

# Voting Theory

Brian Weatherson, Philosophy 444, February 20-22, 2018

So far, we've been looking at the way that an individual may make a decision. In practice, we are just as often concerned with group decisions as with individual decisions. These range from relatively trivial concerns (e.g. Which movie shall we see tonight?) to some of the most important decisions we collectively make (e.g. Who shall be the next President?). So methods for grouping individual judgments into a group decision seem important.

Unfortunately, it turns out that there are several challenges facing any attempt to merge preferences into a single decision. In this chapter, we'll look at various approaches that different groups take to form decisions, and how these different methods may lead to different results. The different methods have different strengths and, importantly, different weaknesses. We might hope that there would be a method with none of these weaknesses. Unfortunately, this turns out to be impossible.

One of the most important results in modern decision theory is the Arrow Impossibility Theorem, named after the economist Kenneth Arrow who discovered it. The Arrow Impossibility Theorem says that there is no method for making group decisions that satisfies a certain, relatively small, list of desiderata. We will set out the theorem, and explore a little what those constraints are.

Then we'll look a bit at real world voting systems, and their different strengths and weaknesses. Different democracies use quite different voting systems to determine the winner of an election. (Indeed, within the United States there is an interesting range of systems used.) And some theorists have promoted the use of yet other systems than are currently used. Choosing a voting system is not quite like choosing a method for making a group decision. We will start off making the assumption that when

we're looking at ways to aggregate individual preferences into a group decision, we'll assume that we have clear access to the preferences of individual agents. A voting system is not meant to tally preferences into a decision, it is meant to tally votes. And voters may have reasons (some induced by the system itself) for voting in ways other than their preferences. For instance, many voters in American presidential elections vote for their preferred candidate of the two major candidates, rather than 'waste' their vote on a third party candidate.

For now we'll put those problems to one side, and assume that members of the group express themselves honestly when voting. Still, it turns out there are complications that arise for even relatively simple decisions.

## Making a Decision

Seven friends, who we'll imaginatively name  $F_1, F_2, \dots, F_7$  are trying to decide which restaurant to go to. They have four options, which we'll also imaginatively name  $R_1, R_2, R_3, R_4$ . The first thing they do is ask which restaurant each person prefers. The results are as follows.

- $F_1, F_2$  and  $F_3$  all vote for  $R_1$ , so it gets 3 votes
- $F_4$  and  $F_5$  both vote for  $R_2$ , so it gets 2 votes
- $F_6$  votes for  $R_3$ , so it gets 1 vote
- $F_7$  votes for  $R_4$ , so it gets 1 vote

It looks like  $R_1$  should be the choice then. It, after all, has the most votes. It has a 'plurality' of the votes - that is, it has the most votes. In most American elections, the candidate with a plurality wins. This is sometimes known as plurality voting, or (for unclear reasons) first-past-the-post or winner-take-all. The obvious advantage of such a system is that it is easy enough to implement.

But it isn't clear that it is the ideal system to use. Only 3 of the 7 friends wanted to go to  $R_1$ . Possibly the other friends are all strongly opposed to this particular restaurant. It seems unhappy to choose a restaurant that a majority is strongly opposed to, especially if this is avoidable.

So the second thing the friends do is hold a 'runoff' election. This is the method used for voting in some U.S. states (most prominently in Georgia and Louisiana) and many European countries. The idea is that if no candidate (or in this case no restaurant) gets a majority of the vote, then there is a second vote, held just between the top two vote getters. Since  $R_1$  and  $R_2$  were the top vote getters, the choice will just be between those two. When this vote is held the results are as follows.

- $F_1, F_2$  and  $F_3$  all vote for  $R_1$ , so it gets 3 votes
- $F_4, F_5, F_6$  and  $F_7$  all vote for  $R_2$ , so it gets 4 votes

This is sometimes called 'runoff' voting, for the natural reason that there is a runoff. Now we've at least arrived at a result that the majority may not have as their first choice, but which a majority are at least happy to vote for.

But both of these voting systems seem to put a lot of weight on the various friends' first preferences, and less weight on how they rank options that aren't optimal for them. There are a couple of notable systems that allow for these later preferences to count. For instance, here is how the polls in American college sports work. A number of voters rank the best teams from 1 to  $n$ , for some salient  $n$  in the relevant sport. Each team then gets a number of points per ballot, depending on where it is ranked, with  $n$  points for being ranked first,  $n-1$  points for being ranked second,  $n-2$  points for being ranked third, and so on down to 1 point for being ranked  $n$ 'th. The teams' overall ranking is then determined by who has the most points.

In the college sport polls, the voters don't rank every team, only the top  $n$ , but we can imagine doing just that. So let's have each of our friends rank the restaurants in order, and we'll give 4 points to each restaurant

that is ranked first, 3 to each second place, etc. The points that each friend awards are given by the following table.

	$F_1$	$F_2$	$F_3$	$F_4$	$F_5$	$F_6$	$F_7$	Total
$R_1$	4	4	4	1	1	1	1	16
$R_2$	1	3	3	4	4	2	2	19
$R_3$	3	2	2	3	3	4	3	20
$R_4$	2	1	1	2	2	3	4	15

Now we have yet a different choice. By this method,  $R_3$  comes out as the best option. This voting method is sometimes called the Borda count. The nice advantage of it is that it lets all preferences, not just first preferences, count. Note that previously we didn't look at all at the preferences of the first three friends, beside noting that  $R_1$  is their first choice. Note also that  $R_3$  is no one's least favourite option, and is many people's second best choice. These seem to make it a decent choice for the group, and it is these facts that the Borda count is picking up on.

