VOTING AND GROUP CHOICE

DYLAN ZWICK

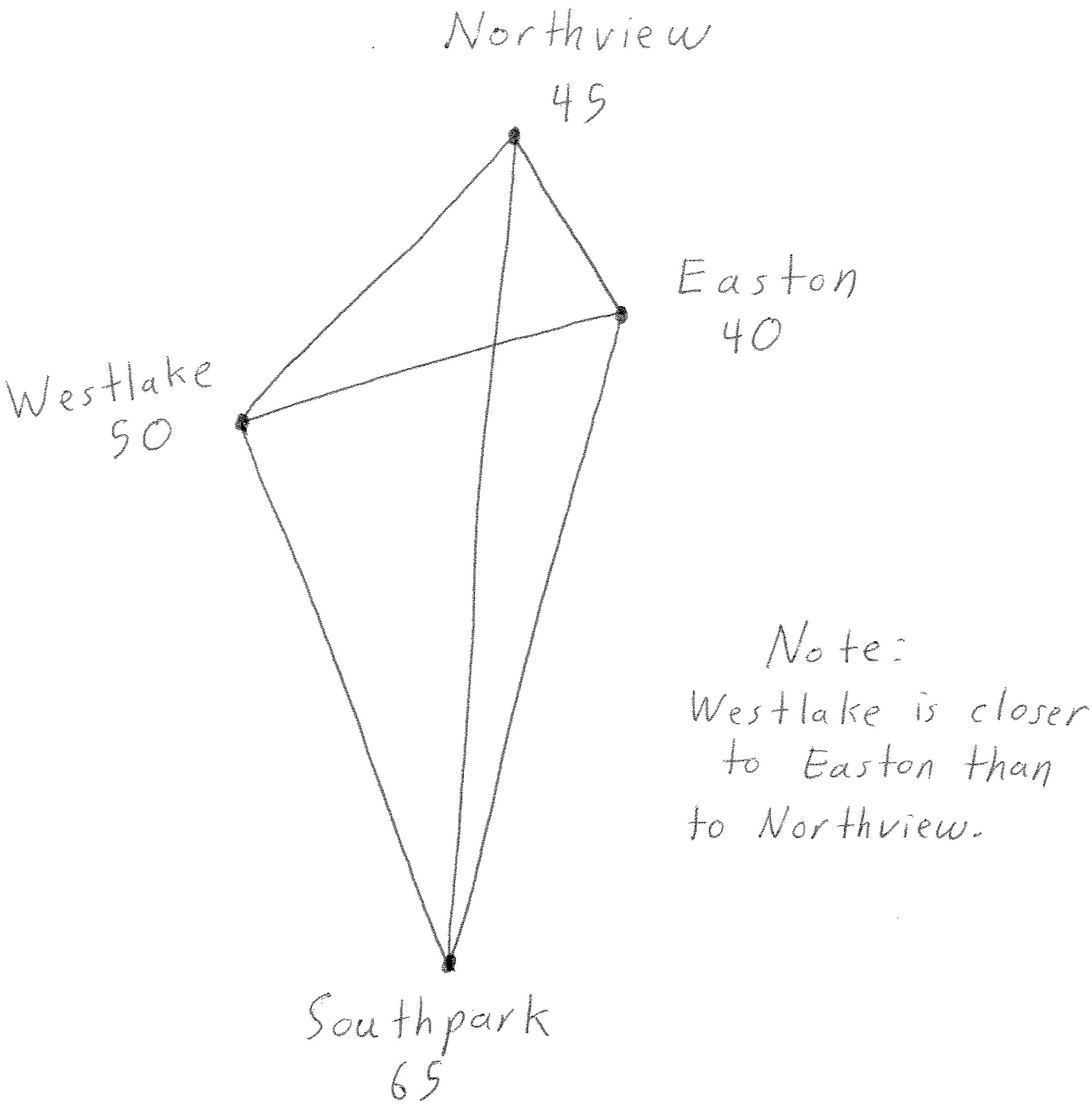
1. Overview

For a mathematician, voting seems like a ridiculous and illogical way to decide upon anything. We don't vote on whether or not the Riemann Hypothesis is true, we prove it or disprove it logically. There is true and false, and that's it. However, when people get involved, things aren't always that nice. In fact, things are very rarely that nice. That's where voting comes in to play as a means of determining what society should do based upon what its people want. Now, voting seems straightforward enough, you just ask everybody what they want, and the majority wins. What could be more straightforward? Well, it's true that things are indeed this easy when dealing with only two options, but when we complicate matters by introducing other options, even if we increase to only three options, things can become very messy indeed. In this lecture we'll discuss some conceptions of voting, their advantages and limitations, and some of the underlying theorems that govern all voting systems. Then, if we have time, we'll mention some of the issues and controversies surrounding voting in the United States, focusing on Presidential voting, and some ideas concerning these issues.

2. The Problem

Let's suppose we have a county with four towns: Northview, Easton, Westlake, and Southpark. The towns are arranged on a map like we have below:

The map


Northview
45
Easton
40
Westlake
50
Southpark
65
Note:
Westlake is closer to Easton than to Northview.

and we want to figure out where to build a hospital. Now, each town wants the hospital to be as close to it as possible, and we're going to say that everybody in each town votes the same way, namely, for whatever option will bring the hospital closer to him or her. In this situation, where do we build the hospital? Well, it depends upon which voting system we use. As we'll see, *all* of the towns are possible in different voting systems, and each of these systems appear reasonable and fair at first blush.

3. Voting Systems

So, how do we figure out where we build the hospital? Well, let's look at a few different voting systems. We'll take a look at majority vote, the Condorcet method, the Borda count, and instant run-off elections. We'll see that each of these methods gives us a different result.

