Gubernatorial Behavior on Twitter

A Dissertation Presented

by

Matthew Duell

to

The Graduate School

in Partial Fulfillment of the

Requirements

for the Degree of

Doctor of Philosophy

in

Political Science

Stony Brook University

August 2021

Stony Brook University

The Graduate School

Matthew Duell

We, the dissertation committee for the above candidate for the

Doctor of Philosophy degree, hereby recommend

acceptance of this dissertation.

John Ryan – Dissertation Advisor
Associate Professor, Political Science

Yanna Krupnikov – Chairperson of Defense
Professor, Political Science

Michael Peress
Associate Professor, Political Science

Yotam Shmargad
Assistant Professor
School of Government and Policy
University of Arizona

This dissertation is accepted by the Graduate School

Eric Wertheimer
Dean of the Graduate School

Abstract of the Dissertation

Gubernatorial Behavior on Twitter

by

Matthew Duell

Doctor of Philosophy

in

Political Science

Stony Brook University

2021

Since its launch in 2006, Twitter has been adopted as a social media platform by 24% of the adult population in the US alone (Pew Research Center, 2019). The platform has become an attractive way for politicians to spread their message, and their posts now attract the attention of major media outlets frequently. Politicians' use of the platform has become so ubiquitous, that Twitter hesitates to enforce its Terms of Service on politicians and legal free speech debates arise when politicians are suspended or banned. The president's Tweets from the White House account are now even considered part of the public record and are archived during the transfer of the @POTUS handle. Political science research on Twitter has primarily focused on congresspeople. In the executive branch, only two presidents to date have made significant use of the platform (Barack Obama and Donald Trump), and governors' tweets have not been systematically studied. There is evidence of some party differences on Twitter, but it is not yet known whether those differences will robustly replicate across different branches of government. It is also unclear whether executive branch candidates and politicians coordinate their messaging on Twitter, and whether such coordination would even help them. This dissertation aims to address these questions with an original data set composed of ten years worth of tweets from gubernatorial candidates. I apply sentiment analysis and machine learning to classify tweets and examine how gubernatorial candidates tweet and what they tweet about. Individual candidates' messaging priorities are compared to others of their own party to determine to what extent these candidates coordinated on messaging. This dissertation is an exploratory effort to quantify the tone and content of gubernatorial tweets which lays the groundwork for further investigation of the use of social media by the executive branch.

Dedication Page

To Jess, without whom I would not have gotten this far. This dissertation would not have been completed without her help editing and proofreading.

Table of Contents

List of Tables

List of Images

List of Figures

CHAPTER 2

Chapter 1

Since its launch in 2006, Twitter has been adopted as a social media platform by 24% of the adult population in the US alone with its signature short post style (Pew Research Center, 2019). The platform has become an attractive way for politicians to spread their message, and their posts now frequently attract the attention of major media outlets. Politicians' use of the platform has become so ubiquitous, that Twitter hesitates to enforce its Terms of Service on politicians and legal free speech debates arise when politicians are suspended or banned. The president's Tweets from the White House account are now even considered part of the public record and are archived during the transfer of the @POTUS handle (McMillen, 2016; Schulman, 2016). Political science research on Twitter has primarily focused on congresspeople. In the executive branch, only two presidents to date have made significant use of the platform (Barack Obama and Donald Trump), and governors' tweets have not been systematically studied. There is evidence of some party differences on Twitter, but it is not yet known whether those differences will robustly replicate across different branches of government. It is also unclear whether executive branch candidates and politicians coordinate their messaging on Twitter, and whether such coordination would even help them. This dissertation aims to address these questions with an original data set composed of ten years' worth of tweets from gubernatorial candidates. I apply sentiment analysis and machine learning to classify tweets and explore how gubernatorial candidates tweet and what they tweet about. Individual candidates' messaging priorities are compared to others of their own party to determine to what extent these candidates coordinated on messaging. This dissertation is an exploratory effort to quantify the tone and content of gubernatorial tweets which lays the groundwork for further investigation of the use of social media by the executive branch.

Background Research

Twitter represents an interesting point of overlap between several areas of political research. It can be understood as a platform for campaign messaging, a medium of mass communication, or a site for network effects. Past research on each of these areas is considered below in relation to Twitter.

Twitter provides a clear platform for campaign messaging, as there is no filter or middleman between what the candidate wants to say and what is communicated in each message. Mobilizing the base and persuading core partisans to turn out and vote are key goals for campaigns, and an area where they can have a significant influence (Holbrook & McClurg, 2005). We know that political messaging can prime voters to use certain values in their political decision-making (Johnston, Blais, Brady, & Crête, 1992; Rosenstone & Hansen, 2001), and that they create lasting impressions of the candidate (Lodge, Steenbergen, & Brau, 1995). Specific candidates may be more or less skilled at political messaging, and some political scientists have argued that this degree of skill makes a difference in close races (Carsey, 2009; Vavreck, 2009, 2014). However, there is also competing evidence that the direct effects of messaging are not strong. Ads which attempt to influence specific groups of voters often fail to do so (Sides & Karch, 2008). Voters are generally not good at identifying what issues the campaign was about, both during and after the fact (Dalager, 1996), and Kalla and Broockman (2018) find that there are no effects of direct contact from a campaign at all, unless the campaign has both isolated persuadable voters and carries very unpopular opinions. Twitter is therefore of interest to political science as a contemporary opportunity to reexamine these findings and controversies on a new platform.

It may also be the case that Twitter has an indirect effect on electoral races through the media; for example, an individual tweet may not have much power until it is picked up by national news and becomes the subject of a story. Shapiro and Hemphill (2017) find that Twitter discussions by members of congress are correlated with New York Times coverage, suggesting that the media amplifies certain comments and positions taken on Twitter by elected officials. The authors find that when there is active debate between Democrats and Republicans on an issue, it is negatively correlated with news coverage (such that extensive back-and-forth arguing isn't as frequently selected for coverage). This being the case, political science benefits from investigating what kinds of political messages the media likes to amplify. Flowers and colleagues find that the media's transmission of political messages depends both on the content (substantive, values-based, etc.) and the candidate's position in the race (Flowers, Haynes, & Crespin, 2003). The media treats messages from challengers and incumbents differently based on these factors; for example, policy messages by challengers are transmitted more than those of incumbents. Furthermore, media companies are more likely to transmit messages from political parties favorable to their consumer base (Haselmayer, Wagner, & Meyer, 2017). It is also well documented that the media tends to talk about issues that are already salient with the public, rather than new issues, and they are more likely to transmit negative messages than positive messages (Ridout & Smith, 2008). Further work has found that media attention of this kind is essential to voters' understanding of politics and of campaigns, and has an effect on moving and defining public opinion (Hill, Lo, Vavreck, & Zaller, 2013; Meyer, Haselmayer, & Wagner, 2020). Candidates can be strategic about what messages they employ to generate more media coverage that will boost their visibility.

Finally, it is useful to consider Twitter from the more obvious perspective of social networking research. Networks alter the political behavior of actors within them, and there is a wealth of research on how a person's network affects their voting behavior. Voting behavior is highly influenced by group norms of political action (Huddy, Mason, & Aarøe, 2015; Thomas & McGarty, 2009), and the political experiences of those you live with (Nickerson, 2008). The presence of political experts or elites in one's network tends to increase surrounding voter participation (McClurg, 2006; Pietryka & DeBats, 2017). All of these effects would be expected to translate to online social networking as well. Bond et al. (2012) find that if people in your online network post a message that they voted, you are also more likely to go vote. Twitter allows its users to self-select into an expanded social network, such that users who are even marginally interested in politics can fill their feed with political experts and elites; even in cases where candidates post non-interactive content, citizens can develop parasocial relationships with these political elites such that they're treated as part of their personal social network (Ballantine & Martin, 2005). This process inducts internet personalities into the user's accepted social network, and the messages these personalities transmit have similar effects as those delivered by friends and trusted sources.

Through each of these three avenues (political messaging, mass media, and social networking), Twitter can have an impact on political discussion and electoral outcomes. It is thus a compelling subject for further political science research as we investigate the specific effects it has on each unique election cycle. In the next section, I review past research on Twitter in political science to identify unanswered questions and outline methodologies which have been used.

Twitter Research

Twitter research in political science has undergone rapid growth since the platform's launch in 2006. Publications have become more common as the technical capabilities of the field have increased; political scientists can now take advantage of machine learning, web scraping, and other methods to standardize collection and coding of Twitter data. The earliest approaches to studying Twitter focus on collecting and coding tweets by hand, using original classification and coding schemes that in turn meant there was little overlap between publications. More recently, methods of classification have advanced as machine learning has been adopted. Hemphill and Schöpke-Gonzalez (2020) coded tweet content using both supervised and unsupervised machine learning; the former entails teaching the computer to code a certain way through examples, while the latter allows the computer to come up with its own sorting system. Hemphill, Culotta, and Heston (2016) also designed an algorithm to take tweets and generate ideology scores based on hashtags. While these technologies have not replaced hand coding (which is still considered a gold standard in the field), machine learning and algorithms save researchers time and effort, and allow for efficient analysis of larger data sets as a result. Unsupervised machine learning also allows for novel interpretations of data which researchers might not initially consider. Even as methods for studying Twitter data advance, Twitter itself continues to grow and change as developers change its features and users adopt new habits on the platform.

To date, most political science research on Twitter has focused on congressional actors due to the sheer number of available candidates. There have been a number of attempts to classify what members of congress discuss on Twitter; most early research suggests that congresspeople use Twitter primarily to announce campaign events or policy positions (M. E.

Glassman, Straus, & Shogan, 2013; Golbeck, Grimes, & Rogers, 2010; Russell, 2018; Straus, Glassman, Shogan, & Smelcer, 2013). These posts were top-down, non-interactive, and purely informational messages. This may suggest that politicians at the time were using Twitter as an outlet for brief press announcements, rather than taking advantage of its potential for back-and-forth interaction with the public. Some of these early studies detected differences in tweet style by gender, race, incumbency, or party, but these differences were fairly small (Evans, Cordova, & Sipole, 2014). Studies of early adoption of Twitter found that Republicans joined Twitter earlier and tweeted more frequently than Democrats (Gainous & Wagner, 2013; Shogan, 2010; Straus et al., 2013). As early as 2010, Twitter adoption began to correlate strongly with election victory (LaMarre & Suzuki-Lambrecht, 2013), suggesting that social media presence is required of a competent campaign. In a rare example of research on gubernatorial candidates on Twitter, Stromer-Galley, Zhang, Hemsley, and Tanupabrungsun (2018) found that challengers are more likely to attack the incumbent on Twitter, but that incumbent attack messages are more likely to circulate on the platform. Attack messages get more retweets on average compared to advocacy messages. Overall, we have a much better idea about the Twitter use of legislators than those in executive positions, and yet much of this work has yet to be updated with the changing standards and expectations of platform use.

Present Project

Given the persistent focus on the legislative branch in Twitter research, examining the executive branch on the platform is a logical next step. However, to date only two presidents (Barack Obama and Donald Trump) and their opponents have used Twitter extensively, limiting the generalizability of the conclusions we might draw from that data. Early research on Twitter in the 2010's would have encountered the same problem with gubernatorial candidates, but as enough time has passed that limitation is no longer a major obstacle. This dissertation aims to fill this gap in the research with an original data set composed of ten years' worth of tweets from gubernatorial candidates. This work is an exploratory effort to quantify the tone and content of gubernatorial tweets. Sentiment analysis and machine learning are used to classify tweets and explore how gubernatorial candidates tweet and what they tweet about, adding to the research which has explored these questions with congressional Twitter data. Individual candidates' messaging priorities are also compared to others of their own party to determine to what extent these candidates coordinated on messaging in a given election cycle.

There are two main goals for this project. The first is to extend existing Twitter research to the executive branch to provide needed context from political actors with different incentive structures. For example, because gubernatorial positions are statewide, governors must appeal to a different constituency, and are held to electoral account for the state economy in a way that federal legislators are not. Second, to assess how use of Twitter by political actors changes over time. Tweets that are considered influential are those with a high number of interactions from other users, and early use of the platform as a means of announcing scheduled events (M. Glassman, 2010; Golbeck et al., 2010; Graham, Broersma, Hazelhoff, & Van'T Haar, 2013) does not reflect the current standard of what success looks like on Twitter for politicians.

Throughout the following chapters, I intend to answer the following four questions. What do gubernatorial candidates tweet about? Are there party differences in Twitter use in the executive branch? How has the use of the platform changed from 2008 to 2018? Finally, how similar are party members to each other?

The second half of this chapter describes the process of data collection, as well as the structure and general content of the data set. Chapter 2 examines the tone and content of the gubernatorial tweets and analyzes the extent of party differences in Twitter use. Chapter 3 examines the extent to which the parties are internally cohesive in their Twitter messaging, and tests whether party coordination helps or hurts candidates.

Data Collection

Below, I describe the collection process for gathering two related datasets. The first is a list of all 314 major-party gubernatorial candidates between 2008 and 2018, as well as a number of standard political variables, such as race, gender, and incumbency. The second is the full collection of 116,647 tweets from these gubernatorial candidates over this time frame. I collect multiple points of information about each tweet, and describe the process I used for assessing both sentiment and content of each tweet.

I began by creating a list of gubernatorial candidates in each election year between 2008 and 2018, using the website Ballotpedia to collect much of the following information. The final list included every major-party gubernatorial candidate in that timeframe, resulting in 314 unique candidate/year combinations. With the candidates' names, I also collected their party affiliation, whether their party was the incumbent party for the seat, whether the candidate was an incumbent themselves, whether each candidate won or lost their race, their vote share in the race, their gender, their primary date, and their election date. -The last two elements were key to collecting Twitter posts in the appropriate timeframe, described below. I limit my search to general -election tweets only, with the expectation that primary election behavior will differ from general election behavior, which would complicate the interpretation of the results. Assessing candidate race manually is difficult, so I code the candidate's race using the 'wru' package in R (Imai & Khanna, 2016). The package compares the surname of the candidate to census data which associates names with race. The result is a probabilistic estimate of the candidate's racial category.

The next step in the process was to assemble a list of each candidate's Twitter handles. This was done using the Twitter API, accessed through the 'twitteR' package in R (Gentry, Gentry, RSQLite, & Artistic, 2016). In the API, I searched for each candidate individually by name, and matched candidates with their accounts by hand. To ensure accuracy, I looked at Twitter account biographical information, samples of the Twitter account's tweets, and filtered out accounts whose content seemed to be satire. Often when it was unclear whether an account belonged to a candidate, I did a web search for the candidate's name and presumed Twitter account to attempt to find news articles which would confirm ownership. One additional complication is that a small group of candidates had deleted their Twitter accounts after the end of their campaign, making their tweets inaccessible to standard searches. I was able to verify only five instances of this occurring, and in some cases, the deleted account was not the candidate's only active account, so I don't consider this an obstacle to analysis.

Here, it's worth noting that there is a key limitation in examining large time periods of Twitter data. The Twitter API, which allows a user to search for, sort, and download tweets, only allows access to tweets from the most recent 7-day period. Normal routes to bypass this limitation require thousands of dollars of payment to either Twitter or to companies which archive and sell Twitter data. In order to access the data, I had to use a web scraper - a program that mimics an internet user's normal access to the platform through the web, and reads the page's source HTML code to copy selected information. Early attempts to build an original web scraper were moderately successful, but ran into difficulties when attempting to gather large amounts of information at once. I ultimately used George Yiannakas's "TwittyJar" program, a self-contained Python module specifically designed to gather old Twitter data, which he published on GitHub (Yiannakas). I modified the module to reference my gathered list of usernames and election dates. In total, the module gathered every tweet for every account in the appropriate date ranges (N = 116,647 tweets). The program pulled the following information for each tweet: a unique identifier, the permanent hyperlink to the tweet, the date and time of the tweet, the tweet's text and image, the number of times the tweet was shared, and whether the tweet was itself original content or a retweet.

Coding Tweets

In order to categorize the text of the tweets, I used the 'gWidgets2' package in R to create a unique user interface (Verzani, 2014). This interface allowed me to examine the text of individual tweets, and create categories which could used to classify further tweets. An example of this interface is presented in Image 1. In order to develop my list of categories, I used a random sample of 500 tweets. I created categories that both matched the text of the tweets, and presented variables which might be interesting to study. After the random sample was coded, I removed categories which appeared in only one or two tweets in the sample, and I combined some categories that were similar. The final list of categories, as well as brief descriptions and examples of their content, are found in Table 1.

To prepare the text for machine learning methods, I had to convert the text of tweets into a useable format which also minimized storage space. Here, I follow the example of Lantz (2013) and transform the text into a unigram data structure called a Document Term Matrix (DTM), where each word becomes a column, each tweet becomes a row, and words are marked as present or not. Though this type of data storage is a sparse matrix, which stores only positive values, and thus much less memory-intensive than a typical R data frame, some adjustments still had to be made. In total, I removed numbers, hyperlinks and pictures, non-alphanumeric characters (with the exception of # and @, which carry semantic meaning on Twitter), stop words such as ‘and’ or ‘the’ (as defined by the ‘tidytext’ package in R), and I converted words to their stems using the ‘SnowballC’ package (Bouchet-Valat & Bouchet-Valat, 2015; Silge, 2016). Finally, because many of the remaining words included attempts to mimic speech by, for example, drawing out vowel sounds (and thus were unique to single tweets or otherwise unhelpful in classification), I removed words which appeared in only one or two tweets. This reduced the dataframe from approximately 48,000 words to 13,714 words.

In order to code the full dataset, I used a machine learning algorithm called a Support Vector Machine (SVM). This algorithm is designed to map data onto a multidimensional space and draw boundaries designed to separate data into binary classifications. The Support Vector Machine algorithm is a supervised model, meaning that it uses a pre-coded training set to develop its classification scheme before it can be used to classify novel data. Therefore, I took another random subsample of 5,000 tweets and coded them by hand using the categories described in Table 1. The SVM model performs relatively well with small numbers of training examples, but some categories were still strongly under-represented in the training set. Categories with five or fewer examples were excluded from further analysis – this included campaign finance, childcare, tradition, and highlighting an individual voter. The first 4,000 tweets were used to train each of the SVM models, while the remaining 1,000 were reserved as a test dataset to assess model accuracy. In order to improve model performance, I used 5-fold repeated cross-validation. This is a method whereby the model, in the course of training, separates the training data into five random partitions, or folds, each representing one-fifth of the total cases. The model trains itself five times, using each fold as reserve of test cases. It then repeats this process three times, each time randomizing which cases belong in which folds. Another popular standard for this process is 10 folds and 10 repeats, which provides substantially more robust results with smaller datasets (Lantz, 2013). I experimented with using this standard, but given the fact that the training set includes 4,000 tweets, accuracy did not noticeably improve, while computation time magnified. Another consideration is that SVM models often attempt to use a linear boundary to create distinctions between classes. Given the complexities of text data, I also experimented with a polynomial specification to explore whether more complicated boundaries were necessary, but again, accuracy was not significantly improved while computation time became unacceptably high.

Table 2 describes the results of these models. Found in the table are the number of examples of each category in the training set, the number of examples in the test dataset, and the accuracy of each model. Accuracy was assessed by using the models compiled on the training data to code the test data, and comparing those codes to my hand-coded values. Table 2 describes accuracy in three ways – on positive cases only, on negative cases only, and overall accuracy. Accuracy of the models ranged from 81.3% to 97.9% overall. This is within normal and acceptable limits on machine learning models (Lantz, 2013; Yu-Wei, 2015). Positive-case accuracy (the proportion of positive cases in the test dataset correctly classified) varies strongly, from 0% to 80%, depending on the model. This is to be expected and not particularly problematic – some categories contain few examples in the test data which might appear more widely in the overall dataset; other categories use more diverse words for the same concepts and therefore should be harder to accurately code. For the categories which are the highlight of my analysis in the following chapters, such as the economy, positive-case accuracy is more than acceptable. After verifying a model's accuracy, it was used to predict classification across the entire dataset of tweets.

Sentiment Scores

Finally, I used the 'tidytext' package's sentiment lexicons to assess the tone of each tweet. I use two of these lexicons to code my data (Silge, 2016). The 'Bing' lexicon is categorical, rating each word as positive or negative, and includes 6,786 different words. The 'AFINN' lexicon is ordinal, rating each word from -5 (most negative) to +5 (most positive). It includes 2,477 words. For each tweet, I create both a Bing score and an AFINN score. The Bing score is generated by subtracting the number of negative words from the number of positive words, and the AFINN score is generated by subtracting the absolute value of all negative words in the tweet from the total value of all positive words in the tweet. Together, these lexicons allow me to investigate both the directionality and the intensity of candidate's tone.

Final Datasets

Combined, the above efforts resulted in two final datasets. The first uses individual tweets as the variable of interest, while the second aggregates tweets into candidate/year averages. The tweet data includes the permanent link to each tweet, a unique identifier for each tweet, the candidate who posted the tweet, the candidate's political party, the username the tweet was posted under, the date and time the tweet was posted, the text of the tweet, the type of tweet (original or retweet), the number of times the tweet was shared, the number of hashtags (#) the tweet used, the number of other users tagged in the post (@), a series of binary classifications for each of the categories outlined above, and the tweet's sentiment scores.

The second dataset includes the year and state in which the campaign occurred, the name of the candidate, the number of Twitter accounts they maintained during the campaign, the total number of tweets across all accounts, the candidate's political party, the incumbent party, whether the candidate was an incumbent, the candidate's victory or loss as well as their vote share, the candidate's gender and probable race, the candidate's average sentiment scores (positive, negative, and total), and the total number of tweets that fall into each of the machine learning categories. This dataset will be especially important in chapter 3, when I look at candidate coordination across content.

Appendix

Image 1. Coding Interface

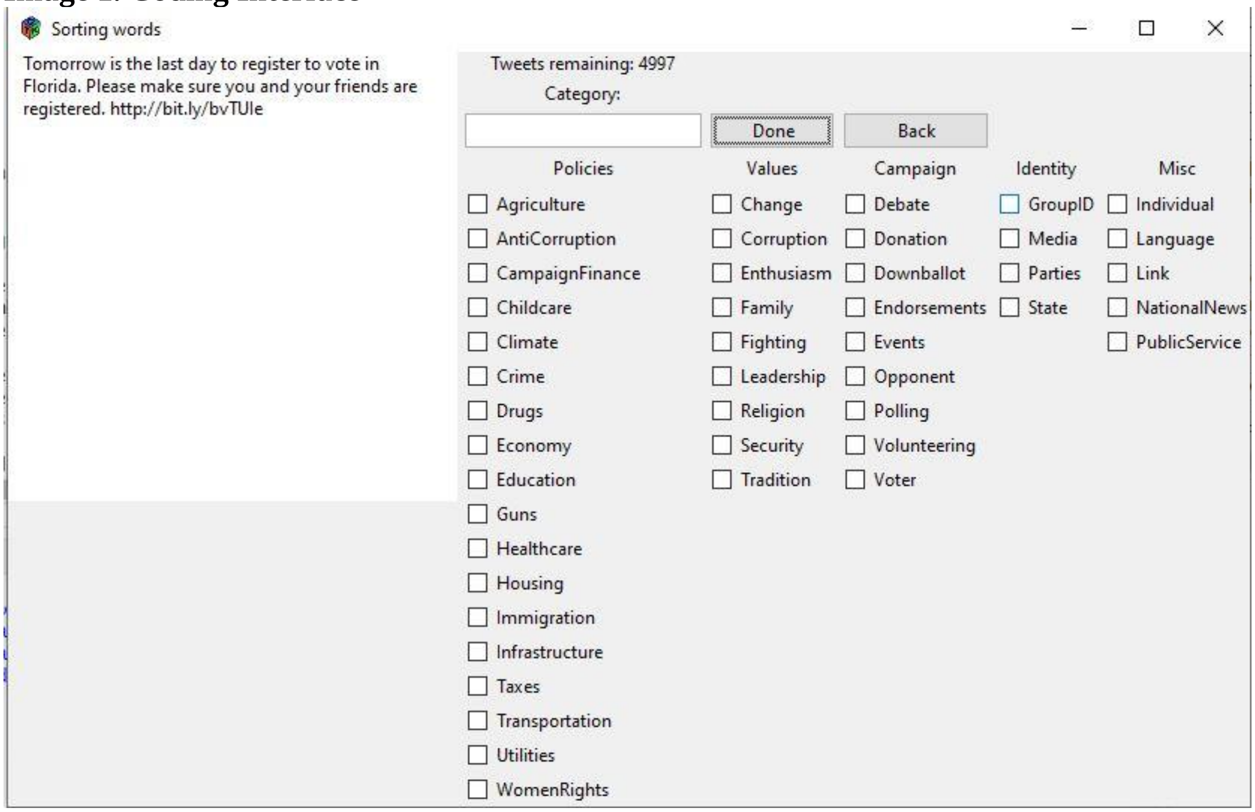


Table 1. Tweet Categories

Type	Category	Description
Policy		
	Agriculture	Agriculture and Farming policy
	Anti-Corruption	Gerrymandering, voter suppression, etc.
	Campaign Finance	Campaign finance policy
	Childcare	Childcare policy
	Climate	Climate change, disaster relief, environmental policy
	Crime	Crime, criminals, police policy
	Drugs	Drug crimes, drug treatment
	Economy	Jobs, employment/unemployment, inflation
	Education	Education policy
	Guns	Guns, gun safety, gun regulations
	Healthcare	Healthcare, disability services, mental health
	Housing	Housing costs, homelessness
	Immigration	Immigration policy
	Infrastructure	Roads, bridges, energy, internet policy
	Taxes	Tax rates, spending, deficits, budget management
	Transportation	Public transportation, trains, cars, cycling
	Utilities	Water, electricity, etc.
	Women’s Rights	Abortion, maternal mortality, pay gap, etc.
Values		
	Corruption	Lying, hiding, special interests, big donors, accountability

Change	Future, moving forward, change
Fighting	Struggle, standing up, fighting back, resist, strength
Religion	Mentions of specific religions or churches
Enthusiasm	Energy, enthusiasm, public support
Tradition	Tradition, referencing the past
Leadership	Vision, planning, ideas, decision-making
Family	Family values, showing their own family
Security	Safety, defending, protection, danger

Campaign

Rallies	Campaign event announcements
Donation	Asking for donations
Downballot	Using the platform to highlight others running for office
Endorsements	Public announcements of endorsements
Polling	Calls to public polling, public support, momentum
Volunteering	Showing volunteers or canvassing, asking followers to take action

Identity

Group ID	Any call to a specific group – farmers, parents, politicians, etc.
Media	Attempts to describe 'the media' as an identity
Parties	Party identification – "Democrats want to" or "Republicans are" etc.
State	State identification – "New Yorkers are…" etc.

