being the alternative receiving the most votes (based on a specified voting method).

- How to take in the information of individual comparisons among the alternatives in the determination of the winner?

- What are the intuitive criteria to judge whether a social choice is "reasonably" acceptable? Is the choice the least unpopular, broadly acceptable, winning in all one-for-one contests, etc?

Example

3 candidates are running for the Senate. By some means, we gather the information on the "preference order" of the voters.

22%	23%	15%	29%	7%	4%
D	D	H	H	J	J
H	J	D	J	H	D
J	H	J	D	D	H

Top choice only	45% for D, 44% for H and 11% for J; D emerges as the "close"' winner.
One-for-one contest between H and D	H scores $(15 + 29 + 7)\% = 51\%$ D scores $(22 + 23 + 4)\% = 49\%$.

General framework

Set A whose elements are called *alternatives* (or candidates); a, b, c, etc.
Set P whose elements are called *people* (or voters); p_1, p_2, p_3, etc.

- Each person p in P has arranged the alternatives in a list according to his preference.
- A social choice procedure is a fixed "receipt" for choosing an alternative based on the preference orderings of the individuals.
- Rational choice assumption: Voters are assumed to make their orderly choices that reflect their personal preferences and desires.

Definition of terms

A "*social choice procedure*" is a function where a typical input is a sequence of individual preference rankings of the alternatives and an output is a single alternative, or a single set of alternatives if we allow ties.

- A sequence of individual preference lists is called a 'profile'.
- The output is called the "social choice" or winner if there is no tie, or the "social choice set" or "those tied for winner" if there is a tie.

Examples of social choice procedures

1. *Plurality voting*

   Declare as the social choice(s) to be the alternative(s) with the largest number of first-place rankings in the individual preference lists.

   1980 US Presidential election: Democrat Jimmy Carter, Republican Ronald Reagan and Independent John Anderson

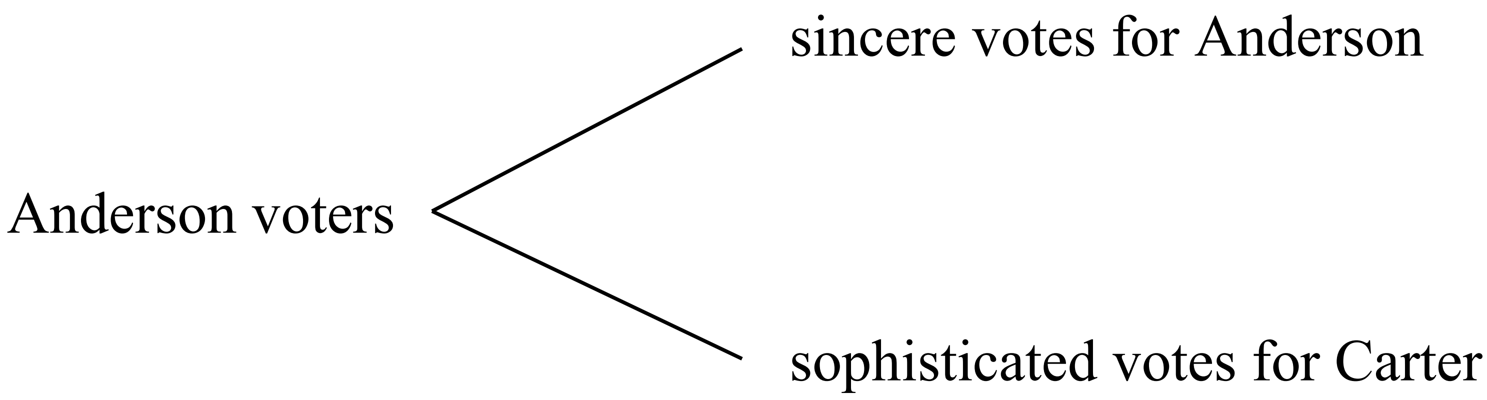
Reagan voters (45%)	*Anderson voters* (20%)	*Carter voters* (35%)
R	A	C
A	C	A
C	R	R

If voters can cast only one vote for their best choice, then Reagan would win with 45% of the vote.

- Reagan was perceived as much more conservative than Anderson who in turn was more conservative than Carter.

Since the chance of Anderson winning is slim, Anderson voters may cast their votes *strategically* to Carter so that their second choice could win.

- A voter's *sincere* strategy is to vote for her first choice.
- Reagan voters have a straightforward strategy: to vote sincerely.
- Adopting an admissible strategy that is not *sincere* is called *sophisticated* voting.



Example

3 voters	2 voters	4 voters	"c" wins with first-choice votes;
a	b	c	but 5-to-4 majority of
b	a	b	voters rank c last.
c	c	a	

Consider pairwise one-for-one contests:-

b beats a by 6 to 3; b beats c by 5 to 4; a beats c by 5 to 4.

Note that b beats the other two in pairwise contests but b is not the winner. Also, c loses to the other two in pairwise contests but c is the winner. This is like Chen in 2000 Taiwan election.

Plurality voting with run-off

Second-step election between the top two vote-getters in plurality election if no candidate receives a majority.

Example

6 voters	*5 voters*	*4 voters*	*2 voters*
a	c	b	b
b	a	c	a
c	b	a	c

"a" with 11 votes beats "b" with 6 votes in the run-off

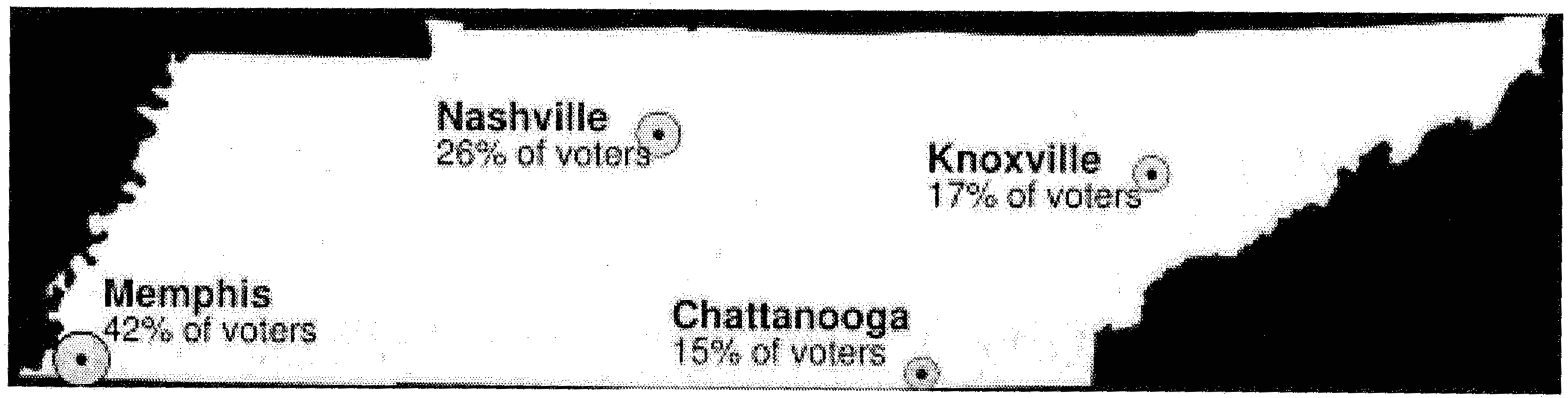
Now, suppose the last 2 voters change their preferences to abc, then "c" beats "a" in the run-off by a vote count of 9 to 8. The moving up of "a" in the last 2 voters indeed hurts "a". (幫他變成害他)

2. *Borda count*

One uses each preference list to award "points" to each of n alternatives: bottom of the list gets zero, next to the bottom gets one point, the top alternative gets $n-1$ points.

The alternative(s) with the highest "scores" is the social choice.

- It sometimes elects broadly acceptable candidates, rather than those preferred by the majority, the Borda count is considered as a consensus-based electoral system, rather than a majoritarian one.



The candidates for the capital of the State of Tennessee are:

- Memphis, the state's largest city, with 42% of the voters, but located far from the other cities
- Nashville, with 26% of the voters, almost at the center of the state and close to Memphis
- Knoxville, with 17% of the voters
- Chattanooga, with 15% of the voters

42% of votors (close to Memphis)	26% of voters (close to Nashville)	15% of voters (close to Chattanooga)	17% of voters (close to Knoxville)
1. Memphis	**1. Nashville**	**1. Chattanooga**	**1. Knoxville**
2. Nashville	2. Chattanooga	2. Knoxvilla	2. Chattanooga
3. Chattanooga	3. Knoxville	3. Nashville	3. Nashville
4. Knoxvilla	4. Memphis	4. Memphis	4. Memphis

City	First	Second	Third	Fourth	Total points
Memphis	42×3	0	0	0	126
Nashville	26×3	42×2	32×1	0	194
Chattanooga	15×3	43×2	42×1	0	173
Knoxville	17×3	15×2	26×1	0	107

- The winner is Nashville with 194 points.

Modification: Voters can be permitted to rank only a subset of the total number of candidates with all unranked candidates being given zero point.

3. *Hare's procedure*

   If no alternative is ranked first by a majority of the voters, the alternative(s) with the smallest number of first place votes is (are) crossed out from all reference orderings, and the first place votes are counted again.

Example 1

5 voters	*2 voters*	*3 voters*	*3 voters*	*4 voters*
a	b	c	d	e
b	c	b	b	b
c	d	d	c	c
d	e	e	e	d
e	a	a	a	a

"b" is eliminated first.

5 voters	*2 voters*	*3 voters*	*3 voters*	*4 voters*
a	c	c	d	e
c	d	d	c	c
d	e	e	e	d
e	a	a	a	a

Next, "d" is eliminated.

5 voters	*2 voters*	*3 voters*	*3 voters*	*4 voters*
a	c	c	c	e
c	e	e	e	c
e	a	a	a	a

There is still no majority winner, so "e" is crossed off. Lastly, "c" is then declared the winner.

- Under plurality with run-off, a and e are the two top vote-getters, ending e as the social choice.

4. *Coombs procedure*

Eliminate the alternative with the *largest* number of *last place* votes, until when one alternative commands the majority support.

Consider Example 1, the steps of elimination are

5 voters	*2 voters*	*3 voters*	*3 voters*	*4 voters*
b	b	c	d	e
c	c	b	b	b
d	d	d	c	c
e	e	e	e	d

"e" is eliminated, leaving

5 voters	*2 voters*	*3 voters*	*3 voters*	*4 voters*
b	b	c	d	b
c	c	b	b	c
d	d	d	c	d

"b", with 11 first place votes, is now the winner.

Example 2

5 voters	*2 voters*	*4 voters*	*2 voters*
a	b	c	c
b	c	a	b
c	a	b	a

- Coombs procedure eliminates "c" and chooses "a".
- If the last two voters change to favor "a" over "b", then "b" will be eliminated and "c" will win.

5. *Dictatorship*

   Choose one of the voters and call her the dictator. The alternative on top of her list is the social choice.

6. *Sequential pairwise voting* (more than 2 alternatives)

   - Two alternatives are voted on first; the majority winner is then paired against the third alternative, etc. The order in which alternatives are paired is called the *agenda* of the voting.

Example

A: Reagan administration – supported bill to provide arms to the Contra rebels.

H: Democratic leadership bill to provide humanitarian aid but not arms.

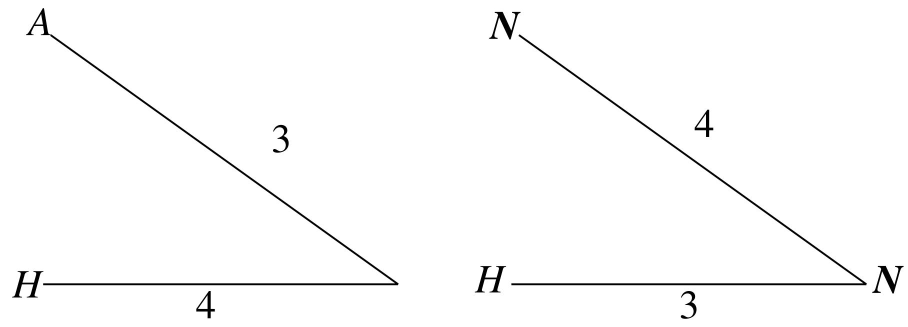
N: giving no aid to the rebels.

In the parliamentary agenda, the first vote was between A and H, with the winner to be paired against N. First, the form of aid is voted, then decide on whether aid or no aid is given to the rebels.

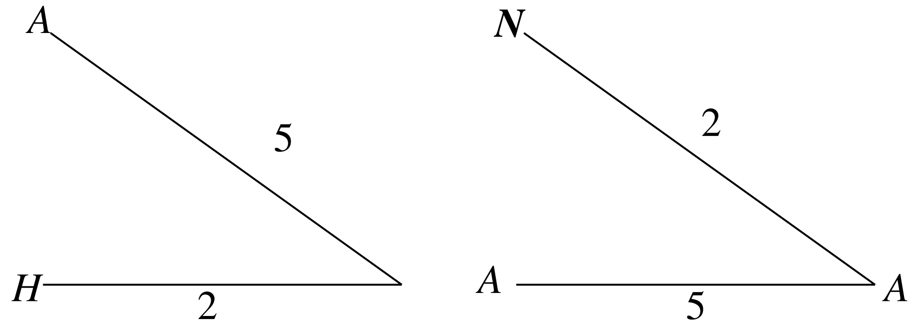
Suppose the preferences of the voters are:

Conservative Republicans	*Moderate Republicans*	*Moderate Democrats*	Liberal Democrats
A	*A*	*H*	*N*
N	*H*	*A*	*H*
H	*N*	*N*	*A*
(2 voters)	(1 voter)	(2 voters)	(2 voters)

- The Conservative Republicans may think that humanitarian aid is non-effective, either no arms or no aid at all. Moderate Democrats may think that some form of aid is at least useful.