But there is something odd about the Borda count. Sometimes when we prefer one restaurant to another, we prefer it by just a little. Other times, the first is exactly what we want, and the second is, by our lights, terrible. The Borda count tries to approximately measure this - if  $X$  strongly prefers  $A$  to  $B$ , then often there will be many choices between  $A$  and  $B$ , so  $A$  will get many more points on  $X$ 's ballot. But this is not necessary. It is possible to have a strong preference for  $A$  over  $B$  without there being any live option that is 'between' them. In any case, why try to come up with some proxy for strength of preference when we can measure it directly?

That's what happens if we use 'range voting'. Under this method, we get each voter to give each option a score, say a number between 0 and 10, and then add up all the scores. This is, approximately, what's used in various sporting competitions that involve judges, such as gymnastics or diving. In those sports there is often some provision for eliminating the extreme scores, but we won't be borrowing that feature of the system.

Instead, we'll just get each friend to give each restaurant a score out of 10, and add up the scores. Here is how the numbers fall out.

	$F_1$	$F_2$	$F_3$	$F_4$	$F_5$	$F_6$	$F_7$	Total
$R_1$	10	10	10	5	5	5	0	45
$R_2$	7	9	9	10	10	7	1	53
$R_3$	9	8	8	9	9	10	2	55
$R_4$	8	7	7	8	8	9	10	57

Now  $R_4$  is the choice! But note that the friends' individual preferences have not changed throughout. The way each friend would have voted in the previous 'elections' is entirely determined by their scores as given in this table. But using four different methods for aggregating preferences, we ended up with four different decisions for where to go for dinner.

I've been assuming so far that the friends are accurately expressing their opinions. If the votes came in just like this though, some of them might wonder whether this is really the case. After all,  $F_7$  seems to have had an outsized effect on the overall result here. We'll come back to this when looking at options for voting systems.

### Desiderata for Preference Aggregation Mechanisms

None of the four methods we used so far are obviously crazy. But they lead to four different results. Which of these, if any, is the correct result? Put another way, what is the ideal method for aggregating preferences? One natural way to answer this question is to think about some desirable features of aggregation methods. We'll then look at which systems have the most such features, or ideally have all of them.

One feature we'd like is that each option has a chance of being chosen. It would be a very bad preference aggregation method that didn't give any possibility to, say,  $R_3$  being chosen.

More strongly, it would be bad if the aggregation method chose an option  $X$  when there was another option  $Y$  that everyone preferred to  $X$ . Using some terminology from the game theory notes, we can express this

constraint by saying our method should never choose a Pareto inferior option. Call this the **Pareto condition**.

We might try for an even stronger constraint. Some of the time, not always but some of the time, there will be an option  $C$  such that a majority of voters prefers  $C$  to  $X$ , for every alternative  $X$ . That is, in a two-way match-up between  $C$  and any other option  $X$ ,  $C$  will get more votes. Such an option is sometimes called a Condorcet option, after Marie Jean Antoine Nicolas Caritat, the Marquis de Condorcet, who discussed such options. The **Condorcet condition** on aggregation methods is that a Condorcet option always comes first, if such an option exists.

Moving away from these comparative norms, we might also want our preference aggregation system to be fair to everyone. A method that said  $F_2$  is the dictator, and  $F_2$ 's preferences are the group's preferences, would deliver a clear answer, but does not seem to be particularly fair to the group. There should be **no dictators**; for any person, it is possible that the group's decision does not match up with their preference.

More generally than that, we might restrict attention to preference aggregation systems that don't pay attention to *who* has various preferences, just to *what* preferences people have. Here's one way of stating this formally. Assume that two members of the group,  $v_1$  and  $v_2$ , swap preferences, so  $v_1$ 's new preference ordering is  $v_2$ 's old preference ordering and vice versa. This shouldn't change what the group's decision is, since from a group level, nothing has changed. Call this the **symmetry condition**.

Finally, we might want to impose a condition that we said is a condition we imposed on independent agents: the **irrelevance of independent alternatives**. If the group would choose  $A$  when the options are  $A$  and  $B$ , then they wouldn't choose  $B$  out of any larger set of options that also include  $A$ . More generally, adding options can change the group's choice, but only to one of the new options.

## Assessing Plurality Voting

It is perhaps a little disturbing to think how few of those conditions are met by plurality voting, which is how Presidents of the USA are elected. (More precisely, it is how the electoral votes in most states are distributed. The President is chosen by a more complicated mechanism.) Plurality voting clearly satisfies the **Pareto condition**. If everyone prefers  $A$  to  $B$ , then  $B$  will get no votes, and so won't win. So far so good. And since any one person might be the only person who votes for their preferred candidate, and since other candidates might get more than one vote, no one person can dictate who wins. So it satisfies **no dictators**. Finally, since the system only looks at votes, and not at who cast them, it satisfies **symmetry**.

But it does not satisfy the **Condorcet condition**. Consider an election with three candidates.  $A$  gets 40% of the vote,  $B$  gets 35% of the vote, and  $C$  gets 25% of the vote.  $A$  wins, and  $C$  doesn't even finish second. But assume also that everyone who didn't vote for  $C$  has her as their second preference after either  $A$  or  $B$ . Something like this may happen if, for instance,  $C$  is an independent moderate, and  $A$  and  $B$  are doctrinaire candidates from the major parties. Then 60% prefer  $C$  to  $A$ , and 65% prefer  $C$  to  $B$ . So  $C$  is a Condorcet candidate, yet is not elected.