	Individual	Calling out an individual voter or volunteer
	Language	Tweets in a foreign language
Misc		
	Public Service	Voting/polling place information, disaster announcements

Table 2. Machine Learning Accuracy

Category	Training#	Test#	Positive	Negative	Overall
Agriculture	39	21	.190	1.00	.979
AntiCorruption	15	8	.000	1.00	.979
CampaignFinance	1	1	NA	NA	NA
Childcare	5	3	.000	1.00	.979
Climate	69	15	.200	1.00	.976
Crime	60	21	.333	.991	.971
Drugs	19	5	.200	1.00	.978
Economy	475	119	.689	.967	.889
Education	253	67	.701	.979	.921
Guns	34	5	.800	1.00	.975
Healthcare	173	55	.618	.984	.944
Housing	19	8	.250	.997	.976
Immigration	28	8	.250	1.00	.977
Infrastructure	47	13	.077	.993	.977
Taxes	189	49	.857	.994	.932
Transportation	38	9	.222	.992	.971
Utilities	29	8	.000	.996	.978
WomensRights	47	13	.462	.992	.967
Change	116	16	.375	.989	.960
Corruption	158	40	.250	.973	.959
Enthusiasm	72	7	.571	.996	.970
Family	126	30	.433	.984	.947
Fighting	86	20	.350	.992	.967
Leadership	100	19	.684	.993	.959
Religion	20	5	.400	1.00	.977
Security	105	17	.529	.991	.958
Tradition	1	0	NA	NA	NA
Debate	127	18	.833	.996	.962
Donation	36	7	.286	.999	.976
Downballot	150	54	.204	.957	.951
Endorsements	178	27	.741	.996	.959
Events	888	223	.619	.909	.813
Opponent	417	103	.495	.942	.897
Polling	34	12	.417	1.00	.974
Volunteering	214	52	.673	.988	.942
Voter	1	0	NA	NA	NA
GroupID	415	106	.358	.925	.902
Media	147	42	.428	.978	.945
Parties	98	27	.407	.983	.953
State	108	15	.200	.988	.970
Individual	80	20	.000	.979	.977
Language	12	1	.000	1.00	.979
Public Service	228	56	.571	.978	.962

References

Ballantine, P. W., & Martin, B. A. (2005). Forming parasocial relationships in online communities. *ACR North American Advances*.

Bond, R. M., Fariss, C. J., Jones, J. J., Kramer, A. D., Marlow, C., Settle, J. E., & Fowler, J. H. (2012). A 61-million-person experiment in social influence and political mobilization. *Nature, 489*(7415), 295-298.

Bouchet-Valat, M., & Bouchet-Valat, M. M. (2015). Package ‘SnowballC’. In.

Carsey, T. M. (2009). *Campaign dynamics: The race for governor*: University of Michigan Press.

Dalager, J. K. (1996). Voters, issues, and elections: Are the candidates' messages getting through? *The Journal of Politics, 58*(2), 486-515.

Evans, H. K., Cordova, V., & Sipole, S. (2014). Twitter style: An analysis of how house candidates used Twitter in their 2012 campaigns. *PS, political science & politics, 47*(2), 454.

Flowers, J. F., Haynes, A. A., & Crespin, M. H. (2003). The media, the campaign, and the message. *American Journal of Political Science, 47*(2), 259-273.

Gainous, J., & Wagner, K. M. (2013). *Tweeting to power: The social media revolution in American politics*: Oxford University Press.

Gentry, J., Gentry, M. J., RSQLite, S., & Artistic, R. L. (2016). Package ‘twitteR’. *Cran. r-project*.

Glassman, M. (2010). *Social networking and constituent communication: Member use of Twitter during a two-week period in the 111th Congress*: DIANE Publishing.

Glassman, M. E., Straus, J. R., & Shogan, C. J. (2013). *Social networking and constituent communications: Members’ use of Twitter and Facebook during a two-month period in the 112th Congress.* Paper presented at the Washington DC: Congressional Research Service, CRS Report for Congress.

Golbeck, J., Grimes, J. M., & Rogers, A. (2010). Twitter use by the US Congress. *Journal of the American society for information science and technology, 61*(8), 1612-1621.

Graham, T., Broersma, M., Hazelhoff, K., & Van'T Haar, G. (2013). Between broadcasting political messages and interacting with voters: The use of Twitter during the 2010 UK general election campaign. *Information, communication & society, 16*(5), 692-716.

Haselmayer, M., Wagner, M., & Meyer, T. M. (2017). Partisan bias in message selection: Media gatekeeping of party press releases. *Political Communication, 34*(3), 367-384.

Hemphill, L., Culotta, A., & Heston, M. (2016). # Polar Scores: Measuring partisanship using social media content. *Journal of Information Technology & Politics, 13*(4), 365-377.

Hemphill, L., & Schöpke-Gonzalez, A. M. (2020). *Two Computational Models for Analyzing Political Attention in Social Media.* Paper presented at the Proceedings of the International AAAI Conference on Web and Social Media.

Hill, S. J., Lo, J., Vavreck, L., & Zaller, J. (2013). How quickly we forget: The duration of persuasion effects from mass communication. *Political Communication, 30*(4), 521-547.

Holbrook, T. M., & McClurg, S. D. (2005). The mobilization of core supporters: Campaigns, turnout, and electoral composition in United States presidential elections. *American Journal of Political Science, 49*(4), 689-703.

Huddy, L., Mason, L., & Aarøe, L. (2015). Expressive partisanship: Campaign involvement, political emotion, and partisan identity. *American Political Science Review, 109*(1), 1-17.

Imai, K., & Khanna, K. (2016). Improving ecological inference by predicting individual ethnicity from voter registration records. *Political Analysis*, 263-272.

Johnston, R., Blais, A., Brady, H. E., & Crête, J. (1992). *Letting the people decide: Dynamics of a Canadian election*: Stanford University Press.

Kalla, J. L., & Broockman, D. E. (2018). The minimal persuasive effects of campaign contact in general elections: Evidence from 49 field experiments. *American Political Science Review, 112*(1), 148-166.

LaMarre, H. L., & Suzuki-Lambrecht, Y. (2013). Tweeting democracy? Examining Twitter as an online public relations strategy for congressional campaigns'. *Public relations review, 39*(4), 360-368.

Lantz, B. (2013). *Machine learning with R*: Packt publishing ltd.

Lodge, M., Steenbergen, M. R., & Brau, S. (1995). The responsive voter: Campaign information and the dynamics of candidate evaluation. *American Political Science Review, 89*(2), 309-326.

McClurg, S. D. (2006). The electoral relevance of political talk: Examining disagreement and expertise effects in social networks on political participation. *American Journal of Political Science, 50*(3), 737-754.

McMillen, D. (2016). Moving Out, Moving In: The National Archives' Important Role When the Presidency Changes Hands. *Prologue Magazine, 48*.

Meyer, T. M., Haselmayer, M., & Wagner, M. (2020). Who gets into the papers? Party campaign messages and the media. *British Journal of Political Science, 50*(1), 281-302.

Nickerson, D. W. (2008). Is voting contagious? Evidence from two field experiments. *American Political Science Review*, 49-57.

Pew Research Center. (2019). *January 2019 Core Trends Survey*. Retrieved from Washington, D.C.: https://pewrsr.ch/2VxJuJ3

Pietryka, M. T., & DeBats, D. A. (2017). It's not just what you have, but who you know: Networks, social proximity to elites, and voting in state and local elections. *American Political Science Review*, 360-378.

Ridout, T. N., & Smith, G. R. (2008). Free advertising: How the media amplify campaign messages. *Political Research Quarterly, 61*(4), 598-608.

Rosenstone, S. J., & Hansen, J. M. (2001). Solving the puzzle of participation in electoral politics. *Controversies in voting behavior*, 69-82.

Russell, A. (2018). US senators on twitter: Asymmetric party rhetoric in 140 characters. *American Politics Research, 46*(4), 695-723.

Schulman, K. (2016). The Digital Transition: How the Presidential Transition Works in the Social Media Age. Retrieved from https://obamawhitehouse.archives.gov/blog/2016/10/31/digital-transition-how-presidential-transition-works-social-media-age

Shapiro, M. A., & Hemphill, L. (2017). Politicians and the policy agenda: Does use of Twitter by the US Congress direct New York Times content? *Policy & internet, 9*(1), 109-132.

Shogan, C. J. (2010). Blackberries, tweets, and YouTube: Technology and the future of communicating with Congress. *PS: Political Science and Politics*, 231-233.

Sides, J., & Karch, A. (2008). Messages that mobilize? Issue publics and the content of campaign advertising. *The Journal of Politics, 70*(2), 466-476.

Silge, J., Robinson, D. (2016). "tidytext: Text Mining and Analysis Using Tidy Data Principles in R. _*JOSS*_, *1*(3). doi:doi: 10.21105/joss.00037

Straus, J. R., Glassman, M. E., Shogan, C. J., & Smelcer, S. N. (2013). Communicating in 140 characters or less: Congressional adoption of Twitter in the 111th Congress. *PS, political science & politics, 46*(1), 60.

Stromer-Galley, J., Zhang, F., Hemsley, J., & Tanupabrungsun, S. (2018). Tweeting the attack: Predicting gubernatorial candidate attack messaging and its spread. *International Journal of Communication, 12*, 22.

Thomas, E. F., & McGarty, C. A. (2009). The role of efficacy and moral outrage norms in creating the potential for international development activism through group-based interaction. *British Journal of Social Psychology, 48*(1), 115-134.

Vavreck, L. (2009). The message matters. *The Economy and Presidential Campaigns, Princeton, Princeton, NJ*.

Vavreck, L. (2014). Want a Better Forecast? Measure the Campaign Not Just the Economy. *PS: Political Science & Politics, 47*(2), 345-347.

Verzani, J. (2014). GWidgets2 [Electronic resource]. *CRAN-Package gWidgets2*.

Yiannakas, G.

Yu-Wei, C. D. C. (2015). *Machine learning with R cookbook*: Packt Publishing Ltd.

Chapter 2

Social media provides a unique opportunity to get the unfiltered messages of political actors. Prior to the advent of Facebook and Twitter, candidates for office would have to rely on the media to accurately transmit their messages to their constituents. With the rise of social media, candidates now have a direct line to their supporters and can bypass the traditional media to deliver their thoughts directly. This makes social media uniquely useful for studying campaign messaging. Though Twitter enforces a rather strict character limit of 140 characters (increased to 280 characters in 2017) this can actually be a boon for the study of campaign messaging, as candidates are forced to distill their messaging into only the most important statements. This allows political scientists to determine what candidates consider the most effective use of direct contact with voters. Twitter data is still relatively uncommon in political science however, leaving room for further investigation into this direct form of candidate messaging.

Though Twitter-based analyses are somewhat rare, one established through-line of research is investigating the differences between political parties, both in the US and abroad. There are a number of reasons why party differences may shine through on a platform like Twitter. For example, research on polarization indicates that the parties have never been more different. Grossmann and Hopkins (2016) find that the American political parties behave as if they are different institutions. Republicans have polarized more than Democrats, retreating further from moderate ideologies and tightly clustering around a single conservative identity. Democrats, meanwhile, act as a coalition of various interests, leading to a wider distribution of in-party ideology. This polarization is associated with higher turnout as well as more engagement with party messages (Abramowitz & Saunders, 2008). Therefore, there should be a clear incentive for party members to drive discussion towards differences between the two parties and amplify polarization with their Twitter messages. Hetherington and Weiler (2018) find that this polarization has encompassed not just ideology, but identity - from the type of food one eats to the type of car one drives, to the way one sees the world. All of these have become wrapped up with partisanship, and so Twitter content may reflect party differences not just in policy, but in tone, strategy, and values. Many of these differences in candidate messaging would be difficult to find with traditional media which passes through journalists and editors before being published, but might be front and center on Twitter.

Early research on the platform hints at some of these party differences being present. Republicans in Congress adopted Twitter earlier than Democrats, registering earlier and posting more often (Gainous & Wagner, 2013; Straus, Glassman, Shogan, & Smelcer, 2013). Republicans also tended to be more negative, and used more partisan language (Straus, Williams, Shogan, & Glassman, 2016). However, the majority of this early content reflected more traditional press-release styles of candidate statements - links to news articles about themselves (often without commentary), updates on daily activities, and event/schedule announcements (Glassman, Straus, & Shogan, 2013; Golbeck, Grimes, & Rogers, 2010). This kind of top-down, non-interactive messaging is hardly what current users of the platform would think of as effective, quality use of its features. As time progresses, however, elected party members may have gotten more adept at using the platform, and the patterns we saw emerge in early research may change.

When candidates do make ideological statements on Twitter, what might we expect? Petrocik (1996) gives some insight with his issue ownership theory. He argues that in public perception, each party has some issues that they exclusively ‘own’, some problems that they are seen as better at solving. Effective campaigns and party messages, he says, should highlight issues that your party owns, and downplay issues that the other party owns. Democrats, he finds, ‘own’ healthcare, education, and other social welfare issues, while Republicans are seen as better dealing with the economy, with foreign defense, and with crime. Other scholars have since attempted to update our understanding of the issue ownership construct, and though there is some debate over whether the original question format truly measures what Petrocik thinks of as issue ownership (Therriault, 2015), the specifics are unimportant for this study. Instead, the underlying idea that parties have different strengths, and that messaging should reflect those strengths leads to a simple hypothesis - that the parties will talk about different issues.

There is some reason to believe that this hypothesis might be falsified. Koch (2008) argues that policies are multidimensional, and that there is room for parties to fight on different dimensions of the same policy. Jerit (2008) investigates this claim using the debate around the Clinton health care bill in 1993, and finds that not only did parties fight over the same policy area, but they argued over the same dimensions - they were actively engaged in debating and defeating each others’ points. If this reflects a wider pattern, then that should appear in this data, and requires an alternate policy hypothesis. In this case, the parties are talking about the same issues, but competing over framing (Jerit, 2008; Vavreck, 2014). If this line of research is correct, then we should expect the opposite of Petrocik’s theory - that parties will talk about the same issues. Though I don’t code the tenor of the exact arguments for or against policies, it’s safe to assume that if parties engage on the same issues, they are focused on creating different framing, or competing over ownership.

This project is hardly the first attempt to code and categorize the campaign messages of party actors. The expectations of this work should therefore be tempered by previous findings.

Glassman et al. (2013) look at tweets sent by Congresspeople in August and September of 2009. They find that the two most common categories of tweet were policy announcements and event announcements within the Congressperson's district or state. However, neither category represented more than 25% of tweets. During this time, 79% of all tweets in the dataset were sent by Republicans. Golbeck et al. (2010) find a similar pattern, with the most common category being information-providing, either about a policy or voting rules or locations. The expansiveness of the category led to over 50% of tweets being classified in this category. The second most common category, roughly 27% of all tweets, was highlighting local visits or events. Graham, Broersma, Hazelhoff, and Van'T Haar (2013), despite studying UK politicians rather than US Congress, find a similar pattern. The most common categories were campaign trail updates and campaign promotions, each representing around 20% of tweets in their dataset. Position-taking around policy was much less common than in the US studies, ~5.6%, and unlike the US, the ideological tint of the most frequent posters flipped, with Labor and the Liberal-Democrats posting more often than the Conservatives. Finally, in a more recent study of party behavior on Twitter, Hemphill, Culotta, and Heston (2016) use a machine learning model to code more than 1 million Congressional tweets, and find that Democrats are more likely to comment on policy generally, and are more likely to demonstrate a diversity of policy opinions. These findings can be summarized as follows: first, non-ideological campaign activities are as or more important than policy messaging. Second, neither of these categories make up a majority of a candidate's tweets. Third, differences in party behavior exist, but aren't clearly established across multiple studies.

Tweet content may not be the only place where party differences can occur. The tone of tweets may also differentiate the parties. The differences between positive and negative campaigning have been well-established (Auter & Fine, 2016; Benoit, 2004a, 2004b), and are often tied to the competitiveness of races. In fact, positioning your opponent as well as yourself is often considered a key component of campaign messaging (Geer & Vavreck, 2014; Spiliotes & Vavreck, 2002). Gainous and Wagner (2013) find that party ID correlates with the tone of messages posted, with Republicans sharing more negative, hostile content. Russell (2018) reinforces this finding, demonstrating that Republicans in the Senate are more likely to participate in negative partisan rhetoric.

Finally, the demographics of the platform itself must be considered. Only 22% of US adults use Twitter, and those are highly concentrated among a younger audience (Pew Research Center, 2019). Due to the way that Twitter networks work, users must self-select into following specific accounts. Therefore, it's likely that only a small percentage of Twitter users will self-select into political networks. Though Twitter offers direct contact with voters, the likelihood of reaching a large number of your *own* voters on Twitter is fairly low. Given that, the behavior incentives for Twitter messaging are more abstract, and more likely to rely on trying to "go viral" by making posts that draw the attention and coverage of traditional media. We might therefore expect patterns not to be unique and to be more consistent with prior research on political parties.

This chapter will focus on attempting to describe the ways that political parties use Twitter. Previous research has identified a series of places to look for party differences in T~~t~~witter use: first, the quantity of use. Previous research finds that Republicans tweet more than Democrats, though this could be an artifact of Republicans being the outparty in Congress at the time the studies were conducted. This dataset includes enough years and electoral contexts to determine whether this is a consistent pattern of party behavior. In addition, I include a preliminary analysis to determine whether language complexity differs between parties. Second, party differences may be found in user sentiment. Previous studies find that Republicans use more negative, more partisan language. I investigate this trend using both the frequency and severity of affective language. Third and finally, the content of the tweets. The studies mentioned above demonstrate that political tweets often include policy references and/or campaign updates. I add to this a further breakdown of these larger categories into specific policy areas and campaign tactics, and additionally include analyses of values-based language and attempts to activate social group identities. This expansive set of categories should provide a clearer image of what politicians consider the best use of social media.

Method

There are a number of places one might look to distinguish message differences between the parties. This chapter will look at Twitter adoption, the quantity of tweets from each party, the complexity of those tweets, the average tone or sentiment expressed by each party's tweets, and the content of each party's tweets. Because the datasets described in chapter 1 represent a nearly-complete set of gubernatorial tweets from 2008 to 2018, I analyze them as population data.

Many of the following subsections are investigated using analyses at two different levels of aggregation, because each represents one possible interpretation of party behavior on Twitter. First, at the level of individual tweets. At this level of aggregation, there is no difference between candidates – one party member's tweet is treated the same as another. Any analysis of party effects, therefore, is more strongly influenced by those who tweet the most. Since citizens turning to Twitter to discover a party's messaging are more likely to encounter and interact with the content of prolific tweeters, they are likely to believe the prolific tweeters represent the party brand. This is one reasonable interpretation of party messaging on Twitter – candidates who tweet the most drive public perception of the party's priorities. Second, at the level of candidate-year aggregates. At this level of aggregation, each candidate's behavior is averaged before showing party effects. This reflects a belief that the party's messaging priorities are the sum of each candidate's behavior. By using both approaches, I can test whether the two definitions of party behavior on Twitter result in different conclusions.

Twitter Adoption

Early Twitter studies of party behavior indicated party differences in Twitter adoption, with Republicans adopting earlier and tweeting more. Figure 2.1 shows the proportion of each party's candidates who posted at least one tweet from one account. Though Barack Obama had adopted Twitter in 2008, gubernatorial candidates did not seem to match his involvement with the platform until later. In 2008, just a single candidate from either party used Twitter. By 2010, however, about 75% of each party had adopted Twitter as a mode of communication. Republicans were slightly more likely to adopt Twitter than Democrats overall, though the differences are small (about 1-3 more candidates per election year). By 2016, over 90% of both parties' candidates had adopted Twitter.

How Much Do Candidates Tweet?

Previous research has also found that Republicans tweet more than Democrats. This data reflects those patterns. Table 1 includes both the means and the medians for the number of tweets sent by gubernatorial candidates from each party. Figure 2.2 offers a visualization of the distributions of these differences. The median candidate in each party tweets between 200 and 300 times between their primary election and general election. A few extremely prolific candidates seem to drag up the average.

Figure 2.3 indicates how these patterns change over time. By 2014, the average number of tweets in both parties increased to around 500. This indicates that members of both parties began to see Twitter as more of an opportunity for messaging. One interesting artifact of the data is that until 2016, very few candidates were prolific tweeters. However, since then, the tail of the distribution has become significantly wider, with a larger proportion of candidates using the platform extensively. With the exception of the bimodal distribution of Democratic tweets in 2016, the patterns exhibited by both parties is remarkably similar, though the average number of tweets for Republicans is almost always slightly higher than for Democrats.

How complex are candidate tweets?

There are a few different ways that this data would allow me to investigate linguistic complexity. However, I focus in on tweet length, in both characters and words. If candidates are consistently reaching the maximum character count per tweet, they are likely trying to send complex messages with as many talking points as they can fit. If tweets are consistently low in character or word count, it would indicate that candidates are attempting to send more simple, direct messages about individual issues. Figure 2.4 shows the distribution of tweet length by party across the entire time frame. These results have one caveat – some tweets are shown to exceed the character limit. Due to the nature of Twitter's encoding, some tweets have 'location tags'. When Twitter users opt in to Twitter's 'enable precise location' service, Twitter includes data from their device's GPS to indicate where the tweet was posted. This text exists outside of the formal character limit on tweets, but was included in the tweet's text when web scraping. There was no clear, systematic way to remove location tags without also removing text from the user's content, so these values are slightly inflated. Regardless, Figure 2.4 shows that Democrats (m = 123.12, sd = 63.96) post slightly longer tweets than Republicans (m = 113.18, sd = 58.36).

Because the maximum character limit changed in 2017, it is necessary to look at how these patterns change over time. Figure 2.5 demonstrates how these averages have changed since 2008. Both parties demonstrate a clear uptick in the length of tweet after the increased character limit in 2017, but the averages remain significantly below that threshold. Democrats continue to prefer slightly longer tweets than Republicans through most of the years in question. An additional point of interest when it comes to changes in tweet length over time is how Twitter's increasing of the character limit changed the distribution of tweet size. Figure 2.6 shows that prior to 2018, both parties peaked between 100 and 150 characters, indicating that many candidates were coming close to the character limit. In 2018, however, rather than the peak of the distribution moving to the right to accommodate more space, the distribution flattens. This suggests that instead of filling as much space as they're given, candidates prefer to send a wide array of message sizes. The cap of 140 characters was likely too limiting to accommodate these preferences, but a 280-character limit seems to free candidates to vary their tweet length.

As an added check, these analyses were run with word count as the measurement of length rather than character size (Figures 2.7 – 2.9), and with candidate-year averages as the unit of interest instead of individual tweets (Figures 2.10 – 2.12). The results are robust across each variation.

Party Sentiment

To investigate the tone of candidate tweets, I use two lexicons available in the "tidytext" package in R (Silge, 2016). Though methods exist to examine bigrams (two-word combinations) trigrams, and more complicated grammatical structures, I focus on unigram methods. That is, each of the two lexicons includes lists of individual words. Though this can lead to some miscodings where a word has one valence in isolation, but another valence as a part of a phrase, it's the simplest analysis to implement. The lexicons are not perfectly overlapping in which words are included, but each serves a useful purpose. The 'Bing' lexicon is categorical, rating each word as positive or negative, and includes 6,786 different words. The 'AFINN' lexicon is ordinal, rating each word from -5 (most negative) to +5 (most positive). It includes 2,477 words. Together, they allow me to examine not just the overall sentiment in a tweet (positive or negative), but the magnitude of that sentiment.

For each tweet, I create both a Bing score and an AFINN score. The Bing score is generated by subtracting the number of negative words from the number of positive words, and the AFINN score is generated by subtracting the total value of all negative words in the tweet from the total value of all positive words in the tweet. Because both measures are sensitive to the number of words in the tweet, I divide both by the tweet's word count. Table 2 shows the range, mean, and standard deviation for each measure at both levels of aggregation.

There are a number of places where party differences in sentiment may appear. First, there may be overall differences between the parties in average sentiment. Second, there may be election-cycle changes in party sentiment. Third, parties may differ in how much they use sentiment language – if a party is both more positive and more negative, averages may hide this difference. Fourth, there may be intra-election cycles in positive and negative language, where candidates alter their tone as their campaign progresses. Finally, there may be interactions between Party and other common independent variables like incumbency.

Overall Party Differences

Figures 2.13 and 2.14 show the Bing score at the tweet and candidate aggregation levels, respectively, while figures 2.15 and 2.16 show the results for the AFINN score. Two patterns worth noting appear. First, at the tweet level, it is clear that the most common level of sentiment for tweets is no sentiment at all. Many tweets are simply links to outside articles, videos, or images with no comment. Others have some text that does not include sentiment-based language, such as in Image 1. Second, the parties have not only nearly-identical means, but also nearly-identical distributions of sentiment language. Both parties prefer slightly positive messages overall, with tweets that are negative being a relative rarity. Though previous research finds tonal differences between the two parties, the finding doesn't seem to be reflected in gubernatorial accounts.

