Quantifying Linguistic Polarization for Congressional Representatives Facing Primary Challengers: A Random Effects Logit Regression Approach

Kelly, Jack (School: Harrison High School)

Linguistic polarization is a change in frequency of policy references in a representative's lexicon. When incumbent representatives face challengers, they may change their speech patterns to distinguish themselves to voters. Since Sulkin (2009) found a strong correlation between how representatives speak and how they vote, linguistic polarization could lead to policy differences. This study used a random effects logit regression to see if progressive primary challengers linguistically polarize Democratic Congressional incumbents. Because Kamarck and Podkul (2018) showed a 151% percentage increase of progressive Democratic House candidates from 2016 to 2018, the 116th Congress was chosen. Twitter was used as the data source because of the brevity and frequency of tweets. A total of 601,304 tweets from 410 representatives were coded for progressive language using a dictionary based approach. The dictionary had 194 terms (identified from 14 websites) and 50,722 total tweets were coded as progressive. 30 progressive challengers were included since they received > 5% of the votes in a primary and a major progressive endorsement. A regression was run for 30 incumbents facing viable challengers that looked at percentage of progressive tweets from incumbents (DV) before and after the challenger was endorsed (IV). Results showed incumbents had a 27% decrease in progressive language after a challenger entered the race ($p<.05$), a decrease 3x greater than a control group of randomized Democratic incumbents. One implication was that progressive challengers might have a moderating effect on Democratic incumbents as they aim to establish a distinct electoral path to victory.

Awards Won:
Fourth Award of $500