A similar example shows that the system does not satisfy the **irrelevance of independent alternatives** condition. If  $B$  was not running, then presumably  $A$  would still have 40% of the vote, while  $C$  would have 60% of the vote, and would win. One thing you might want to think about is how many elections in recent times would have had the outcome changed by eliminating (or adding) unsuccessful candidates in this way.

## Arrow's Theorem

Arrow's Theorem is an important mathematical result about voting systems. It needs a bit of setup to state carefully. The theorem says that there is no way to aggregate individual preferences into a group preference ordering that satisfies some intuitive criteria.

To get started, we'll assume that each agent has a **complete** and **transitive** preference ordering over the options. If we say  $A >_V B$  means that  $V$  prefers  $A$  to  $B$ , that  $A =_V B$  means that  $V$  is indifferent between  $A$  and  $B$ , and that  $A \geq_V B$  means that  $A >_V B \vee A =_V B$ , then these constraints can be expressed as follows.

**Completeness** For any voter  $V$  and options  $A, B$ , either  $A \geq_V B$  or  $B \geq_V A$

**Transitivity** For any voter  $V$  and options  $A, B$ , the following three conditions hold:

- If  $A >_V B$  and  $B >_V C$  then  $A >_V C$
- If  $A =_V B$  and  $B =_V C$  then  $A =_V C$
- If  $A \geq_V B$  and  $B \geq_V C$  then  $A \geq_V C$

More generally, we assume the **substitutivity of indifferent options**. That is, if  $A =_V B$ , then whatever is true of the agent's attitude towards  $A$  is also true of the agent's attitude towards  $B$ . In particular, whatever comparison holds in the agent's mind between  $A$  and  $C$  holds between  $B$  and  $C$ . (The last two bullet points under transitivity follow from this principle about indifference and the earlier bullet point.)

The effect of these assumptions is that we can represent the agent's preferences by lining up the options from best to worst, with the possibility that we'll have to put two options in one 'spot' to represent the fact that the agent values each of them equally.

A **ranking function** is a function from the preference orderings of the agent to a new preference ordering, which we'll call the preference ordering of the group. We'll use the subscript  $G$  to note that it is the group's ordering we are designing. We'll also assume that the group's preference ordering is complete and transitive.

There are any number ranking functions that don't look at all like the *group's* preferences in any way. For instance, if the function is meant to work out the results of an election, we could consider the function

that takes any input whatsoever, and returns a ranking that simply lists the candidates by age, with the oldest first, the second oldest second, etc. This doesn't seem like it is the group's preferences in any way. Whatever any member of the group thinks, the oldest candidate wins. What Arrow called the **citizen sovereignty** condition is that for any possible ranking of the candidates, it should be possible to have the group end up with that ranking.

The citizen sovereignty follows from another constraint we might put on ranking functions. If everyone in the group prefers  $A$  to  $B$ , then  $A >_G B$ , i.e. the group prefers  $A$  to  $B$ . We'll call this the **Pareto** constraint. It is sometimes called the **unanimity** constraint, but we'll call it the Pareto condition.

One way to satisfy the Pareto constraint is to pick a particular person, and make them dictator. That is, the function 'selects' a person  $V$ , and says that  $A >_G B$  if and only if  $A >_V B$ . If everyone prefers  $A$  to  $B$ , then  $V$  will, so this is consistent with the Pareto constraint. But it also doesn't seem like a way of constructing the group's preferences. So let's say that we'd like a non-dictatorial ranking function. More carefully, for any given person, and any preference ordering they may have, there is some

The last constraint is one we discussed in the previous chapter: the **independence of irrelevant alternatives**. Formally, this means that whether  $A >_G B$  is true depends only on how the voters rank  $A$  and  $B$ . So changing how the voters rank, say  $B$  and  $C$ , doesn't change what the group says about the  $A, B$  comparison.

It's sometimes thought that it would be a very good thing if the voting system respected this constraint. Let's say that you believe that if Gary Johnston had not been a candidate in the 2016 U.S. Presidential election, then Hillary Clinton, not Donald Trump, would have won the election. (I have no idea whether this is actually true, but I can see why someone might believe it.) Then you might think it is a little odd that whether Clinton or Trump wins depends on who else is in the election, and not on the voters' preferences between Clinton and Trump. This is a special

case of the independence of irrelevant alternatives - you think that the voting system should end up with the result that it would have come up with had there been just those two candidates. If we generalise this motivation a lot, we get the conclusion that third possibilities should be irrelevant.

Unfortunately, we've now got ourselves into an impossible situation. Arrow's theorem says that any ranking function that satisfies the Pareto and independence of irrelevant alternatives constraints, has a dictator in any case where the number of alternatives is greater than 2. When there are only 2 choices, majority rule satisfies all the constraints. But nothing, other than dictatorship, works in the general case.

### Cyclic Preferences

We can see why three option cases are a problem by considering one very simple example. Say there are three voters,  $V_1, V_2, V_3$  and three choices  $A, B, C$ . The agent's rankings are given in the table below. (The column under each voter lists the choices from their first preference, on top, to their least favourite option, on the bottom.)

