

In this research poster we will explore where money is being spent in congressional races and what can total spending help us to predict? For this we will use a combination of election results data and total spending by campaign sourced from the FEC.

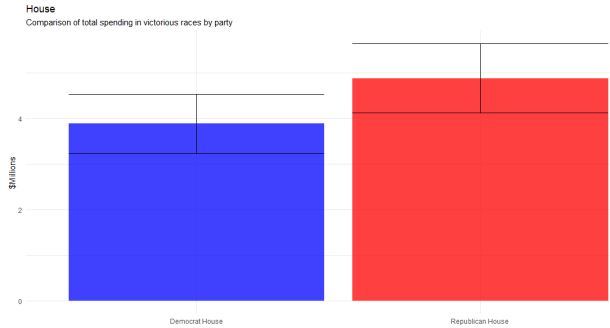
I am hypothesizing that most money will be spent in ‘swing states’ for Senate Races and coastal areas will have the more expensive races in the House elections.

I believe that higher total spending will be indicative of a closer race.

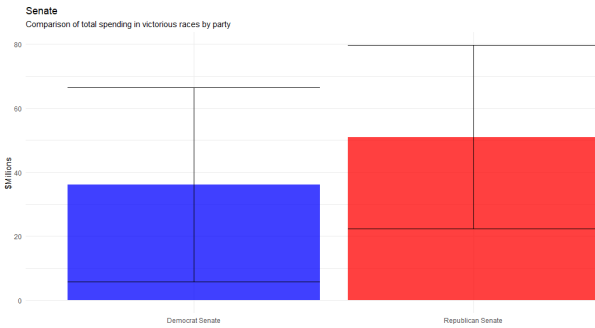
Finally as the Republican Party is often considered to be the party of the wealthy, my hypothesis is that races that Republicans win are likely to have higher total spending than that of Democrats.

The question of the impact of total spending is important because as the cost of elections increase, the dynamics of politics change. If money plays a large role in elections we risk having a political arena defined by those with money not by the will of the people.

SPENDING PATTERNS IN THE 2020 CONGRESSIONAL ELECTIONS

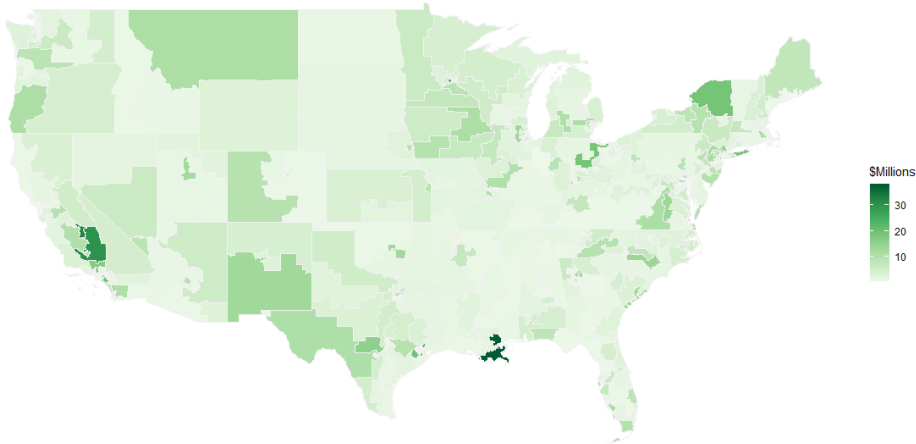


When observing the mean total spending (as the dependent variable) in races won by Democrats in comparison with Republicans we see that at first glance that races won by Republicans seem to be more expensive. However when constructing confidence intervals at the 95% confidence levels we see a small amount of overlap. If we perform a t test to see in the difference is statistically significant we observe a p value of 0.05021. This means there is only around a 5% of observing this difference due to a sampling error. As a result we can conclude with about 95% certainty (just under) that there is a statistically significant difference in mean total spending between races won by Democrats and Republicans.