Sentiment Changes Over Time

One possible explanation for the above findings is that parties change their messaging style in different elections. If parties switch roles, from negative messaging styles to positive or vice versa, then overall party averages may mask this effect. Figures 2.17 – 2.20 explore this hypothesis. Figures 2.17 and 2.18 show the changes in average Bing score at the tweet and candidate aggregation levels, respectively. It's clear from Figure 2.17 that there is some evidence of tonal shifts between elections, with parties alternating between more positive and more negative language, though the differences are not substantively large. When aggregating at the candidate level, the cycles appear less rapid, taking multiple election cycles to occur, but they don't disappear. Figures 2.19 and 2.20 examine the same patterns for the AFINN score, and the same results are present. It's clear that parties do shift their messaging based on electoral context, though the effects are relatively small. That the cycles lengthen when candidate-year averages are the unit of interest suggests that those most likely to participate in this sentiment cycling are those who tweet the most.

Positive/Negative Sentiment

One place where parties might differ is in their use of positive and negative language separately. If one party uses both more positive and more negative language, then the averages will look the same even though the substance is different. Therefore, it's worthwhile to look at the distributions of positive and negative language between parties. For this subsection, I calculate the Bing and AFINN scores for each tweet using only positive or negative language.

Figures 2.21 and 2.22 examine the positive and negative Bing scores at the candidate level of aggregation. The results indicate that parties are remarkably similar in their use of positive and negative language. Both parties use substantially more positive than negative language, and a demonstrate a wider distribution of positive than negative language. Though not shown, these results are robust whether I aggregate at the tweet level or whether I use AFINN scores instead of Bing scores. Figures 2.23 and 2.24 demonstrate that the overall use of positive and negative language does not substantially change over time.

Sentiment Within Elections

Another avenue of investigation is whether parties have different approaches to messaging within election cycles. Some previous work has found that advertising has different effects on vote choice and vote intention at different points of the election cycle (Krupnikov, 2011). This work suggests that towards the end of a campaign cycle, voters have decided who they would like to support and further messaging only influences whether or not they plan to vote. Late in the election cycle, then, we might expect a shift towards more negative language designed to dissuade opponent's voters or drive engagement among one's own voters. For this analysis, I use the dates tweets were posted to calculate how many days remain until the general election. Figures 2.25 and 2.26 show how candidates' Bing and AFINN scores change as election day approaches.

Both parties appear to alter their tone over the course of a campaign by shifting between more positive and more negative messaging. Both parties demonstrate a drop in positive language around 2 months out from election day, and then a sharp uptick in positive language in the final week or two of the campaign. Republicans, especially, spend the middle of the campaign using much more positive language before their shift to more negative language.

Incumbency

In an ideal world, I would use an array of standard covariates to examine subgroup differences in tweeting behavior. However, in this dataset, only 43 candidates out of 314 are women, with 31 of those being Democrats, and only 16 candidates are non-white. The groups simply aren't large enough to make meaningful comparisons. As with the presidency, candidates for governor tend to be overwhelmingly white and male.

One common covariate that I can examine is incumbency. Figures 2.27 and 2.28 look at the distribution of sentiment across party and incumbency. There is little of note in these figures, except that Republican incumbents appear slightly more positive on the whole than the other three groups. Overall, there does not seem to be a significant interaction between incumbency and messaging tone.

Tweet Content

It's important to look not just at how candidates tweet and the tone they use, but what they choose to tweet about. To investigate party differences in tweet content, I use the classification categories described in Chapter 1. Classifications are broken into four over-arching categories: policy, values, campaign tactics, and group identification. Each of these categories is broken into more specific categories, like specific policy domains or campaign strategies. These categories are not exclusive – if, for example, a tweet refers to both agriculture policy and farmers as a group, it is classified both as policy and a group appeal. In some of the below analyses, I also add public service announcements, which include things like hurricane warnings, information for when and how to vote, and other such things. I begin by looking at how many of these categories candidates attempt to address per tweet, and then investigate which types of tweets candidates make a priority.

How Dense are Candidates' Tweets?

First, I look at how many categories candidates try to fit into each tweet. Figure 2.29 shows a histogram which indicates that both parties mostly favor tweets that include either a single category that I code for or none of them. For context, 30.76% of all tweets belong to none of the coded categories. The average number of categories in a tweet is 1.16, with Democrats slightly above average (m = 1.20) and Republicans slightly below average (m = 1.13). Figure 2.30 demonstrates that these trends show some change over time, though the effect is not dramatic, even after the character limit doubles in 2017.

On the whole, candidates seem to be focusing on singular issues in each tweet rather than addressing multiple topics and filling tweets with as many messaging points as possible.

What Are Candidates Discussing?

Figures 2.31 and 2.32 show what proportion of all tweets belong to each meta-category for Democrats and Republicans, respectively. I examine the relative proportions of sub-categories in later sections of this chapter. Of the tweets which belong to any of the categories I use, the majority are campaign tweets, including calls for supporters to appear at a rally, or asking followers to volunteer for phone banks or canvassing. These make up between 30 and 40% of all tweets, depending on election year. The next most common category is policy, then group appeals and values language. These patterns are the same for both parties, and across all years, with the exception of Democrats in 2017 and 2018, where policy messaging ties or beats campaign messaging in overall importance.

Because these patterns might be driven by those who tweet constantly, and might not be representative of party members as whole, I also examine the candidate level of aggregation in Figures 2.33 and 2.34. The overall ordering of meta-categories doesn't change. However, this makes clearer the increase for both parties in policy, group, and values messaging following the character limit increase in 2017. Given the increased number of average categories per tweet outlined above, this is expected, though there is no corresponding increase in campaign tweets.

Campaign Tweets

Figures 2.35 and 2.36 show the proportion of all tweets in each of the campaign subcategories. The most common types of campaign messages for both parties are event notifications for rallies, appearances, and so on. The next most common strategy is attempting to position one's opponent by calling out their behavior, or attacking their character or policy priorities. The other campaign strategies identified are relatively uncommon, including alerting followers to an upcoming or ongoing debate, encouraging followers to volunteer or donate to the campaign, or drawing attention to announced endorsements. Figures 2.37 and 2.38 show the results for the candidate level of aggregation, but the outcomes are mostly unchanged, with the exception that there is a larger gap between event notifications and the next most common strategy. This suggests that the most prolific posters are more likely to attack their opponent.

Policy Tweets

Figures 2.39 and 2.40 outline Democratic and Republican policy priorities at the tweet aggregation level. Both parties highly prioritize messages about the economy, which includes language about jobs, inflation, and unemployment, among other topics. After that, some party differences in policy preferences appear. Democrats have generally favored education as a secondary priority with a sharp increase in healthcare discussion after 2016. Republicans, meanwhile, also discuss education, but talk about taxes substantially more than Democrats. Other issues, like gun policy or women's rights and abortion, though they might appear on any issue ownership list, don't appear to be a common topic for gubernatorial tweets. Figures 2.41 and 2.42 demonstrate that these patterns hold regardless of which level of aggregation I use.

Values Tweets

Figures 2.43 and 2.44 show Democratic and Republican values messaging. The ordering of these values is highly variable over time, with no clear and stable party preferences. Figures 2.45 and 2.46 expand this analysis to the candidate level of aggregation, and still no stable party preferences are revealed. Values language seems to be highly related to election year. The one clear pattern is that the amount of values language used increased in both parties after 2016. Whether this is due to the changing messaging priorities of the Trump era or an artifact of the increased character limit of Twitter allowing candidates to create more elaborative messages is unclear.

Group Tweets

Finally, we can look at the sub-categories for group identification. Figures 2.47 through 2.50 examine Democratic and Republican group identity messaging priorities at both the tweet and candidate levels of aggregation. We might expect that party members would appeal to loyalists using party identity cues – "Democrats should…", "Republicans are…" etc. However, party ID cues represent less than 2% of overall tweets. Gubernatorial candidates are state-wide elections, so we might expect appeals to a state identity, but this is no more common than party identity. Instead, the most common identity messaging for both parties revolves around highlighting specific, political-adjacent identities – parents, union members, farmers, teachers. This ties into existing polarization research, which argues that Americans' political identities are expanding, encompassing previously non-political identities and tying them to politics (Hetherington & Weiler, 2018; Mason, 2018). Here we have some evidence that gubernatorial candidates are explicitly trying to use these politically-adjacent identities to garner political support.

Discussion

This chapter examines many different facets of gubernatorial tweeting behavior to determine whether differences between the parties exist or not. The evidence is mixed. On the one hand, there seems to be no substantial difference between parties in Twitter adoption rates, the complexity of their tweets, in their use of positive and negative language or their overall sentiment. There is no difference between parties in how much gubernatorial candidates tweet about each of the meta-categories, or even in most of the subcategories. The few party differences that exist are limited to:

1) Republicans tweet more than Democrats
2) Republicans show wider sentiment shifts within elections
3) Democrats' tweets are slightly more category-dense than Republicans'
4) Democrats tweet more about healthcare, while Republicans tweet more about taxes

Other than this short list, there are remarkably few differences between Democrats and Republicans. This flies in the face of much of the previous research, and there are a couple possible reasons why differences may fail to appear in this context. First, much of the previous messaging research, especially that which focuses on Twitter as a platform, focuses on members of Congress. Gubernatorial candidates may be more isolated from the national polarization fight, and may use the platform differently. Second, there may be nuances of language that are missed by the specific measures of both sentiment and categorization that I use. A perfect assessment of this hypothesis would rely on hand-coding each of the over 100,000 tweets in this dataset. Finally, these results may just be evidence in favor of the line of reasoning outlined in Koch (2008), Jerit (2008) and Vavreck (2014) – that candidates fight on different dimensions of the same policies and strategies, in which case many of the party differences I test for just don't exist.

Tables

Table 1. Tweet Quantity by Party

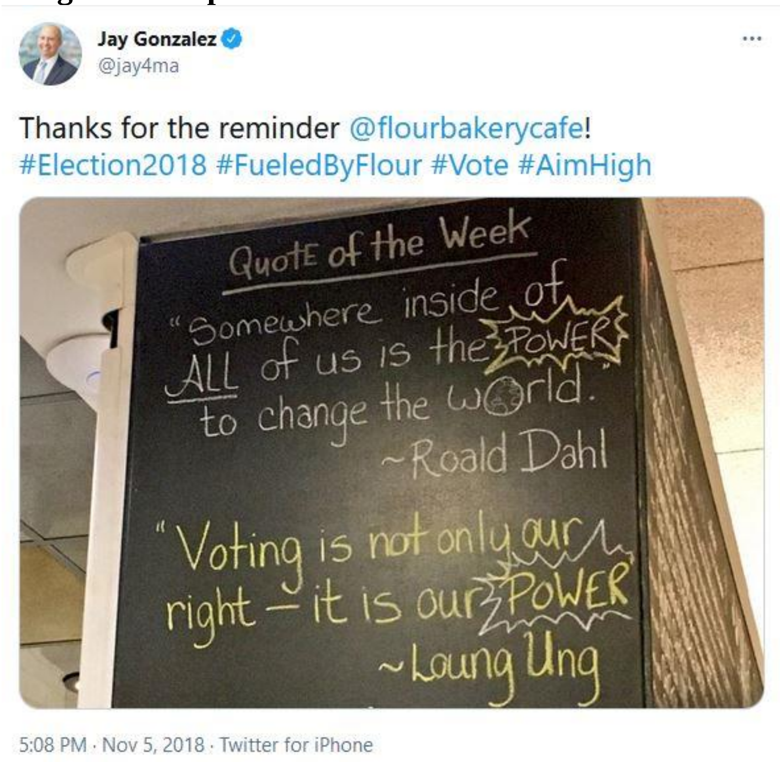
	Mean	Median
Democrats	330.09	214
Republicans	378.81	270

Table 2. Bing & AFINN scores

Aggregation	Min	Max	Mean	SD
Candidate				
Bing	- 0.028	0.117	0.033	0.017
AFINN	-0.022	0.315	0.091	0.046
Tweet				
Bing	-1	1	0.032	0.069
AFINN	-3	4	0.084	0.164

Images

Image 1. Example Tweet



Figures

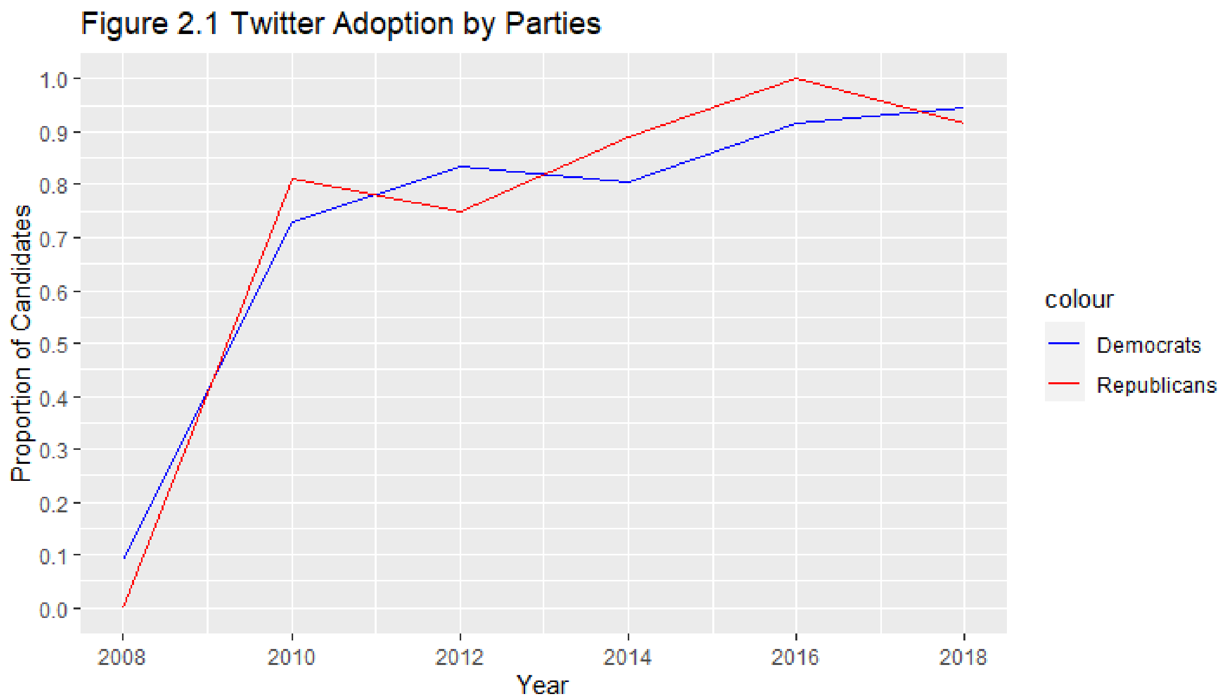


Figure 2.1 Twitter Adoption by Parties
Proportion of Candidates
1.0
0.9
0.8
0.7
0.6
0.5
0.4
0.3
0.2
0.1
0.0
2008
2010
2012
2014
2016
2018
Year
colour
Democrats
Republicans

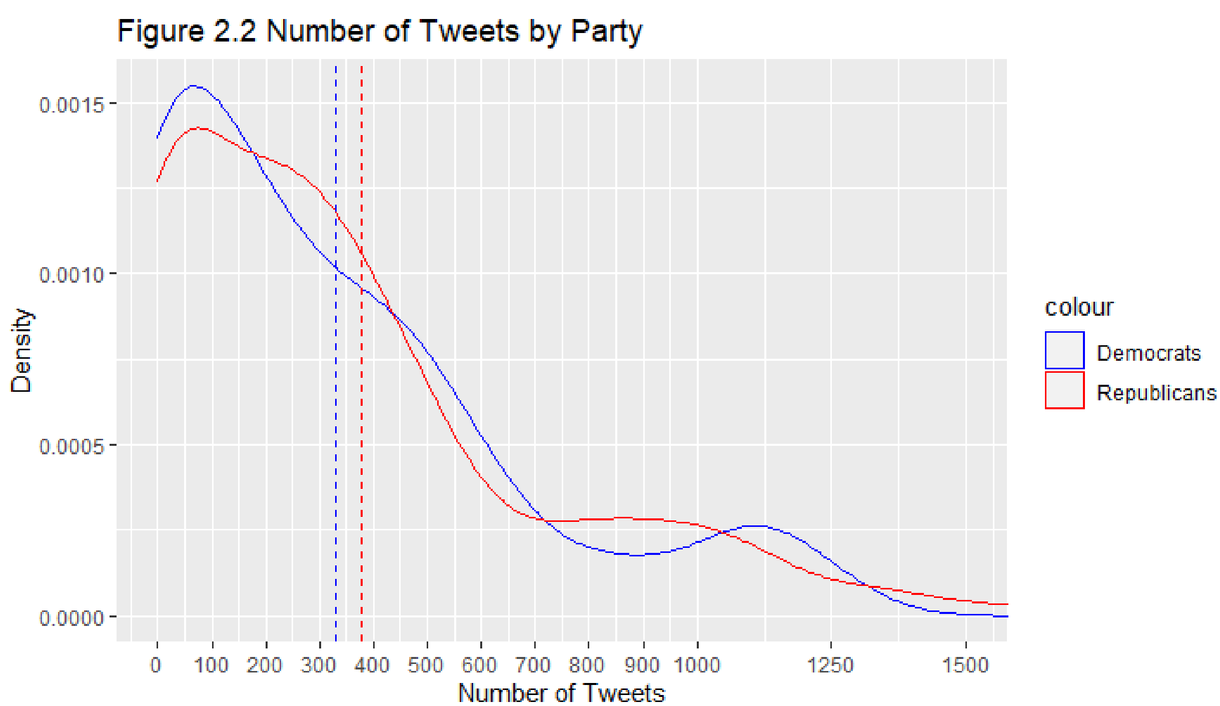


Figure 2.2 Number of Tweets by Party
Density
0.0015
0.0010
0.0005
0.0000
0
100
200
300
400
500
600
700
800
900
1000
1250
1500
Number of Tweets
colour
Democrats
Republicans