$V_1$	$V_2$	$V_3$
$A$	$B$	$C$
$B$	$C$	$A$
$C$	$A$	$B$

If we just look at the  $A/B$  comparison,  $A$  looks pretty good. After all, 2 out of 3 voters prefer  $A$  to  $B$ . But if we look at the  $B/C$  comparison,  $B$  looks pretty good. After all, 2 out of 3 voters prefer  $B$  to  $C$ . So perhaps we should say  $A$  is best,  $B$  second best and  $C$  worst. But wait! If we just look at the  $C/A$  comparison,  $C$  looks pretty good. After all, 2 out of 3 voters prefer  $C$  to  $A$ .

It might seem like one natural response here is to say that the three options should be tied. The group preference ranking should just be that  $A =_G B =_G C$ . But note what happens if we say that and accept independence of irrelevant alternatives. If we eliminate option  $C$ , then we

shouldn't change the group's ranking of  $A$  and  $B$ . That's what independence of irrelevant alternatives says. So now we'll be left with the following rankings.

$V_1$	$V_2$	$V_3$
$A$	$B$	$A$
$B$	$A$	$B$

By independence of irrelevant alternatives, we should still have  $A =_G B$ . But 2 out of 3 voters wanted  $A$  over  $B$ . The one voter who preferred  $B$  to  $A$  is making it that the group ranks them equally. That's a long way from making them a dictator, but it's our first sign that our constraints give excessive power to one voter. One other thing the case shows is that we can't have the following three conditions on our ranking function.

- If there are just two choices, then the majority choice is preferred by the group.
- If there are three choices, and they are symmetrically arranged, as in the table above, then all choices are equally preferred.
- The ranking function satisfies independence of irrelevant alternatives.

I noted after the example that  $V_2$  has quite a lot of power. Their preference makes it that the group doesn't prefer  $A$  to  $B$ . We might try to generalize this power. Maybe we could try for a ranking function that worked strictly by consensus. The idea would be that if everyone prefers  $A$  to  $B$ , then  $A >_G B$ , but if there is no consensus, then  $A =_G B$ . Since how the group ranks  $A$  and  $B$  only depends on how individuals rank  $A$  and  $B$ , this method easily satisfies independence of irrelevant alternatives. And there are no dictators, and the method satisfies the Pareto condition. So what's the problem?

Unfortunately, the consensus method described here violates transitivity, so doesn't even produce a group preference ordering in the formal sense we're interested in. Consider the following distribution of preferences.

$V_1$	$V_2$	$V_3$
$A$	$A$	$B$
$B$	$C$	$A$
$C$	$B$	$C$

Everyone prefers  $A$  to  $C$ , so by unanimity,  $A >_G C$ . But there is no consensus over the  $A/B$  comparison. Two people prefer  $A$  to  $B$ , but one person prefers  $B$  to  $A$ . And there is no consensus over the  $B/C$  comparison. Two people prefer  $B$  to  $C$ , but one person prefers  $C$  to  $B$ . So if we're saying the group is indifferent between any two options over which there is no consensus, then we have to say that  $A =_G B$ , and  $B =_G C$ . By transitivity, it follows that  $A =_G C$ , contradicting our earlier conclusion that  $A >_G C$ .

This isn't going to be a formal argument, but we might already be able to see a difficulty here. Just thinking about our first case, where the preferences form a cycle suggests that the only way to have a fair ranking consistent with independence of irrelevant alternatives is to say that the group only prefers options when there is a consensus in favour of that option. But the second case shows that consensus based methods do not in general produce *rankings* of the options. So we have a problem. Arrow's Theorem shows how deep that problem goes.

### Proofs of Arrow's Theorem

The proofs of Arrow's Theorem, though not particularly long, are a little tricky to follow. So we won't go through them in any detail at all. But I'll sketch one proof due to John Geanakoplos of the Cowles Foundation at Yale.<sup>1</sup> Geanakoplos assumes that we have a ranking function that satisfies Pareto and independence of irrelevant alternatives, and aims to show

<sup>1</sup>The proof is available at <http://ideas.repec.org/p/cwl/cwldpp/1123r3.html>.

that in this function there must be a dictator.

The first thing he proves is a rather nice lemma. Assume that every voter puts some option  $B$  on either the top or the bottom of their preference ranking. Don't assume they all agree: some people hold that  $B$  is the very best option, and the rest hold that it is the worst. Geanakoplos shows that in this case the ranking function must put  $B$  either at the very top or the very bottom.

To see this, assume that it isn't true. So there are some options  $A$  and  $C$  such that  $A \geq_G B$  and  $B \geq_G C$ . Now imagine changing each voter's preferences so that  $C$  is moved above  $A$  while  $B$  stays where it is - either on the top or the bottom of that particular voter's preferences. By Pareto, we'll now have  $C >_G A$ , since everyone prefers  $C$  to  $A$ . But we haven't changed how any person thinks about any comparison involving  $B$ . So by independence of irrelevant alternatives,  $A \geq_G B$  and  $B \geq_G C$  must still be true. By transitivity, it follows that  $A \geq_G C$ , contradicting our conclusion that  $C >_G A$ .

This is a rather odd conclusion I think. Imagine that we have four voters with the following preferences.

$V_1$	$V_2$	$V_3$	$V_4$
$B$	$B$	$A$	$C$
$A$	$C$	$C$	$A$
$C$	$A$	$B$	$B$

By what we've proven so far,  $B$  has to come out either best or worst in the group's rankings. But which should it be? Since half the people love  $B$ , and half hate it, it seems it should get a middling ranking. One lesson of this is that independence of irrelevant alternatives is a very strong condition, one that we might want to question.

