

Final Project: Politics, Advertising, and Facebook

Tina Tucker

Summary

Scholars have speculated that the advent of social media has revolutionized the dynamics of political advertising in the US. In the pre-digital broadcast era, the high cost of television ad slots led to only the wealthiest campaigns being able to afford to advertise. This typically meant campaigns run by long-time incumbent candidates or candidates competing for high-profile, highly competitive seats. But some theorize that social media advertising – with its much lower costs – has allowed longshot challengers and safe seat candidates to advertise more. This project tests this theory. Leveraging an original dataset on Facebook ad spending by US Senate candidates, it uses multiple linear regression analysis to investigate which sorts of candidates spent the most on advertising in the last two elections. It finds that ad spending remains concentrated in the most competitive races and among incumbents. But it also finds a significant partisan divide on Facebook ad spending: Democratic candidates outspend Republican candidates 3-to-1. Conclusions and limitations are discussed.

Introduction

Every election cycle, the use of digital advertising by US political campaigns becomes more commonplace. Yet digital ads have long proven challenging for researchers to study, given their transience and microtargeting capabilities. With no complete or long-lasting record of digital political ads, researchers have had to rely on small, non-random samples in order to theorize about online candidate behavior and campaign messages (e.g. Anstead et al., 2018; Ballard, Hillygus, & Konitzer, 2016). Recently, however, companies like Facebook, Twitter, and Google have begun establishing archives of political ads run on their platforms. For this project, I leverage a new dataset on Facebook ads run by 2018 and 2020 US Senate candidates. I examine which kinds of candidates are more likely to run ad campaigns on Facebook.

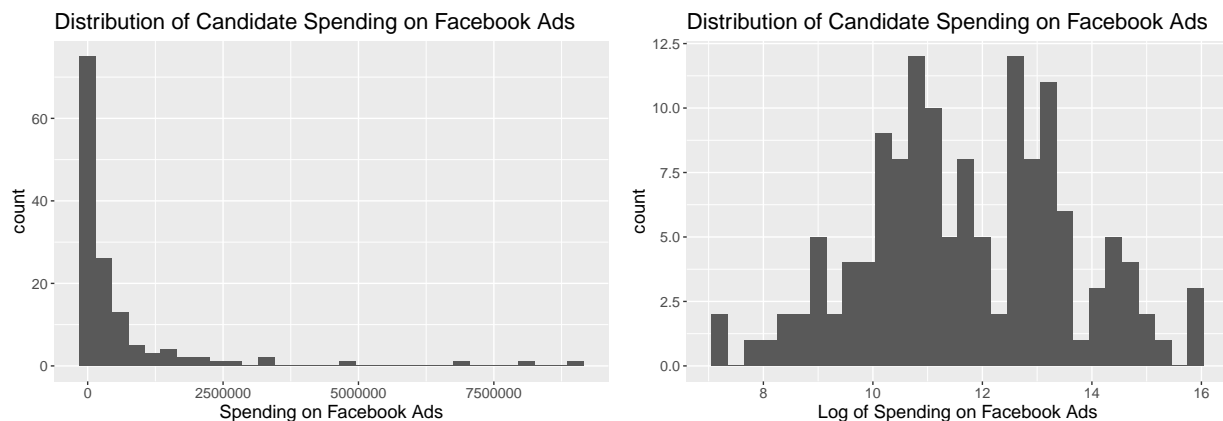
Past research on television political advertising has found that advertising is concentrated among candidates running for highly competitive seats or among long-time incumbents. This is because purchasing TV ad slots is expensive — and only campaigns who have enough cash on hand from large-scale fundraising can do so. Is this still the case in the digital era? Social media has dramatically lowered the costs of political advertising. On average, \$100 spent on a Facebook ad can reach 2,000 voters. In addition, the microtargeting capabilities of social media ads mean that candidates no longer have to advertise to broad, diverse television audiences. Instead, they seek out narrow segments of the electorate online to persuade and mobilize.

Scholars theorize that challengers and candidates in running for safe seats are advertising to voters now via digital means (Baldwin-Philippi, 2017). I investigate whether Facebook advertising has truly “levelled the field” for candidates, or whether advertising is still concentrated in the most competitive races and among incumbent candidates. I also investigate whether the Democratic Party still has a “digital advantage,” as they did in 2008 and 2012, compared to Republicans (Kreiss, 2012) — or whether the two political parties are now on par in terms of their resources devoted to social media ads.