Sincere voting



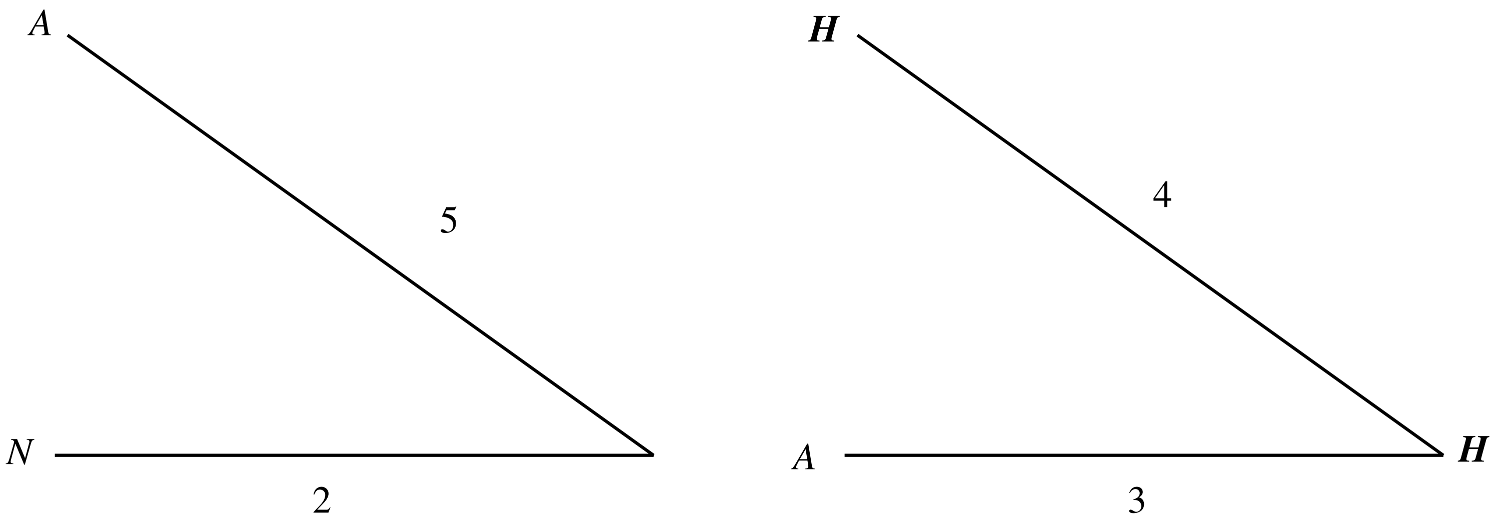
Sophisticated voting

By sophisticated voting, if voters can make A to win first, then A can beat N by 5 to 2.

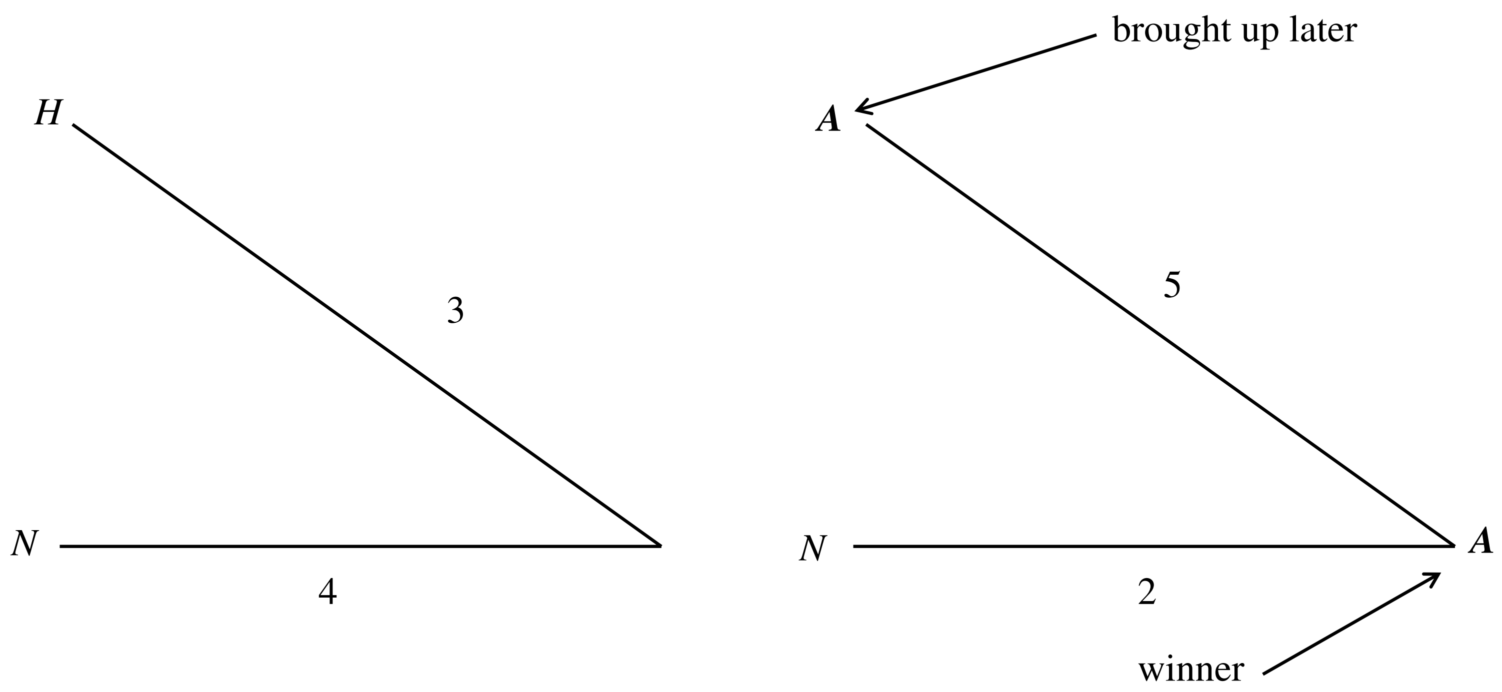
Republicans should vote sincerely for A, the liberal Democrats should vote sincerely for H, but the moderate Democrats should have voted sophisicatedly for A (N is the last choice for moderate Democrats).

Alternative agendas

- produce *any one* of the alternatives as the winner under sincere voting:



Sincere voting

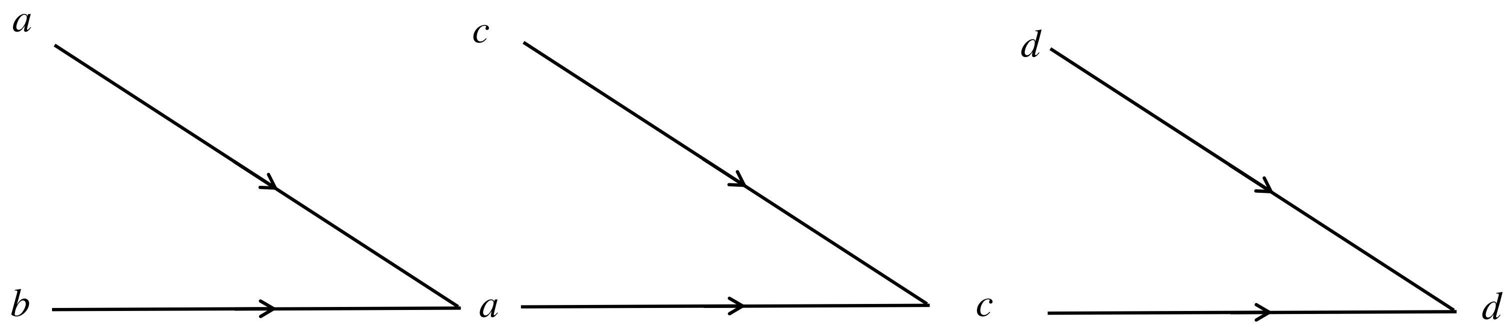


Sincere voting

Remark: The later you bring up your favored alternative, the better chance it has of winning.

Example

Voters are unanimous in preferring b to d.	1 *voter*	1 *voter*	1 *voter*
	a	c	b
	b	a	d
	d	b	c
	c	d	a



This represents a violation of the Pareto condition since all voters prefer b to d but d is the winner.

Voting paradox of Condorcet

Consider the following 3 preference listings of 3 alternatives

list #1	list #2	list #3
a	c	b
b	a	c
c	b	a

If a is the social choice, then #2 and #3 agree that c is better.
If b is the social choice, then #1 and #2 agree that a is better.
If c is the social choice, then #1 and #3 agree that b is better.

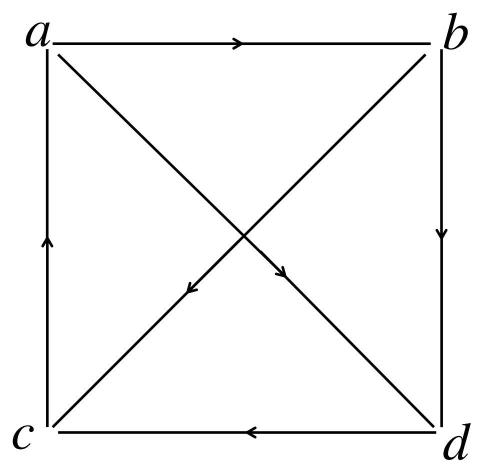
Two-thirds of the people are "constructively unhappy" in the sense of having a single alternative that all agree is superior to the proposed social choice.

Generalization to n alternatives and n people, involving unhappiness of $\dfrac{n-1}{n}$ of the people:

Loss of transitivity in pairwise contest

If a is preferred to b and b is preferred to c, then we expect a to be preferred to c.

1 *voter*	1 *voter*	1 *voter*
a	c	b
b	a	d
d	b	c
c	d	a



a beats b in pairwise contest, b beats c in pairwise contest but a loses to c in pairwise contest.

Chair's paradox

"Apparent power" needs not correspond to control over outcomes.

Consider the same example as in the voting paradox of Condorcet:

$$\begin{array}{ccc} A & B & C \\ a & b & c \\ b & c & a \\ c & a & b \end{array}$$

Here, the preference lists will not be regarded as inputs for the procedure, but only be used to "test" the extent to which each of A, B and C should be happy with the social choice.

The social choice is determined by the plurality voting procedure where voter A (Chair) also has a tie-breaking vote.

Definition

Fix a player P and consider two strategies $V(x)$ and $V(y)$ for P. Here, $V(x)$ denotes "vote for alternative x". $V(x)$ is said to be weakly dominating for player P if

1. For every possible scenario (choice of alternatives for which to vote by the other players), the social choice resulting from $V(x)$ is at least as good for player P as that resulting from $V(y)$.

2. There is at least one scenario in which the social choice resulting from $V(x)$ is strictly better for player P than that resulting from $V(y)$.

A strategy is said to be weakly dominant for player P if it weakly dominates every other available strategy.

How do we determine whether a strategy is a weakly dominant one? List all possible scenarios and compare the result achieved by using this strategy and all other strategies – use of a tree.

Proposition

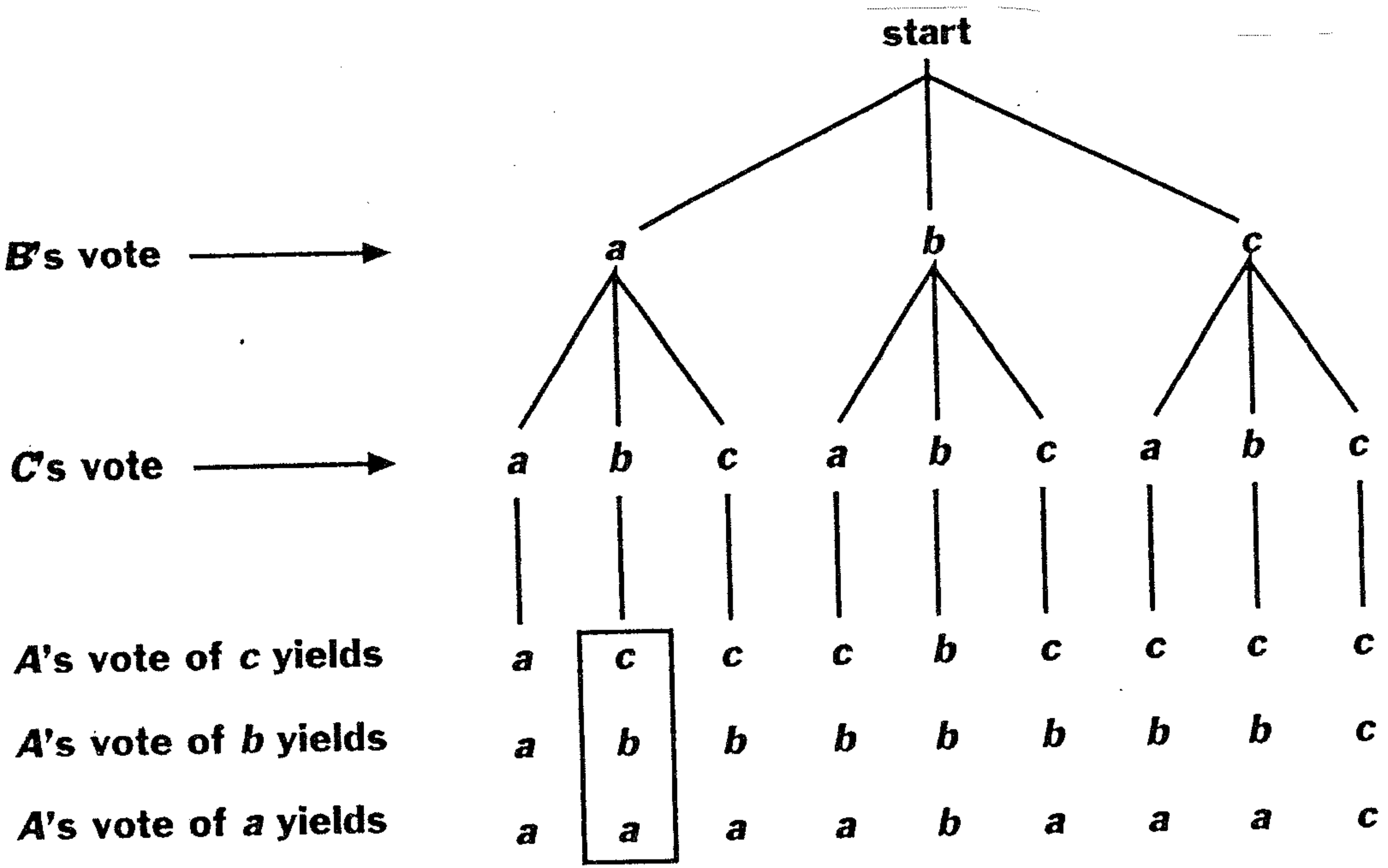
"Vote for alternative a" is a weakly dominant strategy for Chair.

Proof Consider the 9 possible scenarios for the choices of B and C that are listed in a tree.

- Whenever there is a tie, Chair's choice wins.

- In the first case, B's vote is a and C's vote is a, then the outcome is always a, independent of the choice of A.

- In the second case, B's vote is a and C's vote is b, then the outcome matches with A's vote since A is the Chair.