The next stage of Geanakoplos's proof is to consider a situation where at the start everyone thinks  $B$  is the very worst option out of some long list of options. One by one the voters change their mind, with each voter in turn coming to think that  $B$  is the best option. By the result we

proved above, at every stage of the process,  $B$  must be either the worst option according to the group, or the best option.  $B$  starts off as the worst option, and by Pareto  $B$  must end up as the best option. So at one point, when one voter changes their mind,  $B$  must go from being the worst option on the group's ranking to being the best option, simply in virtue of that person changing their mind.

We won't go through the rest, but the proof continues by showing that that person has to be a dictator. Informally, the idea is to prove two things about that person, both of which are derived by repeated applications of independence of irrelevant alternatives. First, this person has to retain their power to move  $B$  from worst to first whatever the other people think of  $A$  and  $C$ . Second, since they can make  $B$  jump all options by changing their mind about  $B$ , if they move  $B$  'halfway', say they come to have the view  $A >_V B >_V C$ , then  $B$  will jump (in the group's ranking) over all options that it jumps over in this voter's rankings. But that's possible (it turns out) only if the group's ranking of  $A$  and  $C$  is dependent entirely on this voter's rankings of  $A$  and  $C$ . So the voter is a dictator with respect to this pair. A further argument shows that the voter is a dictator with respect to every pair, which shows there must be a dictator.

## Voting Systems

The Arrow Impossibility Theorem shows that we can't have everything that we want in a voting system. In particular, we can't have a voting system that takes as inputs the preferences of each voter, and outputs a preference ordering of the group that satisfies these three constraints.

1. **Unanimity:** If everyone prefers  $A$  to  $B$ , then the group prefers  $A$  to  $B$ .
2. **Independence of Irrelevant Alternatives:** If nobody changes their mind about the relative ordering of  $A$  and  $B$ , then the group can't change its mind about the relative ordering of  $A$  and  $B$ .
3. **No Dictators:** For each voter, it is possible that the group's ranking will be different to their ranking

Any voting system either won't be a function in the sense that we're interested in for Arrow's Theorem, or will violate some of those constraints. (Or both.) But still there could be better or worse voting systems. Indeed, there are many voting systems in use around the world, and serious debate about which is best. In these notes we'll look at the pros and cons of a few different voting systems.

The discussion here will be restricted in two respects. First, we're only interested in systems for making political decisions, indeed, in systems for electing representatives to political positions. We're not interested in, for instance, the systems that a group of friends might use to choose which movie to see, or that an academic department might use to hire new faculty. Some of the constraints we'll be looking at are characteristic of elections in particular, not of choices in general.

Second, we'll be looking only at elections to fill a single position. This is a fairly substantial constraint. Many elections are to fill multiple positions. The way a lot of electoral systems work is that many candidates are elected at once, with the number of representatives each party gets being (roughly) in proportion to the number of people who vote for that party. This is how the parliament is elected in many countries around the world (including, for instance, Mexico, Germany and Spain). Perhaps more importantly, it is basically the norm for new parliaments to have such kind of multi-member constituencies. But the mathematical issues get a little complicated when we look at the mechanisms for selecting multiple candidates, and we'll restrict ourselves to looking at mechanisms for electing a single candidate.

### Plurality voting

By far the most common method used in America, and throughout much of the rest of the world, is plurality voting. Every voter selects one of the candidates, and the candidate with the most votes wins. As we've already noted, this is called plurality, or first-past-the-post, voting.

Plurality voting clearly does not satisfy the independence of irrelevant

alternatives condition. We can see this if we imagine that the voting distribution starts off with the table on the left, and ends with the table on the right. (The three candidates are  $A$ ,  $B$  and  $C$ , with the numbers at the top of each column representing the percentage of voters who have the preference ordering listed below it.)

40%	35%	25%	40%	35%	25%
$A$	$B$	$C$	$A$	$B$	$B$
$B$	$A$	$B$	$B$	$A$	$C$
$C$	$C$	$A$	$C$	$C$	$A$

All that happens as we go from left-to-right is that some people who previously favoured  $C$  over  $B$ , come to favour  $B$  over  $C$ . Yet this change, which is completely independent of how anyone feels about  $A$ , is sufficient for  $B$  to go from losing the election 40-35 to winning the election 60-40.

This is how we show that a system does not satisfy independent of irrelevant alternatives - coming up with a pair of situations where no voter's opinion about the relative merits of two choices (in this case  $A$  and  $B$ ) changes, but the group's ranking of those two choices changes.

One odd effect of this is that whether  $B$  wins the election depends not just on how voters compare  $A$  and  $B$ , but on how voters compare  $B$  and  $C$ . One of the consequences of Arrow's Theorem might be taken to be that this kind of thing is unavoidable, but it is worth stopping to reflect on just how pernicious this is to the democratic system.

Imagine that we are in the left-hand situation, and you are one of the 25% of voters who like  $C$  best, then  $B$  then  $A$ . It seems that there is a reason for you to not vote the way your preferences go; you'll have a better chance of electing a candidate you prefer if you vote, against your preferences, for  $B$ . So the voting system might encourage voters to not express their preferences adequately. This can have a snowball effect - if in one election a number of people who prefer  $C$  vote for  $B$ , at future elections other people who might have voted for  $C$  will also vote for  $B$



because they don't think enough other people share their preferences for *C* to make such a vote worthwhile.

Indeed, if the candidate *C* themselves strongly prefers *B* to *A*, but thinks a lot of people will vote for them if they run, then *C* might even be discouraged from running because it will lead to a worse election result. This doesn't seem like a democratically ideal situation.