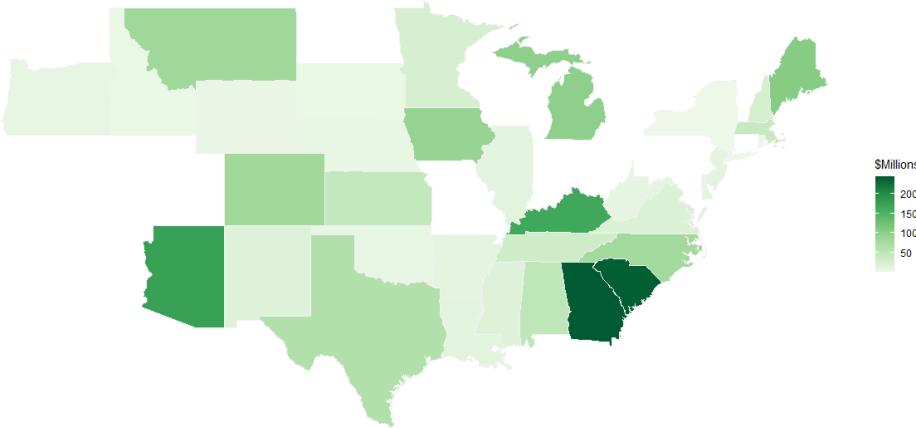
The first striking thing about this graphic is obviously the huge error bars. This owes to the unfortunately small sample size due to the nature of the Senate. There aren't many seats and thus not many races per cycle. nevertheless Republicans do record yet again a higher mean total spending in races they won. Performing the same test as with the house yields much less convincing results however with a p value of 0.4917 meaning it's roughly 50/50 as to whether this difference is due to a sampling error or a true difference in means.

Total Spending in House Races



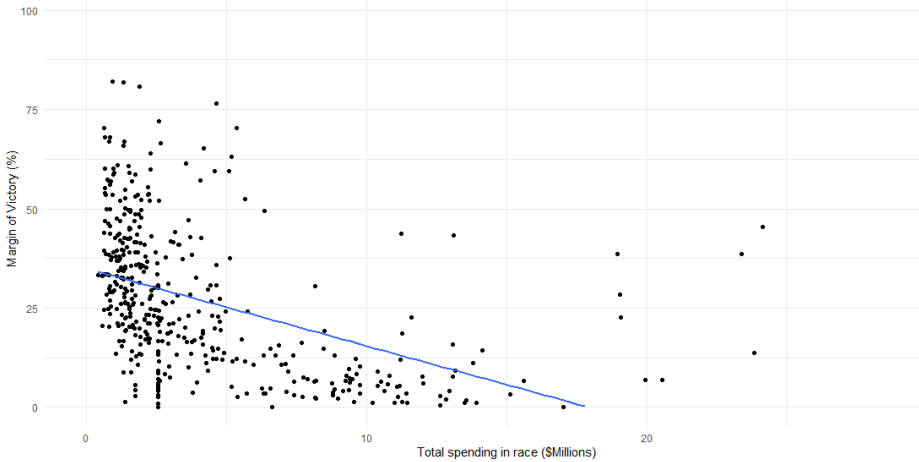
This heat map we are using total sending as the dependent variable. The house elections show some very expensive races that stand out from the rest. In terms of the hypothesis that coastal areas will have the more expensive races, the data do not seem to support this statement as many midwestern areas, southern areas and even Montana appear to have had relatively expensive races. This methodology has its limits however as it only accounts for the spending directly by the campaign and not by PACs or super PACs.

