Campaign Spending Strategies and Election Outcomes

17.835 Final Project | *Hanna Tuomi, Emily Huang, Devin Murphy, Julia Clarke, and Hesham Nawaz*

MIT POLITICAL SCIENCE

Overview

In the US, candidates for the Senate and House who outspend their opponents win approximately 80-90% of the time. Could certain allocation decisions and spending strategies help lower-funded candidates maximize their chances of winning the election? Our project focuses on the **impact of candidate's spending strategies on their election outcome.**

To answer this question, we first identified and characterized spending strategies through a PCA analysis, k-means clustering, and a difference of means visualization. In order to better understand the impact of each spending strategy on election outcome, we matched candidates and calculated the ATT and ATE of each spending strategy.

Data Sets

For the project, we specifically focused on data for candidates for the US House of Representatives in 2010, 2012, and 2014. We selected campaign spending categories and demographic information commonly used for voting behavior analysis. Below are our sources of data and their respective variables.

Congressional Races
Sources: FEC and Rutgers Center for American Women in Politics
Variables: District, Candidates, Political Party, Incumbency, Vote Share, Vote Totals, Candidate Gender

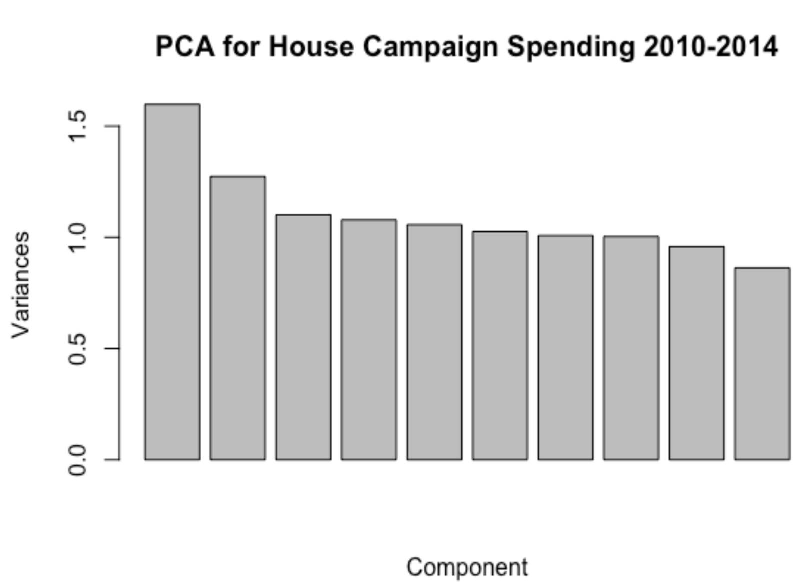
Campaign Spending
Sources: OpenSecrets
Variables: 11 categories of campaign expenditures

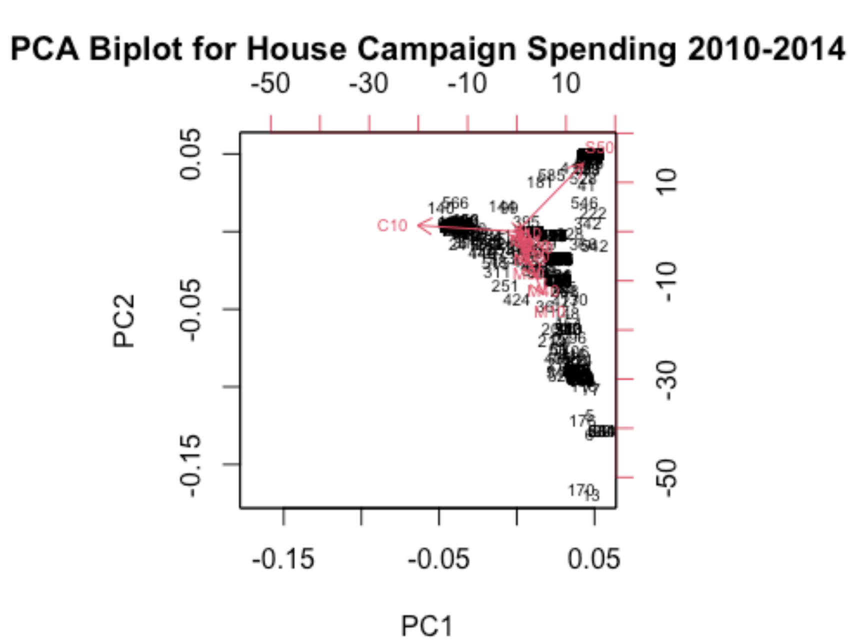
AMERICAN COMMUNITY SURVEY
U.S. CENSUS BUREAU

Demographics
Source: 2019 US Census American Community Survey
Variables: Age, Race, Place of Birth, Employment, Income, Home Value, Education