Some of these consequences are inevitable consequences of a system that doesn't satisfy independence of irrelevant alternatives. And the Arrow Theorem shows that it is hard to avoid independence of irrelevant alternatives. But some of them seem like serious democratic shortcomings, the effects of which can be seen in American democracy, and especially in the extreme power the two major parties have. (Though, to be fair, a number of other electoral systems that use plurality voting do not have such strong major parties. Indeed, Canada seems to have very strong third parties despite using this system, as does the United Kingdom.)

One clear advantage of plurality voting should be stressed: it is quick and easy. There is little chance that voters will not understand what they have to do in order to express their preferences. And voting is, or at least should be, relatively quick. The voter just has to make one mark on a piece of paper, or press a single button, to vote. When the voter is expected to vote for dozens of offices, as is usual in America (though not elsewhere) this is a serious benefit. In many U.S. elections, we have seen queues hours long of people waiting to vote. Were voting any slower than it actually is, these queues might have been worse.

Relatedly, it is easy to count the votes in a plurality system. You just sort all the votes into different bundles and count the size of each bundle. Some of the other systems we'll be looking at are much harder to count the votes in. This can lead to serious delays in even being able to announce the results of an election.

<sup>2</sup>This isn't actually that unusual in the overall scope of American elections, especially in primaries.

## Runoff Voting

One solution to some of the problems with plurality voting is runoff voting, which is used in parts of America (including Georgia and Louisiana and, in practice, California) and is very common throughout Europe and South America. The most notable upcoming election using this system is the Presidential election in France.

The idea is that there are, in general, two elections. At the first election, if one candidate has majority support, then they win. But otherwise the top two candidates go into a runoff. In the runoff, voters get to vote for one of those two candidates, and the candidate with the most votes wins.

This doesn't entirely deal with the problem of a spoiler candidate having an outsized effect on the election, but it makes such cases a little harder to produce. For instance, imagine that there are four candidates, and the arrangement of votes is as follows.

35%	30%	20%	15%
<i>A</i>	<i>B</i>	<i>C</i>	<i>D</i>
<i>B</i>	<i>D</i>	<i>D</i>	<i>C</i>
<i>C</i>	<i>C</i>	<i>B</i>	<i>B</i>
<i>D</i>	<i>A</i>	<i>A</i>	<i>A</i>

In a plurality election, *A* will win with only 35% of the vote.<sup>2</sup> In a runoff election, the runoff will be between *A* and *B*, and presumably *B* will win, since 65% of the voters prefer *B* to *A*. But look what happens if *D* drops out of the election, or all of *D*'s supporters decide to vote more strategically.

35%	30%	20%	15%
<i>A</i>	<i>B</i>	<i>C</i>	<i>C</i>
<i>B</i>	<i>C</i>	<i>B</i>	<i>B</i>
<i>C</i>	<i>A</i>	<i>A</i>	<i>A</i>

Now the runoff is between *C* and *A*, and *C* will win. *D* being a candidate means that the candidate most like *D*, namely *C*, loses a race they could have won.

In one respect this is much like what happens with plurality voting. On the other hand, it is somewhat harder to find real life cases that show this pattern of votes. That's in part because it is hard to find cases where there are (a) four serious candidates, and (b) the third and fourth candidates are so close ideologically that they eat into each other's votes and (c) the top two candidates are so close that these third and fourth candidates combined could leapfrog over each of them. Theoretically, the problem about spoiler candidates might look as severe, but it is much less of a problem in practice.

The downside of runoff voting of course is that it requires people to go and vote twice. This can be a major imposition on the time and energy of the voters. More seriously from a democratic perspective, it can lead to an unrepresentative electorate. In American runoff elections, the runoff typically has a much lower turnout than the initial election, so the election comes down to the true party loyalists. At least, this is true in the southern states that have traditional runoff elections. In the Californian system, where the primary becomes a de facto first round of voting it is the other way around. In Europe, the first round has sometimes had a very low turnout, with fringe candidates making it to the final round on the strength of a small but loyal supporter base.

### Instant Runoff Voting

One approach to this problem is to do, in effect, the initial election and the runoff at the same time. In instant runoff voting, every voter lists their preference ordering over their desired candidates. In practice, that means marking '1' beside their first choice candidate, '2' beside their second choice and so on through the candidates.

When the votes are being counted, the first thing that is done is to count how many first-place votes each candidate gets. If any candidate

has a majority of votes, they win. If not, the candidate with the lowest number of votes is eliminated. The vote counter then distributes each ballot for that eliminated candidate to whichever candidate receives the '2' vote on that ballot. If that leads to a candidate having a majority, that candidate wins. If not, the candidate with the lowest number of votes at this stage is eliminated, and their votes are distributed, each voter's vote going to their most preferred candidate of the remaining candidates. This continues until a candidate gets a majority of the votes.

This avoids the particular problem we discussed about runoff voting. In that case, *D* would have been eliminated at the first round, and *D*'s votes would all have flowed to *C*. That would have moved *C* about *B*, eliminating *B*. Then with *B*'s preferences, *C* would have won the election comfortably. But it doesn't remove all problems. In particular, it leads to an odd kind of strategic voting possibility. The following situation does arise, though rarely. Imagine the voters are split the following way.

45%	28%	27%
<i>A</i>	<i>B</i>	<i>C</i>
<i>B</i>	<i>A</i>	<i>B</i>
<i>C</i>	<i>C</i>	<i>A</i>

As things stand, *C* will be eliminated. And when *C* is eliminated, all of *C*'s votes will be transferred to *B*, leading to *B* winning. Now imagine that a few of *A*'s voters change the way they vote, voting for *C* instead of their preferred candidate *A*, so now the votes look like this.

