PARTISAN METRICS: NOTES

MATH 19, MATH OF SOCIAL CHOICE
SPRING 2018

This unit is about metrics of partisan gerrymandering, especially two kinds: *partisan symmetry* and the *efficiency gap*.

Let's standardize some notation. Suppose that a state has districts numbered 1 through m (i.e., the state has m seats in the House of Representatives). Then an election is conducted. Simplifying assumptions: there are only two parties, R and D; and each district has the same number of voters. Suppose the Republican vote share in each district is $V_1, \ldots, V_m$.

- The R vote share in the whole state is the mean (or average) of these numbers, denoted $\overline{V} = \frac{1}{m}\sum_{i=1}^{m} V_i$.
- The R vote share in the median district, denoted V_{med}, is the middle value of the V_i if you listed them in order.
- The number of districts with R vote shares above average is denoted $m_{\uparrow}$.

(Note: vote shares always range from 0 to 1, since they're just a proportion of the total.) Let's also write $\overline{S}$ for the Republican seat share, which is their fraction of the m House seats.

When an election has been conducted, you can write

$$(V_1, V_2, \ldots, V_m)$$

to record the R vote share in each district.

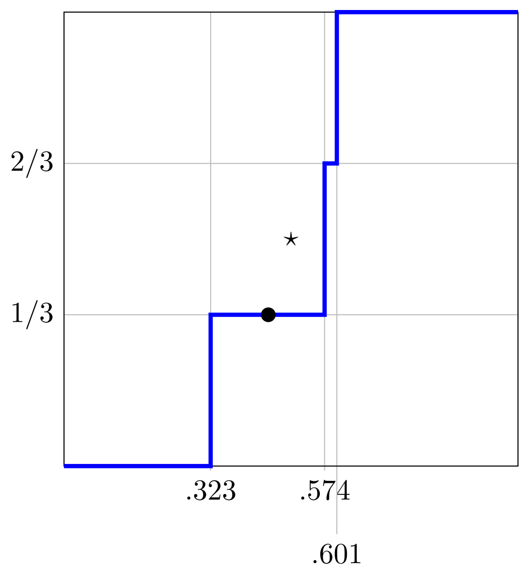
Many political scientists use the hypothesis of *uniform partisan swing* to study elections. That means that you consider what would happen if you add or subtract the same amount from the vote share in each district. You can generate the *seats-votes curve* (or SV curve for short) by varying the amount you adjust the votes by and seeing how that changes S.

Partisan symmetry scores

The mean-median score MM of an election is defined as the horizontal distance of the SV curve from the center point $\star = (.5, .5)$. The partisan bias score PB is the vertical distance of that curve from the center point $\star$.

Let's do an example. In 2016, New Mexico's congressional results were $(.349, .627, .376)$, from a Republican point of view. That means Rs got a vote share of $\overline{V} = (.349 + .627 + .376)/3 \approx .45$ and a seat share of one out of three seats, or $\overline{S} = 1/3$. The heavy dot in the plot marks the outcome of this election: $(\overline{V}, \overline{S}) = (.45, .33)$.

Here is the seats-votes curve generated by uniform partisan swing:



So now we can see that this isn't perfectly symmetric about the center point $(.5, .5)$ (which is marked with a $\star$), because in particular the curve doesn't go through that point! We define the *mean-median score* to be the horizontal distance of the curve from the center point, and the *partisan bias* to be the vertical distance of the curve from the center point. In this case, the horizontal distance goes from the star to the curve, which passes it at the value $V = .574$. So the distance from .5 to .574 gives us a mean-median score of $MM = .074$. This means that Republicans would have needed 57.4% of the vote in order to get half of the representation, which is a gap of 7.4 percentage points in favor of Democrats.

How about the vertical distance? That one is easy: the center point is at height $1/2$ and the curve passes below it at height $1/3$, so the difference is $PB = 1/2 - 1/3 = 1/6 \approx .167$. The interpretation is that with half of the vote, Rs would be receiving only $1/3$ of the representation, and PB records the gap in those numbers.