PCA

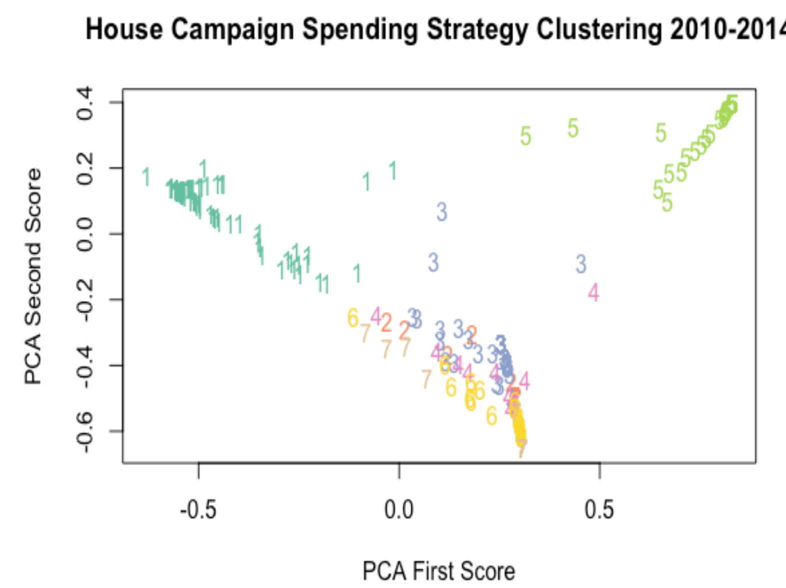
Running a PCA analysis, the principal components explain a good amount of the data variance. The biplot breaks down the influence of each spending category in 3 main directions characterized by C10 (Campaign Materials/Mailings), S50 (Campaign Consulting), and then general media/ad type spending.

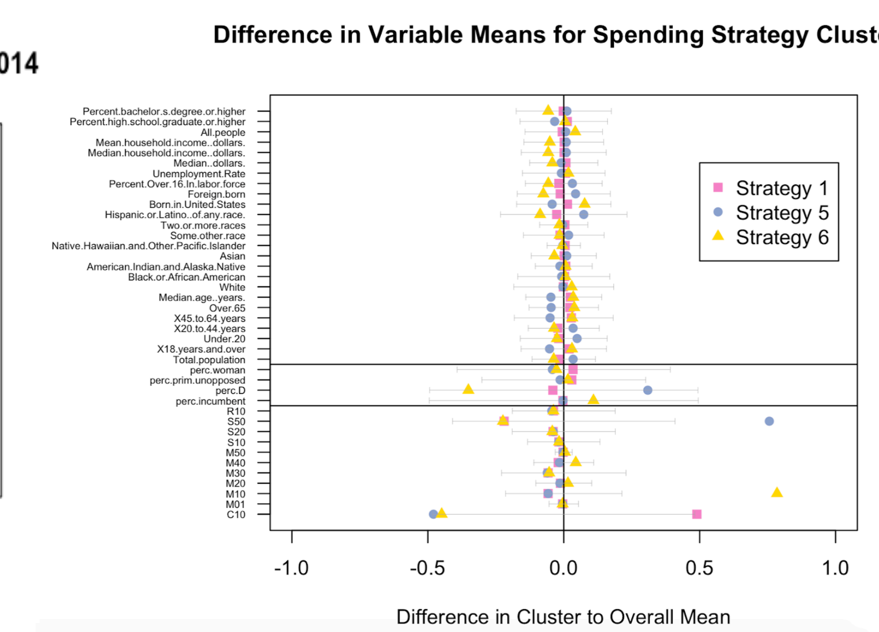




Identifying Spending Strategies

A K-means clustering identified 7 distinct spending categories. To characterize the clusters, we used a difference in means of the spending, candidate, and district variables between each cluster and the overall data. This means analysis is shown for the 3 major clusters, 1, 5, and 6.