Figure 2.3 Number of Tweets by Party and Year

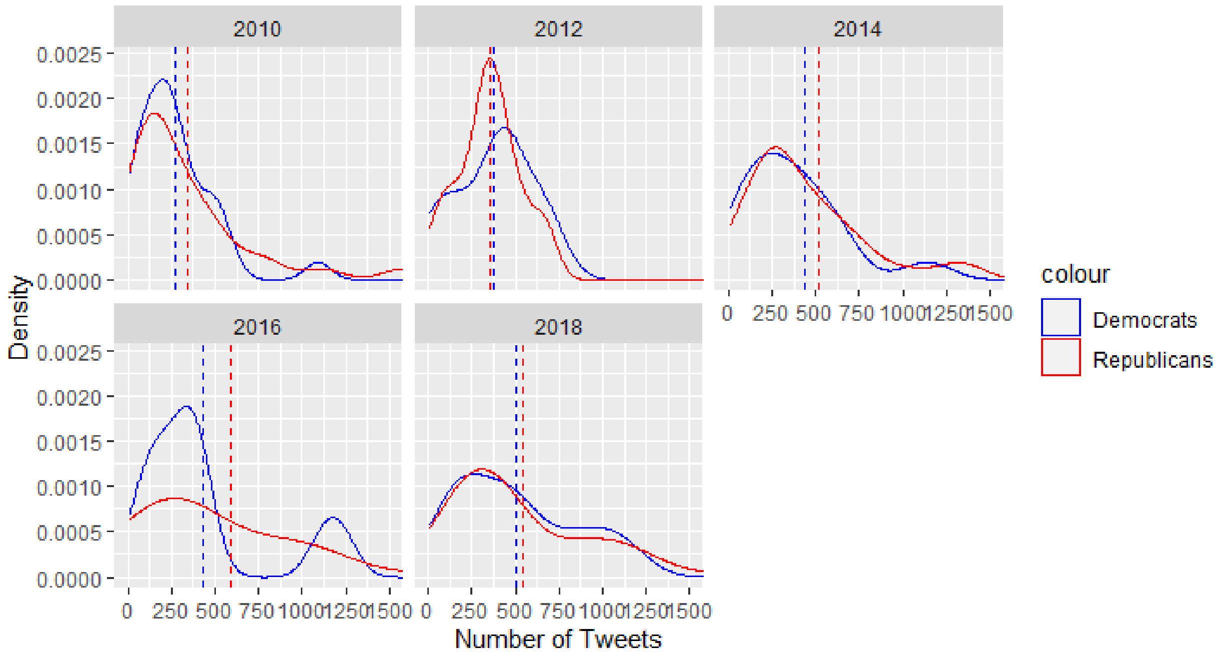


Figure 2.4 Number of Characters per Tweet

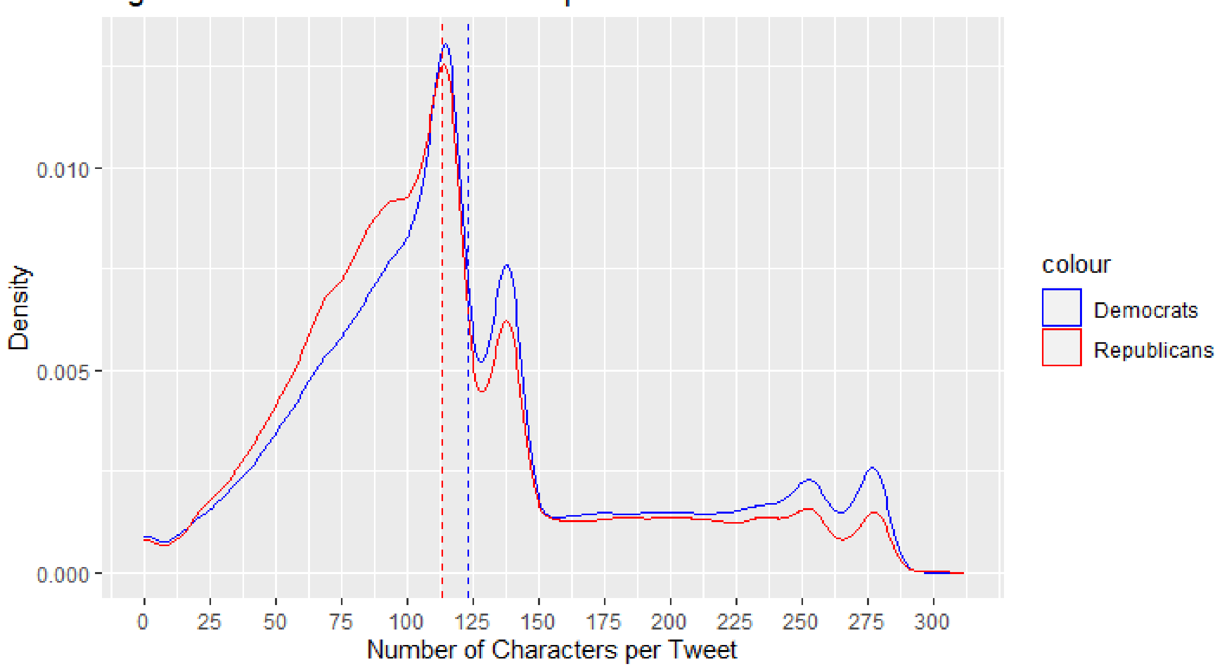


Figure 2.5 Number of Characters per Tweet by Year

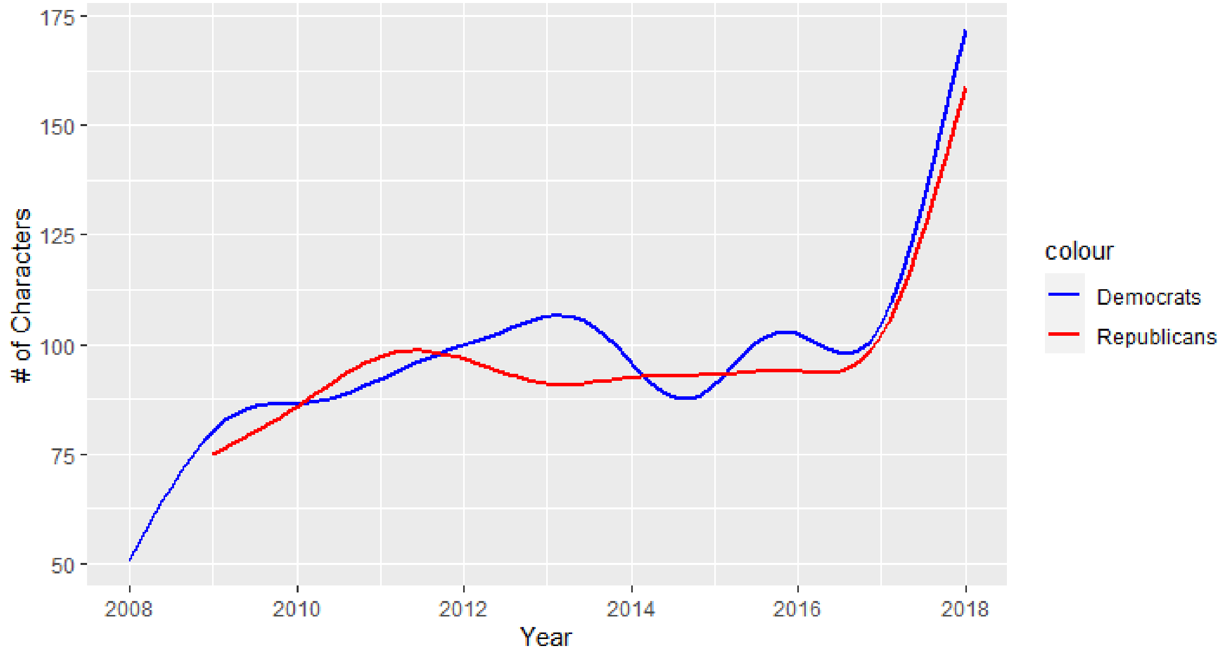


Figure 2.6 Distribution of Tweet Length over time

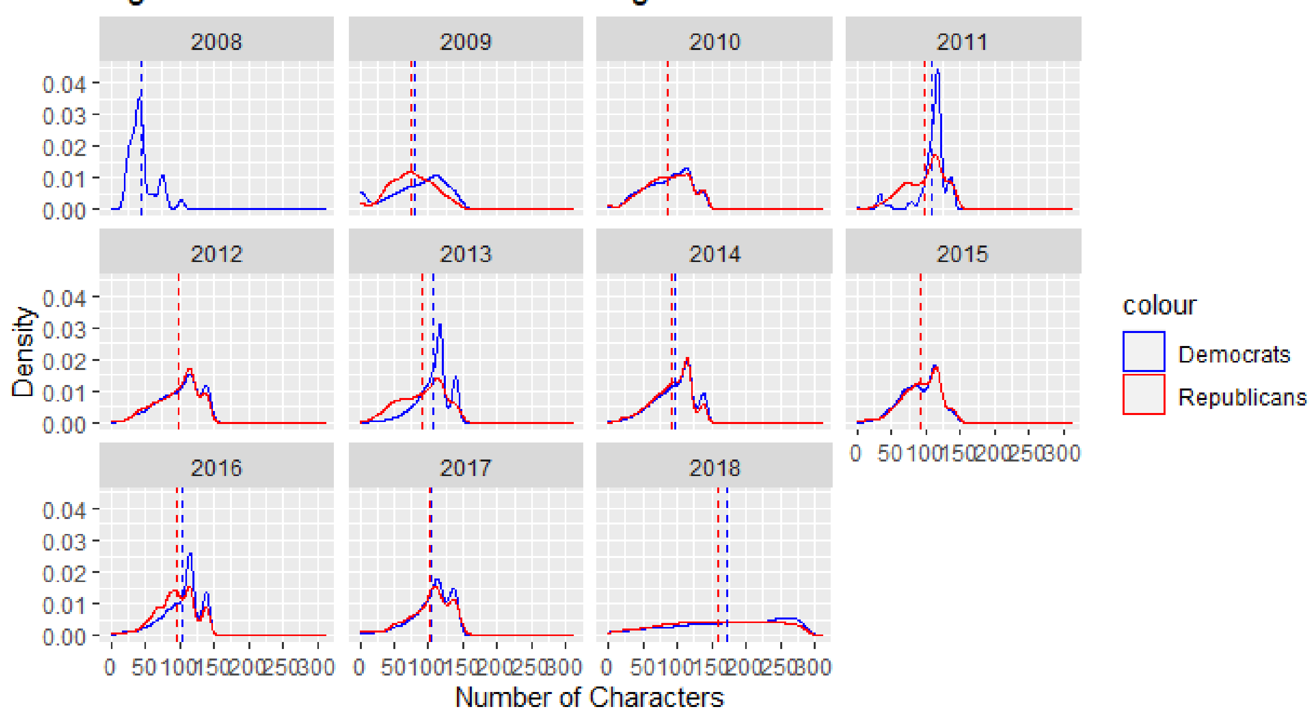


Figure 2.7 Number of Words per Tweet

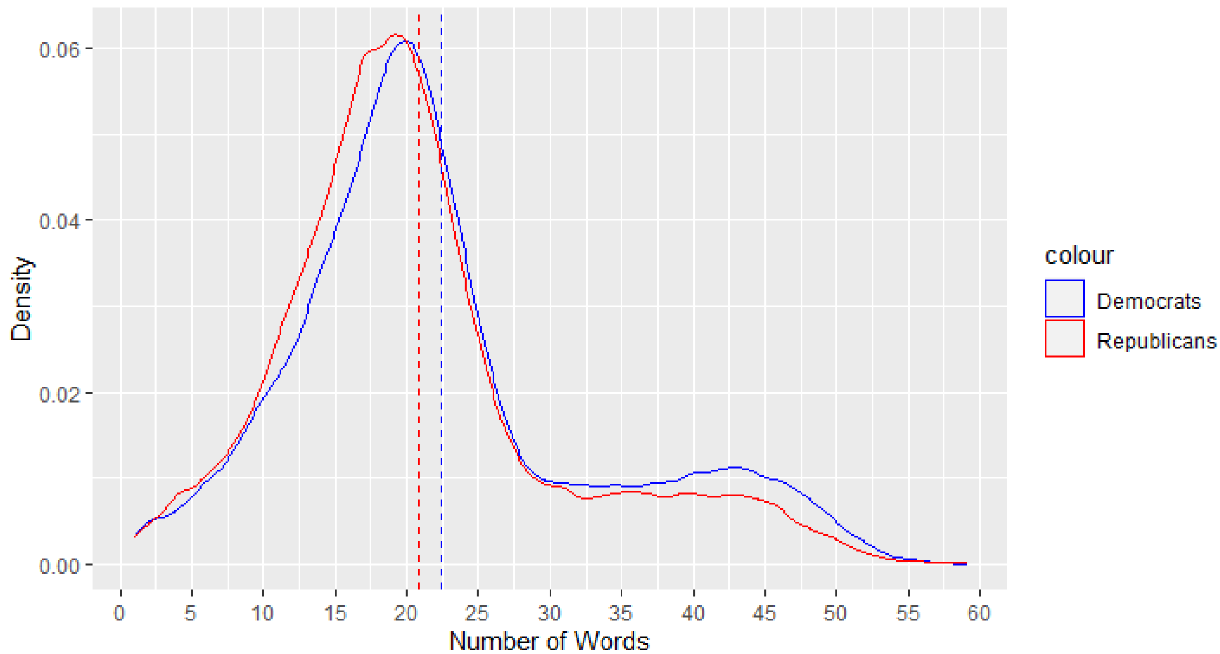


Figure 2.8 Words Per Tweet

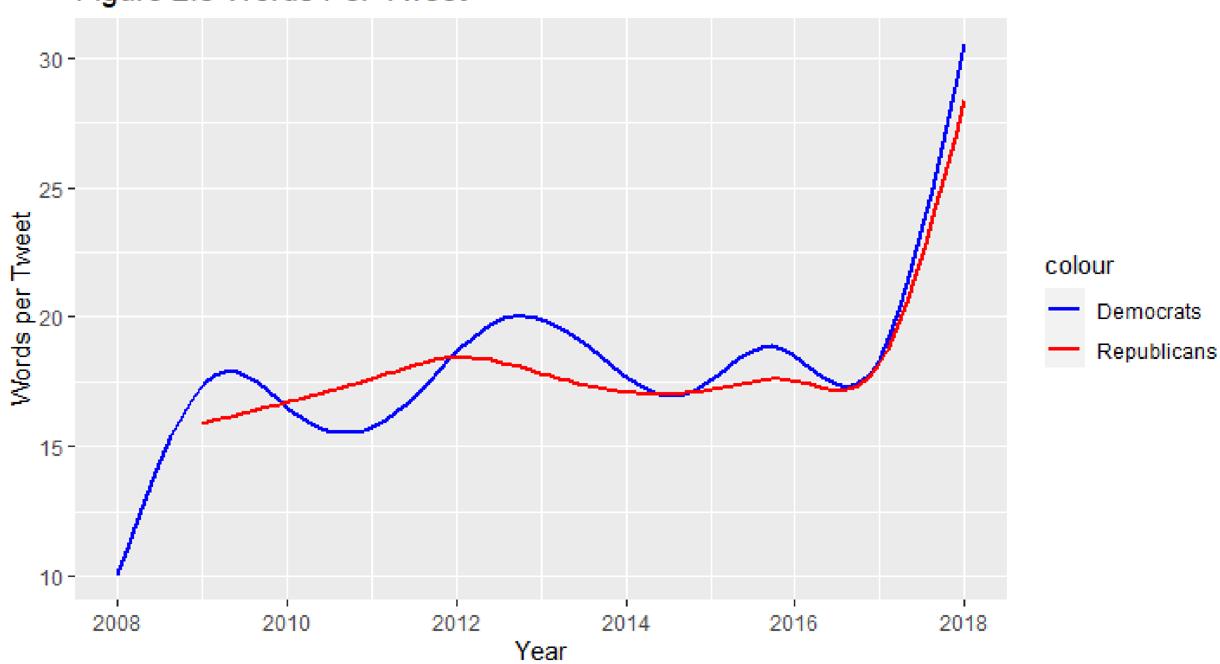


Figure 2.9 Words per Tweet by Year

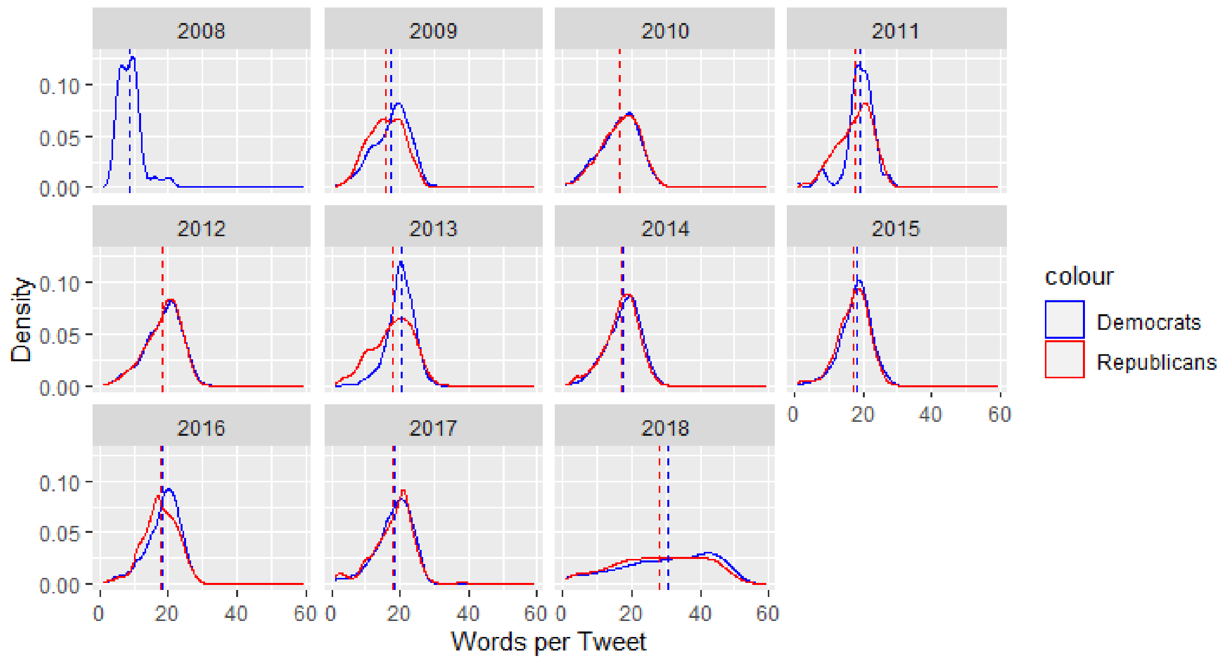


Figure 2.10 Characters per Tweet - Aggregate

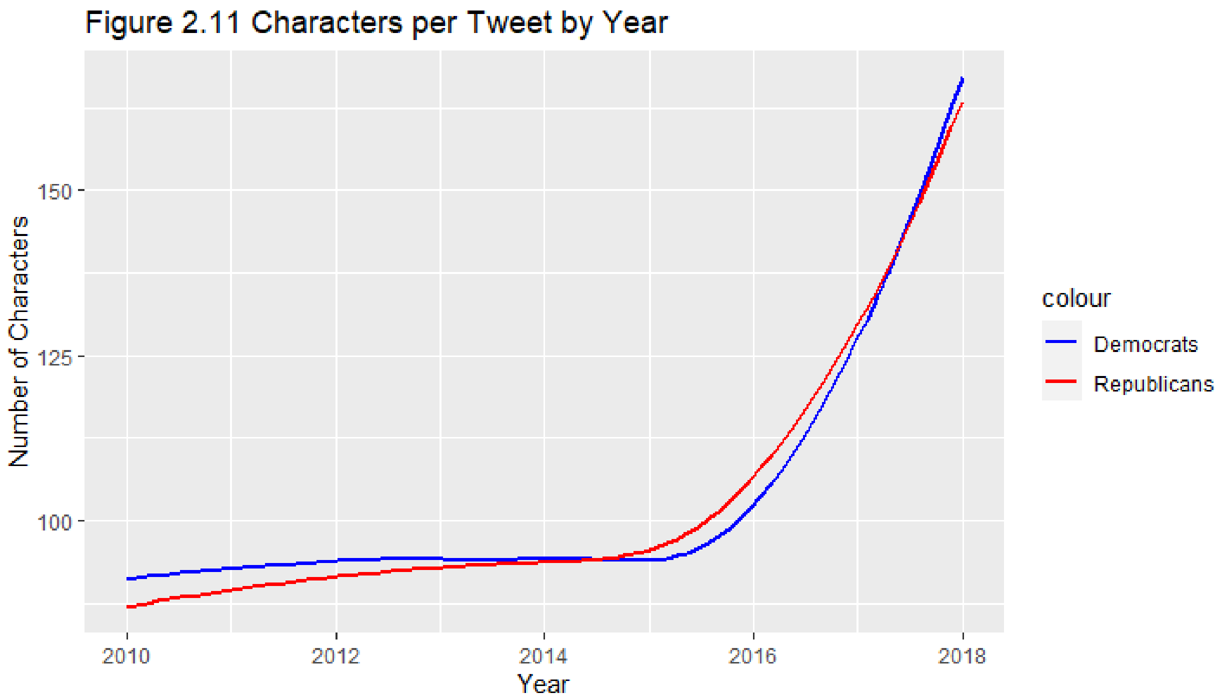


Figure 2.11 Characters per Tweet by Year
Number of Characters
150
125
100
2010
2012
2014
2016
2018
Year
colour
Democrats
Republicans

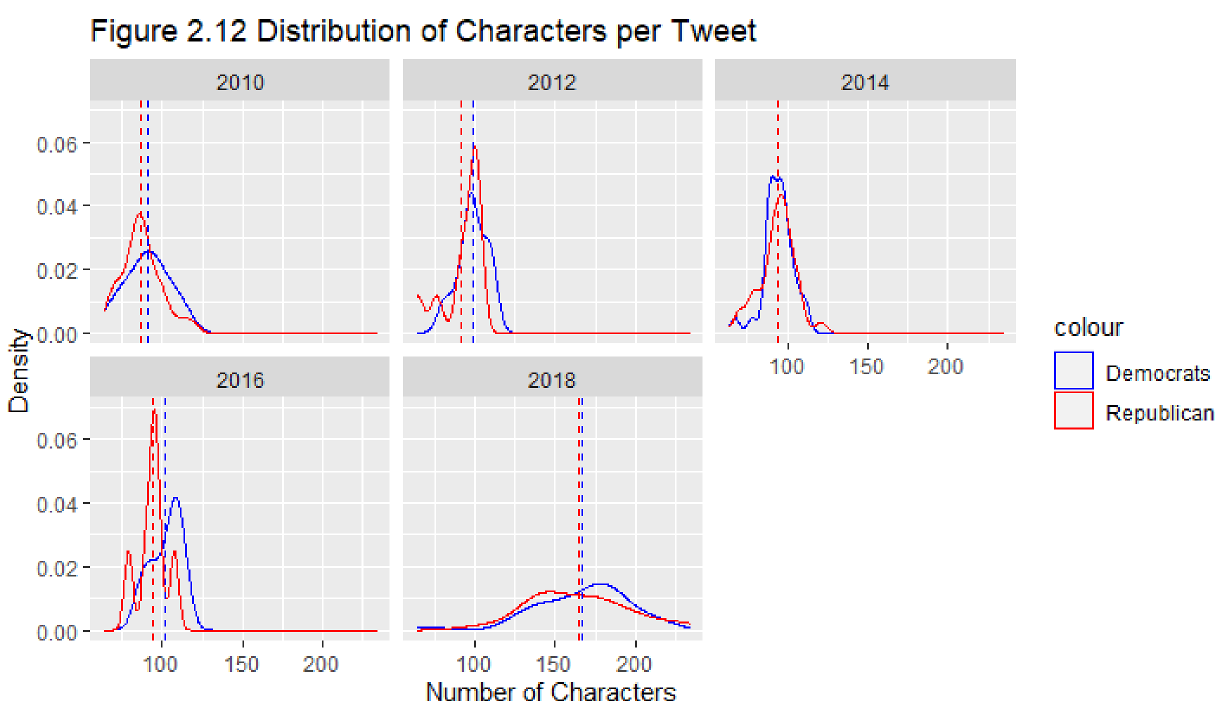


Figure 2.12 Distribution of Characters per Tweet
2010
2012
2014
2016
2018
Density
0.06
0.04
0.02
0.00
100
150
200
Number of Characters
colour
Democrats
Republican