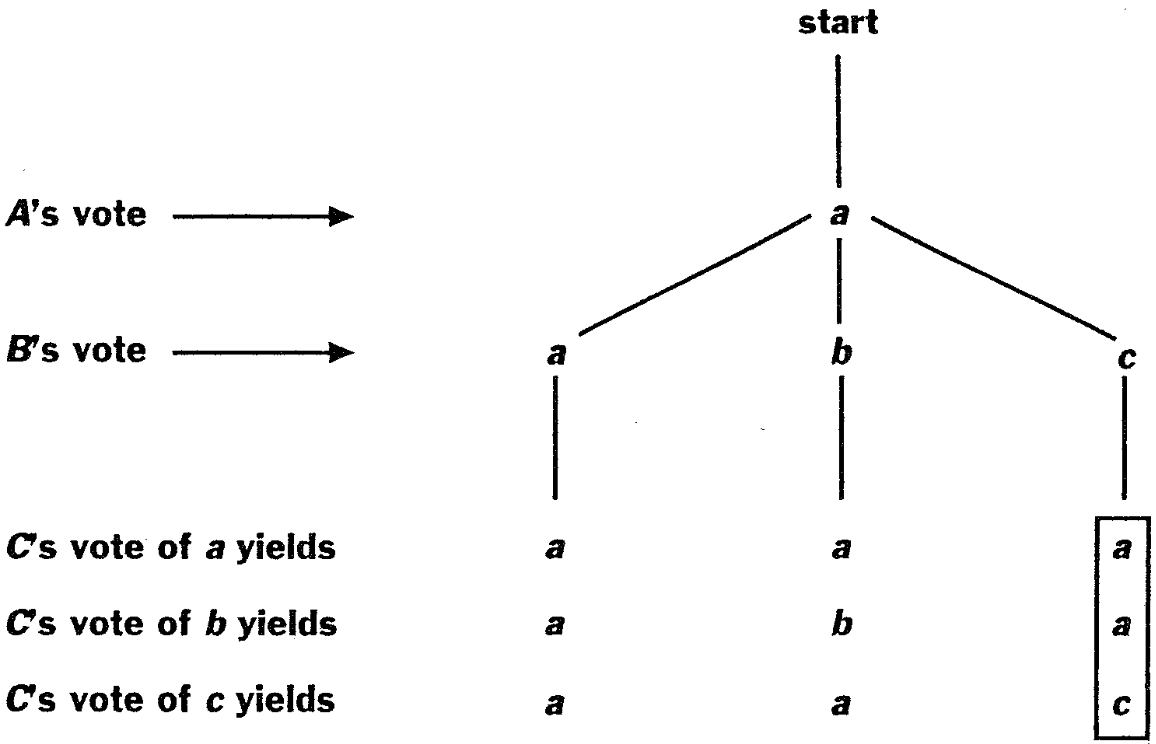

start
B's vote
C's vote
A's vote of c yields
A's vote of b yields
A's vote of a yields
a b c
a b c a b c a b c
a c c c b c c c c
a b b b b b b b c
a a a a b a a a c

The outcome at the bottom of each column (corresponding to A's vote of a) is never worse for A than either of the outcomes (corresponding to A's vote of either b or c) above it, and that it is strictly better than both in at least one case.

- Player A appears to have no rational justification for voting for anything except a.

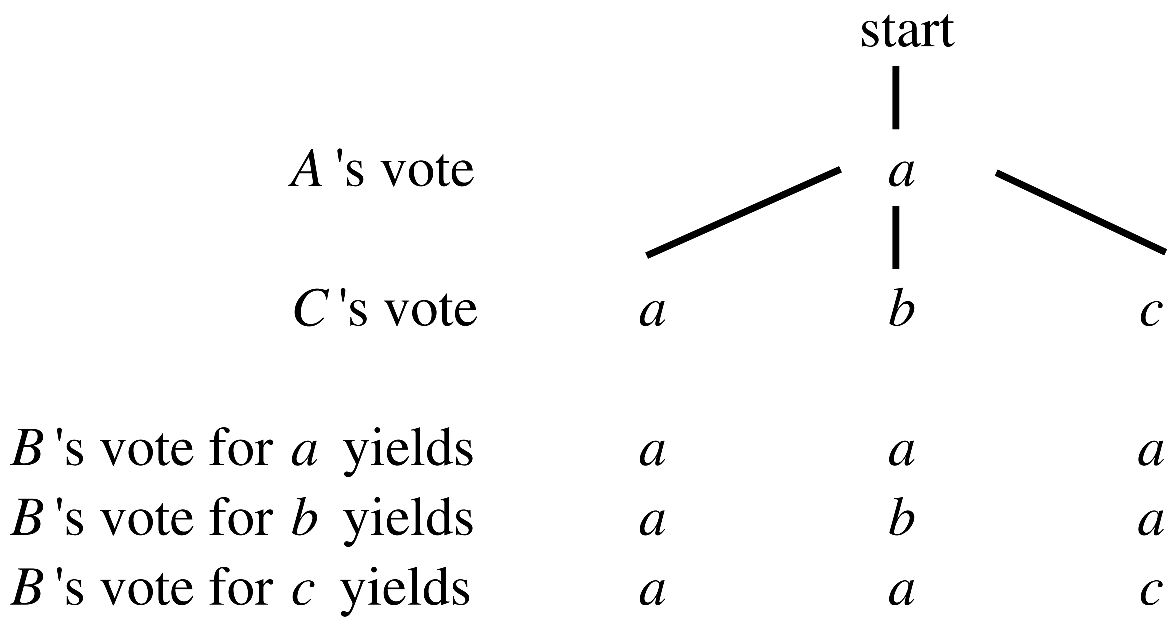
- If we assume that A will definitely go with his weakly dominant strategy, then we analyze what rational self-interest will dictate for the other 2 players in the new game.

For player C: In the last column, C's vote of b yields a since A is the Chair (tie-breaker).



"Vote for c" is a weakly dominating strategy for C since C's preference is $(c \quad a \quad b)$.

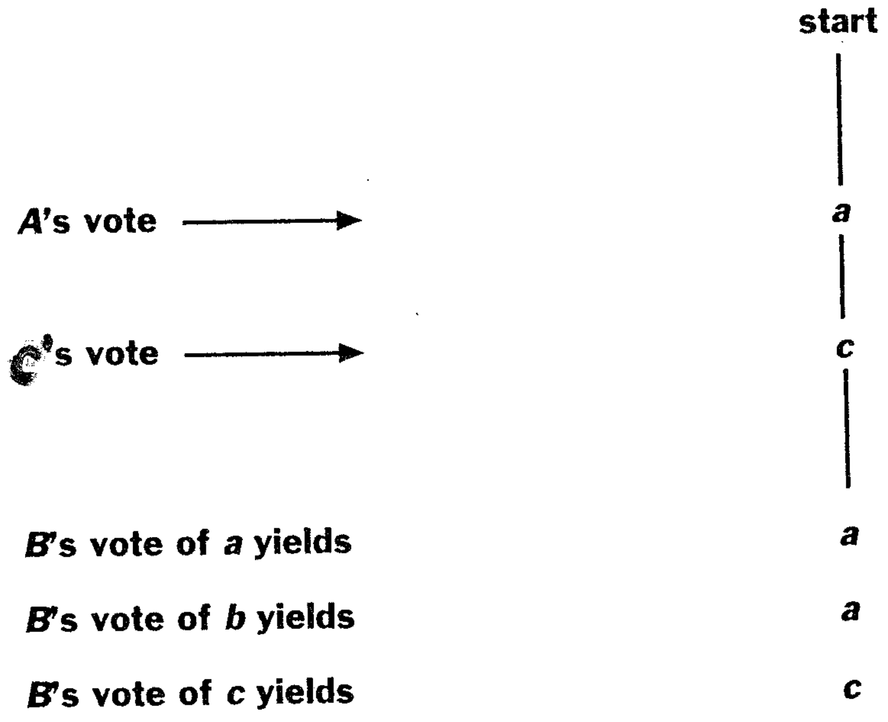
For player B:



B's preference: $(b \quad c \quad a)$

"Vote for b" is not a weakly dominant strategy for B.

In the new game where Player A definitely votes for a and Player C definitely votes for c, the strategy "vote for c" is a weakly dominant strategy for Player B.



Sophisticated voting: A votes for a, B votes for c and C votes for c yield c. Alternative c is A's least preferred alternative even though A had the additional "tie-breaking" power. The additional power as Chair forces the other two votes to vote sophisticatedly.

4.2 Analysis of voting methods

Some properties that are, at least intuitively, desirable.

- If ties were not allowed, then we could have said "the" social choice instead of "a" social choice.

Pareto condition

If everyone prefers x to y, then y cannot be a social choice.

Condorcet Winner Criterion (Condorcet winner may not exist)

If there is an alternative x which could obtain a majority of votes in pairwise contests against *every* other alternative, a voting rule should choose x as the winner.

Condorcet Loser Criterion (Condorcet loser may not exist)

If an alternative y would lose in pairwise majority contests against every other alternative, a voting rule should *not* choose y as a winner.

Monotonicity Criterion (幫他不會導致害他)

If x is a winner under a voting rule, and one or more voters change their preferences in a way favorable to x (without changing the order in which they prefer any other alternatives), then x should still be a winner.

Independence of irrelevant alternatives

For any pair of alternatives x and y, if a preference list is changed but the relative positions of x and y to each other are not changed, then the new list can be described as arising from upward and downward shifts of alternatives other than x and y. Changing preferences toward these other alternatives should be irrelevant to the social preference of x to y.

As a corollary, suppose we start with x a winner while y is a non-winner, people move some other alternative z around, then we cannot guarantee that x is still a winner. However, the independence of irrelevant alternatives at least claims that y should remain a non-winner.

Positive results

1. The plurality procedure satisfies the Pareto condition.

   Proof: If everyone prefers x to y, then y is not on the top of any list (let alone a plurality), and thus y is certainly not a social choice.

2. The Borda count satisfies the Pareto condition.

   Proof: If everyone prefers x to y, then x receives more points from each list than y. Thus, x receives a higher total than y and so y cannot be a winner.

3. The Hare system satisfies the Pareto condition.

Proof: If everyone prefers x to y, then y is not on the top of any list. Thus, either we have immediate winner and y is not among them or the procedure moves on and y is eliminated at the very next stage. Hence, y is not a winner.

4. Sequential pairwise voting satisfies the Condorcet winner criterion.

Proof: A Condoret winner (if exists) always wins the kind of one-on-one contest that is used to produce the winner in sequential pairwise voting.

5. The plurality procedure satisfies monotonicity.

   Proof: If x is on the top of the most lists, than moving x up one spot on some list (and making no other changes) certainly preserves this.

6. The Borda count satisfies monotonicity

   Proof: Swapping x's position with the alternative above x on some list adds one point to x's score and subtracts one point from that of the other other alternative; the scores of all other alternatives remain the same.

7. Sequential pairwise voting satisfies monotonicity.

   Proof: Moving x up on some list only improves x's chances in one-on-one contests.

8. The dictatorship procedure satisfies the Pareto condition.

   Proof: If everyone prefers x to y, then, in particular, the dictator does. Hence, y is not on top of the dictator's list and so is not a social choice.

9. A dictatorship satisfies monotonicity.

   Proof: If x is the social choice then x is already on top of the dictator's list. Hence, the exchange of x with some alternative immediately above x must be taking place on some list other than that of the dictator. Thus, x is still the social choice.

10. A dictatorship satisfies independence of irrelevant alternatives.

    Proof: If x is the social choice and no one — including the dictator — changes his or her mind about x's preference to y, then y cannot come up on top of the dictator's list. Thus, y is not the social choice.

Negative results

1. Sequential pairwise voting with a fixed agenda does not satisfy the Pareto condition.

   Proof:

Voter 1	Voter 2	Voter 3
a	c	b
b	a	d
d	b	c
c	d	a

   Everyone prefers b to d. But with the agenda a b c d, a first defeats b by a score of 2 to 1, and then a loses to c by this same score. Alternative c now goes on to face d, but d defeats c again by a 2 to 1 score. Thus, alternative d is the social choice even though everyone prefers b to d. Alternative d has the advantage that it is bought up later.

2. The plurality procedure fails to satisfy the Condorcet winner criterion.

Proof: Consider the three alternatives a, b, and c and the following sequence of nine preference lists grouped into voting blocs of size four, three, and two.

Voters 1–4	**Voters 5–7**	**Voters 8–9**
a	b	c
b	c	b
c	a	a

- With the plurality procedure, alternative a is clearly the social choice since it has four first-place votes to three b and two for c.
- b is a Condorcet winner, b would defeat a by a score of 5 to 4 in one-on-one competition, and b would defeat c by a score of 7 to 2 in one-on-one competition.

3. The Borda count does not satisfy the Condorcet winner criterion.

4. A dictatorship does not satisfy the Condorcet winner criterion.

   Proof: Consider the three alternatives a, b and c, and the following three preference lists:

Voter 1	**Voter 2**	**Voter 3**
a	c	c
b	b	b
c	a	a

   Assume that Voter 1 is the dictator. Then, a is the social choice, although c is clearly the Condorcet winner since it defeats both others by a score of 2 to 1.

5. The Hare procedure does not satisfy the Condorcet winner criterion.

Proof:

Voters 1–5	**Voters 6–9**	**Voters 10–12**	**Voters 13–15**	**Voter 16–17**
a	e	d	c	b
b	b	b	b	c
c	c	c	d	d
d	d	e	e	e
e	a	a	a	a

- b is the Condorcet winner: b defeats a (12 to 5), b defeats c (14 to 3), b defeats d (14 to 3), b defeats e (13 to 4).
- On the other hand, the social choice according to the Hare procedure is definitely not b. That is, no alternative has the nine first place votes required for a majority, and so b is deleted from all the lists since it has only two first place votes.

6. The Hare procedure does not satisfy monotonicity.

Proof

Voters 1–7	**Voters 8–12**	**Voters 13-16**	**Voter 17**
a	c	b	b
b	a	c	a
c	b	a	c

Since no alternative has 9 or more of the 17 first place votes, we delete the alternatives with the fewest first place votes. In this case, that would be alternatives c and b with only five first place votes each as compared to seven for a. But now a is the only alternative left, and so it is obviously on top of a majority (in fact, all) of the lists. Thus, a is the social choice when the Hare procedure is used.