Total Spending in Senate Races

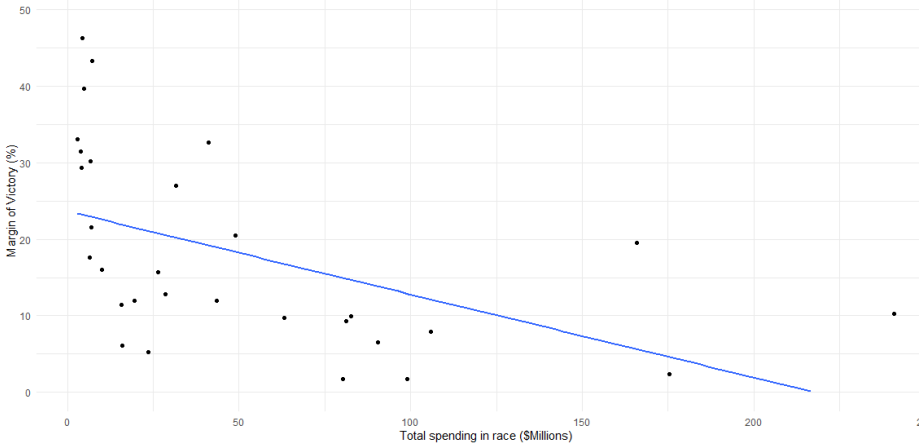


As for the senate races, we see two pretty standout elections, Georgia and South Carolina. I corrected for the fact that Georgia had two races so this does not impact its high levels of spending. However it also had two runoff elections therefore leading to a longer race, this could certainly have had an impact on the levels of total spending. As for South Carolina, this was just a very expensive race, perhaps owing to the national attention (mostly negative) payed to Senator Lindsey Graham. The hypothesis of higher spending in swing states, although a pretty uncontroversial statement, seems to be supported by the data. One way in which the Senate data differs from the House is that their don't seem to be any very expensive single party races. By this I mean a race in which the party of the winner is all but predetermined and the primary elections are far more important. This could be for any number of reasons, one such explanation could be that coordinated party expenditures are far more efficiently spent on competitive races.

House



Senate



In this graph and subsequent regression analysis we are using total spending as the independent variable and margin of victory as the dependent.

The First thing we notice when looking at this graph is that most of the data are clustered towards the left of the graph. This likely reflects the fact that the majority of House races are (relatively) inexpensive and rarely exceeding a few million dollars. The Senate exhibits the same phenomenon but with elevated levels of spending reflective of their generally larger constituences. The variation in the margin of victory for these less expensive races demonstrates how the majority of house races cost between 0 and 2.5 million dollars, 0-15 for the Senate, and within this range there seems to be no correlation between spending and margin of victory. One interpretation of this could be that this is the 'entry cost' or minimum spending for these elections that regardless of how competitive a race is. As we move to more expensive elections the margin of victory seems to fall except for some outliers who clearly do not follow the trend. The explanation for outliers could be very competitive primaries which could drive up costs but then lead to a relatively decisive general. When looking at the regression analysis for this relationships we see the following:

House:

Residuals:

Min 1Q Median 3Q Max
-29.341 -12.668 -3.892 10.888 62.314

Coefficients:

Estimate Std. Error t value Pr(>|t|)
(Intercept) 32.5277 1.0766 30.213 < 2e-16 ***
spending_millions -1.2562 0.1608 -7.812 4.33e-14 ***

Senate:

Residuals:

Min 1Q Median 3Q Max
-15.836 -7.197 -4.491 8.005 23.092

Coefficients:

Estimate Std. Error t value Pr(>|t|)
(Intercept) 23.64927 2.70509 8.743 1.71e-09 ***
spending_millions -0.10850 0.03496 -3.103 0.00434 **

The data from the house is more reliable due to the higher sample size but both data sets lead to significance at atleast the 95% confidence level for all coefficients. Some interesting differences seem to occur between the two branches of congress. The intercept (the value for margin of victory we would expect to see if there was 0 spending in the race) is much lower for the senate (23.6%) than that of the house (32.5%). This could be reflective of the more competitive nature of senate races, neither party is content to allow for easy victories in the senate as each seat carries far more weight than that of the house.

The coefficients of 'spending_millions' (slope) tell a very interesting story. The value observed for the senate is -0.109, meaning that for an extra 1 million in total spending we would expect the margin of victory to fall by 0.109%. For the house however we see a value of -1.26, an extra 1 million of total spending would lead to a fall in the margin of victory of around 1.26%. This is interesting because the mean total spending for the senate was around \$49.9 million, for the house it was \$4.31 million. This is a ratio of about 11.6. If we multiply the house coefficient of -0.109 by 11.6 we get a value of -1.26 which to 3 significant figures is exactly the value of the coefficient observed in the regression analysis. This means that when we remove the scale of spending, total spending has almost exactly the same relationship with margin of victory in both branches of congress.

References for external sources:

election data https://github.com/kjhealy/us_elections_2020_csv
spending data <https://www.fec.gov/data/browse-data/?tab=bulk-data>
state translator <https://worldpopulationreview.com/states/state-abbreviations>

BY STUART BLADON