Figure 2.13 - Bing Score by Party

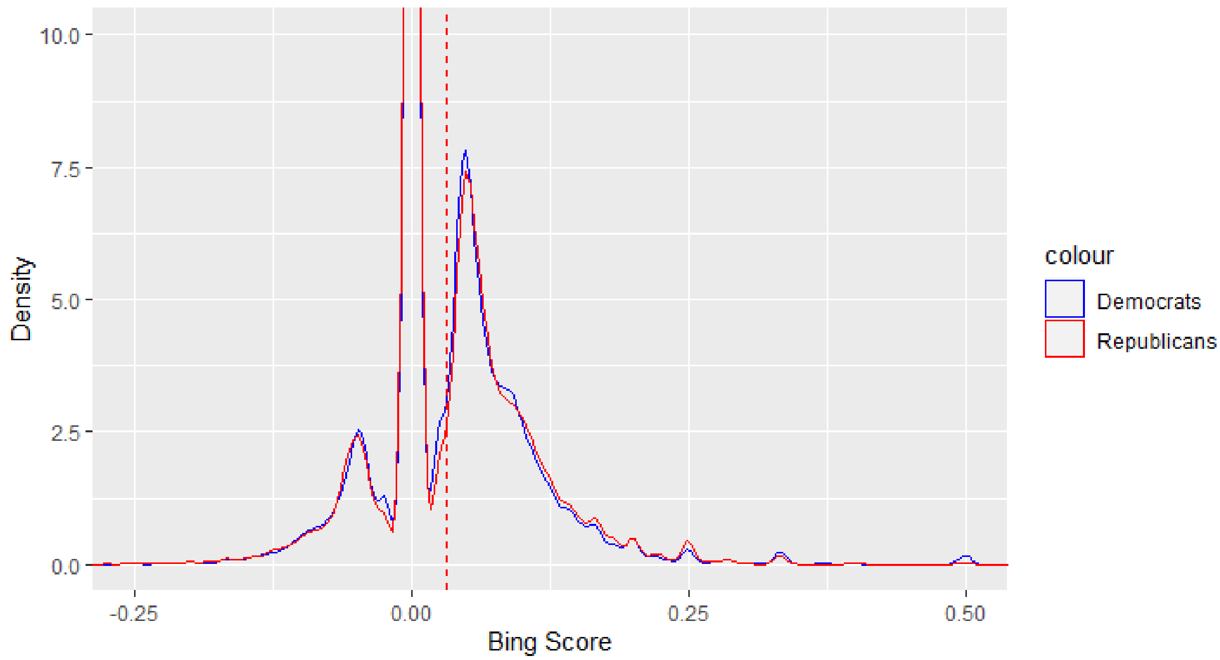


Figure 2.14 - Bing Score by Party - Aggregate

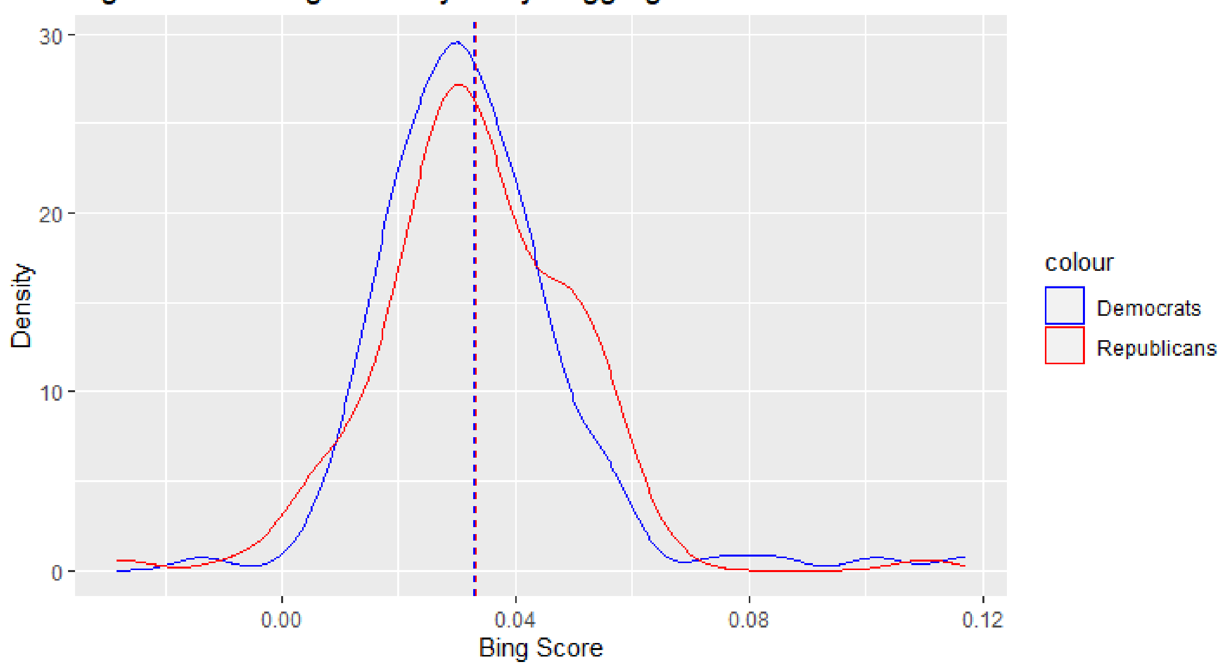


Figure 2.15 - Afinn Score by Party

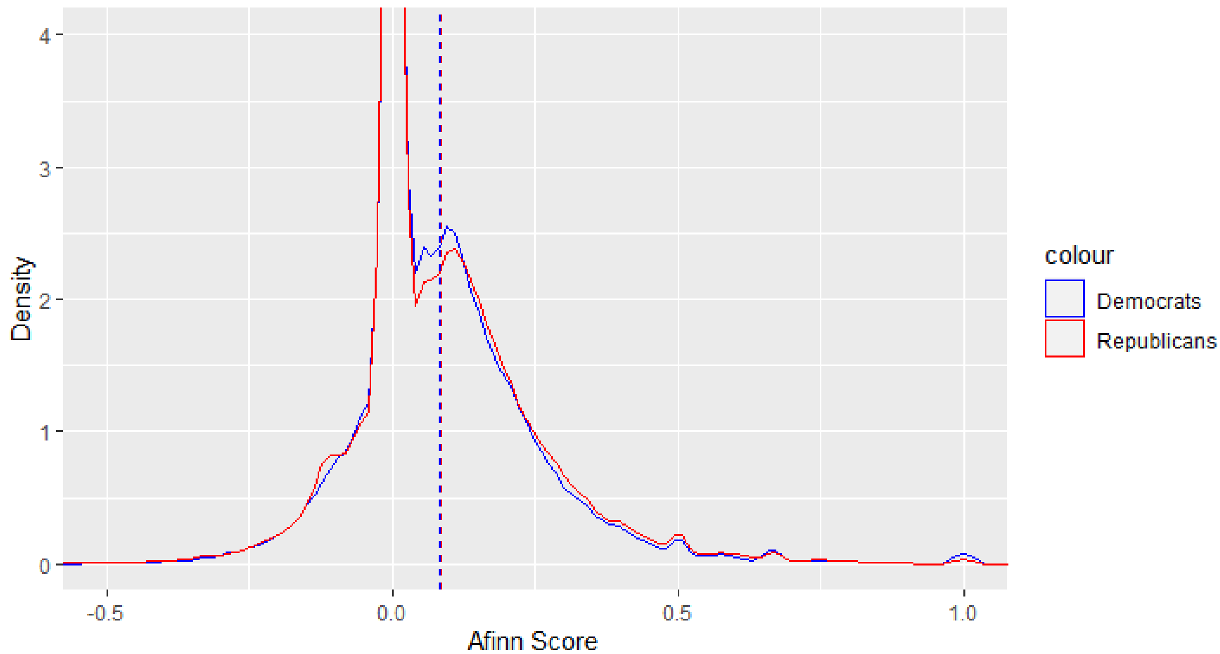


Figure 2.16 - Afinn Score by Party - Aggregate

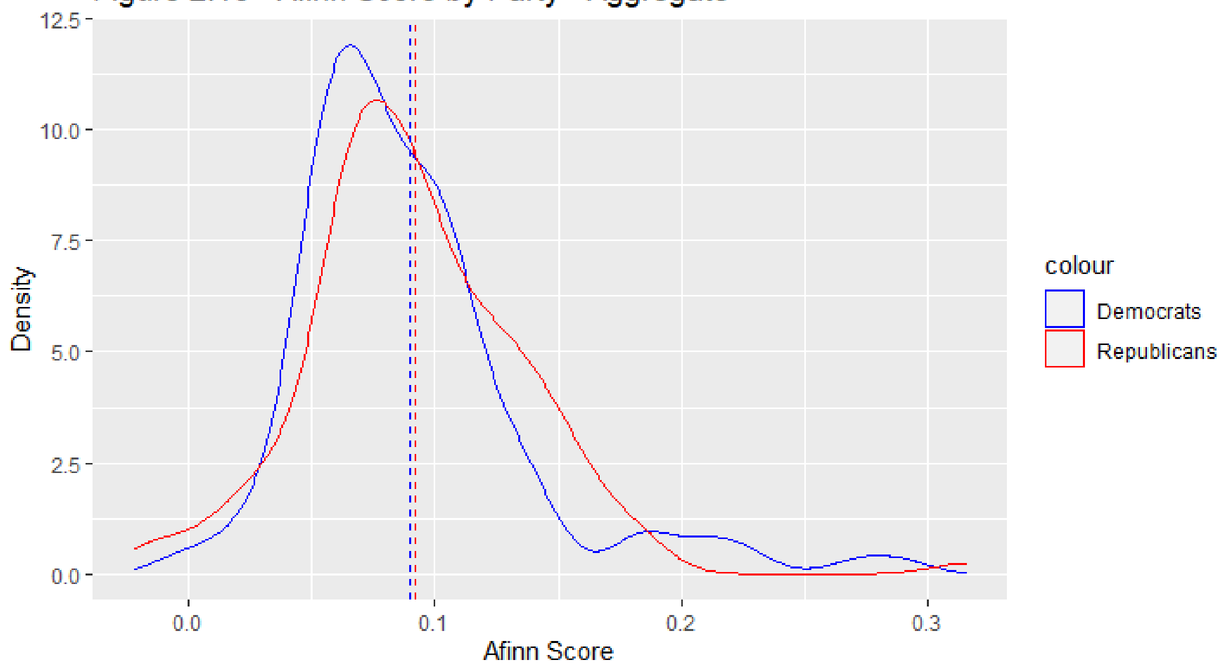


Figure 2.17 - Bing Score by Year

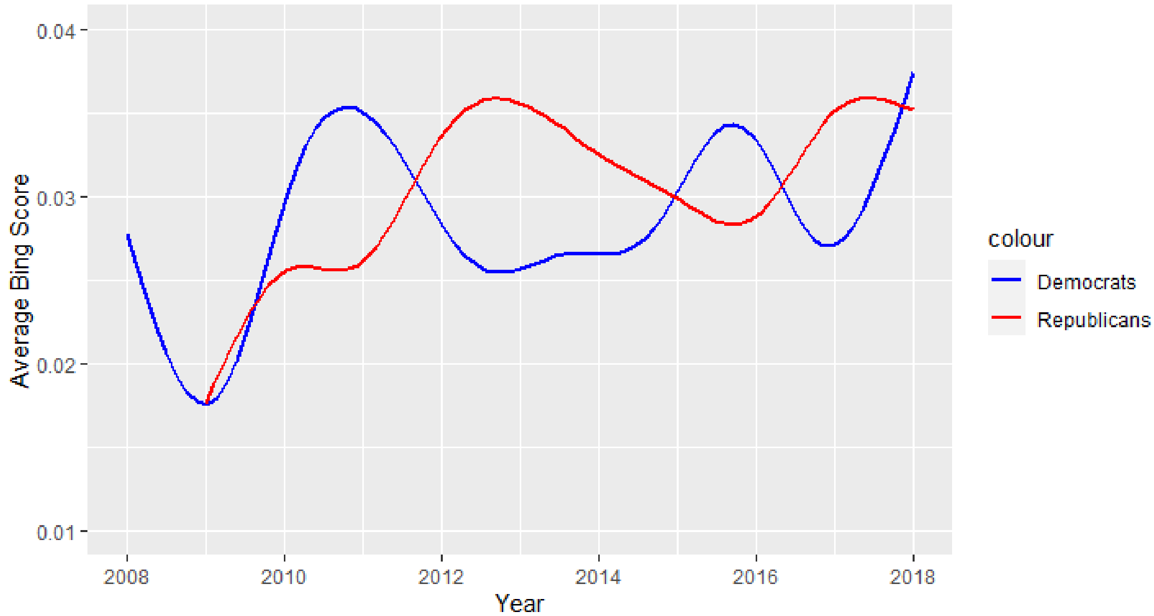


Figure 2.18 - Bing Score by Year - Aggregate

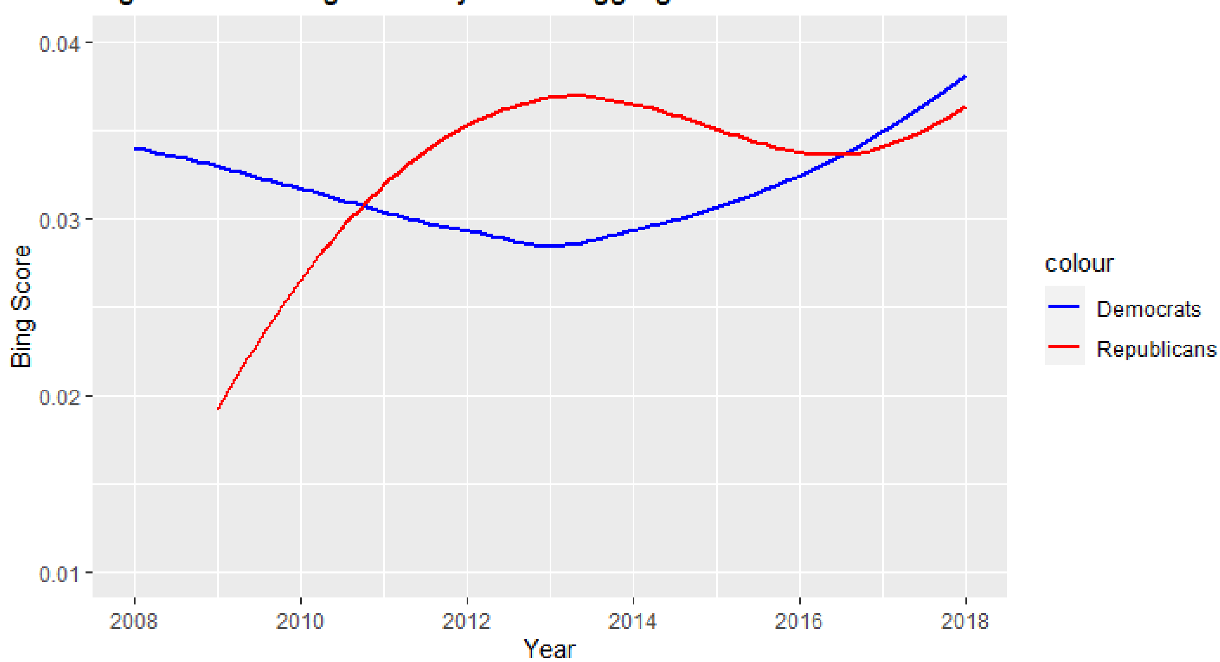


Figure 2.19 - Afinn Score by Year

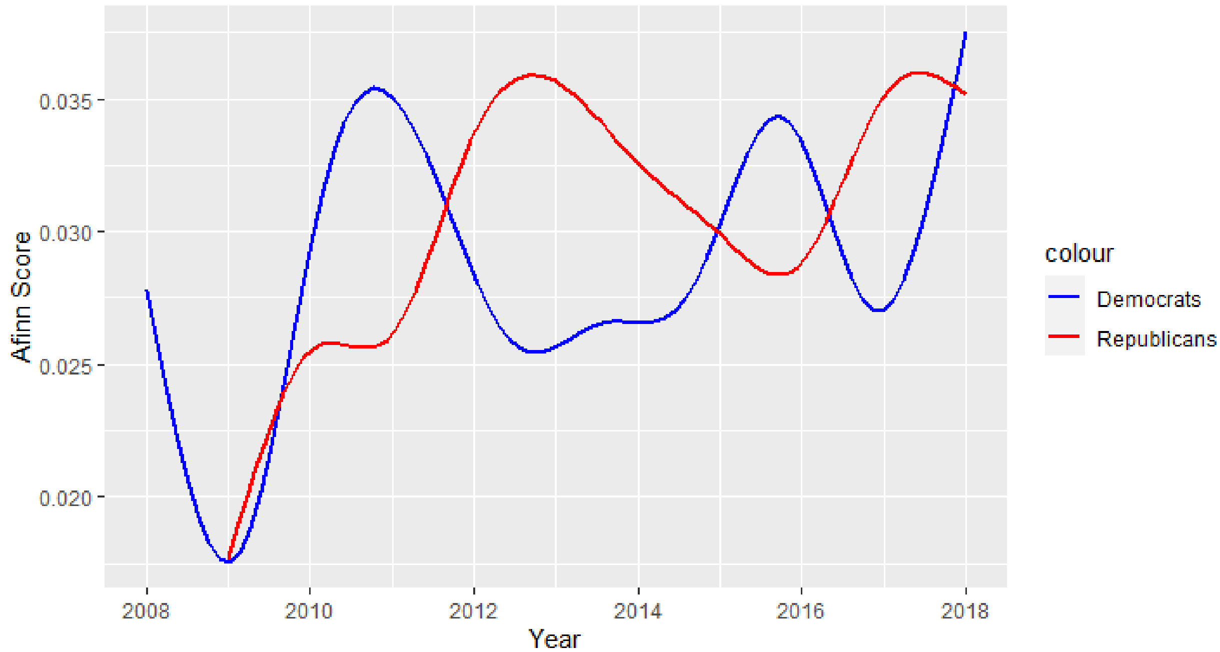


Figure 2.20 - Afinn Score by Year - Aggregate

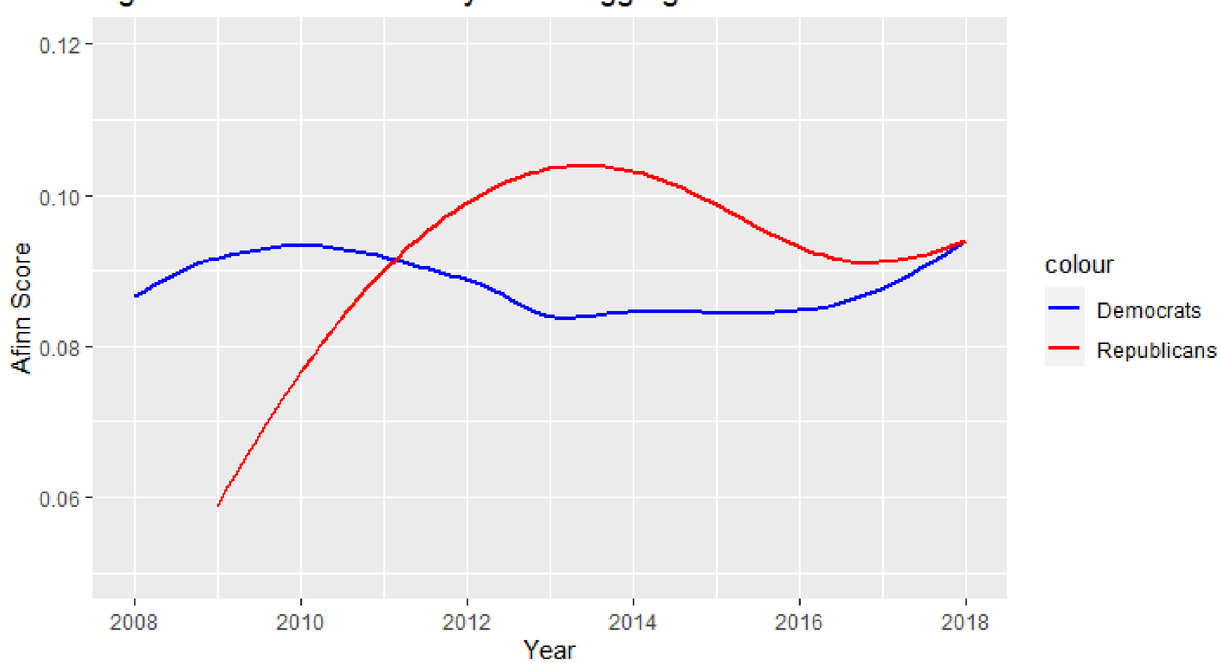


Figure 2.21 - Positive Bing Score by Party

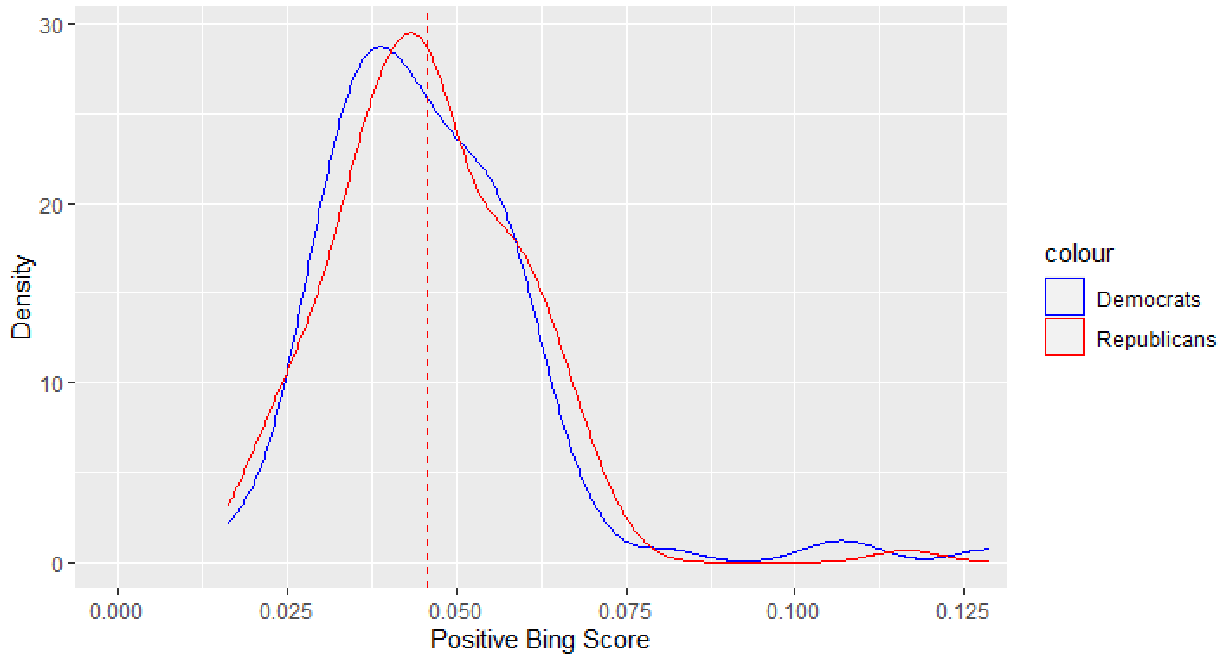


Figure 2.22 - Negative Bing Score by Party

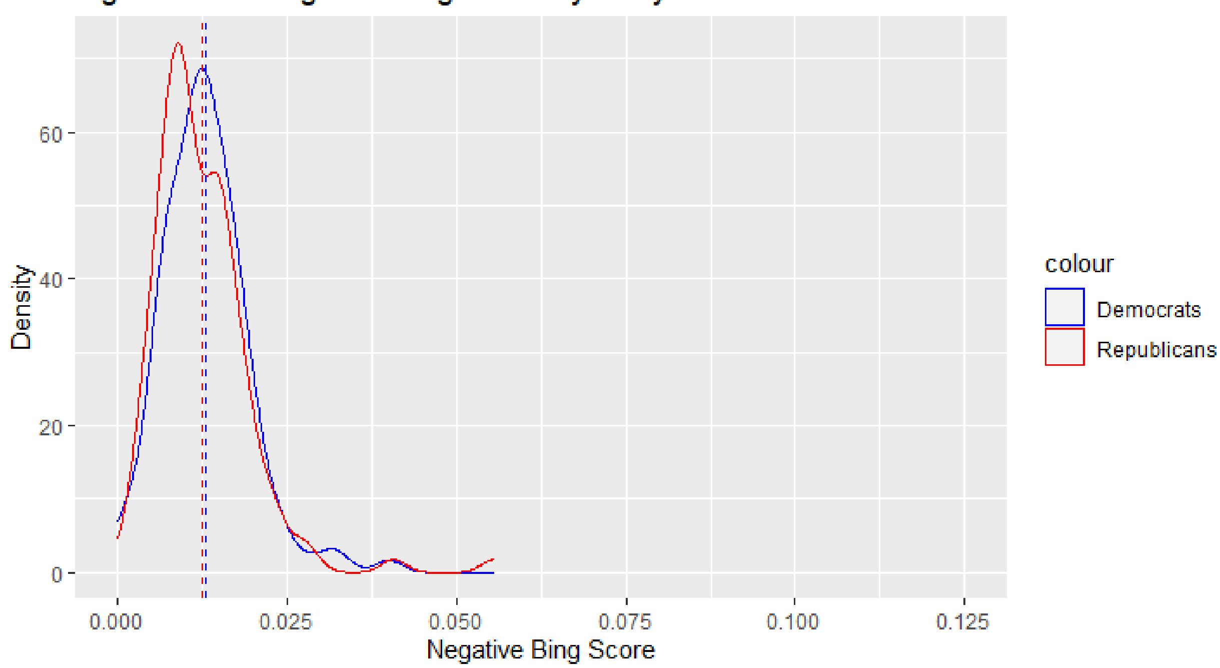


Figure 2.23 - Positive Bing Scores by Year and Party

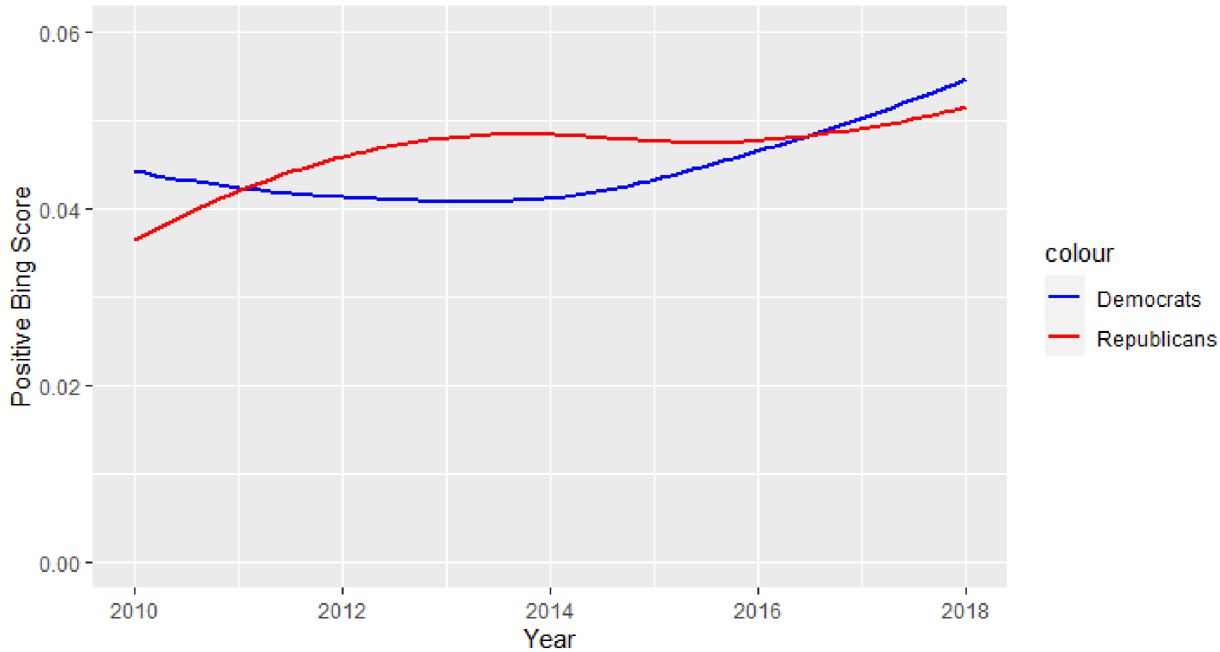


Figure 2.24 - Negative Bing Scores by Year and Party

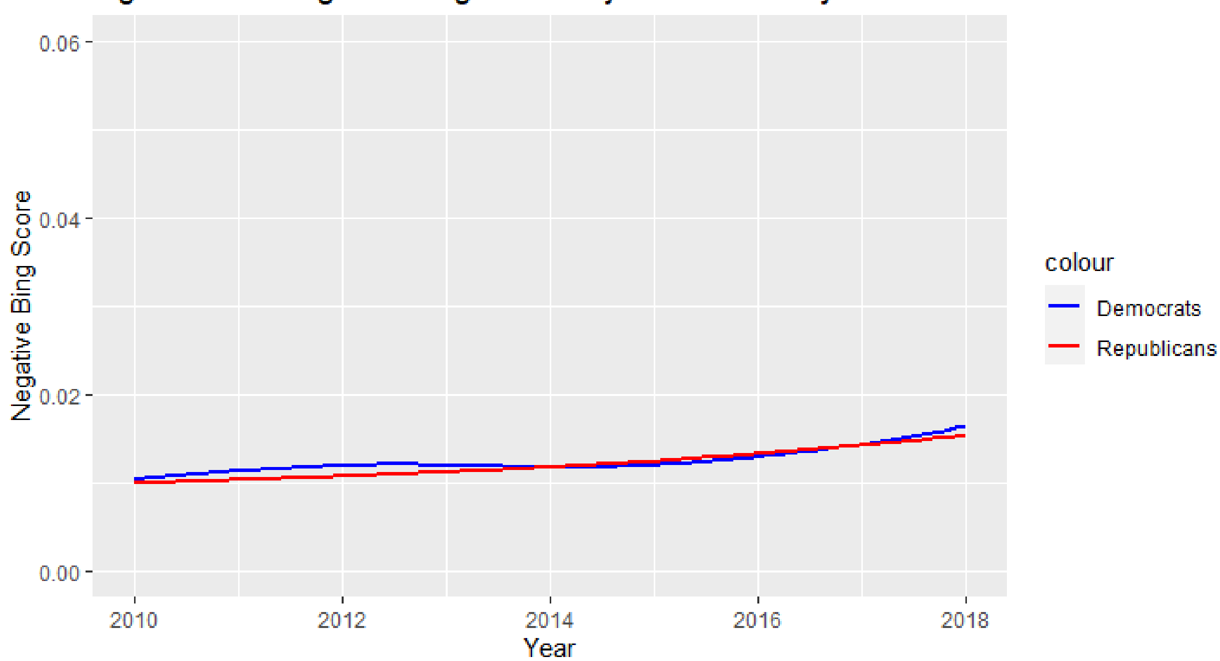


Figure 2.25 - Bing score by Distance to Election Day

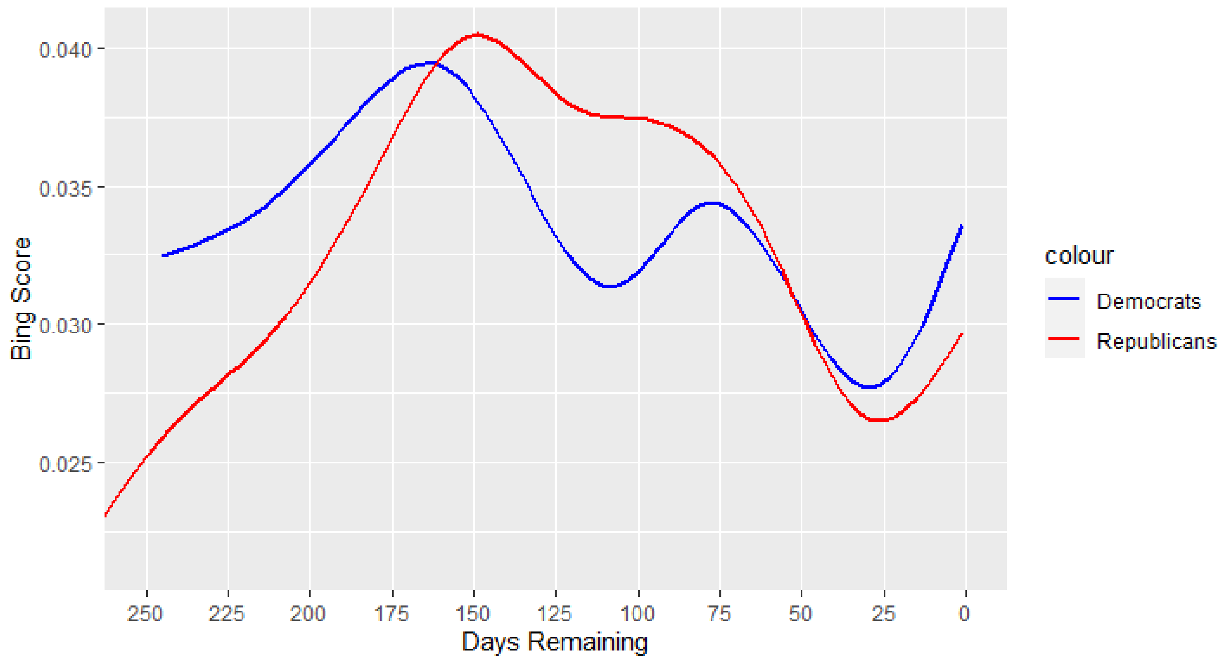