Favorable-to-a-change yields the following sequence of preference lists:

Voters 1–7	**Voters 8–12**	**Voters 13-16**	**Voter 17**
a	c	b	a
b	a	c	b
c	b	a	c

If we apply the Hare procedure again, we find that no alternative has a majority of first place votes and so we delete the alternative with the fewest first place votes. In this case, that alternative is b with only four. But the reader can now easily check that with b so eliminated, alternative c is on top of 9 of the 17 lists. This is a majority and so c is the soical choice.

7. The plurality procedure does not satisfy independence of irrelevant alternatives.

Voter 1	**Voter 2**	**Voter 3**	**Voter 4**
a	a	b	c
b	b	c	b
c	c	a	a

When the plurality procedure is used, a is a winner and b is a non-winner. Suppose that Voter 4 changes his or her list by moving the alternative c down between b and a. The lists then become:

Voter 1	**Voter 2**	**Voter 3**	**Voter 4**
a	a	b	b
b	b	c	c
c	c	a	a

Notice that we still have b over a in Voter 4's list. However, plurality voting now has a and b tied for the win with two first place votes each. Thus, although no one changed his or her mind about whether a is preferred to b or b to a, the alternative b went from being a non-winner to being a winner.

8. The Borda count does not satisfy independence of irrelevant alternatives.

Proof:

Voters 1–3	**Voters 4 and 5**
a	c
b	b
c	a

The Borda count yields a as the social choice since it gets 6 points $(2+2+2+0+0)$ to only five for $b\,(1+1+1+1+1)$ and four for $c\,(0+0+0+2+2)$.

Voter 1–3	**Voter 4 and 5**
a	b
b	c
c	a

The Borda count now yields b as the social choice with 7 points to only 6 for a and 2 for c.

9. The Hare procedure fails to satisfy independence of irrelevant alternatives.

Proof:

Voter 1	**Voter 2**	**Voter 3**	**Voter 4**
a	a	b	c
b	b	c	b
c	c	a	a

Alternative a is the social choice when the Hare procedure is used because it has at least half the first place votes, a is a winner and b is a non-winner.

Voter 1	**Voter 2**	**Voter 3**	**Voter 4**
a	a	b	b
b	b	c	c
c	c	a	a

Notice that we still have b over a in Voter 4's list. Under the Hare procedure, we now have a and b tied for the win, since each has half the first place votes. Thus, although no one changed his or her mind about whether a is preferred to b or b to a, the alternative b went from being a non-winner to being a winner.

10. Sequential pairwise voting with a fixed agenda fails to satisfy independence of irrelevant alternatives.

Proof:

Consider the alternative c, b and a, and assume this reverse alphabetical ordering is the agenda. Consider the following sequence of three preference lists:

Voter 1	**Voter 2**	**Voter 3**
c	a	b
b	c	a
a	b	c

In sequential pairwise voting, c would defeat b by the score of 2 to 1 and then lose to a by this same score. Thus, a would be the social choice (and thus a is a winner and b is a non-winner).

Suppose that Voter 1 moves c down between b and a, yielding the following lists:

Voter 1	**Voter 2**	**Voter 3**
b	a	b
c	c	a
a	b	c

Now, b first defeats c and then b goes on to defeat a. Hence, the new social choice is b. Thus, although no one changes his or her mind about whether a is preferred to b or b to a, the alternative b went from being a non-winner to being a winner. This shows that independence of irrelevant alternatives fails for sequential pairwise voting with a fixed agenda.

Summary

	Pareto	Condorcet Winner Criterion	Monotonicity	Independence of Irrelevant Alternatives
Plurality	Yes	No	Yes	No
Borda	Yes	No	Yes	No
Hare	Yes	No	No	No
Seq pairs	No	Yes	Yes	No
Dictator	Yes	No	Yes	Yes

Query: The stated properties appear to be quite reasonable. Why haven't we presented a number of natural procedures that satisfy all of these properties and more?

Condorcet voting methods

Recall that only the sequential pairwise voting satisfies the Condorcet winner criterion. However, Borda count does not satisfy the Condorcet winner criterion.

3 voters	*2 voters*	*Borda count:*
a	b	"a" is 6
b	c	"b" is 7
c	a	"c" is 2.

"b" is the Borda winner but "a" is the Condorcet winner. Worse, "a" has an absolute majority of first place votes. [Majority criterion: If a majority of voters have an alternative x as their first choice, a voting rule should choose x.]

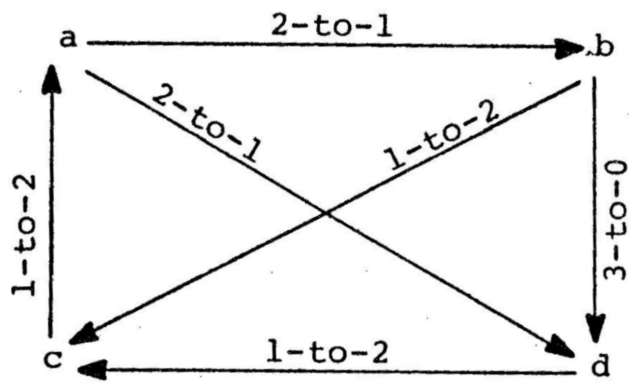
Why "b" wins in the Borda count? The presence of "c" enables the last 2 voters to weigh their votes for "b" over "a" more heavily than the first 3 voters' votes for "a" over "b". If "c" is put to the lowest choice, then "a" is chosen as the Borda winner. This shows a violation of "Independence of Irrelevant Alternatives".

Black method

Value the Condorcet criterion, but also believe that the Borda count has advantages.

- In cases where there is a Condorcet winner, choose it; otherwise, choose the Borda winner.

voter A	*voter* B	*Voter* C
a	c	b
b	a	d
d	b	c
c	d	a



- We check to see if one alternative beats all the other in pairwise contests. If so, that alternative wins. If not, we use the numbers to compute the Borda winner.

- Black method satisfies the Pareto, Condorcet loser, Condorcet winner and Monotonicity criteria. However, it does not satisfy

Generalized Condorcet criterion: If the alternatives can be partitioned into two sets A and B such that every alternative in A beats every alternative in B in pairwise contests, then a voting rule should **not** select an alternative in B.

The above criterion implies both the Condorcet winner and Condorcet loser criteria (take A to be the set which consists of only the Condorcet winner, or B to be the set which consists of only the Condorcet loser).

The following example shows that Black's rule violates this criterion:

1 Voter	*1 Voter*	*1 Voter*
a	b	c
b	c	a
x	x	x
y	y	y
z	z	z
w	w	w
c	a	b

- If we partition the alternatives as $A = [a, b, c]$ and $B = [x, y, z, w]$, then every alternative in A beats every alternative in B by a 2-to-1 vote.

- Furthermore, there is no Condorcet winner, since alternatives a and b and c beat each other cyclically.

- When we compute Borda counts, we get:

a	b	c	x	y	z	w
11	11	11	12	9	6	3

By the Black rule, x is the winner.

Nanson method

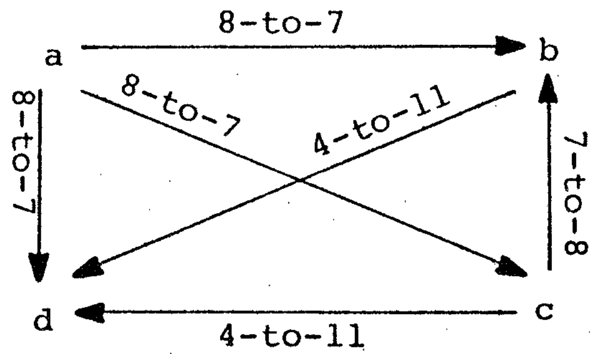
- It is a Borda elimination scheme which sequentially eliminates the alternative with the lowest Borda count until only one alternative or a collection of tied alternatives remains.

- This procedure indeed always select the Condorcet winner, if there is one. Note that the Condorcet winner must gather more than half the votes in its pairwise contests with the other alternatives. Though there is no guarantee that the Condorcet winner wins in Borda count in each pairwise contest, by comparing the sum of the total number of alternatives that are below or above the Condorcet winner, it must always have a *higher than average* Borda count. It would never have the *lowest* Borda count and can never be eliminated in all steps.

3 Voters	4 Voters	4 Voters	4 Voters
b	b	c	d
c	a	a	a
d	c	b	c
a	d	d	b

The sum among all votes of all alternatives that are above a is $3 \times 3 + 4 + 4 + 4 = 21$ while those below a is $2 \times 4 + 2 \times 4 + 2 \times 4 = 24$.

The pairwise voting diagram is:



so that alternative a is the Condorcet winner. The Borda counts are a : 24, b : 25, c : 26 and d : 15. Hence, alternative c would be the Borda winner, and alternative a would come in next-to-last.

Under Nanson's procedure, alternative d is eliminated and new Borda counts are computed:

3 Voters	*4 Voters*	*4 Voters*	*4 Voters*	
b	b	c	a	Borda a : 16
c	a	a	c	counts b : 14
a	c	b	b	c : 15

Alternative b is now eliminated, and in the final round alternative a beats c by 8-to-7.

- Nanson's procedure so cleverly reconciles the Borda count with the Condorcet criterion. It is a shame, but perhaps not surprising, to find that it shares the defect of other elimination schemes: it is not monotonic.

8 Voters	*5 Voters*	*5 Voters*	*2 Voters*
a	c	b	c
b	a	c	b
c	b	a	a

- The Borda counts are $a : 21, b : 20$, and $c : 19$. Hence c is eliminated, and then alternative a beats b by 13-to-7.

- If the last two voters change their minds in favor of alternative a over b, so that their preference ordering is *cab*, the new Borda counts will be $a : 23, b : 18$ and $c : 19$. Hence b will be eliminated and then c beats a by 12-to-8. The change in alternative a's favor has produced c as the winner.

Nanson method always observes monotonicity since the Borda count always increases when the position is moved up in a preference list.

Copeland method

- One looks at the results of pairwise contests between alternatives. For each alternative, compute the number of pairwise wins it has minus the number of pairwise losses it has. Choose the alternative(s) for which this difference is largest.

- It is clear that if there is a Condorcet winner, Copeland's rule will choose it: the Condorcet winner will be the only alternative with all pairwise wins and no pairwise losses. The Copeland rule also satisfies *all* of the other criteria we have considered.

- This method is more likely than other methods to produce ties. If its indecisiveness can be tolerated, it seems to be a very good voting rule indeed.

- It may come into spectacular conflict with the Borda count.

1 Voter	4 Voters	1 Voter	3 Voters
a	c	e	e
b	d	a	a
c	b	d	b
d	e	b	d
e	a	c	c

Copeland scores:	a : 2	Borda scores:	a : 16
	b : 0		b : 18
	c : 0		c : 18
	d : 0		d : 18
	e : -2		e : 20

- Alternative a is the Copeland winner and e comes in last, but e is the Borda winner and a comes in last. The two methods produce diametrically opposite results.

- If we try to ask directly whether a or e is better, we notice that the Borda winner e is preferred to the Copeland winner, alternative a, by *eight* of the nine voters!

Summary

- Sequential pairwise voting is bad because of the agenda effect and the possibility of choosing a Pareto dominated alternative.

- Plurality voting is bad because of the weak mandate（來自選民的授權）it may give. In particular, it may choose an alternative which would lose to *any* other alternative in a pairwise contest. This is a violation of the Condorcet Loser criterion.

- Plurality with run-off and the elimination schemes due to Hare, Coombs and Nanson all fail to be monotonic: changes in an alternative's favor can change it from a winner to a loser.

- Of these four elimination schemes, Coombs and Nanson are better than the others. They generally avoid disliked alternatives, the Nanson rule always detects a Condorcet winner when there is one, and the Coombs scheme almost always does.

- The Borda count takes positional information into full account and generally chooses a non-disliked alternative. Its major difficulty is that it can directly conflict with majority rule, choosing another alternative even when a majority of voters agree on what alternative is best. Thus, the Borda count would only be appropriate in situations where it is acceptable that an alternative preferred by a majority not be chosen if it is strongly disliked by a minority.

- The voting rules due to Copeland and Black appear to be quite strong. The Black rule directly combines the virtues of the Condorcet and Borda approaches to voting. The Copeland rule emphasizes the Condorcet approach. How can it be modified to avoid the most violent of conflicts with the Borda approach?

4.3 Arrow's Impossibility Theorem

Glimpse of Impossibility

There is no social choice procedure for three or more alternatives that satisfies both independence of irrelevant alternatives and the Condorcet winner criterion.

Proof by contradiction: Suppose we have a social choice procedure that satisfies both independence of irrelevant alternatives and the Condorcet winner criterion. We then show that if this procedure is applied to the profile that constitutes Condorcet's voting paradox, then it produces no winner.

Proof

Assume that we have a social choice procedure that satisfies both independence of irrelevant alternatives and the Condorcet winner criterion. Consider the following profile from the voting paradox of Condorcet:

$$\begin{array}{ccc} a & c & b \\ b & a & c \\ c & b & a \end{array}$$

Claim 1 The alternative a is a non-winner.

Consider the following profile (obtained by moving alternative b down in the third preference list from the voting paradox profile):

$$\begin{array}{ccc} a & c & c \\ b & a & b \\ c & b & a \end{array}$$

- Notice that c is a Condorcet winner (defeating both other alternatives by a margin of 2 to 1). Thus, our social choice procedure must produce c as the only winner. Thus, c is a winner and a is a non-winner for this profile.

- Suppose now that the third voter moves b up on his or her preference list. The profile then becomes that of the voting paradox. But no one changed his or her mind about whether c is preferred to a or a is preferred to c. By "independence of irrelevant alternatives", and because we had c as a winner and a as a non-winner in the profile with which we began the proof of the claim, we can conclude that a is still a non-winner when the procedure is applied to the voting paradox profile.

Claim 2 The alternative b is a non-winner.

- Consider the following profile (obtained by moving alternative c down in the second preference list from the voting paradox profile):

$$\begin{array}{ccc} a & a & b \\ b & c & c \\ c & b & a \end{array}$$

Notice that a is a Condorcet winner (defeating both other alternatives by a margin of 2 to 1). Thus, our social choice procedure (which we are assuming satisfies the Condorcet winner criterion) must produce a as the only winner. Thus, a is a winner and b is a non-winner for this profile.

- Suppose now that the second voter moves c up on his or her preference list. The profile then becomes that of the voting paradox. But no one changed his or her mind about whether a is preferred to b or b is preferred to a. By "independence of irrelevant alternatives", and because we had a as a winner and b as a non-winner in the profile with which we began the proof of the claim, we can conclude that b is still a non-winner when the procedure is applied to the voting paradox profile.

Claim 3 It can be shown similarly that the alternative c is a non-winner.