Data

For this project, I focus on Facebook advertising by US Senate candidates during the 2018 and 2020 elections. I create a new dataset of 138 observations, combining ad data from the [Facebook Ad Library](#) with contextual data on the Senate races from [Ballotpedia](#). The Facebook Ad Library is an online archive of every political ad run on the platform. It provides an estimate of the dollar amount spent on each Facebook ad. I used this to calculate the *total* estimated dollar amount spent by each Senate candidate on Facebook ads during the four-month period prior to the 2018 and 2020 elections.

The **dollar amount spent** on Facebook ads per candidate serves as my response variable. When checking its distribution, I found it was heavily skewed to the right. Thus, I chose to transform it into the log. This resulted in a more normal distribution.

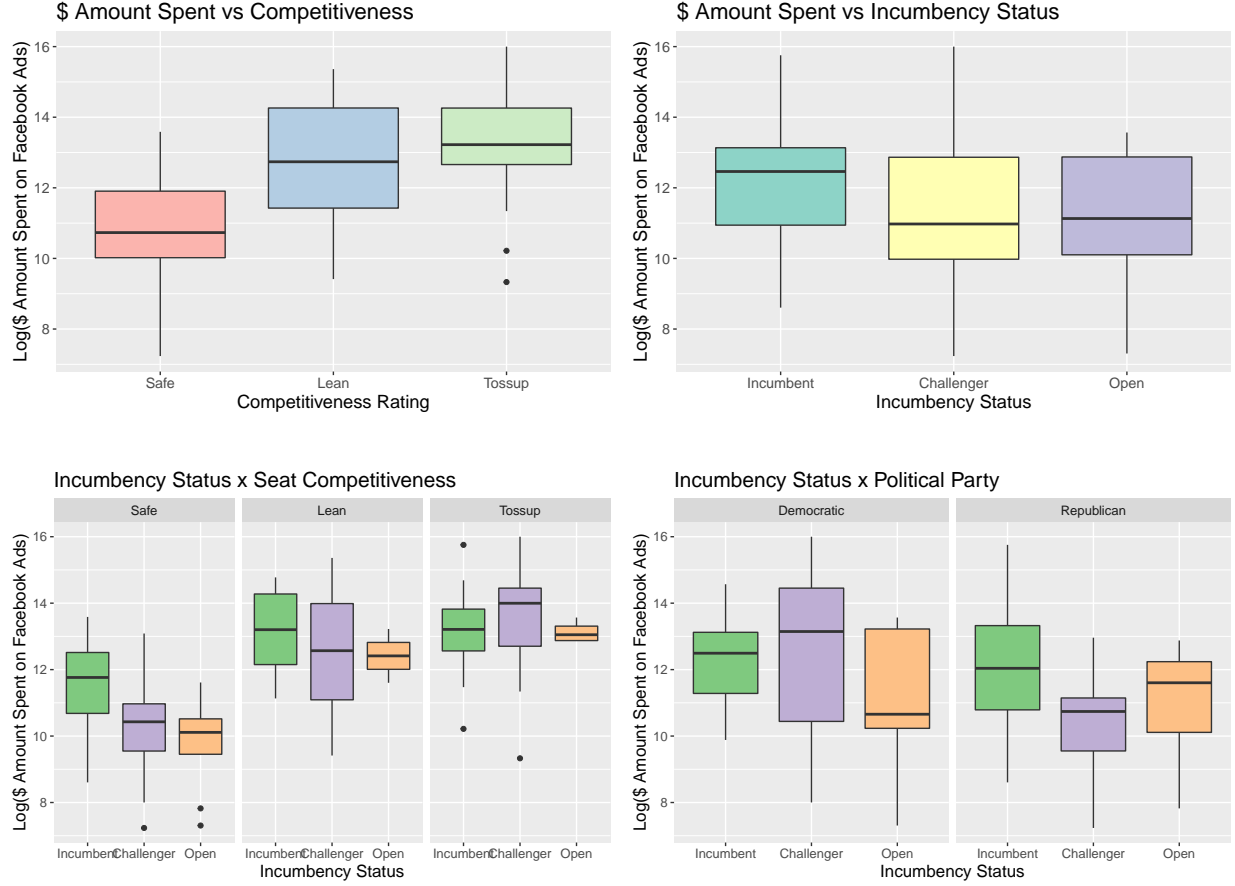


The dataset includes our three main predictors of interest: the candidate's **political party** (Republican or Democratic), **incumbency status**, and **seat competitiveness**. **Incumbency status** refers to whether the candidate was an incumbent, a challenger to the incumbent, or competing in an open seat. **Seat competitiveness** was taken from [Cook Political Report](#), an organization that rates, based on polling, how close every electoral race is. The main categories are "safe," meaning that the incumbent is highly likely to win re-election; "lean," meaning that the race is more likely, but not certain, to be won by a particular candidate over the other; and "tossup," meaning the race is so close that it could be won by either candidate.