3.1. **Majority Vote.** This is the easiest one to figure out, and it's the one that most resembles reflects the way we vote in the United States. Under this system, each town would vote for itself, and the town with the highest population, namely Southpark, would get the hospital. This seems reasonable, but is it really that desirable? Southpark is the farthest town away from Northview, Westlake, and Easton. It's the least preferred option of 135 people. In other words, more than twice the number of people who put Southpark first put it last. Is this really the optimal way of deciding things?

We do note that if there are only two options then majority vote is definitely the way to go, as the option with the most votes will win, which is (or at least should be) the desired outcome here.

3.2. **The Condorcet Method.** This method is named after the 18th century philosopher and mathematician the Marquis de Condorecet, who wrote about it. This method is a head-to-head method that pits every contestant against every other in an election, and then bases the decision on the one that wins all the contests. In other words, the option that is preferred in a head-to-head match against all other options is considered the best. This is the way a champion is determined in most sports leagues. The exception is, as University of Utah fans know all so well, the ridiculous Bowl system in college football.

Using the Condorcet method for out towns we get the following table:

Town 1	**Votes**	**Town 2**	**Votes**	***Winner***
Northview	85	Westlake	115	*Westlake*
Northview	45	Easton	155	*Easton*
Northview	135	Southpark	65	*Northview*
Westlake	115	Easton	85	*Westlake*
Westlake	135	Southpark	65	*Westlake*
Easton	135	Southpark	65	*Easton*

Now, for our towns Southpark loses when it goes up against *any* other town. In fact, in our town competition Westlake is the town that wins every head-to-head against all other towns. So, Westlake would be the winner and the town in which we'd build the hospital according to the Condorecet method. Seems pretty reasonable to me, and given Westlake is situated towards the middle it seems like a good compromise between competing interests.