Figure 2.26 - Afinn score by Distance to Election Day

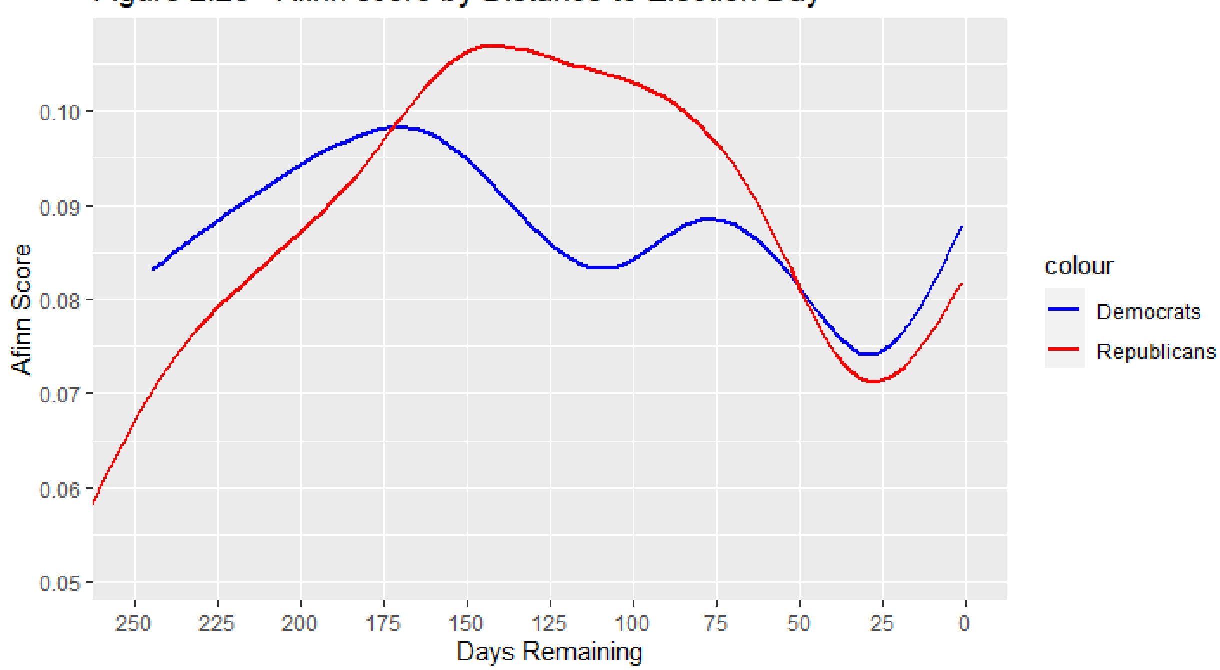


Figure 2.27 - Bing Score by Party and Incumbency

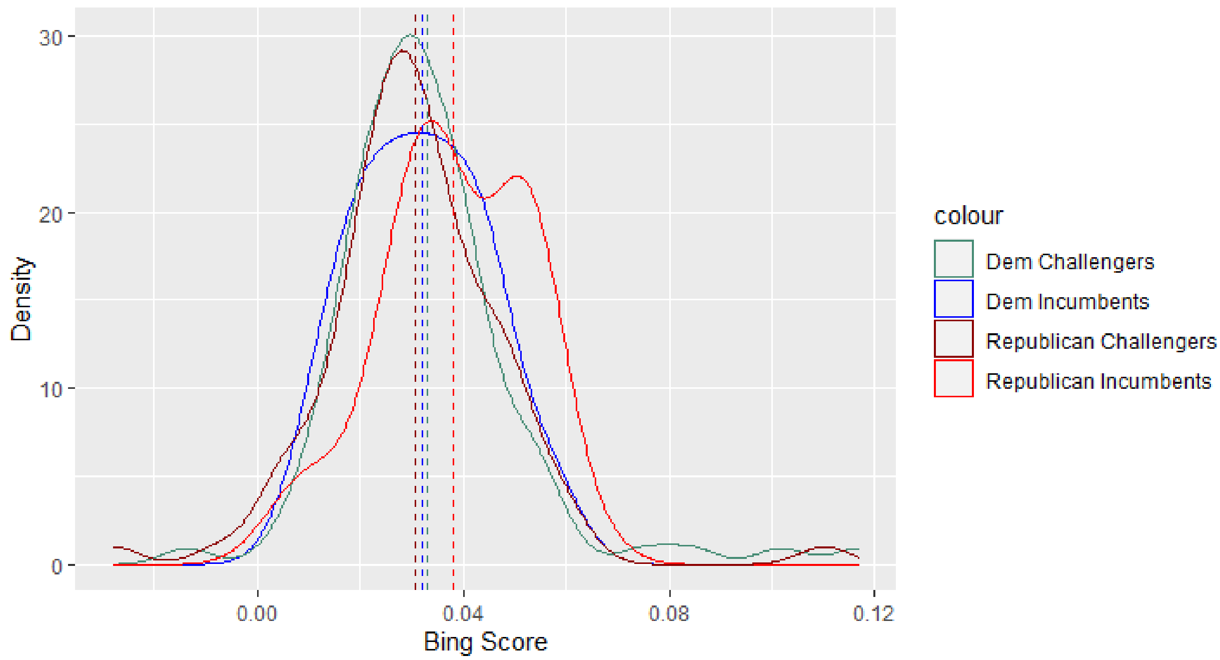


Figure 2.28 - Afinn Scores by Party and Incumbency

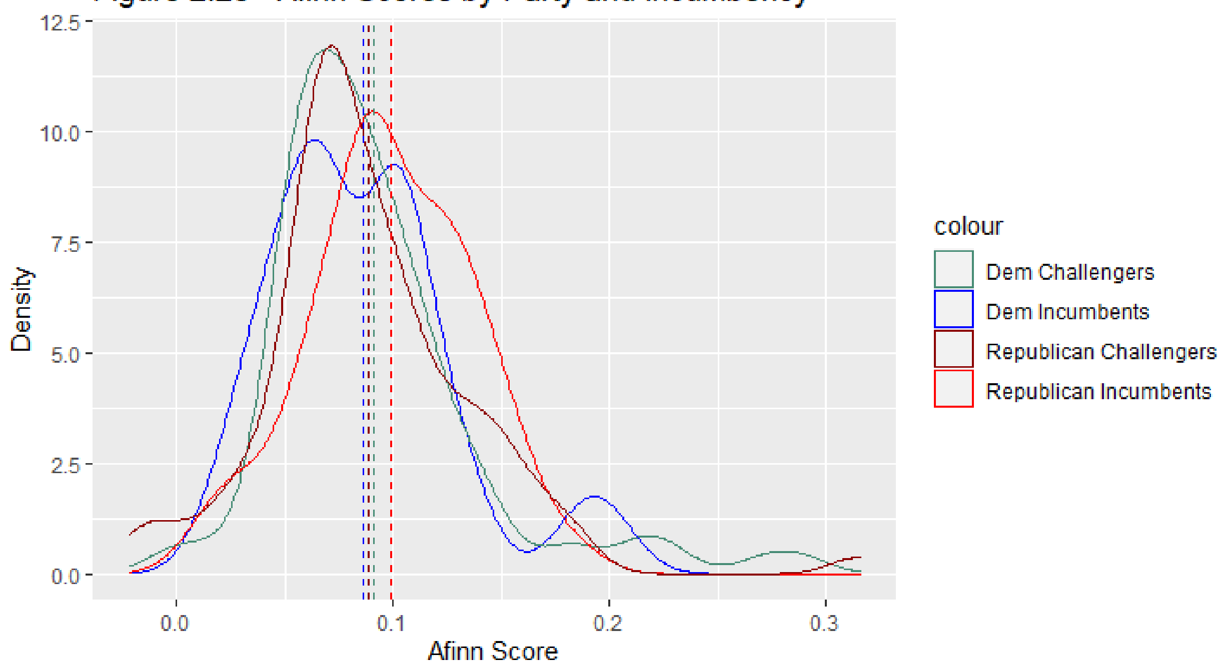


Figure 2.29 - Categories per Tweet

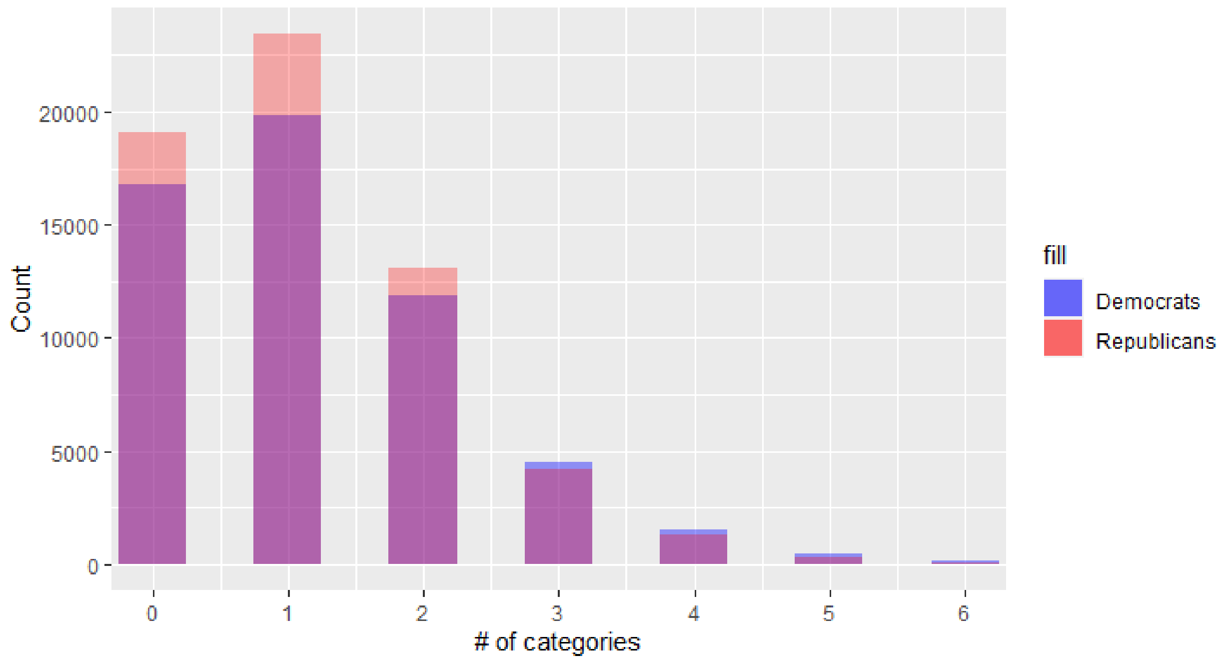


Figure 2.30 - Categories per Tweet by Year

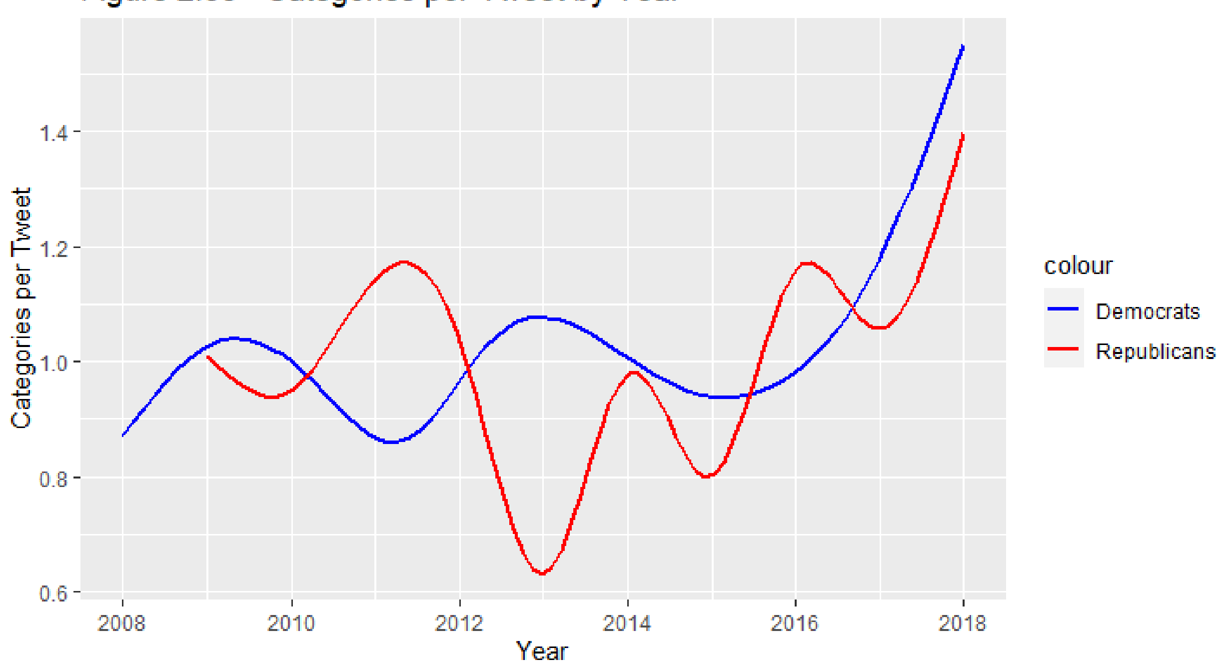


Figure 2.31 - Democratic Party Content

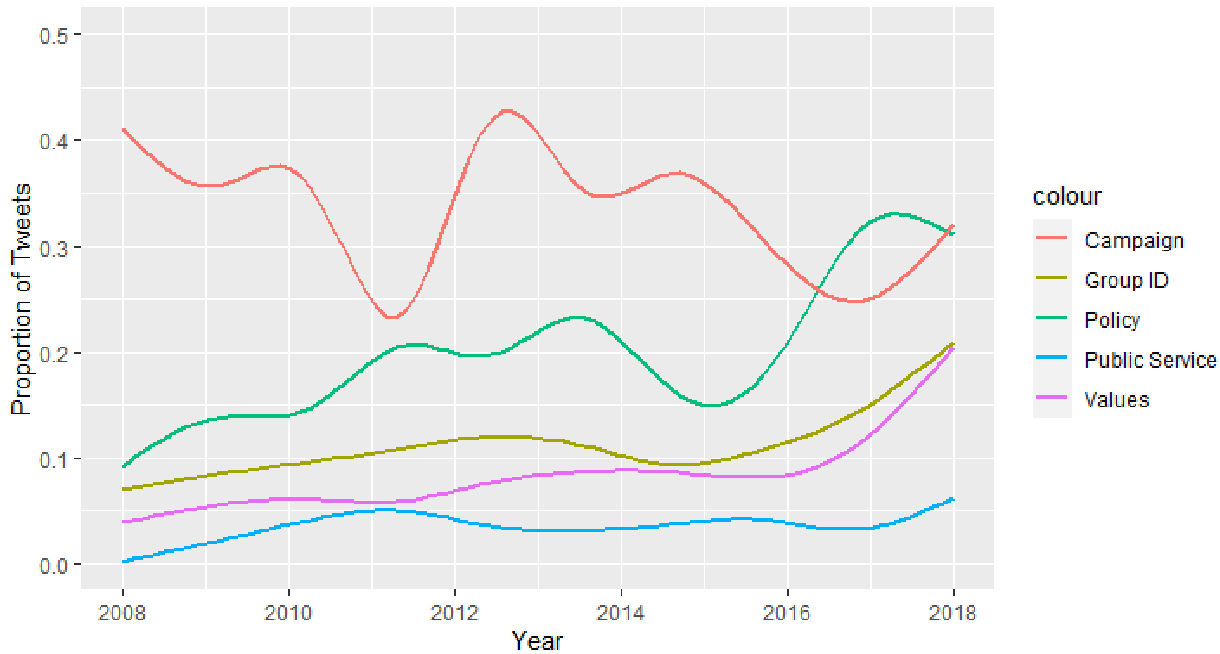
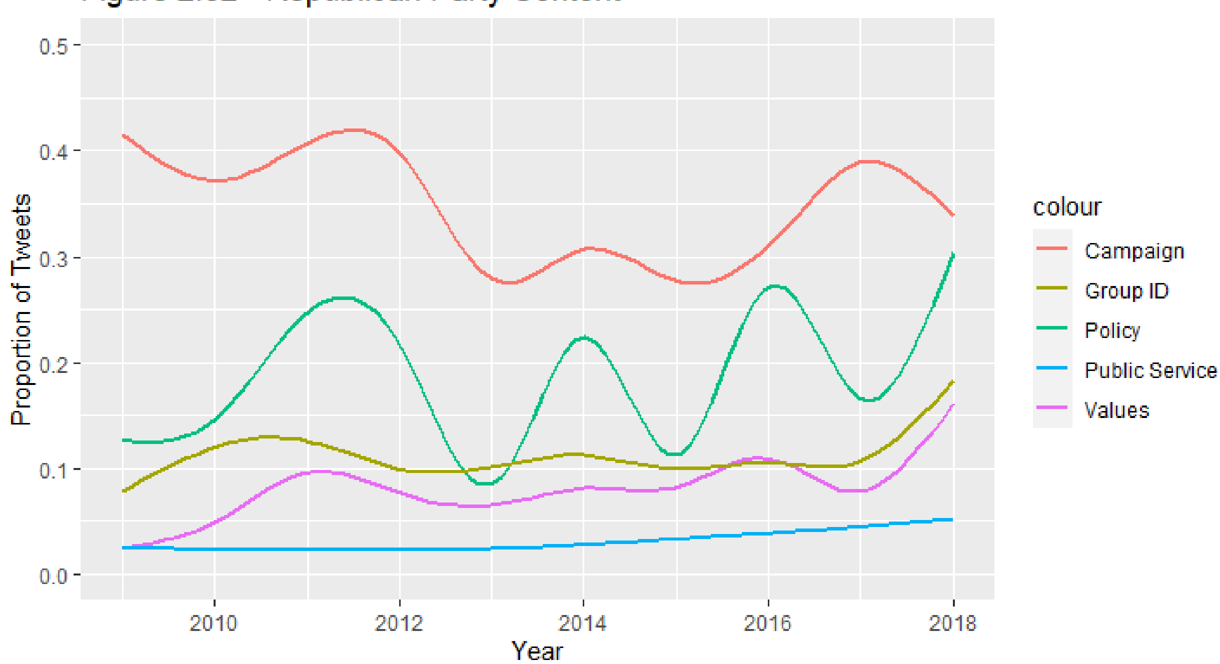


Figure 2.32 - Republican Party Content



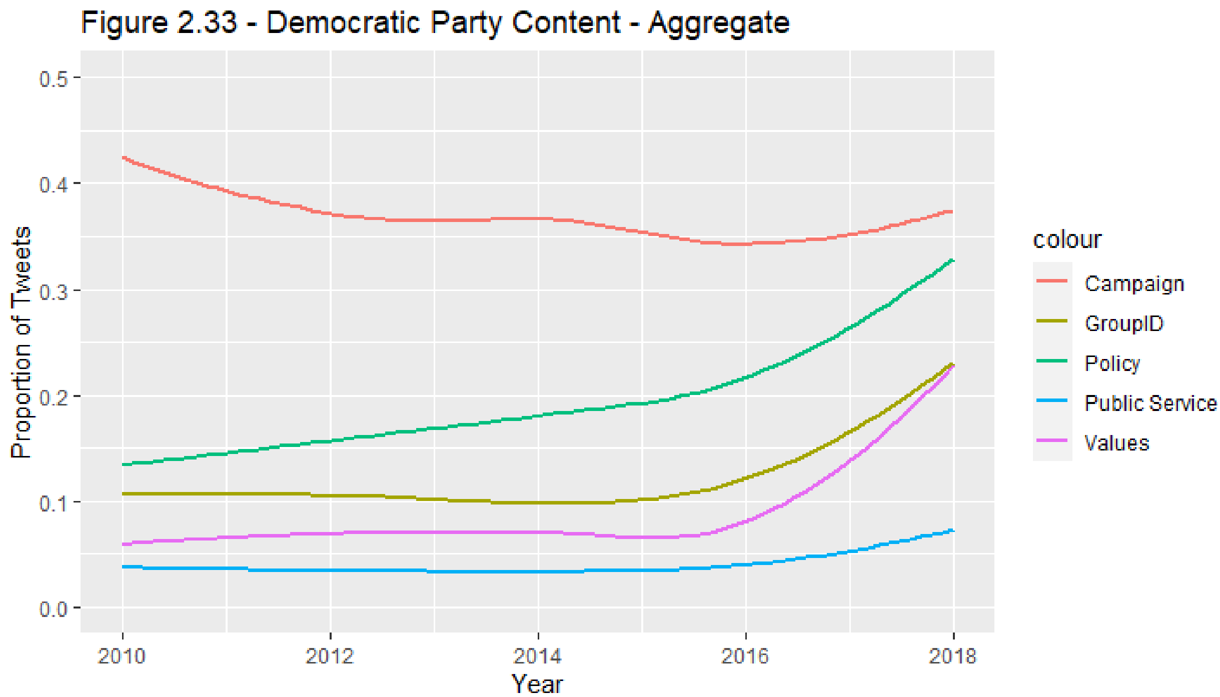


Figure 2.33 - Democratic Party Content - Aggregate
Proportion of Tweets
0.5
0.4
0.3
0.2
0.1
0.0
2010
2012
2014
2016
2018
Year
colour
Campaign
GroupID
Policy
Public Service
Values

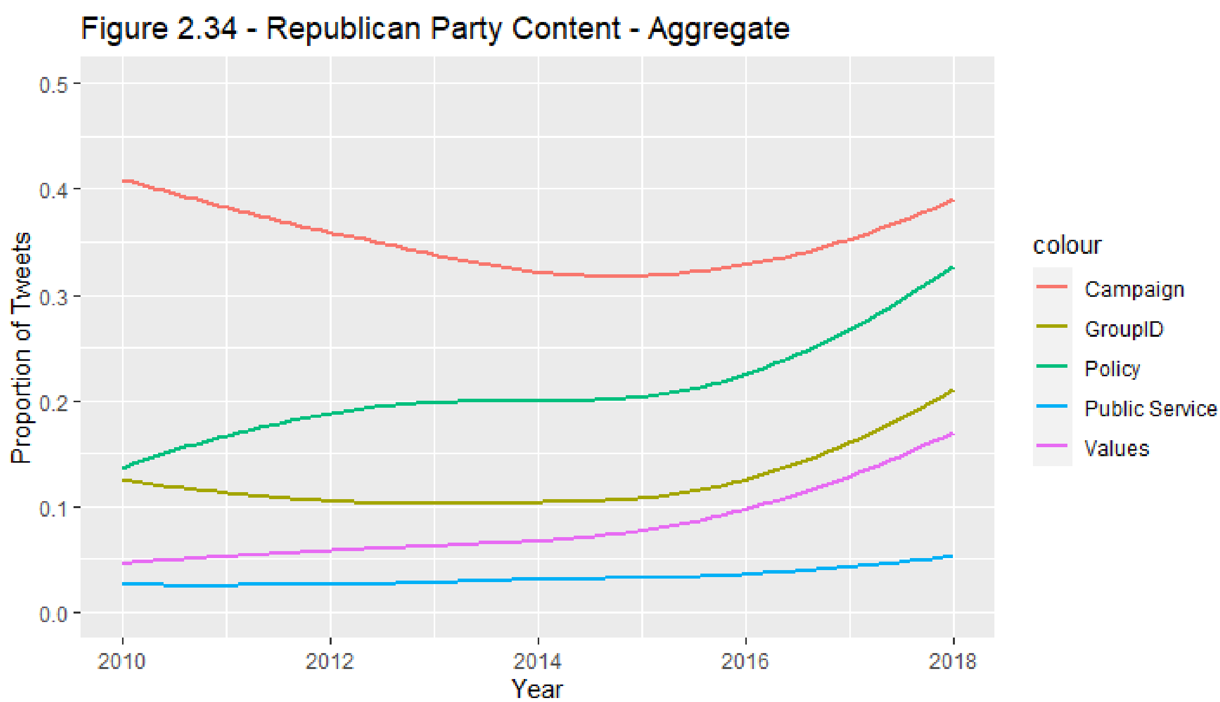


Figure 2.34 - Republican Party Content - Aggregate
Proportion of Tweets
0.5
0.4
0.3
0.2
0.1
0.0
2010
2012
2014
2016
2018
Year
colour
Campaign
GroupID
Policy
Public Service
Values

Figure 2.35 - Democratic Campaign Content

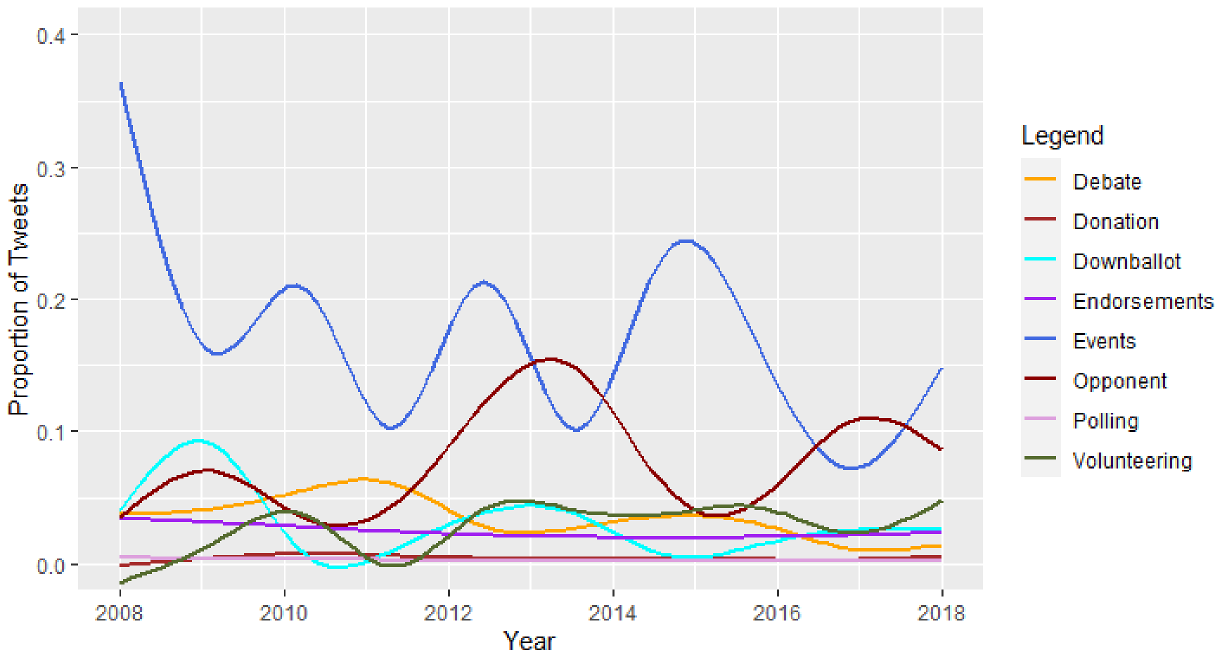
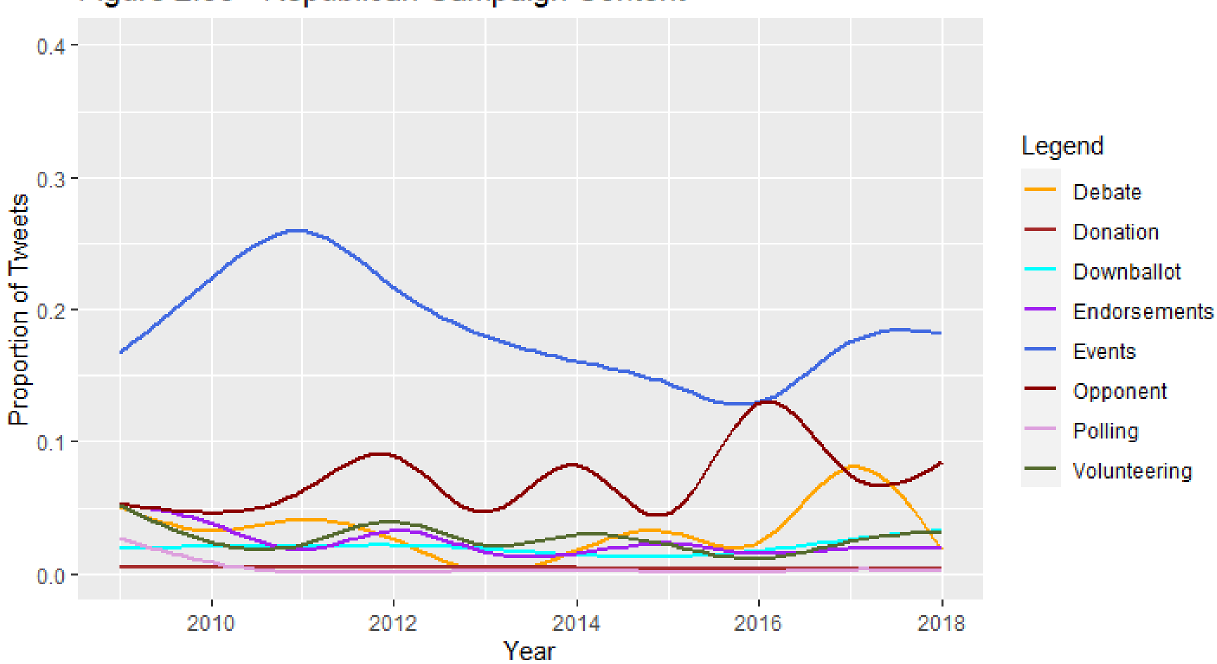