43%	28%	27%	2%
<i>A</i>	<i>B</i>	<i>C</i>	<i>C</i>
<i>B</i>	<i>A</i>	<i>B</i>	<i>A</i>
<i>C</i>	<i>C</i>	<i>A</i>	<i>B</i>

Now *C* has more votes than *B*, so *B* will be eliminated. But *B*'s voters have *A* as their second choice, so now *A* will get all the new votes, and *A* will easily win. Some theorists think that this possibility for strategic voting is a sign that instant runoff voting is flawed.

Perhaps a more serious worry is that the voting and counting system is more complicated. This slows down voting itself, though this is a problem can be partially dealt with by having more resources dedicated to making it possible to vote. The vote count is also somewhat slower. A worse consequence is that because the voter has more to do, there is more chance for the voter to make a mistake. In some jurisdictions, if the voter does not put a number down for each candidate, their vote is invalid, even if it is clear which candidate they wish to vote for. It also requires the voter to have opinions about all the candidates running, and this may include a number of frivolous candidates. But it isn't clear that this is a major problem if it does seem worthwhile to avoid the problems with plurality and runoff voting.

The other systems we'll look at are not in wide use around the world.

### Borda Count

In a Borda Count election, each voter ranks each of the candidates, as in Instant Runoff Voting. Each candidate then receives  $n$  points for each first place vote they receive (where  $n$  is the number of candidates),  $n-1$  points for each second place vote, and so on through the last place candidate getting 1 point. The candidate with the most points wins.

One nice advantage of the Borda Count is that it eliminates the chance for the kind of strategic voting that exists in Instant Runoff Voting, or for that matter any kind of Runoff Voting. It can never make it more likely that  $A$  will win by someone changing their vote away from  $A$ . Indeed, this could only lead to  $A$  having fewer votes. This certainly seems to be reasonable.

Another advantage is that many preferences beyond first place votes count. A candidate who is every single voter's second best choice will not do very well under any voting system that gives a special weight to first preferences. But such a candidate may well be in a certain sense the best representative of the whole community.

And a third advantage is that the Borda Count includes a rough ap-

proximation of voter's strength of preference. If one voter ranks  $A$  a little above  $B$ , and another votes  $B$  many places above  $A$ , that's arguably a sign that  $B$  is a better representative of the two of them than  $A$ . Although only one of the two prefers  $B$ , one voter will be a little disappointed that  $B$  wins, while the other would be very disappointed if  $B$  lost.

These are not trivial advantages. But there are also many disadvantages which explain why no major electoral system has adopted Borda Count voting yet, despite its strong support from some theorists.

First, Borda Count is particularly complicated to implement. It is just as difficult for the voter to as in Instant Runoff Voting; in each case they have to express a complete preference ordering. But it is much harder to count, because the vote counter has to detect quite a bit of information from each ballot. Getting this information from millions of ballots is not a trivial exercise.

Second, Borda Count has a serious problem with 'clone candidates'. In plurality voting, a candidate suffers if there is another candidate much like them on the ballot. In Borda Count, a candidate can seriously gain if such a candidate is added. Consider the following situation. In a certain electorate, of say 100,000 voters, 60% of the voters are Republicans, and 40% are Democrats. But there is only one Republican, call them  $R$ , on the ballot, and there are 2 Democrats,  $D1$  and  $D2$  on the ballot. Moreover,  $D2$  is clearly a worse candidate than  $D1$ , but the Democrats still prefer the Democrat to the Republican. Since the district is overwhelmingly Republican, intuitively the Republican should win. But let's work through what happens if 60,000 Republicans vote for  $R$ , then  $D1$ , then  $D2$ , and the 40,000 Democrats vote  $D1$  then  $D2$  then  $R$ . In that case,  $R$  will get  $60,000 \times 3 + 40,000 \times 1 = 220,000$  points,  $D1$  will get  $60,000 \times 2 + 40,000 \times 3 = 240,000$  points, and  $D2$  will get  $60,000 \times 1 + 40,000 \times 2 = 140,000$  points, and  $D1$  will win. Having a 'clone' on the ticket was enough to push  $D1$  over the top.

On the one hand, this may look a lot like the mirror image of the 'spoiler' problem for plurality voting. But in another respect it is much

worse. It is hard to get someone who is a lot ideologically like your opponent to run in order to improve your electoral chances. It is much easier to convince someone who already wants you to win to add their name to the ballot in order to improve your chances. In practice, this would either lead to an arms race between the two parties, each trying to get the most names onto the ballot, or very restrictive (and hence undemocratic) rules about who was even allowed to be on the ballot, or, most likely, both.

The third problem comes from thinking through the previous problem from the point of view of a Republican voter. If the Republican voters realise what is up, they might vote tactically for *D2* over *D1*, putting *R* back on top. In a case where the electorate is as partisan as in this case, this might just work. But this means that Borda Count is just as susceptible to tactical voting as other systems; it is just that the tactical voting often occurs downticket. (There are more complicated problems, that we won't work through, about what happens if the voters mistakenly judge what is likely to happen in the election, and tactical voting backfires.)