However, there's a flaw in this system, and it's a big one. Namely, it doesn't always produce a clear answer. Let's look at an example. Suppose we have a group of 9 people and they're deciding where to go for dinner. They can eat French, Italian, or Chinese. Suppose the order of preference breaks down as follows:

$$\begin{array}{l} \text{Chinese} < \text{Italian} < \text{French (4 people)} \\ \text{French} < \text{Chinese} < \text{Italian (3 people)} \\ \text{Italian} < \text{French} < \text{Chinese (2 people)} \end{array}$$

Well, what type of food wins? It depends upon the voting method. If we take a simple majority vote then these people go with French. On the other hand, if we use the Condorcet method, we run into a big problem. In the choice between French and Italian, French wins. In the choice between Italian and Chinese, Italian wins. In the choice between French and Chinese, well, Chinese wins! So, you've got a situation where French beats Italian, Italian beats Chinese, and Chinese beats French. It's a rock-scissors-paper type situation. In other words, the Condorcet method is not a *transitive* method[1] for determining group choice. There isn't always a winner, and this can become a big problem if people start voting strategically instead of honestly.

3.3. **The Borda Count.** So, how can we get around this problem of a lack of transitivity? Well, there's another method introduced by yet another Frenchman called the Borda count. This technique is based upon the idea that a candidate that is liked by a lot of people is a good

[1] If A beats B, and B beats C, then A beats C.

choice. The idea is that each person takes all their candidates and ranks them from least favored to most favored. Then, according to his or her preferences, a person awards points. A person's least favored option is awarded 0 points by him or her, the second least favored is awarded one point, and so on. Each person does this, and we then take the sum of the points awarded the various candidates. The candidate with the most points then wins. For our situation this would give us the following table of votes, with the following total sums:

Candidate	Easton	Westlake	Northview	Southpark	**Total**
Easton	120	100	90	65	375
Westlake	40	150	45	130	365
Northview	80	50	135	0	265
Southpark	0	0	0	195	195

So, under the Borda count the most favored town is Easton, and so Easton gets the hospital. The nice thing about the Borda count is that it's transitive. The potentially bad thing is illustrated by this example. Suppose that Northview knows that it isn't going to win, and so drops out of the race. The people of Northview still vote, but the town of Northview is no longer a candidate. In this situation the votes are then between three towns. We'd figure that because Easton won while Northview was in the competition, Easton should still win now that Northview is out of the competition. That is, the most favored choice shouldn't depend upon what are called "irrelevant alternatives." However, if we run the contest again with Northview out of the competition, here's the table we get:

Candidate	Easton	Westlake	Northview	Southpark	**Total**
Easton	80	50	90	0	220
Westlake	40	100	45	65	250
Northview	0	0	0	0	0
Southpark	0	0	0	130	130

That's right, Westlake beats Easton. This is kind of like if I asked you where you wanted to go on vacation: London, Paris, or Rome, and you said London. I then go to book the tickets and I say "Well, that was lucky. It turns out we couldn't go to Rome anyways." to which you reply "Oh, well then I want to go to Paris." It's nonsensical. However, this problem of irrelevant alternatives turns out to be particularly pernicious and difficult[2] to eliminate.

[2]In fact, as we'll see, essentially impossible.

3.4. **Instant Run-Off Election.** We'll take a look at one more voting system. This system is called an instant run-off system, and it's based upon a sequence of ballots. On the first ballot, all the candidates are an option, and people vote for their most favored candidate. The candidate who receives the least votes is dropped, and then for the second ballot people vote again, with one less option. This continues until all candidates but one have dropped out, and the last candidate standing is the winner. How would this system play out for our scenario?

Well, on the first ballot we'd have:

Ballot 1	Easton	Westlake	Northview	Southpark
Votes	40	50	45	65

and so Easton would drop out. On the second ballot we'd have:

Ballot 2	Westlake	Northview	Southpark
Votes	50	85	65

and so Westlake would drop out. One the third and final ballot we'd have:

Ballot 3	Northview	Southpark
Votes	135	65

and so Southpark would drop out, leaving Northview as the winner.