Theorem 1. *Recall that $\overline{V}$ denotes the mean R vote share, $V_{\rm med}$ denotes the median, and $m_\uparrow$ denotes the number of seats whose R vote share is greater than average. Under uniform partisan swing, the SV curve goes through the points*

$$\left(\frac{1}{2} + \overline{V} - V_{\rm med}\ ,\quad \frac{1}{2}\right) \qquad \text{and} \qquad \left(\frac{1}{2}\ ,\ \frac{m_\uparrow}{m}\right).$$

Proof. The middle seat for Republicans is the one with vote share $V_{\rm med}$. We want to find the level at which the control of that seat flips one way or the other. So we have to add just enough to move that vote share to $1/2$. To move $V_{\rm med}$ to $1/2$, we must add $1/2 - V_{\rm med}$. Since we're assuming uniform partisan swing, we add that to all districts, and the statewide vote average shifts by the same number of percentage points. Since the statewide vote share was $\overline{V}$, it becomes $V = \overline{V} + \frac{1}{2} - V_{\rm med}$ after shifting, which will cause us to arrive at split control $S = \frac{1}{2}$. That means that $\left(\frac{1}{2} + \overline{V} - V_{\rm med}\ ,\quad \frac{1}{2}\right)$ is on the curve.

For the other part, we ask ourselves how both parties would fare if each one got half of the vote. Since Republicans initially got $\overline{V}$ share of the vote, we need to shift that by adding $\frac{1}{2} - \overline{V}$ in all districts. This produces vote levels of

$$(V_1 + \frac{1}{2} - \overline{V}, \ldots, V_m + \frac{1}{2} - \overline{V}).$$

How many of those numbers are greater than 1/2? The answer to that tells us how many R seats will be produced when $V = 1/2$. But note that $V_i + \frac{1}{2} - \overline{V} \geq \frac{1}{2}$ exactly if $V_i \geq \overline{V}$. So we need to simply count how many districts have vote share greater than $\overline{V}$, and that is what $m_\uparrow$ counts. This produces a seat share of $S = m_\uparrow/m$, as desired. This means that $\left(\frac{1}{2}\ ,\ \frac{m_\uparrow}{m}\right)$ is on the curve. □

Corollary 1. *Unless the SV curve has a step at $S = 1/2$ or $V = 1/2$, we have*

$$|MM| = \overline{V} - V_{\text{med}} \qquad \text{and} \qquad |PB| = \frac{m_\uparrow}{m} - \frac{1}{2}.$$

Efficiency gap

Let's say that a *wasted vote* is any winning vote in excess of half the votes in a district, or any losing vote at all. Then let's let W^R be the total votes wasted by Republicans statewide, and W^D be the votes wasted by Democrats. Then the efficiency gap is defined as $EG = \frac{W^R - W^D}{Tot}$, which is the difference in wasted votes, divided by the total turnout of the election.

Example: let's do the New Mexico 2016 election from above, with R vote shares $(.349, .627, .376)$. For instance, suppose that each district had $T = 1000$ voters. Then we'd get the following outcomes.

i	R votes	D votes	W_i^R	W_i^D
1	349	651	349	151
2	627	373	127	373
3	376	624	376	124
total	1352	1648	852	648

so we get $$EG = \frac{852 - 648}{3000} = .068.$$

The popularizers of efficiency gap, Nick Stephanopoulos and Eric McGhee, proposed that any election with $|EG| > .08$ is probably gerrymandered! So this one is skewed but not officially gerrymandered. It's positive, meaning that Republicans wasted more votes, so according to EG it's tilted in favor of Democrats.

Theorem 2. *Under the assumption that every district has equal turnout, simplification gives us*

$$EG = 2\overline{V} - \overline{S} - \frac{1}{2}.$$

Proof. Let's suppose every district has turnout T so that the total turnout in the election is $Tot = Tm$. Suppose Rs won k out of m districts, so that $\overline{S} = k/m$. That means that Ds won the other $m - k$ districts. In each district won by Ds, the Rs waste all their votes, which is $V_i T$ votes, while the Ds waste their excess votes, which is $((1 - V_i) - 1/2)T$ votes. In each district won by Rs, the Ds waste all their votes, which is $(1 - V_i)T$ votes, while the Rs waste their excess votes, which is $(V_i - 1/2)T$ votes. Adding these up we get total R wastage equal to $W^R = \overline{V}Tm - \frac{k}{2}T$, and D wastage of $-\overline{V}Tm + T\frac{k}{2} + T\frac{m}{2}$.

Note that if we add these together we get $W^R + W^D = \frac{1}{2}Tm$, which is correct—that's half the total number of votes!

But the efficiency gap is computed by subtracting:

$$EG = \frac{W^R - W^D}{Tot} = \frac{2\overline{V}Tm - kT - T\frac{m}{2}}{Tm} = 2\overline{V} - \frac{k}{m} - \frac{1}{2} = 2\overline{V} - \overline{S} - \frac{1}{2}. \quad \square$$

This is kind of surprising! It means that in the end, the efficiency gap doesn't really depend on what's happening in each district; it only depends on the statewide vote share vs the statewide seat share.