- The above three claims show that when our procedure produces *no* winner. But a social choice procedure must always produce at least one winner. Thus, we have a contradiction and the proof is complete.

Social welfare function

1. Accepts as input a sequence of individual preference lists of some set A (the set of alternatives), and,

2. Produces as output a listing (perhaps with ties) of the set A; this list is called the *social preference list*.

   * Allow ties in the output but not in the input.

Universality (Unrestricted domain) – The social welfare function should account for all preferences among all votes to yield a unique and complete ranking of societal choices.

Note that unlike a social choice procedure, the output is a "social preference listing" of the alternatives.

A social welfare function produces a listing of all alternatives. We can take alternative (or alternatives if tied) at the top of the list as the social choice.

Proposition

Every social welfare function (obviously) gives rise to a social choice procedure (for that choice of voters and alternatives). Moreover (and less obviously), every social choice procedure gives rise to a social welfare function.

- We have a social choice procedure, how to use this procedure to produce a listing of all the alternatives in A.

Iteration procedure

- Simply delete from each of the individual preference lists those alternatives that we've already chosen to be on top of the social preference list.

- Now, input these new individual preference lists to the social choice procedure at hand. The new group of "winners" is precisely the collection of alternatives that we will choose to occupy the second place on the social preference list.

- Continuing this, we delete these "second-round winners" and run the social choice procedure again to obtain the alternatives that will occupy the third place in the social preference list, and so on until all alternatives have been taken care of.

Individual preference lists

b	c	c	$\cdots$
d	a	f	
c	b	a	
e	f	g	
$\vdots$	$\vdots$	$\vdots$	$\cdots$

$\longrightarrow$ **Social Welfare Function or Social Choice Mechanism** $\longrightarrow$

Social preference list

a
d
e
f
$\vdots$

A social welfare function aggregates individual preference lists into a social preference list.

Definition

If A is a set (of alternatives) and P is a set (or people), then a *social welfare function* for A and P that it accepts as inputs only those sequences of individual preference listings of this particular set A that correspond to this particular set P.

- Assume for the moment that we have a fixed set A of three or more alternatives and a fixed finite set P of people. Our goal is to find a social welfare function for A and P that is "reasonable" in the sense of reflecting the will of the people.

Social choice functions for two alternatives

- n people and two alternatives: x and y.
- In this case of having only two alternatives, we may simply vote for one of the alternatives instead of providing a preference list.
- Majority rule declares the social choice to be whichever alternative which has more than half the votes.

Some examples of social welfare functions

1. Designate one person as the dictator.
2. Alternative x is always the social choice.
3. The social choice is x when the number of votes for x is even.

Desirable properties of social welfare functions

1. Anonymity (identity of the voter is irrelevant)

   anonymous（不具名）if the social welfare function is invariant under permutation of the people

   – Dictatorship does not satisfy anonymity

   That is, anonymity implies non-dictatorship.

2. Neutrality (identity of the alternative is irrelevant)

   neutral if it is invariant under permutations of the alternatives

For example, if $(H \quad L \quad H \quad L \quad L)$ yields L; by swapping H for L,

then $(L \quad H \quad L \quad H \quad H)$ should yield H.

If $\begin{pmatrix} a & b & c \\ c & a & b \\ b & c & a \end{pmatrix}$ produces $\begin{pmatrix} c \\ b \\ a \end{pmatrix}$, then $\begin{pmatrix} c & b & a \\ a & c & b \\ b & a & c \end{pmatrix}$ produces $\begin{pmatrix} a \\ b \\ c \end{pmatrix}$. Note that we have swapped a for c and vice versa.

– "Fixing a particular alternative as always the social choice" does not satisfy neutrality.

3. Monotonicity (winning status will not be altered when more votes are received)

   If outcome is L, and one or more votes are changed from H to L, then the outcome is still L.

Quota system

n people and 2 alternatives; fix a number q that satisfies

$$\frac{n}{2} < q \leq n + 1.$$

Consider the procedure wherein the outcome is a *tie* when both alternatives have less than q votes. If one of the alternatives has q or more votes, then it *alone* is the social choice.

1. If n is odd and $q = \dfrac{n+1}{2}$, then the quota system is just the majority rule.

2. What would happen when n is even and $q = \frac{n}{2} + 2$? One alternative may receive $\frac{n}{2} + 1$ while the other receives $\frac{n}{2} - 1$. It leads to a tie since none of the alternatives has q or more votes. In this case, the majority rule is not observed.

3. If $q = n + 1$ and there are only n people, then the outcome is always a tie. This corresponds to the procedure that declares the social choice to be a tie between the two alternatives regardless of how the people vote.

4. If we do not impose $q > \frac{n}{2}$, then it is possible that both alternatives achieve quota. This violates the condition for "lone winner".

All quota systems satisfy anonymity, neutrality, and monotonicity. The first two properties are seen to be automatically satisfied by any quota system since the procedure performs the direct votes counting. The last property is also obvious since adding more votes should not move from winner to "non-winner".

Theorem

Suppose we have a social welfare function for two alternatives that is anonymous, neutral, and monotone. Then that procedure is a quota system.

Proof

It suffices to prove the following 2 conditions:

1. The alternative L alone is the social choice precisely when q or more people vote for L.
2. $\dfrac{n}{2} < q \leq n + 1$.

- The procedure is invariant under permutations of the people, so the outcome depends on the number of people who vote for, say, L.

- Let G denote the set of all numbers k such that L alone (獨贏) is the social choice when exactly k people vote for L.

  (a) When $G = \phi$, this implies that L alone never wins. Also, H alone never wins by neutrality. In this case, the outcome is always a tie.

  (b) If G is not empty, then we let q be the smallest number in G.

  It is easily seen that Monotonicity $\Rightarrow$ (1)

Remark Case (a) corresponds to $q = n + 1$. It is superfluous to take q to be larger than $n + 1$.

- By neutrality, if k is in G, then $n - k$ is definitely not in G. Otherwise, we would have H alone as the social choice when exactly $n - k$ people voted for H (occurring automatically as k people voted for L). This leads to a contradiction that L wins alone.

For example, take $n = 11$ and $q = 8$. Now, $k = 9$ is in G but $n - k = 2$ cannot be in G. Otherwise, if 2 votes are sufficient for L to win, then 2 votes are also sufficient for H to win (neutrality property). However, when L receives 9 votes, then H receives 2 votes automatically. Both H and L win and this is contradicting to L wins alone when it receives 9 votes.

- By invoking monotonicity as a further step, if k is in G, then $n - k$ cannot be as large as k. Thus, $n - k < k$ or $n < 2k$. Hence, $n/2 < k$ for any number that is in G.

- Lastly, $q \leq n$ when G is non-empty and it suffices to take q to be $n+1$ when $G = \phi$. Thus,

$$n/2 < q \leq n + 1.$$

Remark

When n is odd and we choose $q > \dfrac{n+1}{2}$, it is possible that the votes of both alternatives cannot achieve the quota. In this case, we have a tie. For example, we take $n = 11$ and $q = 7$, suppose L has 6 votes and H has 5 votes, then a tie is resulted.

May Theorem

If the number of people is odd and *ties are excluded*, then the only social welfare function for two alternatives that satisfies anonymity, neutrality and monotonicity is majority rule.

Note that at least one of the alternatives must receive number of votes to be $\dfrac{n+1}{2}$ or above. That is, when n is odd and $q = \frac{n+1}{2}$, we can always find a social choice that is alone (no tie). Note that when we choose q to be higher than $\dfrac{n+1}{2}$, then tie occurs.

Weakly reasonable social welfare function

A social welfare function (for A and P) is called *weakly reasonable* if it satisfies the following three conditions:

1. Pareto: also called unanimity (一致同意). Society put alternative x strictly above y whenever every individual puts x strictly above y. As a consequence, suppose the input consists of a sequence of identical lists, then this single list should also be the social preference list produced as output.

   Therefore, Pareto condition implies the surjective property of a social welfare function. That is, every possible societal preference order should be achievable by some set of individual preference lists.

2. Independence of irrelevant alternatives (IIA): Suppose we have our fixed set A of alternatives and our fixed set P of people, but two different sequences of individual preference lists. Also, exactly the same people have alternative x over alternative y in their list.

For example, in the set of 6 voters, the 1st and the 4th voters place x above y while others place y above x. If we move other alternatives around to produce a new sequence, the social preference ordering between x and y remains unchanged.

$$\begin{pmatrix} \cdot \\ \cdot \\ x \\ \cdot \\ \cdot \\ \cdot \\ y \\ \cdot \\ \cdot \end{pmatrix} \quad \cdots \quad \begin{pmatrix} \cdot \\ \cdot \\ \cdot \\ x \\ \cdot \\ \cdot \\ y \\ \cdot \\ \cdot \end{pmatrix} \quad \cdots \quad \xrightarrow[\text{alternatives around}]{\text{moving other}} \quad \begin{pmatrix} x \\ \cdot \\ \cdot \\ \cdot \\ \cdot \\ \cdot \\ \cdot \\ \cdot \\ y \end{pmatrix} \quad \cdots \quad \begin{pmatrix} \cdot \\ \cdot \\ \cdot \\ x \\ y \\ \cdot \\ \cdot \\ \cdot \\ \cdot \end{pmatrix} \quad \cdots$$

Interpretation of Independence of Irrelevant Alternatives

Then we either get x over y in both social preference lists, or we get y over x in both social preference lists. The positioning of alternatives other than x and y in the individual preference lists is irrelevant to the question of whether x is socially preferred to y or y is socially preferred to x. In other words, the social relative ranking (higher or lower) of two alternatives x and y depends only on their relative ranking by every individual.

3. Monotonicity: If we get x over y in the social preference list, and someone who had y over x in his individual preference list interchanges the position of x and y in his list, then we still should get x over y in the social preference list.

Non-dictatorship

There is no individual whose preference always prevails, that is, no individual's preference list is always the social preference list.

Proposition

If A has at least three elements, then any social welfare function for A that satisfies both IIA and the Pareto condition will never produce ties in the output.

Proof

- Assume, for contradiction, some sequence of individual preference lists result in a social preference list in which the alternatives a and b are tied, even though we are not allowing ties in any of the individual preference lists.

- Because of IIA, we know that a and b will remain tied as long as we don't change any individual preference list in a way that reverses that voter's ranking of a and b.

Let c be any alternative that is distinct from a and b. Let X be the set of voters who have a over b in their individual preference lists, and let Y be the rest of the voters (who therefore have b over a in their lists).

$$\overbrace{\begin{matrix} a & \dots & a \\ b & & b \end{matrix}}^{X} \qquad\qquad \overbrace{\begin{matrix} b & \dots & b \\ a & & a \end{matrix}}^{Y}$$

yields

$$ab \text{ (tied).}$$

- Suppose we now insert c between a and b in the lists of the voters in X, and we insert c above a and b in the lists of the voters in Y. Then we will still get a and b tied in the social preference list (by independence of irrelevant alternatives), and we will get c over b by Pareto, since c is over b in every individual preference list. Thus, we have:

$$\overbrace{\begin{array}{ccc} a & & a \\ c & \cdots & c \\ b & & b \end{array}}^{X} \qquad\qquad \overbrace{\begin{array}{ccc} c & & c \\ b & \cdots & b \\ a & & a \end{array}}^{Y}$$

yields

$$\begin{array}{c} c \\ ab. \end{array}$$

- Independence of irrelevant alternatives guarantees us that, as for as a versus c goes, we can ignore b. Thus, we can conclude that if everyone in X has a over c and everyone in Y has c over a, then we get c over a in the social preference list.

- To get our desired contradiction, we will go back and insert c differently from what we did before. We insert c under a and b for the voters in X, and between a and b for the voters in Y. Using Pareto as before shows that we now get:

$$\overbrace{\begin{array}{ccc} a & & a \\ b & \dots & b \\ c & & c \end{array}}^{X} \qquad \overbrace{\begin{array}{ccc} b & & b \\ c & \dots & c \\ a & & a \end{array}}^{Y}$$

yields

$$\begin{array}{c} ab \\ c. \end{array}$$

- Independence of irrelevant alternatives guarantees us that, as far as a versus c goes, we can ignore b. Thus, we can now conclude that if everyone in X has a over c and everyone in Y has c over a, then we get a over c in the social preference list. This is the opposite of what we concluded above, and thus we have the desired contradiction.

Question

Are there any weakly reasonable social welfare functions for A and P?

Yes—appoint a dictator. Taking the dictator's entire individual preference listing of A and declaring it to be the social preference list. Why?

Dictatorship satisfies Pareto condition (if x is preferred to y by all, including the dictator, then x is socially preferred to y), IIA (moving other alternatives would not change the social ranking of x and y) and monotonicity (interchanging the relative order of x and y in lists other than that of the dictator is irrelevant).

Theorem (Arrow, 1950). If A has at least three elements and the set P of individuals is finite, then the only social welfare function for A and P satisfying the Pareto condition, independence of irrelevant alternatives, and monotonicity is a dictatorship.

Remark

The reference to monotonicity is completely unnecessary. It is included simply because it makes the proof conceptually easier. Monotonicity can be removed by an additional lemma.

(Restatement of Arrow's Theorem). If A has at least three elements and the set P of individuals is finite, then it is impossible to find a social welfare function for A satisfying the Pareto condition, independence of irrelevant alternatives, and non-dictatorship.

Setup of the Proof

Under the assumption of Pareto, IIA, and monotonicity, we would like to establish that there always exists a particular singleton voter where the social preference list is the same as the preference of this singleton voter – a dictator. (逃不過“只有獨裁者”的命運)

Definition 某組人能足夠保證把a放在b之上

X is a set of people, a and b are alternatives. "X can force a over b" means

"We get a over b in the social preference list whenever everyone in X places a over b in their individual preference lists."

- Our secret weapons are IIA and monotonicity. In order to show that X forces a over b it suffices to produce a *single sequence* of individual preference lists for which the following all hold.

1. Everyone in X has a over b in their lists.

2. Everyone not in X has b over a in their lists.

3. The resulting social preference list has a over b.

- IIA says that whether or not we get a over b in the social preference list does not depend in any way on the placement of other alternatives in the individual preference lists. Hence, in showing that X forces a over b, it suffices to consider a single sequence of individual preference lists with the property that everyone in X places a over b. Other sequences with the same property would also give a over b in the social preference list.

- By virtue of monotonicity, it suffices to consider the "worst scenario" where those not in X place b above a.

- An empty set cannot force a above b. Why? By (2) suppose every one has b over a, by virtue of the Pareto condition, the resulting social preference list *cannot* have a over b.