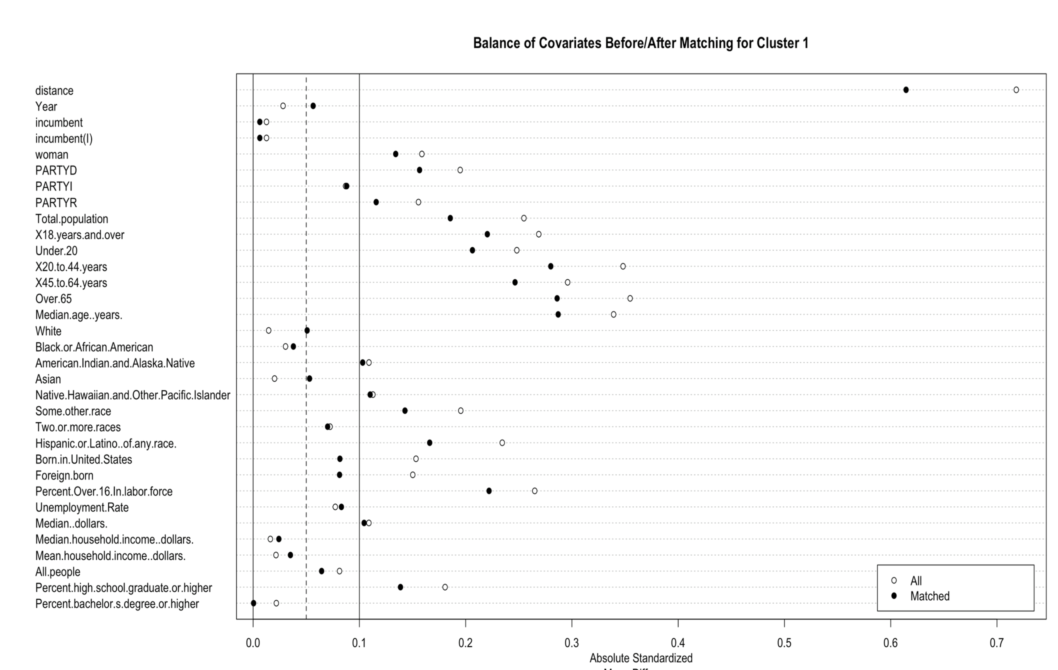


Cluster #	Members	Defining Characteristics
Cluster 1	322	+C10 (Campaign Mailings and Materials)
Cluster 2	27	+S20 (Campaign Data and Technology)
Cluster 3	55	+M20 (Print Ads), +M40 (Media Production), +S10 (Polling and Surveys)
Cluster 4	29	+R10 (National Party Cont.), -% Dem, -% incumbent
Cluster 5	149	-C10, +S50 (Campaign Consulting), +% Dem
Cluster 6	43	-C10, +M10 (Broadcast Ads) , -% Dem, +% incumbent
Cluster 7	40	-C10, +M30 (Web Ads), +% incumbent, -%D

District demographics seem to have little influence on any spending strategies – almost all clusters had demographic variable means within ~0.5 sd of the overall mean. Each strategy was most clearly identified by 1-2 spending categories, party, and incumbency.

Candidate Matching

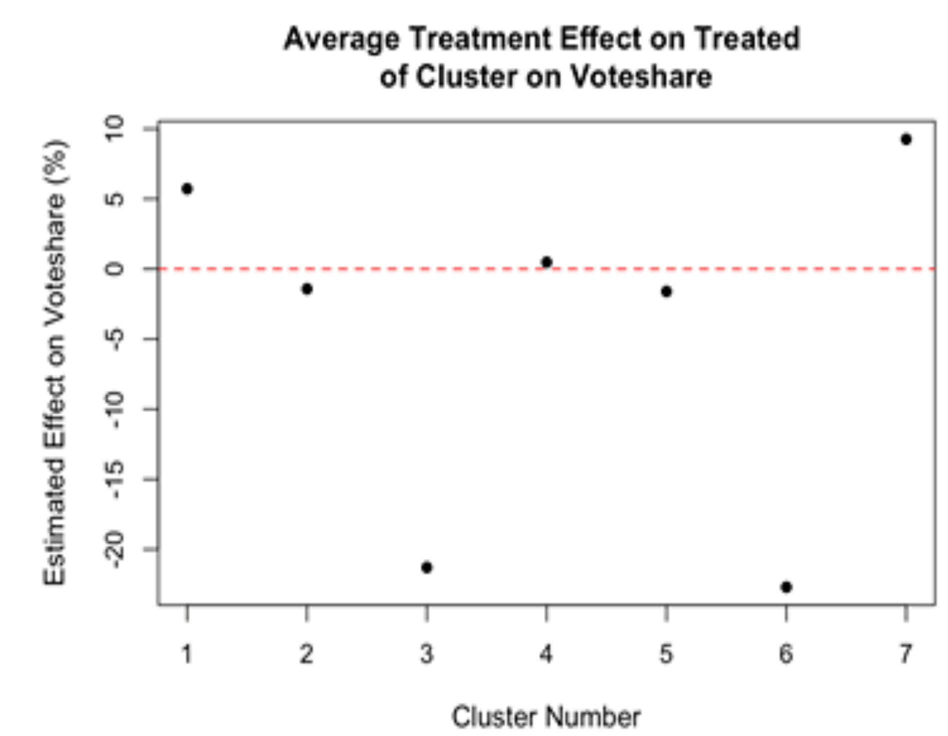
To investigate the causal relationship between a candidate's campaign spending strategy and their election outcome, we started by matching candidate campaigns. To control for confounders, we matched races based on a combination of candidate characteristics and district demographics using a nearest-neighbors approach with replacement. We show the balance of the covariates as well as an example match for spending strategy 1 as identified above.