Figure 2.36 - Republican Campaign Content



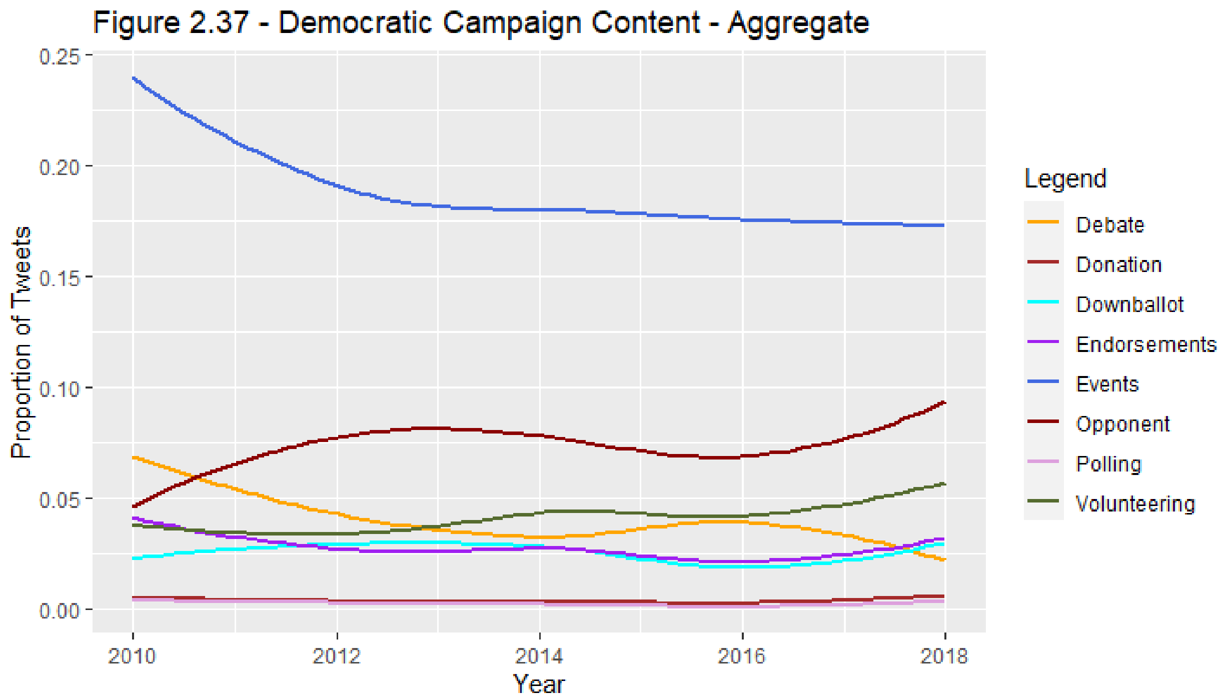


Figure 2.37 - Democratic Campaign Content - Aggregate
Proportion of Tweets
0.25
0.20
0.15
0.10
0.05
0.00
2010
2012
2014
2016
2018
Year
Legend
Debate
Donation
Downballot
Endorsements
Events
Opponent
Polling
Volunteering

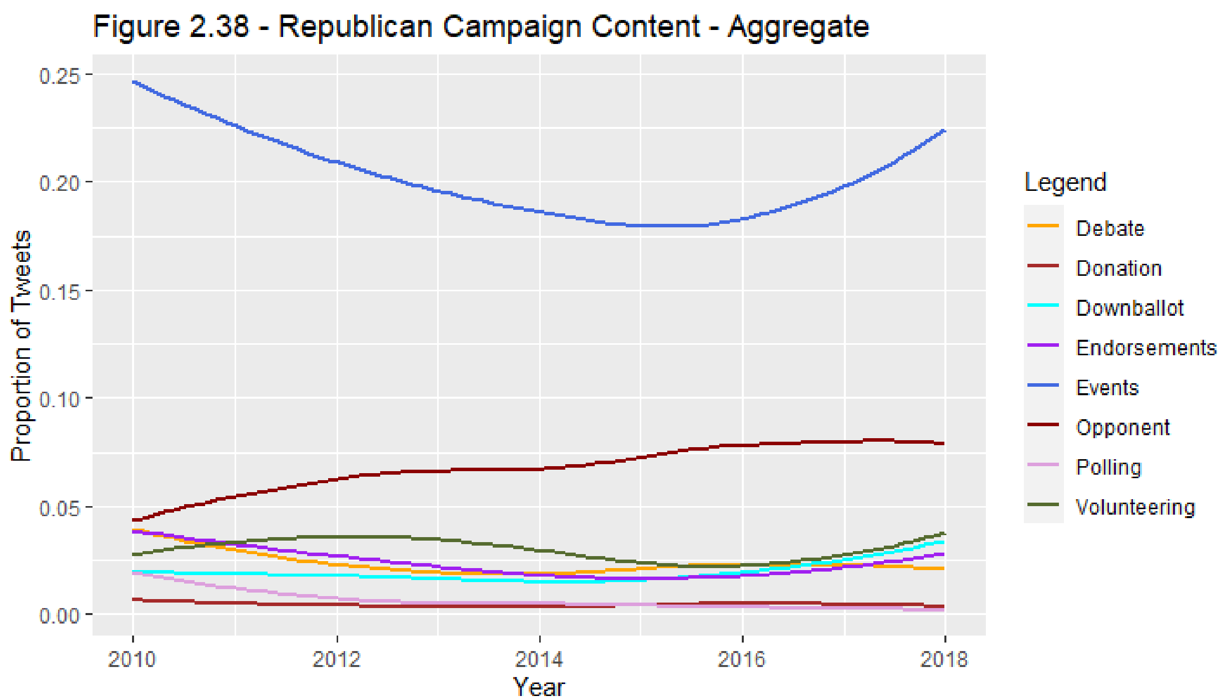


Figure 2.38 - Republican Campaign Content - Aggregate
Proportion of Tweets
0.25
0.20
0.15
0.10
0.05
0.00
2010
2012
2014
2016
2018
Year
Legend
Debate
Donation
Downballot
Endorsements
Events
Opponent
Polling
Volunteering

Figure 2.39 - Democratic Policy Content

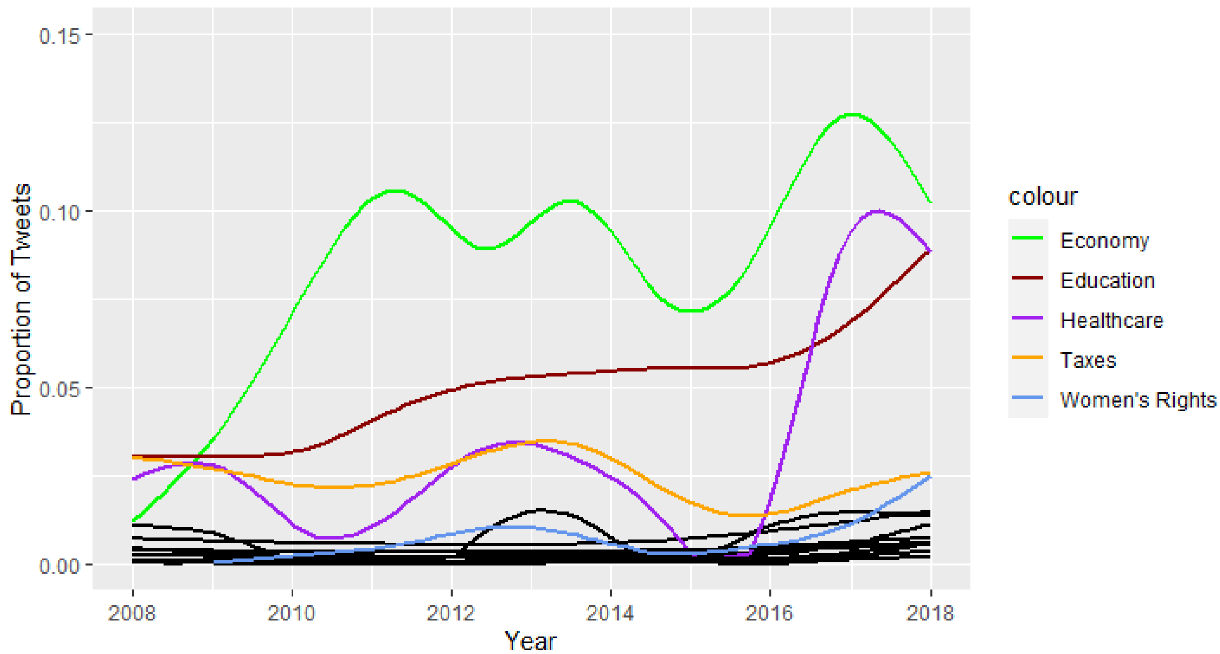


Figure 2.40 - Republican Policy Content

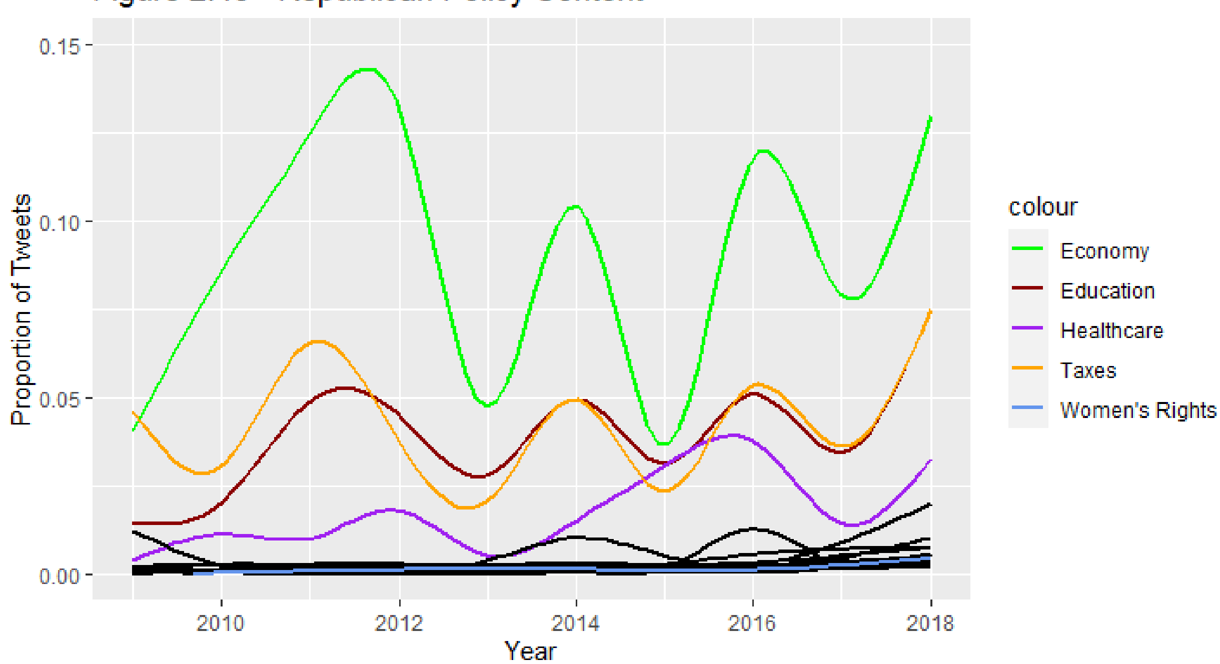


Figure 2.41 - Democratic Policy Content - Aggregate

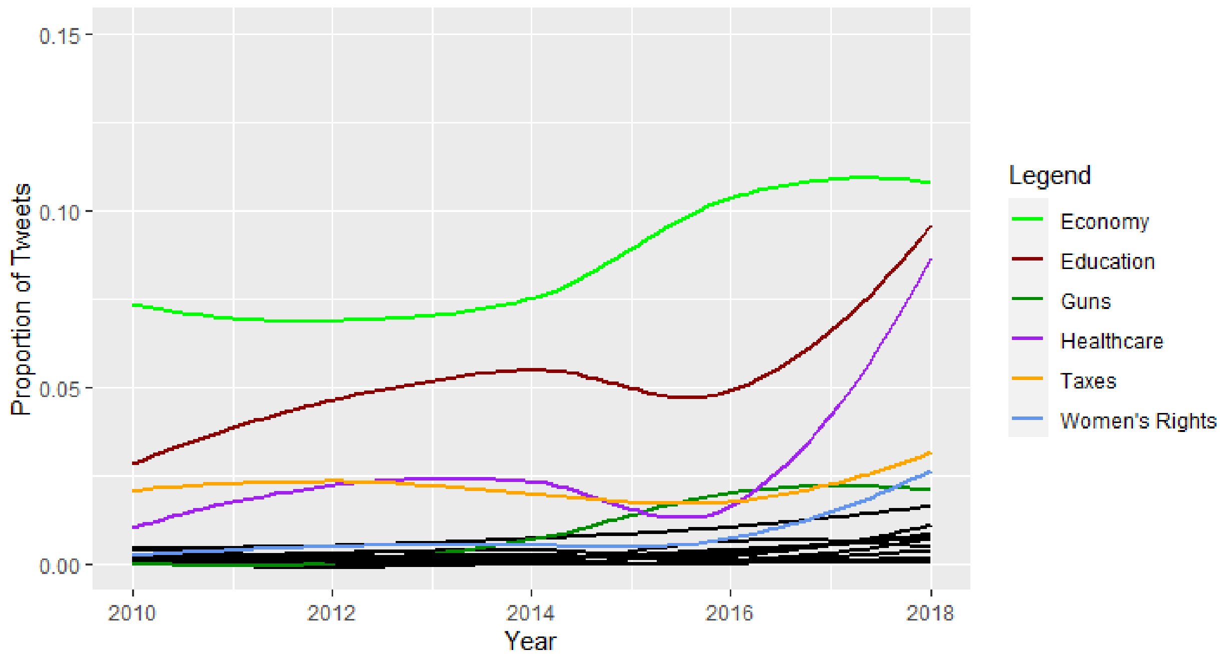


Figure 2.42 - Republican Policy Content - Aggregate

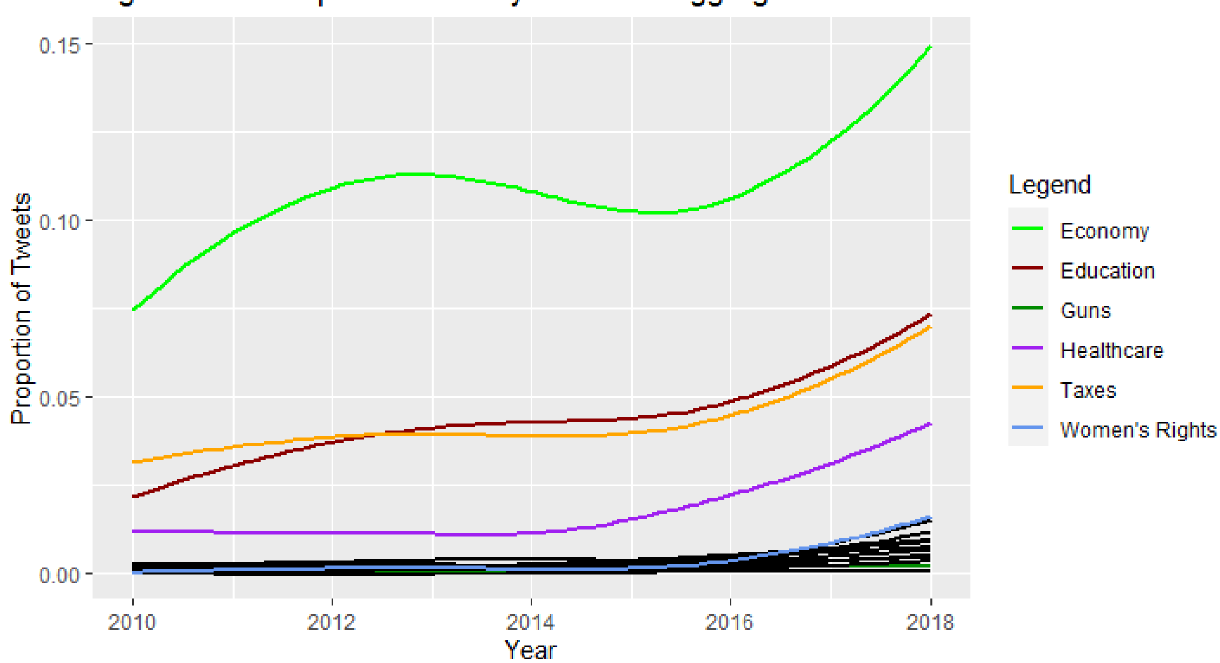


Figure 2.43 - Democratic Values Content

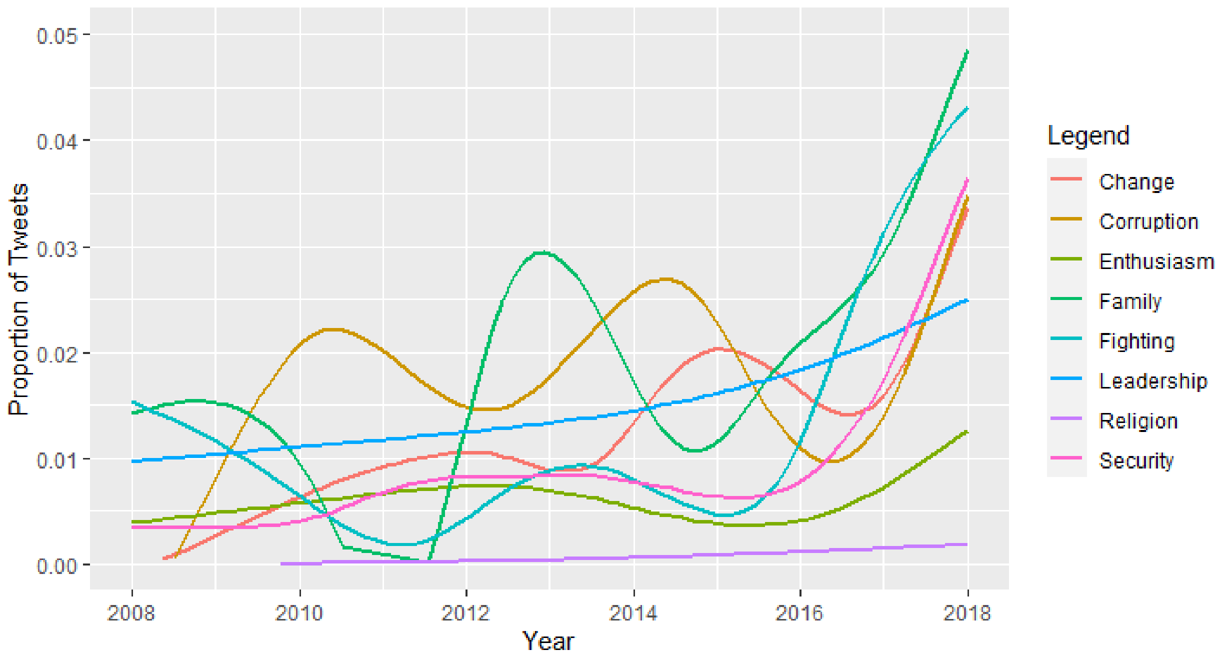
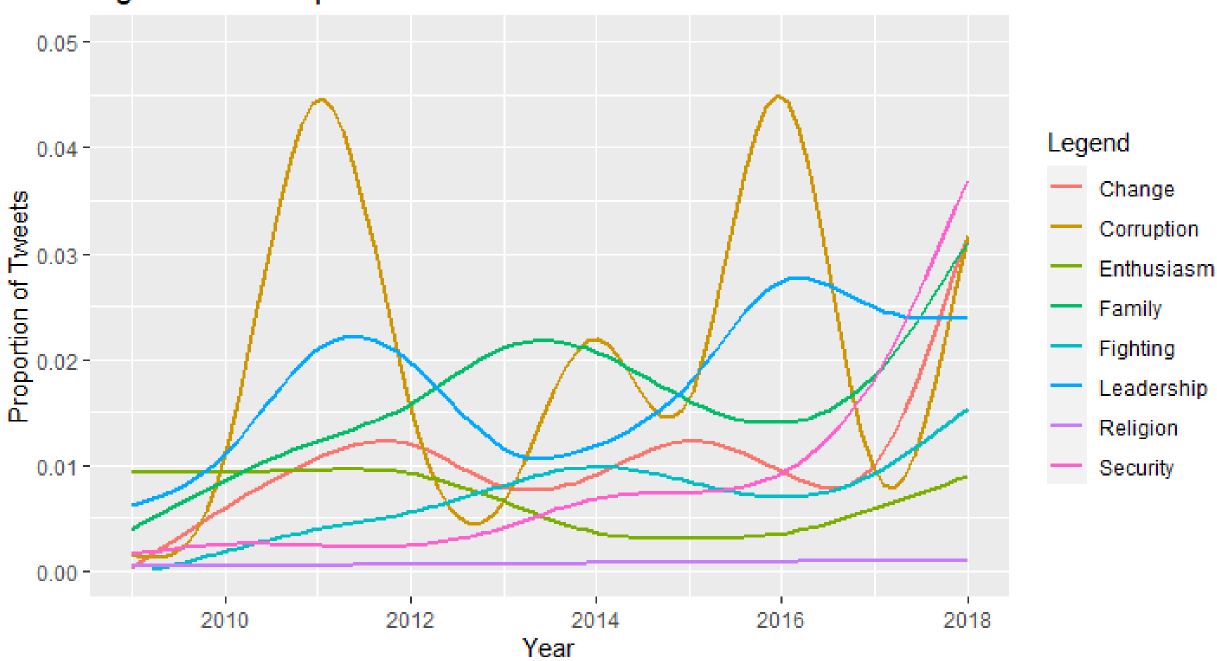


Figure 2.44 - Republican Values Content



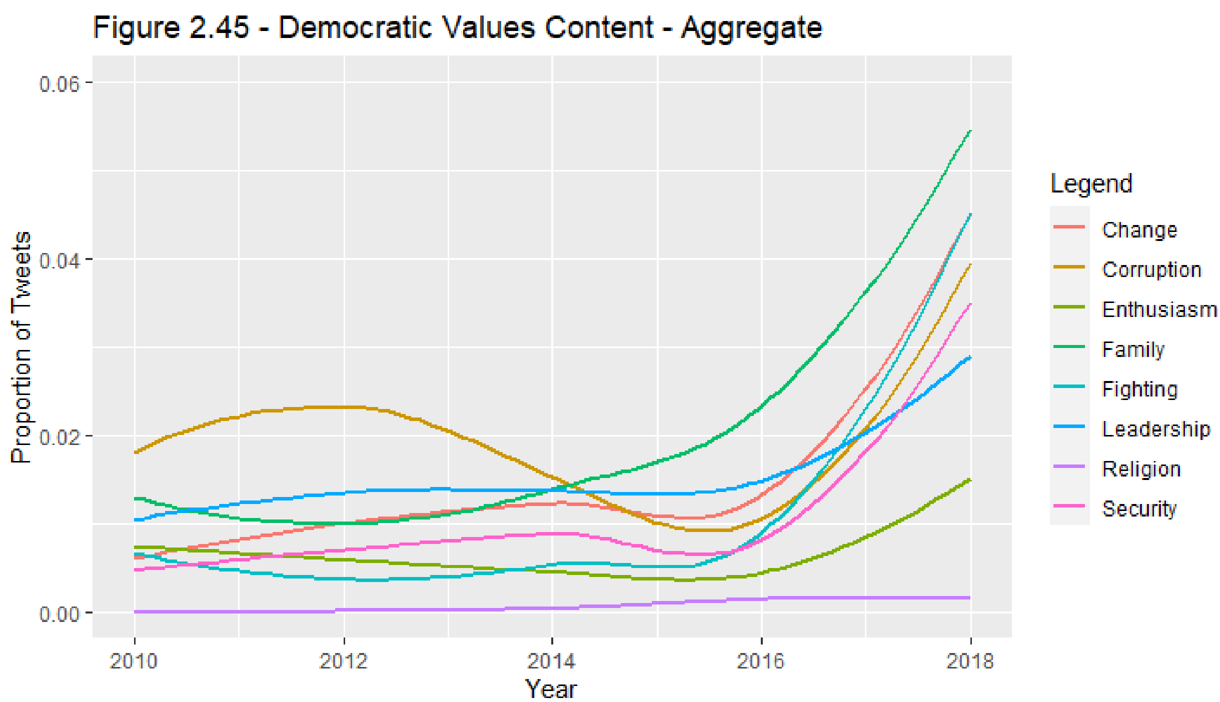


Figure 2.45 - Democratic Values Content - Aggregate
0.06
0.04
0.02
0.00
Proportion of Tweets
2010
2012
2014
2016
2018
Year
Legend
Change
Corruption
Enthusiasm
Family
Fighting
Leadership
Religion
Security