Definition of a "dictating set"

Given a social welfare function, a set X is called a dictating set if X can force a over b whenever a and b are two distinctive alternatives in A.

1. If X is the set of all individuals, then X is a dictating set. This follows directly from the Pareto condition. It is guaranteed to have a dictating set once the Pareto condition is satisfied.

2. Let p be one of the individuals. X is a dictating set with single individual p if and only if p is a dictator.

   Dictatorship $\Rightarrow$ "force a over b" is obvious. On the other hand, if p as the only single individual in the dictating set that can always force a over b for any pair of alternatives, the social preference list must coincide with his own preference list, then p is a dictator.

- Dictator may not exist. If a set contains the dictator, then it is a dictating set. If a dictator exists, then only dictating set must contain the dictator.

The strategy for passing from the very large dictating set P where we are starting to the very small dictating set $\{p\}$ involves the following:

Show that if X is a dictating set, and if we split X into any two sets Y and Z of disjoint partitions (so that everyone in X is in exactly one of the two sets), then either Y is a dictating set or Z is a dictating set.

Under the assumption of Pareto, IIA and monotonicity, we would like to establish that there always exists a particular singleton voter where the social preference list is the same as the preference list of this singleton voter – a dictator. This is deduced from the result that there always exists a dictating set with only one element.

Five lemmas yielding Arrow's Theorem

Lemma 1

Suppose X forces a over b and c is an alternative distinct from a and b. Suppose now that X is split into two sets Y and Z (either of which may be the empty set) so that each element of X is in exactly one of the two sets. Then either Y forces a over c or Z forces c over b.

Intuition: If X has the power to force a high and b low, then either Y inherits the power to force a high or Z inherits the power to force b low.

Proof

Suppose X forces a over b under a given social welfare function. Consider what happens when the social welfare function under consideration is applied to the following sequence of individual preference lists as input into the social welfare function:

$$\begin{array}{ccc}
\overbrace{\begin{array}{c} a \\ b \\ c \end{array}}^{\text{Everyone in } Y} \qquad & \qquad \overbrace{\begin{array}{c} c \\ a \\ b \end{array}}^{\text{Everyone in } Z} \qquad & \qquad \overbrace{\begin{array}{c} b \\ c \\ a \end{array}}^{\text{Everyone else}}
\end{array}$$

Alternatives other than a, b, and c can be placed arbitrarily in the individual preference lists. By virtue of IIA, the irrelevant alternatives do not affect the relative ordering of a, b and c in the social preference choice. Notice that everyone in both Y and Z (and thus everyone in X) has a over b.

Since we are assuming that X forces a over b, this means that we get a over b in the social preference list.

Given that a is over b, the three possibilities of ranking a, b and c in the social preference list are

$$\begin{array}{ccc} a & a & c \\ b & c & a. \\ c & b & b \end{array}$$

We have either a over c or c over b in the social preference list.

(i) We get a over c in the social preference list

In this case, we have produced a *single sequence* of individual preference lists for which everyone in Y has a over c in their lists, everyone not in Y has c over a in their lists, and the resulting social preference list has a over c. This suffices to show that Y forces a over c.

(ii) We get c over b in the social preference list.

Proceed in a similar manner for Z.

Query: Can we have both Y forces a over c and Z forces c over b? This corresponds to the case where the societal ranking is a over c and c over b.

Lemma 2

Suppose X forces a over b and c is an alternative distinct from a and b. Then X forces a over c and X forces c over b.

Intuition: If X can force a over b, equivalently, X can force b under a, then X can force a over anything and X can force b under anything.

Proof

- Using Lemma 1, set $Y = X$ and $Z = \phi$. The conclusion is then that either X forces a over c (as desired) or the empty set forces c over b (which is ruled out by the Pareto condition). Thus X forces a over c.

- In a completely analogous way, a consideration of the special case of Lemma 1 where Y is the empty set and Z is the whole set X shows that X forces c over b.

Lemma 3

If X forces a over b, then X forces b over a.

Intuition: The forcing relation is symmetric.

Proof

Choose an alternative c distinct from a and b. (This is possible since we are assuming that we have at least three alternatives.) Assume that X forces a over b. Then, by Lemma 2, X forces a over anything. In particular, X forces a over c. But Lemma 2 now also guarantees that X forces c under anything — in particular, X forces c under b. This is the same as saying X forces b over c. Thus, by Lemma 2 one more time, we have that X forces b over anything, and so X forces b over a as desired. Briefly,

$$X \text{ forces } \begin{matrix} a \\ b \end{matrix} \Rightarrow X \text{ forces } \begin{matrix} a \\ c \end{matrix} \Rightarrow X \text{ forces } \begin{matrix} b \\ c \end{matrix} \Rightarrow X \text{ forces } \begin{matrix} b \\ a \end{matrix}.$$

Lemma 4

Suppose there are two alternatives a and b so that X can force a over b. Then X is a dictating set.

Intuition: If X has a little local power, then X has complete global power.

Proof

Assume X can force a over b, and assume x and y are two arbitrary alternatives. We must show that X can force x over y. Notice that Lemma 3 guarantees that X can also force b over a. Thus, Lemma 2 now lets us conclude that X can force a over or under anything and X can force b over or under anything.

(i) $a = y$

Here, we want to show that X can force x over a. But since we know X can force a under anything, we have that X can force a under x. Equivalently, X can force x over a, as desired.

(ii) $a \neq y$

Since X forces a over b and $a \neq y$, we know that X can force a over y. Equivalently, X can force y under a, and thus X can force y under anything. In particular, X can force y under x. Thus, X can force x over y as desired. Briefly,

$$X \text{ forces } \begin{matrix} a \\ b \end{matrix} \Rightarrow X \text{ forces } \begin{matrix} a \\ y \end{matrix} \Rightarrow X \text{ forces } \begin{matrix} x \\ y \end{matrix}.$$

Lemma 5

Suppose that X is a dictating set and suppose that X is split into two sets Y and Z so that each element of X is in exactly one of the two sets. Then either Y is a dictating set or Z is a dictating set.

Proof

Choose three distinct alternatives a, b, and c. Since X is a dictating set, we have that X can force a over b. Lemma 1 now guarantees that either Y can force a over c (in which case Y is a dictating set by Lemma 4), or Z can force c over b (in which case Z is a dictating set by Lemma 4 again).

Final statement

- We split a given dictating set (at least P is a dictating set) based on splitting a single element off the set at each step. We can always obtain a dictating set which is a singleton. The single element in that dictating set is a dictator.

- We obtain a sequence of dictating sets, the smaller sets are obtained by deleting some players from the larger ones. Actually, all these dictating sets contain the dictator.

4.4 Cumulative voting and proportional representation

Plurality voting

- In *single-winner* plurality voting, each voter is allowed to vote for only one candidate; and the winner of the election is whichever candidate represents a plurality of voters.

- In *multi-member* constituencies, referred to as an *exhaustic counting system*, one member is elected at a time and the process repeated until the number of vacancies is filled.

  Example

  With 8,000 voters and 5 to be elected, under plurality voting, a coalition C of 4001 members can elect 5 candidates of its choice by giving each of the 5 candidates 4,001 votes.

Cumulative voting

Cumulative voting is a multiple-winner voting system intended to promote proportional representation while also being simple to understand.

You may offer up to 3 votes			
1	2	3	
○	○	○	Chan
●	●	○	Lee
○	○	○	Cheung
○	○	●	Wong
○	○	○	Ho

2 votes for Lee and 1 vote for Wong

Voters can 'plump' their votes, conferring all n votes on a single candidate or distributing their n votes as they please.

In cumulative voting, each voter is allotted the same number of votes, while allowing for expression of intensity of candidate preference.

Use of cumulative voting system in the US electoral systems

- Under the usual one-member district system (winner-take-all), voters can elect just one representative from that district, even if another candidate won a substantial percentage of votes.

- Between 1870 and 1980, voters of a state congressional district were able to elect 3 candidates for the Illinois House of Representatives. This allowed for the election of "political minorities". Voters did not understand the cumulative voting system. In 1960s, nearly 45% of Illinois House elections involved only 3 candidates for 3 seats.

- New York City ended cumulative voting in the 1950s because of the election of a communist from Harlem.

"Pros" of cumulative voting systems

- Since 1980, Illinois tried "redrawing political districts" in order to guarantee election of political minorities. This takes power away from the people and gives it to politicians and to the courts.
  - There is nothing in the Illinois Constitution or the US Constitution that requires single-member districts.
  - Proportional voting is the system in most European countries. If 7% of the voters support the Green Party, the Green Party gets 7% of the seats.

- Minority group voters do not have to be made into majorities of voters in order to elect a candidate. The need to manipulate *district lines* is largely, if not completely, eliminated.

Assuring a certain representation

- Voting literature frequently mentions "thresholds", which designate a fraction of population for which a cohesive group whose population fraction is above the threshold can assure itself a certain level of representation under a method of voting.

- For example, a like-minded grouping of voters that is 20% of a city would be well positioned to elect one out of five seats.

- Let P be the total number of voters (population) and n the number of seats to be elected, $P > n$.

- We want the fraction of population x/P over which the group can elect k of n, if the group desires to do so and if they vote strategically. Everybody has n votes.

Negative remarks

It does usually provide proportional representation. However, it may promote factional strife and thus seriously affect the efficiency of the company. It also paves the way for "extremists".

Fair apportionment of seats

- Cumulative voting can guarantee a minority the opportunity to elect representatives in the same number that they would receive by one of the apportionment methods.

- A minority can never guarantee itself greater representation by cumulative voting than that would be allotted and deemed fair by Webster or Jefferson apportionment.

Theorem

Assume that there are P voters and n seats. Under cumulative voting, a coalition C of x voters can guarantee the election of $\left\lfloor \frac{x}{P} n \right\rfloor$ candidates.

Example

Suppose $x = 46, P = 81, n = 8$, a coalition of 46 voters can elect $\left\lfloor \frac{46}{81} \times 8 \right\rfloor = 4$ candidates by giving each of its four candidates $\frac{46 \times 8}{4} = 92$ votes.

Actually, the coalition can elect 5 candidates by giving each of them $\frac{368}{5}$ votes.

Proof

Let $k = \left\lfloor \frac{x}{P} n \right\rfloor$. Coalition C may cast $\left\lfloor \frac{x}{k} n \right\rfloor$ votes for each of these k candidates. It suffices to show that it is impossible to have $n - k + 1$ candidates to receive at least $\frac{x}{k} n$ votes.

Since $k \leq \frac{x}{P} n$, so

$$\frac{n - k + 1}{k} \geq \frac{n - \frac{x}{P} n + 1}{\frac{x}{P} n}.$$

Rearranging, we obtain

$$(n - k + 1)\frac{x}{k} n \geq \left(n - \frac{x}{P} n + 1 \right) \frac{xn}{\frac{x}{P} n} = Pn - xn + P > (P - x)n.$$

where $(P - x)n$ is the maximum number of votes that can be casted by voters outside the coalition. The number of votes required to win $n - k + 1$ candidates is beyond the maximum number of votes held.

Recall that $\frac{x}{k}$ is the number of voters represented by each candidate for the minority if k candidates are chosen, and similarly, that for the majority is $\frac{P-x}{n-k+1}$ if $n-k+1$ candidates are chosen. There is a threshold head counts x required in order to guarantee the election of k candidates.

Lemma

Under cumulative voting, a coalition C of x voters can guarantee the election of k candidates if and only if

$$\frac{x}{k} > \frac{P-x}{n-k+1} \quad \Leftrightarrow \quad \frac{x}{P} > \frac{k}{n+1}.$$

Example

Let $P = 81$ and $n = 8$. A coalition of size $x = 46$ can guarantee the election of 5 candidates since $46 \times 9 > 5 \times 81$.

Proof

(i) $\dfrac{x}{k} > \dfrac{P-x}{n-k+1} \Rightarrow$ election of k candidates.

A coalition of x voters can give each of k candidates $\dfrac{xn}{k}$ votes. The least popular of $n-k+1$ other candidates could receive no more than $\dfrac{(P-x)n}{n-k+1}$ votes. Thus the coalition of x voters can guarantee the election of k candidates if

$$\frac{xn}{k} > \frac{(P-x)n}{n-k+1} \;\Leftrightarrow\; \frac{x}{k} > \frac{P-x}{n-k+1} \;\Leftrightarrow\; \frac{x}{P} > \frac{k}{n+1}.$$

(ii) election of k candidates $\Rightarrow \dfrac{x}{k} > \dfrac{P-x}{n-k+1}$

By contradiction, suppose $\dfrac{x}{k} \leq \dfrac{P-x}{n-k+1}$, then the other $P-x$ voters can block the election of the k^{th} candidate of coalition C. This is because $\dfrac{(P-x)n}{n-k+1}$ votes is more than $\dfrac{xn}{k}$ votes.

- The commonly cited "threshold of exclusion" for cumulative voting $\dfrac{1}{n+1}$ above which a minority can assure itself representation is just a special case with $k = 1$.

- How do we compare with the generalized plurality multimember voting, where every voter has n votes but no plumping is allowed? The most votes that each of a coalition's k candidates receives is x. However, the $(n-k+1)^{\text{st}}$ candidate can receive $P - x$ votes. To elect k candidates, the coalition needs

$$x > P - x \quad \text{or} \quad \frac{x}{P} > \frac{1}{2}.$$

  This result is independent of k, so to assure any representation under generalized plurality voting, a coalition must be a population majority.

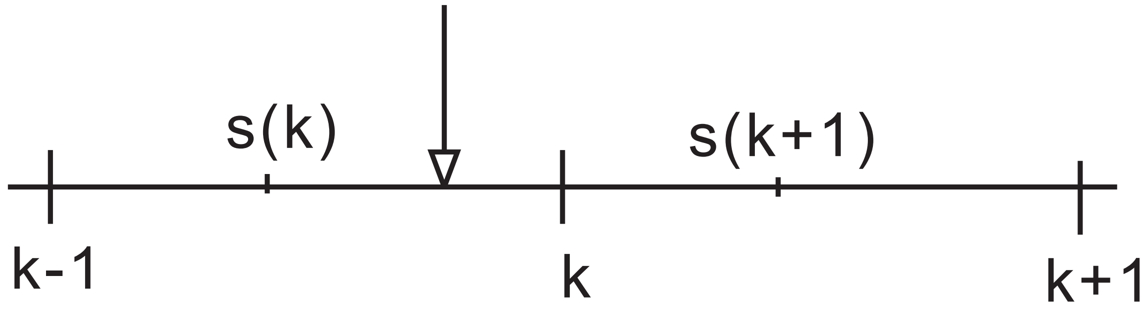
Fair representation

- Webster's method minimizes the absolute difference between all pairs of states, in the numbers of representatives per person, known as "per capita representation". That is, $\left|\frac{a_i}{p_i} - \frac{a_j}{p_j}\right|$ is minimized between any pair of states.