The dataset includes the following variables as controls: **year** of the Senate race (2018 or 2020), whether it was a general or **special** election, and whether the eventual **outcome** of the race was a win or loss for the candidate. I create an additional control variable, **region**, by collapsing the **state** variable into five regions of the US – northeast, midwest, southeast, southwest, and west.

Exploratory data analysis revealed some clear trends right off the bat. Average **dollar amount spent** on Facebook ads was noticeably higher in races rated "tossup" and "lean" than in "safe" seats. It was also higher among incumbent candidates versus challenger and open seat candidates. In addition, Democratic candidates appeared to spend more on average than Republican candidates. Finally, ad spending was higher in the 2020 election than the 2018 election, and ad spending was higher in the southwest and midwest where there are more 'swing' states.

Further examination also revealed some potential interactions. One is that spending on Facebook ads appears to be higher among challengers — but only in "*tossup*" races. Another is that spending on Facebook ads appears to be higher among challengers — but only among *Democratic* challengers. Among Republican candidates, incumbent ad spending is actually higher. We turn to linear regression modelling to determine if these trends are significant and meaningful.



Model

Building the Model

Given that our response variable consists of continuous integers, I opt for building a multiple linear regression model. Recall that the response variable **dollar amount spent** is transformed into the log. Initially, I included in the model the three main predictors — categorical variables for **party**, **incumbency**, and **competitiveness** — plus the main controls, i.e. indicator variables for **year**, **special** election, **outcome** of the election, and a categorical variable for **region** of the US. I also included the potential interactions identified during exploratory data analysis: **incumbency x competitiveness** and **incumbency x party**.

Many of these variables proved statistically insignificant: The **outcome** of the election, the **region** of the US, and whether it was a general or **special** election had a limited associational relationship with Facebook ad spending. The interactions were also not significant — and their inclusion even worsened model fit. The model's AIC was slightly lower with the interactions (492.4) versus without the interactions (486.3), although this difference was not statistically significant according to an F-test.

Thus, I engaged in stepwise model selection to determine which variables are statistically irrelevant and which together make the best model. This produced the following final model:

$$y_i = x_i\beta + \epsilon_i; \epsilon_i \sim \mathcal{N}(0, \sigma^2)$$

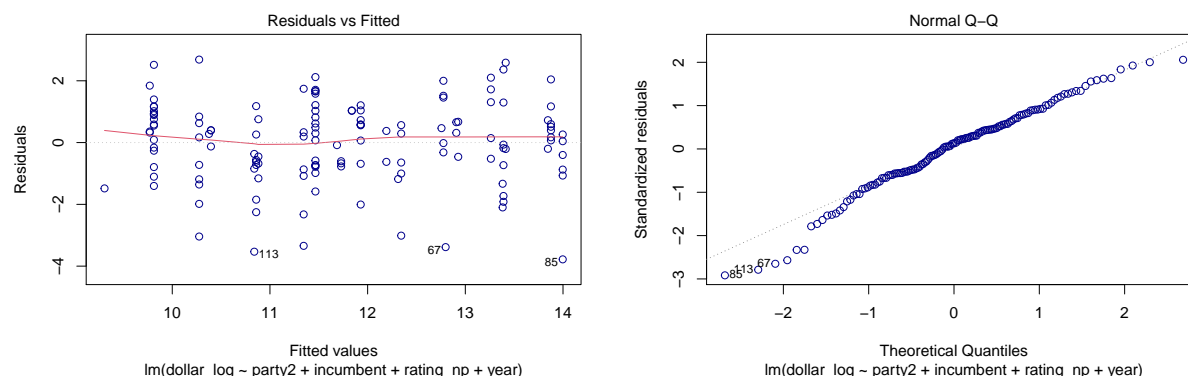
where y_i is the log of dollar amount spent on Facebook ads for observation i , and x_i is the vector containing the corresponding values for political party, incumbency status, seat competitiveness, and year of election.

Assessing the Model

The final model's AIC is 478.9 and the adjusted R^2 is 0.497. This R^2 value indicates that the model performs decently well, explaining about 50 percent of the variation in Facebook ad spending. However, given the low number of observations in the dataset, we should be cautious about model overfitting. We need to collect more data to be able to assess how well this model performs in other contexts, i.e. in other electoral races or other election years.