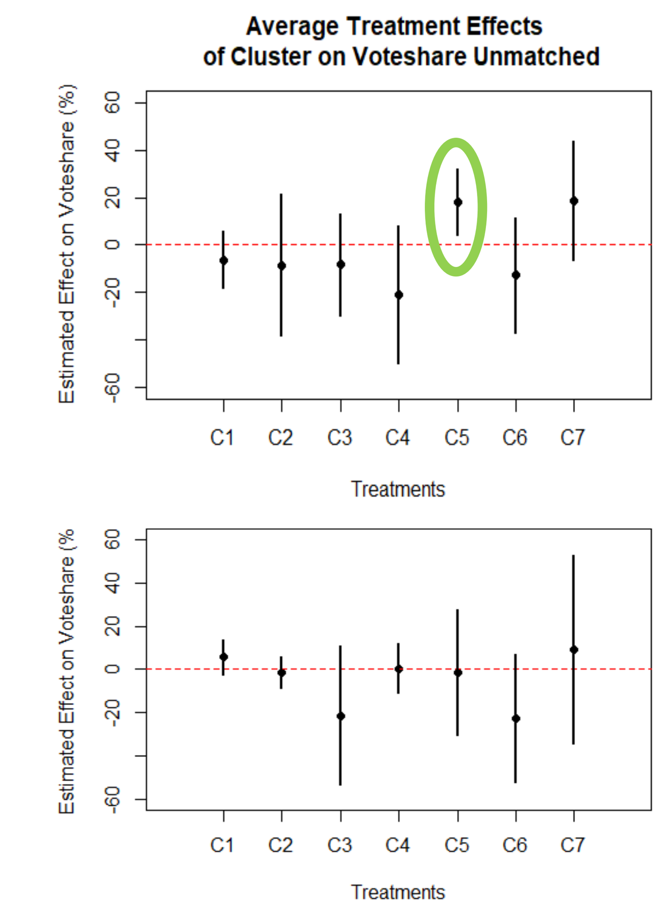


	PARTY	STATE	candidate.name	Total.population	Under.20	X20.to.44.years	X45.to.64.years	Over.65	Median.age..years.
144	D	Rhode Island	Cicilline, David N.	0.001429146	0.4694009	0.40078	0.7121672	0.298709	0.3882784
147	D	Rhode Island	Segal, David A.	0.001429146	0.4694009	0.40078	0.7121672	0.298709	0.3882784

ATT and ATE

We analyzed the Average Treatment Effect for the Treated (ATT) and Average Treatment Effect (ATE) for the seven campaign spending strategies. ATT provides insight into campaign strategy impacts on vote share percentage within a treatment cluster and ATE shows how spending in a certain cluster would affect vote share for a typical house candidate. To estimate the ATEs, we used an OLS regression of vote share on the treatment variables for each cluster. Shown on the right are the 95% confidence interval for the Beta term of the linear regression for both unmatched and matched candidates by cluster.





The confidence intervals for ATE are large, and most of them touch 0. Thus, our data does not necessarily show a causal relationship between cluster classification and vote-share. However, cluster 5 does have a rather high ATE with a confidence interval not including 0. The Beta for this cluster's regression was at 18.03. This suggests that allocating spending towards campaign consulting, which characterizes cluster 5, could possibly help a candidate win a higher percentage of votes in their election.

Results

Using our data, we were able to classify candidates into 7 campaign spending strategies, each characterized by a few spending categories and the candidate's party and incumbency, and notably not district demographics. Through matching, ATT, and ATE analysis, we found that while most of these strategies have little to no impact on vote share, more targeted spending on campaign consulting may help a candidate increase vote share in their election.

Next Steps

- Refine matching process with spending strategies with better controls
- Causal inference with individual components of spending strategies
- Repeat analysis with spending data that are more extensive and supplemented for accuracy
- Account for confounding factors such as likeability/track record/turnout, etc

A big thank you to Professor In Song and the 17.835 Teaching Staff!