So, let's pause for a second to think about what we've done so far. We've examined four different voting systems, each of which, at first glance, seems to be a fair and reasonable way of making a decision. However, *all four* of our possible options can be the winner, depending upon which of these voting systems we use. Is there an optimal voting system that gets rid of these problems? We'll examine that question soon, but first let's look a little more closely at this problem of irrelevant alternatives.

3.5. **Irrelevant Alternatives.** The problem of irrelevant alternatives turns out to be incredibly difficult to eliminate, and in fact it's a problem in our own Presidential elections. William Poundstone in his book *Gaming the Vote* makes a credible case that in no less than five of our Presidential elections: 1844, 1848, 1884, 1912, and 2000, a third party spoiler candidate led to the election of a President when more voters would have preferred one of the alternatives. In two other elections: 1892 and 1992 it's debatable. And in fact, in 1860 the confusion concerning *four* major national candidates and the attendant electoral college difficulties created such a lack of legitimacy that it is considered one of the causes of our Civil War. Lincoln, it should be noted, won that election, so I don't mean to say that divided elections always produce suboptimal outcomes, but that's a different discussion.

In each of the five cases where a spoiler affected the outcome of an election the end result was that the viable candidate that would be the least favored by the spoiler actually took office. An abolitionist caused the election of a slaveholder, a prohibitionist caused the election of the candidate most friendly to the saloons, a former Republican caused the election of a Democrat, and a consumer advocate caused the election of the candidate most favored by large corporations. Is there a better way? Perhaps, but as we'll see shortly getting rid of this problem turns out to be very, very difficult, and many would argue impossible.

4. Arrow's Theorem

This is a math discussion, so let's introduce some mathematical ideas here. We're going to define an election by letting a set A be the set of outcomes (you can view this as the set of candidates for a particular office), and letting the number N be the number of voters. We shall denote the set of all full linear orders of A by $L(A)$, and any given linear ordering can represent a list, in order of preference, of the outcome of any individual. This set is equivalent to the set $S_{|A|}$ of permutations of the outcome set A.

We define a strict social welfare function as a function:

$$F : L(A)^N \to L(A)$$

as a map from the list of preferences of all the voters to the list of preferences of the "society". One element of the list of the preferences of all the individuals, a set of lists $(R_1, \ldots, R_N)$, is called a "preference profile". Now, Arrow's Theorem begins with the assumption that we want for this social welfare function to satisfy three criteria:

Transitivity: - This requirement is incorporated in our definition of a social welfare function, but it's worth pointing out in that the Condorcet method violated this requirement, and so would not qualify as a social welfare function. Transitivity simply states that if society prefers A to B, and B to C, then society must prefer A to C.

Independence of Irrelevant Alternatives: - This requirement can be stated formally that for two preference profiles $(R_1, \ldots, R_N)$

and $(S_1, \ldots, S_N)$ if the relative position (ahead or behind) of a and b is the same in R_i as it is for S_i for all i, then the relative position of a and b in $F(R_1, \ldots, R_N)$ should be the same as in $F(S_1, \ldots, S_N)$.

Pareto Efficiency: - Pareto efficiency is an almost common sense requirement that if outcome A is preferred by *everybody* to outcome B then society should prefer outcome A to outcome B. In fact, this requirement can even be weakened to requiring that if one outcome is the most desired outcome of *everybody* in society, then it should be the top outcome from that society. This, for example, eliminates a completely randon social welfare function.

Now, it turns out that there is indeed an excellent voting method that satisfies all of these requirements, and it's one that I personally feel all societies should adapt immediately. This method is that everybody votes, and then at the end of the day we throw away all the votes except mine, and we do what I say! Yes, that's right, a dictatorship. There is no other solution. If we add another requirement to our function, the "no dictatorship" requirement:

No Dictatorship - These is no individual i whose preference always prevails. That is, there is no $i \in \{1, \ldots, N\}$ such that $F(R_1, \ldots, R_N) = R_i$ no matter what the other R_j do.

then Arrow's Impossibility Theorem states that no such function exists! The proof is actually quite ingenious.