Finally, it's worth thinking about whether the supposed major virtue of Borda Count, the fact that it considers all preferences and not just first choices, is a real gain. The core idea behind Borda Count is that all preferences should count equally. So the difference between first place and second place in a voter's affections counts just as much as the difference between third and fourth. But for many elections, this isn't how the voters themselves feel. I suspect many people reading this have strong feelings about who was the best candidate in the past Presidential election. I suspect very few people had strong feelings about who was the third best versus fourth best candidate. This is hardly a coincidence; people identify with a party that is their first choice. They say, "I'm a Democrat" or "I'm a Green" or "I'm a Republican". They don't identify with their third versus fourth preference. Perhaps voting systems that give primary weight to first place preferences are genuinely reflecting the desires of the voters.

---

<sup>3</sup>There is one circumstance when something like Approval Voting is used in the United States. It is not uncommon when there are  $n > 1$  candidates to be elected to a position, like school board members, that each voter can select  $n$  candidates, and the top  $n$  vote getters are elected. This isn't true Approval Voting, since there is still a cap on how many people you can vote for.

## Approval Voting

In plurality voting, every voter gets to vote for one candidate, and the candidate with the most votes wins. Approval voting is similar, except that each voter is allowed to vote for as many candidates as they like. The votes are then added up, and the candidate with the most votes wins. Of course, the voter has an interest in not voting for too many candidates. If they vote for all of the candidates, this won't advantage any candidate; they may as well have voted for no candidates at all.

The voters who are best served by approval voting, at least compared to plurality voting, are those voters who wish to vote for a non-major candidate, but who also have a preference between the two major candidates. Under approval voting, they can vote for the minor candidate that they most favor, and also vote for the the major candidate who they hope will win. Of course, runoff voting (and Instant Runoff Voting) also allow these voters to express a similar preference. Indeed, the runoff systems allow the voters to express not only two preferences, but express the order in which they hold those preferences. Under approval voting, the voter only gets to vote for more than one candidate, they don't get to express any ranking of those candidates.

But arguably approval voting is easier on the voter. The voter can use a ballot that looks just like the ballot used in plurality voting. And they don't have to learn about preference flows, or Borda Counts, to understand what is going on in the voting. Currently there are many voters who vote for, or at least appear to try to vote for, multiple candidates. This is presumably inadvertent, but approval voting would let these votes be counted, which would refranchise a number of voters. Approval voting has never been used as a mass electoral tool, so it is hard to know how quick it would be to count, but presumably it would not be incredibly difficult.<sup>3</sup>

One striking thing about approval voting is that it is not a function

from voter preferences to group preferences. Hence it is not subject to the Arrow Impossibility Theorem. It isn't such a function because the voters have to not only rank the candidates, they have to decide where on their ranking they will 'draw the line' between candidates that they will vote for, and candidates that they will not vote for. Consider the following two sets of voters. In each case candidates are listed from first preference to last preference, with stars indicating which candidates the voters vote for.

40%	35%	25%	40%	35%	25%
* <i>A</i>	* <i>B</i>	* <i>C</i>	* <i>A</i>	* <i>B</i>	* <i>C</i>
<i>B</i>	<i>A</i>	<i>B</i>	<i>B</i>	<i>A</i>	* <i>B</i>
<i>C</i>	<i>C</i>	<i>A</i>	<i>C</i>	<i>C</i>	<i>A</i>

In the election on the left-hand-side, no voter takes advantage of approval voting to vote for more than one candidate. So *A* wins with 40% of the vote. In the election on the right-hand-side, no one's preferences change. But the 25% who prefer *C* also decide to vote for *B*. So now *B* has 60% of the voters voting for them, as compared to 40% for *A* and 25% for *C*, so *B* wins.

This means that the voting system is not a function from voter preferences to group preferences. If it were a function, fixing the group preferences would fix who wins. But in this case, without a single voter changing their preference ordering of the candidates, a different candidate won. Since the Arrow Impossibility Theorem only applies to functions from voter preferences to group preferences, it does not apply to Approval Voting.

### Range Voting

In Range Voting, every voter gives each candidate a score. Let's say that score is from 0 to 10. The name 'Range' comes from the range of options

the voter has. In the vote count, the score that each candidate receives from each voter is added up, and the candidate with the most points wins.

In principle, this is a way for voters to express very detailed opinions about each of the candidates. They don't merely rank the candidates, they measure how much better each candidate is than all the other candidates. And this information is then used to form an overall ranking of the various candidates.

In practice, it isn't so clear this would be effective. Imagine that a voter *V* thinks that candidate *A* would be reasonably good, and candidate *B* would be merely OK, and that no other candidates have a serious chance of winning. If *V* was genuinely expressing their opinions, they might think that *A* deserves an 8 out of 10, and *B* deserves a 5 out of 10. But *V* wants *A* to win, since *V* thinks *A* is the better candidate. And *V* knows that what will make the biggest improvement in *A*'s chances is if they score *A* a 10 out of 10, and *B* a 0 out of 10. That will give *A* a 10 point advantage, whereas they may only get a 3 point advantage if the voter voted sincerely.

It isn't unusual for a voter to find themselves in *V*'s position. So we might suspect that although Range Voting will give the voters quite a lot of flexibility, and give them the chance to express detailed opinions, it isn't clear how often it would be in a voter's interests to use these options.

And Range Voting is quite complex, both from the perspective of the voter and of the vote counter. There is a lot of information to be gleaned from each ballot in Range Voting. This means the voter has to go to a lot of work to fill out the ballot, and the vote counter has to do a lot of work to process all that information. This means that Range Voting might be very slow, both in terms of voting and counting. And if voters have a tactical reason for not wanting to fill in detailed ballots, this might mean it's a lot of effort for not a lot of reward, and that we should stick to somewhat simpler vote counting methods.