- Consider representation that is apportioned to reflect minority and majority subsets of a population, Dean's method would be more favorable to the minority than Hill's method, which would be more favorable than Webster's method. Recall biases toward larger states: Dean (harmonic mean) $<$ Hill (geometric mean) $<$ Webster (arithmetic mean).

- Suppose that there are 2 groups: minority with population x and majority with population $P - x$. The eligible quota for the minority is $\frac{x}{P}n$.

If the quota falls within $[s(k), s(k+1)]$, then the minority wins k seats.



Recall that $s(k)$ is some chosen form of mean of $k-1$ and k.

For example, the population threshold x for the Webster-fair representation is given by

$$\frac{x}{P} > \frac{s_{\text{Web}}(k)}{n} = \frac{k - \frac{1}{2}}{n}.$$

Reference

"The potential of cumulative voting to yield fair representation", by Duane A. Cooper, *Journal of Theoretical Politics*, vol.19, (2007) p.277-295.

In summary, to deserve k of n seats, the group's quota (as derived from the population threshold x) must be greater than the mean of $k-1$ and k.

Hill-fair representation

$$\frac{x}{P} > \frac{\sqrt{(k-1)k}}{n}$$

Dean-fair representation

$$\frac{x}{P} > \frac{\frac{2}{\frac{1}{k-1}+\frac{1}{k}}}{n} = \frac{k(k-1)}{\left(k-\frac{1}{2}\right)n}.$$

The above means observe the following order: HM $<$ GM $<$ AM

$$\frac{k(k-1)}{\left(k-\frac{1}{2}\right)n} < \frac{\sqrt{k(k-1)}}{n} < \frac{k-\frac{1}{2}}{n}$$

On one hand, minority coalition of population fraction $\frac{x}{P}$ can win k of n seats under cumulative voting method if and only if

$$\frac{1}{2} > \frac{x}{P} > \frac{k}{n+1}.$$

On the other hand, Webster-fair representation requires $\frac{x}{P} > \frac{k-\frac{1}{2}}{n}$.

Comparing $\frac{k-\frac{1}{2}}{n}$ and $\frac{k}{n+1}$, we deduce the algebraic property:

$$\frac{k-\frac{1}{2}}{n} < \frac{k}{n+1} \Leftrightarrow \frac{k}{n+1} < \frac{1}{2}.$$

$$\frac{k-\frac{1}{2}}{n} < \frac{k}{n+1} < \frac{x}{P} < \frac{1}{2}$$

For any minority, cumulative voting can be deemed more favorable to the majority than Webster's method in that a greater threshold is required for the cumulative voting electoral possibilities than is necessary in the measure of Webster-fairness. This counter claims that cumulative voting would be unfairly advantageous to minority populations.

Fairness of cumulative voting

- How often does cumulative voting yield the opportunity for a minority to elect its fair share against a majority?
- When cumulative voting does not make it possible for minority voting strength to elect a fair share, it is possible to elect only one less representative than the Webster-fair amount.

Theorem

In an election for n representatives of the population under cumulative voting, the probability that the minority is unable to elect its Webster-fair share of the n seats is

$$\begin{cases} \frac{1}{4}\frac{n}{n+1}, \text{if } n \text{ is even} \\ \\ \frac{1}{4}\frac{n-1}{n}, \text{if } n \text{ is odd.} \end{cases}$$

Moreover, if the minority's Webster-fair share is $k_w \geq 1$, then it has the voting strength to elect either k_w or $k_w - 1$ representatives.

Proof

Under the scenario of winning k out of n seats for minority $\left(\frac{x}{P} < \frac{1}{2}\right)$, the Webster threshold $\frac{k-\frac{1}{2}}{n}$ is less than the cumulative voting threshold $\frac{k}{n+1}$.

1. The minority cannot elect any more than the Webster-fair number of representation, say, $k_w + 1$. If otherwise, the Webster-fair representation would be at least $k_w + 1$.

2. Also, a minority is able to elect at least $k_w - 1$ representatives. If otherwise, we could have

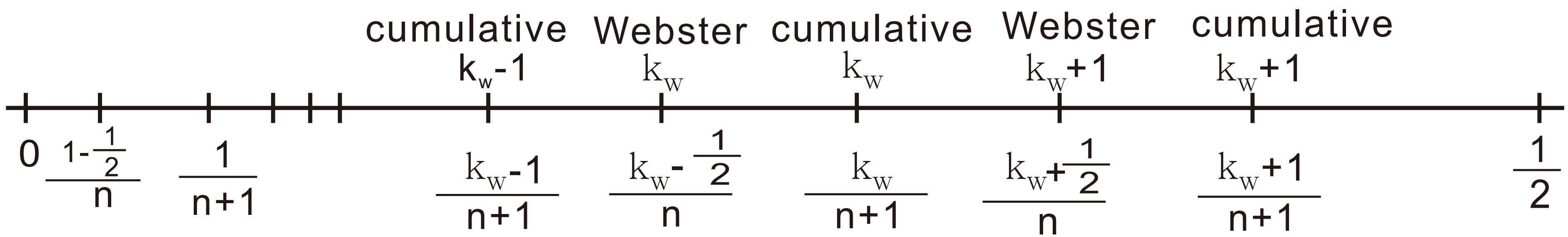
$$\frac{k_w - \frac{1}{2}}{n} < \frac{x}{P} < \frac{k_w - 1}{n + 1}.$$

   (a) The left inequality arises since the Webster-fair representation is k_w;

   (b) The right inequality arises when cumulative voting is assumed to elect less than $k_w - 1$ representatives.

This is impossible since

$$\frac{k_w - 1}{n+1} < \frac{k_w - 1}{n} < \frac{k_w - \frac{1}{2}}{n}.$$



By virtue of the above inequality and $\frac{k-\frac{1}{2}}{n} < \frac{k}{n+1}$, the interval $\left(0, \frac{1}{2}\right)$ can be partitioned by an alternating sequence of Webster- and cumulative voting thresholds as follows:

$$0, \frac{1-\frac{1}{2}}{n}, \frac{1}{n+1}, \frac{2-\frac{1}{2}}{n}, \frac{2}{n+1}, \dots, \frac{\lfloor\frac{n}{2}\rfloor-\frac{1}{2}}{n}, \frac{\lfloor\frac{n}{2}\rfloor}{n+1}, \frac{1}{2},$$

where

$$\lfloor\frac{n}{2}\rfloor = \begin{cases} \frac{n}{2} & \text{if } n \text{ is even} \\ \frac{n-1}{2} & \text{if } n \text{ is odd.} \end{cases}$$

Consider a population of size P. Consider a minority fraction of the population $\frac{x}{P}$ chosen from the uniform distribution on $\left(0, \frac{1}{2}\right) \cap Q$, where Q is the set of rational numbers. The remaining $\frac{P-x}{P}$ constitutes the population's majority.

The probability that cumulative voting does not make it possible for the minority to attain its Webster-fair representation is the probability that the minority has the voting strength to elect $k_w - 1$ representatives but not k_w, which is just the probability that $\dfrac{x}{P}$ belongs to one of the subintervals

$$\left(\frac{k-\frac{1}{2}}{n}, \frac{k}{n+1}\right)$$

of $\left(0, \dfrac{1}{2}\right)$, where $1 \leq k \leq \dfrac{n}{2}$. This probability is just

$$\begin{aligned} & \left|\bigcup_k \left(\frac{k-\frac{1}{2}}{n}, \frac{k}{n+1}\right)\right| \Big/ \left|\left(0, \frac{1}{2}\right)\right| \\ = & \sum_{k=1}^{\lfloor \frac{n}{2} \rfloor} \left(\frac{k}{n+1} - \frac{k-\frac{1}{2}}{n}\right) \Big/ \frac{1}{2}. \end{aligned}$$

Case 1: n is even.

$$\begin{aligned}
\sum_{k=1}^{\lfloor\frac{n}{2}\rfloor}\left(\frac{k}{n+1}-\frac{k-\frac{1}{2}}{n}\right) &= \sum_{k=1}^{\frac{n}{2}}\Big(\frac{k}{n+1}-\frac{k}{n}+\frac{1}{2n}\Big) \\
&= \frac{\frac{\frac{n}{2}(\frac{n}{2}+1)}{2}}{n+1}-\frac{\frac{\frac{n}{2}(\frac{n}{2}+1)}{2}}{n}+\frac{1}{2n}\cdot\frac{n}{2} \\
&= \frac{(n^2+2n)-(n^2+n)}{8(n+1)} \\
&= \frac{n}{8(n+1)}.
\end{aligned}$$

Case 2: n is odd.

$$
\begin{aligned}
\sum_{k=1}^{\lfloor \frac{n}{2} \rfloor} \left(\frac{k}{n+1} - \frac{k-\frac{1}{2}}{n} \right) &= \sum_{k=1}^{\frac{n-1}{2}} \left(\frac{k}{n+1} - \frac{k}{n} + \frac{1}{2n} \right) \\
&= \frac{\frac{\frac{n-1}{2}\left(\frac{n-1}{2}+1\right)}{2}}{n+1} - \frac{\frac{\frac{n-1}{2}\left(\frac{n-1}{2}+1\right)}{2}}{n} + \frac{1}{2n} \cdot \frac{n-1}{2} \\
&= \frac{n-1}{8} - \frac{n^2-1}{8n} + \frac{2n-2}{8n} \\
&= \frac{n-1}{8n}.
\end{aligned}
$$

Therefore, the probability that cumulative voting does not make it possible for the minority to attain its Webster-fair representation is

$$\begin{cases} \frac{n}{8(n+1)} \Big/ \frac{1}{2} = \frac{n}{4(n+1)} & \text{if } n \text{ is even} \\ \frac{n-1}{8n} \Big/ \frac{1}{2} = \frac{n-1}{4n} & \text{if } n \text{ is odd.} \end{cases}$$

Conclusion

Under cumulative voting, a minority of arbitrary size is able, if it chooses, to elect its Webster-fair share of n seats against the majority more than 75% of the time. In the remaining instances, the minority can do no worse than one less than its Webster-fair share.

Example

Consider a population of 500, divided into a polarized majority and minority of 340 and 160 people, respectively, and suppose a five-member representative body is to be elected. The minority – at 32 per cent – has more that $\frac{1}{6}$, but less than $\frac{2}{6}$, of the population; thus under cumulative voting the minority has the electoral strength to elect one, but not two, representatives.

Recall the population threshold for the cumulative voting method to elect k out of n is $\frac{k}{n+1}$. With $n = 5$, the threshold values are $\frac{1}{6}, \frac{2}{6}, \frac{3}{6}, \frac{4}{6}, \frac{5}{6}$.

If the actual population fraction falls within $\frac{k}{n+1}$ and $\frac{k+1}{n+1}$, $k < n$, then k is elected out of n.

- *Webster's appointment*

  Were the five-member body apportioned by Webster's method, the minority's quota of $\frac{160}{500} \cdot 5 = 1.6$ would be rounded up to deserve 2 seats, and the majority's quota of $\frac{340}{500} \cdot 5 = 3.4$ would be rounded down to deserve three seats. It can be readily verified that the absolute difference in per capita representation, $\frac{2}{160} - \frac{3}{340} \approx 0.00368$, is the minimum value for all possible apportionments.

- The population fraction $\frac{160}{500} = 0.32$ exceeds the threshold $\frac{2 - 1/2}{5} = 0.3$ for deserving two of five seats by Webster's method but fails to attain the threshold $\frac{2}{5+1} = \frac{1}{3} \approx 0.333$ to assure two of five seats under cumulative voting.

- Continuing with the consideration of a total population of 500, a minority in the range of 151 to 166 people in a polarized electorate would have to settle for one less than its Webster-fair share of two representatives.

  (i) 151 people can attain 2 seats under Webster apportionment.

  (ii) 167 people are required to attain 2 seats under cumulative voting method.

- Likewise, a minority of size from 51 to 83 would deserve one of five seats by Webster but would not reach the threshold of exclusion necessary for representation by cumulative voting.

- Minorities of sizes 1–49, 84–149, or 167–249 could earn their Webster-fair share of representatives under cumulative voting, comprising about 80 per cent of the possible minority sizes for total population $P = 500$. This is consistent with the theorem's predicted result, where $\left.\dfrac{n-1}{4n}\right|_{n=5} = 20\%$.

Theorem – Cumulative voting and Jefferson's method

A population of size P is partitioned into 2 subgroups of x and $P - x$, with n seats. The number of seats each group can be assured under cumulative voting is equivalent to the number of seats each group would be assigned by Jefferson's method of apportionment.

Numerical example - Jefferson's apportionment

- To apportion the seats under Jefferson's method, again with a majority of 340 and a minority of 160, we would start with divisor $d = \dfrac{500}{5} = 100$, divide that into the populations, and round down, repeating until an appropriate divisor is determined to allocate five seats.

- At first, we have

$$\left\lfloor \frac{340}{100} \right\rfloor = \lfloor 3.4 \rfloor = 3 \quad \text{and} \quad \left\lfloor \frac{160}{100} \right\rfloor = \lfloor 1.6 \rfloor = 1,$$

  but $3 + 1 = 4 \neq 5$.

  We see that $d = 85$ works (as will any d satisfying $80 < d \leq 85$), yielding

$$\left\lfloor \frac{340}{85} \right\rfloor = \lfloor 4.0 \rfloor = 4 \quad \text{and} \quad \left\lfloor \frac{160}{85} \right\rfloor = \lfloor 1.88 \rfloor = 1,$$

  with $4 + 1 = 5$, so the majority is allotted four seats and the minority gets one, the same result achieved by cumulative voting for these subpopulations.

Comparison between Jefferson's and Webster's apportionment

The total over-representation of this Jefferson (4–1) apportionment, $\frac{4}{340} - \frac{5}{500} \approx 0.00176$, is the minimum for all possible apportionments; in particular, it is less than the over-representation $\frac{2}{160} - \frac{5}{500} = 0.00250$ of the Webster (3–2) apportionment.

Concurrently, the total under-representation of the Jefferson apportionment, $\frac{5}{500} - \frac{1}{160} = 0.00375$, is greater than the under-representation $\frac{5}{500} - \frac{3}{340} \approx 0.00118$ of the Webster apportionment.