4.1. **Informal Proof. First Part**

Say there are three choices for society, call them A, B, and C. Suppose first that everyone prefers option B the least. That is, *everyone* prefers *every* other option to B. By Pareto efficiency, this means that society must prefer every option to B. Call this situation *Profile 1.*

On the other hand, if everyone preferred B to everything else, then society would have to prefer B to everything else again by Pareto efficiency. So, what we do is that we take Profile 1 and, running through the members in the society in some arbitrary but specific order, move B from the bottom of each person's preference list to the top. At some point in this process B must move from the bottom of society's preference as well, since we know that if we switch B to the top of everybody's preferences then B must be at the top of society's preference as well. No problem.

Now, here's the interesting part. During this process, at the point when the pivotal voter n moves B off the bottom of her preferences to the top, the society's B moves to *the top(!)* of its preferences as well, not to an intermediate point.

To prove this, consider what would happen if it were not true. Then society would have some option it prefers to B, say A, and one less preferable than B, say C. (We can rename our options if we must.)

Now if each person moves his preference for C above A, then society would prefer C to A again by Pareto efficiency. By the fact that A is already preferred to B, C would now be preferred to B as well in the rankings. But, and here's where we see just how powerful this independence of irrelevant alternatives requirement is, moving C above A shouldn't change anything about how B and C compare. That is, since B is either at the very top or bottom of each person's preferences, moving C or A around doesn't change how either compares with B. We have a contradiction, and so it must be that B moves from the bottom to the top when n switches.

Second Part

Now, we show that this voter n in fact must be a dictator. Call the case with all voters *up to* n having B at the top of their preference and the rest with B at the bottom *Profile 2*. Call the case with all voters *up through* n having B on the top and the rest having B on the bottom *Profile 3*.

Now suppose everyone up to n ranked B at the bottom, n ranks B below A but above C, and everyone else ranks B at the top. As far as the A-B decision is concerned, this organization is just as in Profile 2, which we proved puts B below A. The new position of C in voter n's list is irrelevant to the A-B order. Likewise, the relation between B and C in n's new order is just as in Profile 3, and so B is above C. So, transitivity tells us that society must put A above B above C.

Now, and here again we see just how powerful independence of irrelevant alternatives is, the position of B must be *irrelevant* to the decision society makes between A and C, and so the fact that we assume particular profiles for B doesn't matter as far as the A-C decision is concerned. So, what does this tell us. It tells us that voter n is the dictator over the A-C decision. In other words, it *doesn't matter* what any of the other voters do. The only decision that matters, as far as whether A is ahead of C or not, is what voter n decides.

Finally, we note that there was nothing special about our choice of B for this proof, the it would work just as well with A or C replacing B, and so we know there must be a dictator for the A-B decision and the B-C decision as well. Now, all three of these dictators must in fact be the same voter, as any two can overrule the third. That is, if the A-C dictator chooses A over C, and the B-C dictator chooses C over B, then transitivity says we must have A over B, meaning that that A-B dictator is overriden, a contradiction unless all dictators are in fact the same voter.

5. How We Vote

In the United States our Presidential voting system I think can fairly be described as, if not broken, then at least Baroque. Just so we're all on the same page, here's how it works. The President is elected by a group of special voters known as the *electoral college*. Each member of the electoral college represents either a state or the District of Columbia. The number of representatives a state receives is equal to the number of representatives it has in the United States House, plus two more representing the number of Senators. The District of Columbia gets the number of electors given to the smallest state. So, for example, Utah has three representatives and, of course, two senators, and so we have five representatives in the electoral college.

How these electors are chosen is left entirely up to the state. Almost every state follows a "winner takes all" allocation strategy, in which the winner of the state's popular vote, no matter how close that vote is, gets every elector. Two states, Maine and Nebraska, follow a strategy in which the winner of the popular vote in each congressional district gets (effectively) that district's vote, while the winner of the statewide popular vote wins (effectively) the two votes provided by the senators. So, for example, in our most recent Presidential election, John McCain received 4 electoral votes from Nebraska, while Barack Obama picked up one from Omaha.