I next check our regression assumptions for independence, equal variance, and normality. There is some evidence that some of our assumptions may have been violated. Looking the residual-fitted values plot below, there is some clustering of points around certain fitted values, indicating an issue with our independence assumption. However, the points do seem fairly equally spread out around zero, indicating that our equal variance assumption holds. Meanwhile, according to a Q-Q plot, there is some deviation from the 45 degree line, indicating a potential issue with our normality assumption for extreme values.

Finally, I check for outliers. Although there appears to be several outliers with high leverage scores, I find no evidence that these outliers actually influence the regression model. No data points fall above a Cook's distance value of 0.12. Therefore I opt not to remove outliers from the dataset.



Interpreting the Model

Overall, I find that political advertising on Facebook is dominated by campaigns for competitive seats over safe seats, by incumbents over challengers, and by Democrats over Republicans.

Seat Competitiveness. There is a strong association between Senate seat competitiveness and Facebook ad spending. Compared to candidates running for a safe seat, candidates running for a tossup seat are associated with a $\exp(2.53) = 12.55$ increase in Facebook ad spending, with 95% CI = (7.32, 21.54). Similarly, compared to candidates running for a safe seat, candidates running for a “lean” seat are associated with a $\exp(1.92) = 6.82$ increase in Facebook ad spending, with 95% CI = (3.63, 12.68). These are both statistically significant at $p < 0.001$.

Incumbency Status. There is also an association between incumbency status and Facebook ad spending. Compared to incumbent candidates, being a challenger candidate is associated with a $\exp(-0.58) = 0.56$ decrease in Facebook ad spending, with 95% CI = (0.35, 0.90). This is statistically significant at $p < 0.05$. Meanwhile, being a candidate running for an open seat is associated with a $\exp(-1.09) = 0.34$ decrease in Facebook ad spending, with 95% CI = (0.16, 0.73). This is statistically significant at $p < 0.01$.

Political Party. Interestingly, the model finds a strong association between political party and Facebook ad spending. Compared to Republican candidates, being a Democratic candidate is associated with a $\exp(1.07) = 2.92$ increase in Facebook ad spending, with 95% CI = (1.86, 4.57). This is statistically significant at $p < 0.001$.

Conclusion

This project finds little evidence that the arrival of social media has revolutionized political advertising. Similar to television advertising, spending on Facebook advertising remains dominated by campaigns for highly competitive seats and by incumbent candidates. However, it is clear that a partisan gap has emerged — Democratic candidates are spending significantly more on Facebook ads than Republican candidates. More research needs to be done to figure out why.

There are many limitations to this project. The main issue is the small size of the dataset at only 138 observations. We need to collect more data from other electoral races and other election years in order to find out if the model's results are consistent across time and level of government. Another potential next step would be to include in the dataset spending totals on television political advertising by the same Senate candidates. It may be the case that digital advertising is indeed dominated by competitive and incumbent campaigns, but that it nonetheless widens the pool of candidates who can advertise in the first place. In other words, longshot challengers and safe seat candidates — normally dissuaded from advertising by the high cost of television ad slots — may opt for a limited digital advertising campaign over doing nothing.

Table 1: Multiple Linear Regression Model for Predicting Facebook Ad Spending by US Senate Candidates

	Dollar Amount Spent on Facebook Ads
	dollar_log
Democratic	1.07*** (0.62, 1.52)
Challenger	-0.58* (-1.06, -0.11)
Open Seat	-1.09** (-1.86, -0.32)
'Lean' Race	1.92*** (1.29, 2.54)
'Tossup' Race	2.53*** (1.99, 3.07)
Year	0.46* (0.01, 0.92)
Constant	10.40*** (9.89, 10.91)
Observations	138
R ²	0.52
Adjusted R ²	0.50
Residual Std. Error	1.33 (df = 131)
F Statistic	23.55*** (df = 6; 131)
Note:	*p<0.05; **p<0.01; ***p<0.001

Note: Data and R code can be found [here](#).

References

1. Anstead, N., Magalhães, J. C., Stupart, R., & Tambini, D. (2018). Political advertising on Facebook: The case of the 2017 United Kingdom general election. In *2018 American Political Science Association Annual Conference*.
2. Baldwin-Philippi, J. (2017). The myths of data-driven campaigning. *Political Communication*, 34(4), 627-633.
3. Ballard, A. O., Hillygus, D. S., & Konitzer, T. (2016). Campaigning online: Web display ads in the 2012 presidential campaign. *PS: Political Science and Politics*, 49(3), 414.
4. Hale, J. F., Fox, J. C., & Farmer, R. (1996). Advertisements in US Senate campaigns: The influence of campaign context. *Social Science Quarterly*, 329-343.
5. Kreiss, D. (2012). *Taking our country back: The crafting of networked politics from Howard Dean to Barack Obama*. Oxford University Press.