Proof

By Jefferson's method, we apportion the n seats by finding a divisor d such that $\left\lfloor \frac{x}{d} \right\rfloor + \left\lfloor \frac{P-x}{d} \right\rfloor = n$. We begin by considering $d = \frac{P}{n}$. If $\left\lfloor \frac{x}{P/n} \right\rfloor + \left\lfloor \frac{P-x}{P/n} \right\rfloor = n$, then the population subgroups occur in a ratio that can precisely be represented proportionally among the n seats. Cumulative voting would give the same proportional representation to the subpopulations, if they choose, with appropriate strategy in this case.

For example, suppose we take $x = 100$, $P = 400$, so $P - x = 300$; also, we take $n = 12$. Minority and majority receive 3 and 9 seats, respectively. Minority (majority) puts all 1,200 (3,600) votes into 3 (9) candidates.

Otherwise, and more commonly, we have $\left\lfloor \frac{x}{P/n} \right\rfloor + \left\lfloor \frac{P-x}{P/n} \right\rfloor < n$. Thus, some $d < \frac{P}{n}$ must be determined to get $\left\lfloor \frac{x}{d} \right\rfloor + \left\lfloor \frac{P-x}{d} \right\rfloor = n$.

In order for the subpopulation of x people to be allotted exactly k of the n seats under Jefferson's apportionment, the following two inequalities must be satisfied:

$$k \leq \frac{x}{d} < k+1 \quad \text{and} \quad n-k \leq \frac{P-x}{d} < (n-k)+1.$$

Rearranging the inequalities to solve for d, we obtain

$$\frac{x}{k+1} < d \leq \frac{x}{k} \quad \text{and} \quad \frac{P-x}{(n-k)+1} < d \leq \frac{P-x}{n-k}.$$

Now combining these results, we must have $\dfrac{P-x}{(n-k)+1} < \dfrac{x}{k}$; solving for $\dfrac{x}{P}$, we find the equivalent inequality, $\dfrac{x}{P} > \dfrac{k}{n+1}$.

Similarly, the statements imply that

$$\frac{x}{k+1} < \frac{P-x}{n-k} \Leftrightarrow \frac{x}{P} < \frac{k+1}{n+1}.$$

Putting the two results together, we obtain

$$\frac{k}{n+1} < \frac{x}{P} < \frac{k+1}{n+1}.$$

Interpretation: When there are minority and majority groups only (two states), the Jefferson apportionment gives k seats out of n seats if the fraction of population satisfies the above pair of inequalities.

The subpopulation of size x has the electoral strength to win k of n seats under cumulative voting, but not $k+1$ seats. The k seats are the same as the allotment from Jefferson's method.

- The only remaining consideration is what happens when the population fraction $\frac{x}{P}$ equals a threshold value $\frac{k}{n+1}$. In this instance, both the electoral result of cumulative voting and the apportionment of Jefferson's method are indeterminate.

- When $\frac{x}{P} = \frac{k}{n+1}$, if the two polarized subpopulations of size x and $P-x$ vote perfectly strategically, a tie breaker would be necessary to determine whether the x voters get k or $k-1$ seats and, correspondingly, whether the $P-x$ voters receive $n-k$ or $(n-k)+1$ seats.

Can the result be extended to more than 2 subgroups?

1. Jefferson apportionment results cannot always be guaranteed by cumulative voting. As a counterexample, consider subpopulations X_1, X_2, X_3 of size $x_1 = 350, x_2 = 350, x_3 = 200$, respectively. Using a divisor of 180, we realize that X_1, X_2, X_3 are awarded one seat apiece, as $\left\lfloor \frac{350}{180} \right\rfloor + \left\lfloor \frac{350}{180} \right\rfloor + \left\lfloor \frac{200}{180} \right\rfloor = 1 + 1 + 1 = 3$. However, X_3 does *not* have the electoral strength to elect one of three representatives by cumulative voting, as its population does not exceed the threshold of exclusion, that is, $\frac{200}{900} \leq \frac{1}{3+1}$.

2. We can prove for more than two population subgroups that a subpopulation can never use cumulative voting to guarantee more seats than would be assigned to it by Jefferson apportionment.

Theorem

Consider a population of size P partitioned into subsets $X_1, X_2, \cdots, X_m$ of size $x_1, x_2, \cdots, x_m$, respectively, with a representative body of n seats to be determined. For $i = 1, \cdots, m$, if X_i has the electoral strength to guarantee at least k seats under cumulative voting, then X_i would receive at least k seats by Jefferson apportionment.

Proof

Suppose population subgroup X_i has the electoral strength to guarantee at least k seats under cumulative voting. Recall that this means their fraction of the population must exceed the necessary threshold, that is,

$$\frac{x_i}{P} > \frac{k}{n+1}.$$

By contradiction, let us suppose that X_i receives fewer than k seats by Jefferson apportionment. This means that for the divisor d that achieves the Jefferson apportionment, we have

$$\left\lfloor \frac{x_i}{d} \right\rfloor \leq k - 1.$$

Therefore, $\dfrac{x_i}{d} < k$ and so $d > \dfrac{x_i}{k}$.

The remaining seats are alloted to the remaining $m-1$ population subgroups, so $\displaystyle\sum_{j \neq i} \left\lfloor \frac{x_j}{d} \right\rfloor \geq n - (k-1)$. Therefore,

$$
\begin{aligned}
n-k+1 \;&\le\; \sum_{j\neq i}\left\lfloor \frac{x_j}{d}\right\rfloor \le \left\lfloor \sum_{j\neq i}\frac{x_j}{d}\right\rfloor \\
&=\; \left\lfloor \frac{\sum_{j\neq i} x_j}{d}\right\rfloor = \left\lfloor \frac{P-x_i}{d}\right\rfloor \le \frac{P-x_i}{d}.
\end{aligned}
$$

Thus, $d \le \dfrac{P-x_i}{n-k+1}$ which, in conjunction with the already established $d > \dfrac{x_i}{k}$, implies that

$$
\frac{x_i}{k} < \frac{P-x_i}{n-k+1}.
$$

It follows that

$$
\frac{n-k+1}{k} < \frac{P}{x_i} - 1 \quad \Leftrightarrow \quad \frac{x_i}{P} < \frac{k}{n+1}.
$$

But this contradicts the hypothesis that X_i has the electoral strength to guarantee at least k seats under cumulative voting! Hence, X_i must receive at least k seats by Jefferson's apportionment.

Conclusion

- Cumulative voting's electoral potential is never more advantageous than apportionment by Jefferson's apportionment method and would favor a majority over a minority in some situations.

- Cumulative voting might still be considered quite good and preferable to the status quo, allowing Webster-fair representation more often than not.

- Since cumulative voting's potential is "bounded above" in a sense by Jefferson apportionment, we know that cumulative voting would provide no incentives for groups to splinter into smaller factions.

- Groups may find it advantageous to join forces in coalition. Jefferson's method is the one method of its type that invariably encourages coalitions: subgroups who join forces could gain but could never lose seats; Dean's, Hill's, and Webster's methods do not share this property.

- Cumulative voting might prove more palatable and practicable for use in the United States, with its two-party domination, where rigorous proportional representation methods would be generally unpopular as a means of assuring or bolstering representation by race.

- The nature of cumulative voting, with each voter having n votes, allows individual freedom to express multiple preferences that transcend a single party, race, or political issue. For example, a voter might not strategically vote to maximize the race's chances of electability, choosing instead to distribute votes for all competing interests, such as race, environmental policy, and candidate locality.

4.5 Fair majority voting - eliminate Gerrymandering

- "Districting determines elections, not votes."

- District boundaries are likely to be drawn to maximize the political advantage of the party temporarily dominant in public affairs (谁人掌权).

On one hand, every member of the House of Representatives represents a district.

On the other hand, representatives should represent their districts, their states, and their parties.

Rationale behind fair majority voting (FMV)

Voters cast ballots in single-member districts. In voting for a candidate, each gives a vote to the *candidate's* party.

1. The requisite number of representatives each party receives is calculated by Jefferson's method of apportionment on the basis of the total party votes.

2. The candidates elected, exactly one in each district, and the requisite number from each party are determined by a biproportional procedure.

2004 Connecticut congressional elections: votes.

District	1st	2nd	3d	4th	5th	Total
Republican	73,273	165,558	68,810	149,891	165,440	622,972
Democratic	197,964	139,987	199,652	136,481	105,505	779,589

- The Democratic candidates as a group out-polled the Republican candidates by over $156,000$ votes. However, only 2 were elected to the Republican's 3.

- By the method of Jefferson, the Republicans should have elected only 2 representatives while the Democratic 3.

- In the FMV approach, the 5 Republicans compete for their 2 seats while the 5 Democrats compete for their 3 seats.

Difficulty

- Among the Republicans, the 2 with the most votes have the strongest claims to seats; and similarly for the 3 Democrats with the most votes.

- However, some of these "party-winners" may be in the same district. Who, then, should be elected? (Consider the 4th district where the race is very competitive.)

Method One

- All the Democratic votes should be scaled up until one more of the Democrats' justified-votes exceeds that of his/her Republican opponent.

- This happens when the scaling factor f or the Democratic Party is

$$\frac{149,892}{136,481} \approx 1.0983.$$

2004 Connecticut congressional elections: justified-votes (Democratic candidates' votes all scaled up, district-winners in bold).

District	multiplier	1st	2nd	3d	4th	5th
Republican	1	73,273	**165,558**	68,810	149,891	**165,440**
Democratic	1.0983	**217,416**	153,743	**219,270**	**149,892**	115,872

- Now, the Democratic Party wins the seat in the 4th district.

Method Two

- If every column (district) has exactly one party-winner, they are elected. In Connecticut, the second district has 2 party-winners, the fourth district none.

- Those in districts with more than one winner should be decreased, while the relative votes between the candidates in each district must remain the same.

2004 Connecticut congressional elections: justified-votes (2nd district's candidates' votes both scaled down, party-winners in bold). The scale down makes the Democratic candidate in the 4th district to emerge as the party-winner.

District	1st	2nd	3d	4th	5th
Republican	73,273	$\mathbf{161,410}$	68,810	149,891	$\mathbf{165,440}$
Democratic	$\mathbf{197,964}$	136,480	$\mathbf{199,652}$	$\mathbf{136,481}$	105,505
multiplier	1	0.9749	1	1	1

Multiply the votes of the 2nd district by $136,480/139,987 \approx 0.9749$.

When there are *exactly 2 parties*, a very simple rule yields the FMV result.

(a) Compute the *percentage* of the votes for each of the 2 candidates in each district.

(b) Elect for each party the number of candidates it deserves, taking those with the highest percentages.

2004 Connecticut congressional elections: percentage of votes in districts (FMV winners in bold). Look at the percentages, rather than the actual vote count.

District	1st	2nd	3d	4th	5th
Republican	27.0%	**54.2%**	25.6%	52.3%	**61.1%**
Democratic	**73.0%**	45.8%	**74.4%**	**47.7%**	38.9%

- It eliminated the possibility of defining electoral districts for partisan political advantage. The great loss in district 1 for the Republicans leads to the loss of the seat in the 4th district.

Pros of FMV

- Since parties are allocated seats on the basis of their total votes in all districts, the necessity of strict equality in the number of inhabitants per district is attenuated (less important). This permits districting boundaries to be drawn that respect traditional political, administrative, natural frontiers, and communities of common interest.

- *FMV makes every vote count*. A state like Massachusetts has no Republican representatives at all seems ridiculous. Certainly at least 10% of the potential voters in Massachusetts have preferences for the Republican party, and should be represented by at least one of the state's 10 representatives.

- FMV would prevent a minority of voters from electing a majority in the House.

- If FMV becomes the electoral system, it is inconceivable that a major party would not present a candidate in every district. Even as little as 10% or 20% of the votes against a very strong candidate would help the opposition party to elect one of its candidates in another district. The anomaly of large numbers of unopposed candidates would disappear.

Cons of FMV

It is possible that a district's representative could have received fewer votes than her opponent in the district.

- California's last redistributing is particularly comfortable: every one of its districts has returned a candidate of the same party since 2002. Fifty were elected by a margin of at least 20% in 2002.

Results of 2002, 2004 and 2006 congressional elections.

	2002	2004	2006
Incumbent candidates	386	392	394
Incumbent candidates reelected	380	389	371
Incumbent candidates who lost to outsiders	4	3	23
Elected candidates ahead by $\geq$ 20% of votes	356	361	318
Elected candidates ahead by $\geq$ 16% of votes	375	384	348
Elected candidates ahead by $\leq$ 10% of votes	36	22	56
Elected candidates ahead by $\leq$ 6% of votes	24	10	39
Candidates elected without opposition	81	66	59
Republicans elected	228	232	202
Democrats elected	207	203	233

"Without opposition" means without the opposition of a Democrat or a Republican. "Democrats elected" includes one independent in 2002 and 2004 who usually votes as a Democrat.

Mathematical formulation

Let $x = (x_{ij})$, with $x_{ij} = 1$ if the candidate of party i is elected in district j and $x_{ij} = 0$ otherwise.

FMV selects a $(0, 1)$-valued matrix **x** that satisfies

$$\sum_i x_{ij} = 1,\ j = 1, 2, \ldots, n,\ \ \sum_j x_{ij} = a_i,\ i = 1, 2, \ldots, m.$$

Does a feasible delegation always exist?

	1st	2nd	3d	4th	5th	6th	7th	seats
party 1	+	+	+	+	+	+	+	2
party 2	+	+	+	+	+	+	+	1
party 3	+	+	+	O	O	O	O	4

- 4 districts (4th to 7th) cast all their votes for parties 1 and 2 that together deserve only 3 seats.

- Party 3 deserves 4 seats but receives all its votes from only 3 districts.

Feasible apportionment $\underset{\sim}{\boldsymbol{a}}$ for a given vote matrix V

A *problem* $(V, \boldsymbol{a})$ defined by an $m \times n$ matrix of votes V and an apportionment $\boldsymbol{a}$ satisfying $\sum a_i = n$ is said to be *feasible* if it has at least one feasible delegation x.

Justified-votes

Given row-multipliers $\lambda = (\lambda_i) > 0$ and column-multipliers $\rho = (\rho_j) > 0$, the matrices $\lambda \circ v = (\lambda_i v_{ij})$, $v \circ \rho = (v_{ij}\rho_j)$, and $\lambda \circ v \circ \rho = (\lambda_i v_{ij} \rho_j)$ are the *justified-votes* of the candidates of the different parties in the various districts.