Now, this system is not without its controversies. I'll highlight the three major ones. First, given how most states choose to set up how they allocate delegates, during Presidential elections the candidates focus almost all of their time and energy on what are called "swing states", that is, states with almost evenly politically divided populations. As a result, for example, the three largest states of California, Texas, and New York receive very little attention even though they have huge populations, because their respective populations are more politically homogeneous. On the other hand, Nevada, with its relatively small population, receives an exceptional amount of attention because it's so evenly politically divided.

Second, the "power" of a person's vote, if measured by the number of voters a given member of the electoral college represents, varies from state to state. Citizens of smaller states, by virtue of each state having the same number of senators, have more "power" than citizens of larger states.

Third, and perhaps most striking, is that the winner of the national popular vote doesn't necessarily win the election. In fact you could have an (admittedly ridiculous) situation in which only one person in the 15 most populous states voted for candidate A, while everybody in the other 35 states voted for candidate B, and candidate A would win, even though the popular vote was 15 to over 100 million! Of course that doesn't happen, but it has definitely happened twice in our history, in 1876 and 2000, that the unquestioned winner of the popular vote lost the election.

6. Election Math

I'll end with a discussion of two ideas, one mathematical and the other just an interesting thought that might appeal to those who like mathematical thinking.

The first concerns a different measure of voter power, and what this measure implies about its relation to the size of a population. It's based upon the premise that in a winner takes all election, an individual voter almost always has no real power at all, in that however she votes won't affect the outcome. For example, if there are 10 people voting, and 7 vote for candidate A while 2 vote for candidate B, then the vote of the tenth person is irrelevant. The election is already decided, and so we may argue that the voter has no actual power. So the question then is what is the probability that your vote will matter, in that it will determine the outcome of the election, and how does that probability depend upon population?

Well, let's simplify the situation a little bit and more precisely define our terms. Suppose you're a voter in a state with population $2n + 1$, and everybody votes. The election is close, and so we'll model the way any individual except you votes by a fair coin toss. If the other $2n$ voters all toss a coin then your vote will only matter is there are an equal number of heads and tails. The probability that for $2n$ coin tosses we'll get n heads is:

$$\frac{(2n)!}{n!n!2^n}$$

So, to approximate this value we can make use of Sterling's approximation, which states that for large n:

$$n! \approx n^{n+1/2}e^{-n}\sqrt{2\pi}$$

and so our probability is approximately:

$$\frac{(2n)!}{n!n!2^n} \approx \frac{(2n)^{2n+1/2}e^{-2n}\sqrt{2\pi}}{k^{2n+1}e^{-2n}(2\pi)2^{2n}} = \frac{1}{\sqrt{n\pi}}$$

So, the probability that your vote will matter, in the sense described here, is inversely proportional to the square root of the population. Now, if the power of your state is directly proportional to the population, then the number of electors you get is ck for some constant c, then one could argue that the average power of a vote is:

$$\frac{nc}{\sqrt{n\pi/2}} = c\sqrt{2n/\pi}$$

So, under this analysis, it is in fact the states with a higher population that have the most power! This just shows that there are many ways of looking at the same problem.

Finally, I'll just end with an idea. There has been a movement recently to do away with the electoral college by creating a de facto popular vote. The idea is this; the larger states (or at least those large states that are not swing states) all get together and decide that they're going to appropriate their electoral votes based upon whichever candidate wins the popular vote. If enough states sign on to this idea, then we have a de facto popular vote. Right now a number of states are debating enacting legislation that says they will put such a policy in play whenever enough states sign on so as to make it effective, and a few states have adopted such laws. The merits of such an approach are open to debate, but from a mathamatical perspective I think the idea is quite ingenious.