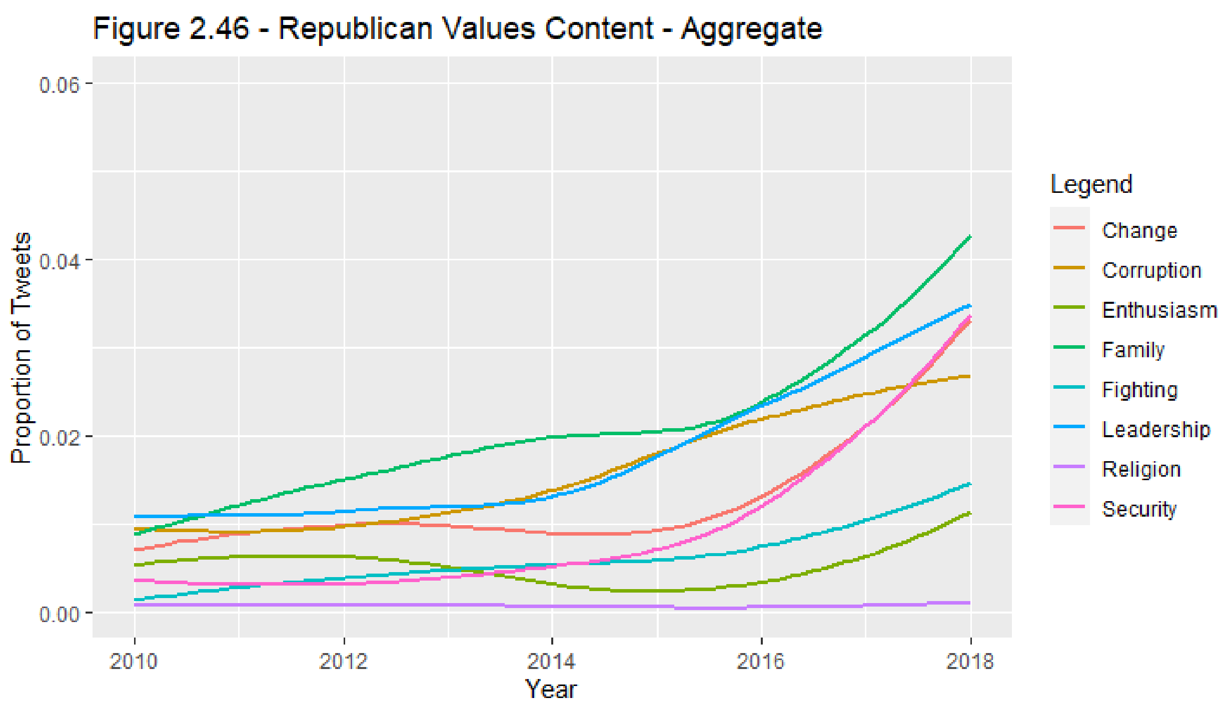


Figure 2.46 - Republican Values Content - Aggregate
0.06
0.04
0.02
0.00
Proportion of Tweets
2010
2012
2014
2016
2018
Year
Legend
Change
Corruption
Enthusiasm
Family
Fighting
Leadership
Religion
Security

Figure 2.47 - Democratic Group Content

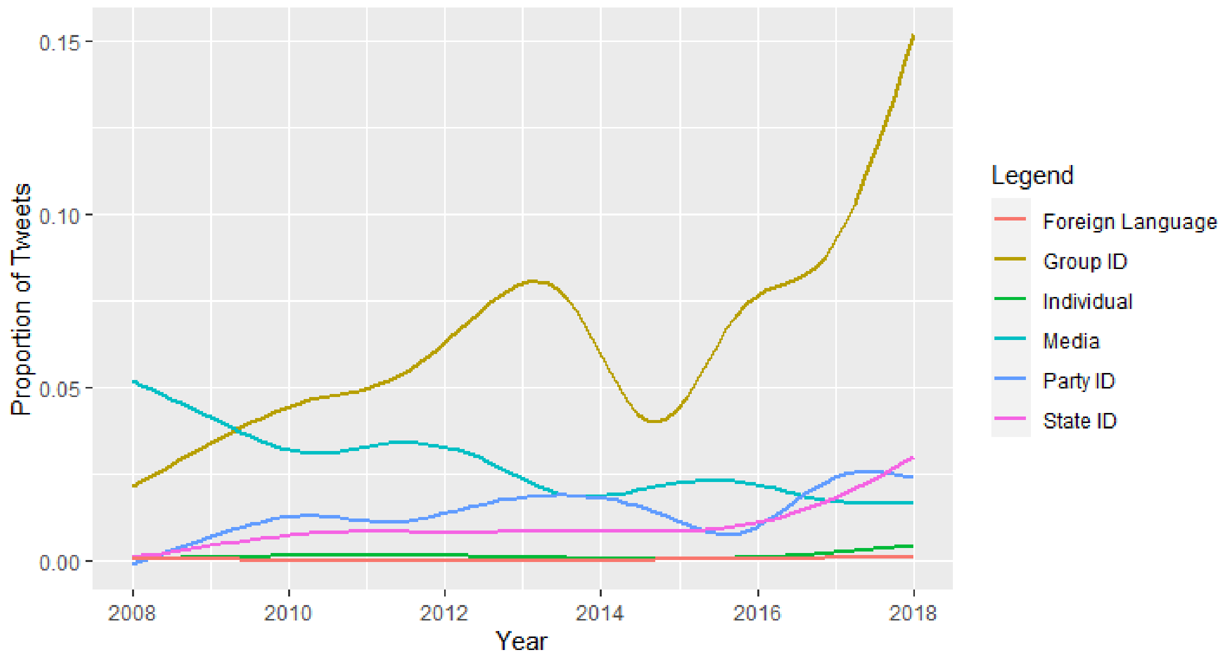


Figure 2.48 - Republican Group Content

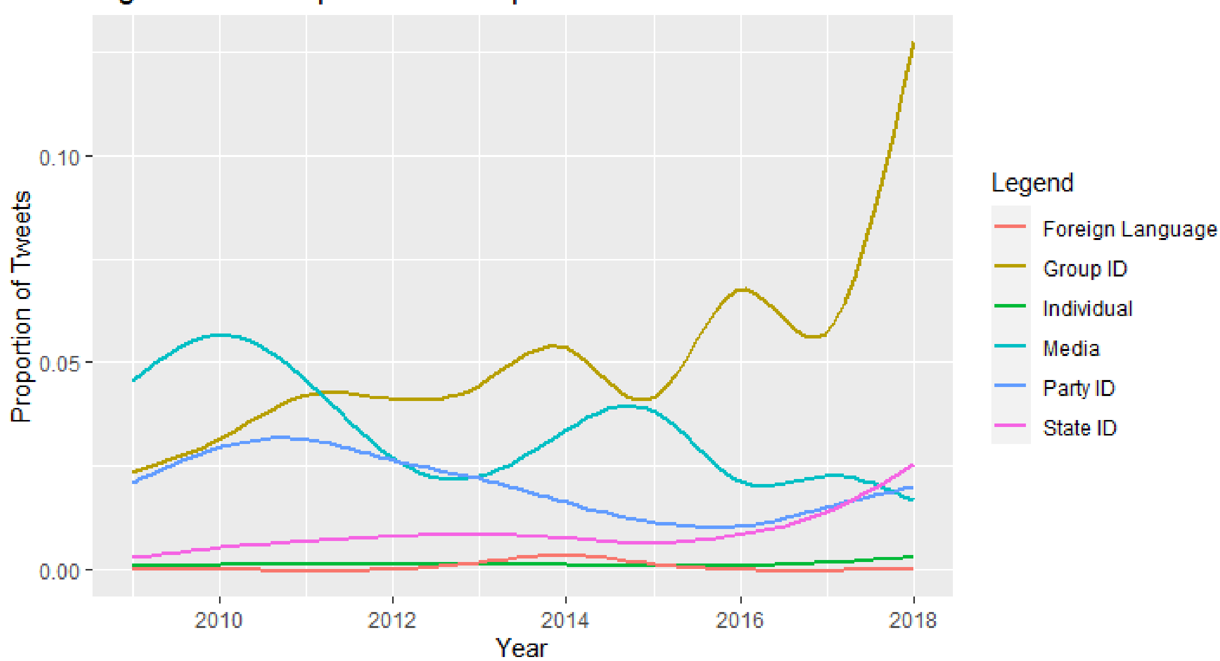


Figure 2.49 - Democratic Group Content - Aggregate

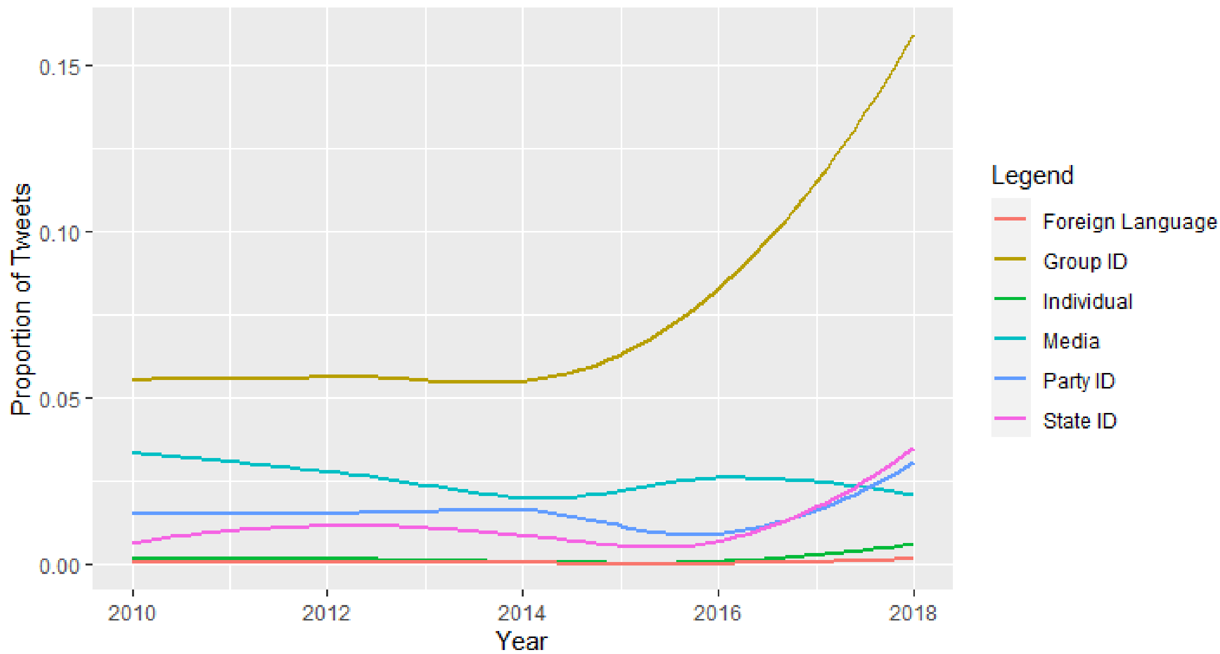
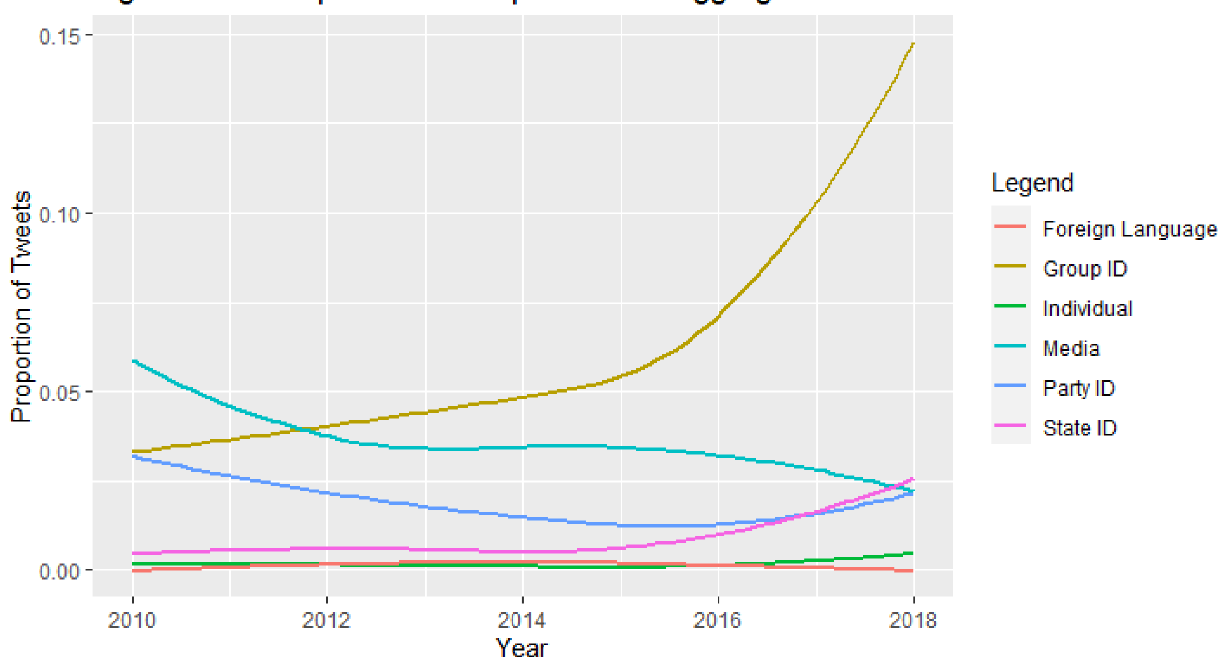


Figure 2.50 - Republican Group Content - Aggregate



References

Abramowitz, A. I., & Saunders, K. L. (2008). Is polarization a myth? *The Journal of Politics, 70*(2), 542-555.

Auter, Z. J., & Fine, J. A. (2016). Negative campaigning in the social media age: Attack advertising on Facebook. *Political Behavior, 38*(4), 999-1020.

Benoit, W. L. (2004a). Election outcome and topic of political campaign attacks. *Southern Communication Journal, 69*(4), 348-355. doi:10.1080/10417940409373305

Benoit, W. L. (2004b). Political party affiliation and presidential campaign discourse. *Communication Quarterly, 52*(2), 81-97. doi:10.1080/01463370409370183

Gainous, J., & Wagner, K. M. (2013). *Tweeting to power: The social media revolution in American politics*: Oxford University Press.

Geer, J. G., & Vavreck, L. (2014). Negativity, information, and candidate position-taking. *Political Communication, 31*(2), 218-236.

Glassman, M. E., Straus, J. R., & Shogan, C. J. (2013). *Social networking and constituent communications: Members' use of Twitter and Facebook during a two-month period in the 112th Congress.* Paper presented at the Washington DC: Congressional Research Service, CRS Report for Congress.

Golbeck, J., Grimes, J. M., & Rogers, A. (2010). Twitter use by the US Congress. *Journal of the American society for information science and technology, 61*(8), 1612-1621.

Graham, T., Broersma, M., Hazelhoff, K., & Van'T Haar, G. (2013). Between broadcasting political messages and interacting with voters: The use of Twitter during the 2010 UK general election campaign. *Information, communication & society, 16*(5), 692-716.

Grossmann, M., & Hopkins, D. A. (2016). *Asymmetric politics: Ideological Republicans and group interest Democrats*: Oxford University Press.

Hemphill, L., Culotta, A., & Heston, M. (2016). # Polar Scores: Measuring partisanship using social media content. *Journal of Information Technology & Politics, 13*(4), 365-377.

Hetherington, M., & Weiler, J. (2018). *Prius or pickup?: How the answers to four simple questions explain America's great divide*: Houghton Mifflin.

Jerit, J. (2008). Issue framing and engagement: Rhetorical strategy in public policy debates. *Political Behavior, 30*(1), 1-24.

Koch, J. W. (2008). Campaign advertisements' impact on voter certainty and knowledge of house candidates' ideological positions. *Political Research Quarterly, 61*(4), 609-621.

Krupnikov, Y. (2011). When does negativity demobilize? Tracing the conditional effect of negative campaigning on voter turnout. *American Journal of Political Science, 55*(4), 797-813.

Mason, L. (2018). *Uncivil agreement: How politics became our identity*: University of Chicago Press.

Petrocik, J. R. (1996). Issue ownership in presidential elections, with a 1980 case study. *American Journal of Political Science*, 825-850.

Pew Research Center. (2019). *January 2019 Core Trends Survey*. Retrieved from Washington, D.C.: https://pewrsr.ch/2VxJuJ3

Russell, A. (2018). US senators on twitter: Asymmetric party rhetoric in 140 characters. *American Politics Research, 46*(4), 695-723.

Silge, J., Robinson, D. (2016). "tidytext: Text Mining and Analysis Using Tidy Data Principles in R. *_JOSS_, 1*(3). doi:doi: 10.21105/joss.00037

Spiliotes, C. J., & Vavreck, L. (2002). Campaign advertising: Partisan convergence or divergence? *The Journal of Politics, 64*(1), 249-261.

Straus, J. R., Glassman, M. E., Shogan, C. J., & Smelcer, S. N. (2013). Communicating in 140 characters or less: Congressional adoption of Twitter in the 111th Congress. *PS, political science & politics, 46*(1), 60.

Straus, J. R., Williams, R. T., Shogan, C. J., & Glassman, M. E. (2016). Congressional social media communications: Evaluating Senate Twitter usage. *Online Information Review*.

Therriault, A. (2015). Whose issue is it anyway? A new look at the meaning and measurement of issue ownership. *British Journal of Political Science, 45*(4), 929-938.

Vavreck, L. (2014). Want a Better Forecast? Measure the Campaign Not Just the Economy. *PS: Political Science & Politics, 47*(2), 345-347.

Chapter 3

The frequent discussion of party differences in political science is based on the general assumption that members of the same party share some degree of similarity amongst themselves. Trying to find and explain these party member similarities has been a driving force in political science, and scholars have isolated party differences in voting behavior, use of political power, policy preferences, and political messaging, among other domains. Though there is debate about the extent to which the similarity is due simply to like-mindedness between candidates or due to intentional enforcement by the party elite (Snyder Jr & Ting, 2002), coordination at the party level is a common occurrence, and an important one. In this chapter, I use my dataset to examine party coordination in electoral messaging, and attempt to answer questions about whether gubernatorial candidates coordinate on message, which topics parties consider part of their brand, and whether and how strongly that brand changes over time.

In previous research, party coordination is often interpreted as a way to establish and defend a party ‘brand’. A brand is a marketing tool designed to create strong feelings and a sense of attachment to a product - in this case, a political party (Bruns, Langner, & Fischer, 2017; Smith & French, 2009). Parties create brands by developing reputations for pursuing different policies and tactics. When a citizen votes, they call on this reputation as a heuristic to streamline their decision-making process (Lau & Redlawsk, 2001). In short, they use the party’s reputation as a replacement for specific candidates and issue knowledge that would otherwise take time and effort to develop. Parties and their members should therefore be interested in working together to create a unified message, to capitalize on the power of this heuristic-forming. Many previous conceptions of party branding focus exclusively on policy. For example, according to Smith and French (2009), a candidate makes promises on the campaign trail and attempts to follow through on those promises once elected. The policies promised and the promises kept combine into a personal brand, a reputation for success on some issues and failure on others. Parties, over time, form a brand in much the same way: by making promises to pass specific policy and then keeping them. Over a number of years, as parties handle some issues and let others go, they develop these reputations. For example, issue ownership literature finds that Democrats are seen as strong on social welfare issues, while Republicans are seen as strong on the economy and national defense, as a result of this kind of brand development. These policy brands are not static - as promises go unfulfilled or new policies become salient, the brand can shift, another party can usurp control of a previously ‘owned’ issue (Pope & Woon, 2009). It is unclear, however, how quickly and in what patterns these shifts occur. How quickly do parties’ reputations shift and how quickly do candidates adapt? Chapter 2 gives some insight into this question, indicating that there is some change over time in how much parties mention specific issues or values, but not much change in the ordering of topics. Where Chapter 2 focuses on aggregate patterns of party behavior, this chapter will further explore this question by determining how tightly party members coordinate on these messaging strategies, and how this coordination shifts over time.

A few scholars argue that a party’s brand goes beyond policy. Butler and Powell (2014) argue that in addition to the policy element, party brands also include a valence element, consisting of things like efficiency and ethical behavior. Certainly, findings from literature on Social Identity Theory would suggest that groups like a political party can hold norms for efficacy and action, or even for emotion (Huddy, Mason, & Aarøe, 2015; Thomas & McGarty, 2009). All of this has the potential to become part of a party’s reputation, and as a result, part of the party’s intentional messaging. This chapter will also investigate the extent to which parties treat non-policy variables, like values and group identity activation, as part of the party brand. If a party demonstrates more coordination over these variables, it would suggest that they believe these variables to be an important part of the party's messaging.

There are a number of reasons why we might expect these party brands to appear in Twitter messaging. First and foremost, Twitter bypasses the middleman of the media, allowing candidates and politicians to say exactly what they want to say directly to voters. This provides a window into party brand, as well as personal brand, that is valuable for analysis, as messages have higher fidelity. Furthermore, in order to be exposed to a politician's message on Twitter, a citizen must self-select into a political Twitter network. This makes it more likely that a politician's Twitter followers are only those who are already most engaged and knowledgeable about politics. These people are most likely part of the party's already-polarized base. Given that party is a strong voting heuristic (Lau & Redlawsk, 2001), and polarization leads to higher turnout (Abramowitz & Saunders, 2008), there should be ample incentive to appeal to party loyalists on Twitter by appealing to the party brand in the runup to an election. If the audience is already engaged and partisan, then messaging designed to appeal to the party base has more potential to drive increased turnout. In addition, the consequences for diluting the party brand can be severe. First, Republicans in particular are likely to challenge and replace incumbents who don't tack to a specific ideological position (Grossmann & Hopkins, 2016; Mann & Ornstein, 2016). Second, parties which don't maintain coherent brands are in danger of dissolution. In a study of Latin American countries, Lupu (2016) finds that the breakdown of political parties is predictably preceded by passing policy which runs counter to a party's brand. This leads to a drop in partisan attachment to the party, and finally to the collapse of the party itself. So, candidates at every level should have a strong incentive to tailor their messaging to maintain party branding, and we should expect evidence of this party brand to appear in any platform, Twitter included.

We know that parties care about these brands because they use power to maintain them. In congress and state legislatures, parties use the power of agenda control to prevent policy from coming to a vote which would pass without the support of the majority of the party, thus giving credit to the minority party (Cox, Kousser, & McCubbins, 2010; Cox & McCubbins, 2005). However, there are reasons to temper our expectations of finding strong coordination in this data set. Although parties use agenda control to protect themselves from a majority roll, agenda control is a power of absence rather than a power of presence; it ensures that certain policies don't pass, not that certain policies do. A vast literature has tried, with little success, to find evidence that party leadership can coerce votes out of members (Cox & McCubbins, 2005; Koger & Lebo, 2017; Krehbiel, 1996; Snyder Jr & Groseclose, 2000), suggesting that coordination is not effectively enforced in legislative bodies. The data under examination in this dissertation focuses on gubernatorial races, and if there are few ways for a party to coerce cohesion at the national level from Senators and Congresspeople, the increased individual power and responsibility of a gubernatorial position ensures that there are even fewer ways to enforce this branding among gubernatorial candidates.

There is also significant evidence that individual members of a party have incentive to *distance* themselves from the party brand. Canes-Wrone, Brady, and Cogan (2002) found that legislative coordination, voting too extensively along party lines, has the potential to hurt legislators going into re-election. Gubernatorial candidates especially, who carry more individual responsibility for state-level policy outcomes, may benefit more from creating a personal reputation separate from the party's. We know that some legislators try to do this -

Grimmer, Messing, and Westwood (2012) find that legislators try to create a personal brand by claiming credit for bringing money home to the district, while Fenno (1978) argues that members of Congress rely on their 'home style' of district service in order to get re-elected. Beyond individual candidates, even local parties attempt to create distance between themselves and the national party, depending on their district's specific concerns. According to Brown (2017), a full 2-3% of local county-level political parties maintain a platform that is distinct and different from the national party platform. Finally, there are identifiable subfactions within national political parties which have an interest in creating their own brand, such as the Tea Party (Clarke, 2020). Therefore, not only do parties lack enforcement mechanisms to make sure candidates stick to the party brand, candidates face pressure to actively break from party branding. Which of these competing concerns wins out in Twitter communication is unclear, but this chapter will begin to answer this question.

In this chapter, I look for evidence of party coordination in the data set to see how similar gubernatorial candidates in the same party are to one another, and I investigate both similarities between party members overall, and similarities between candidates of the same electoral cohort. If candidates of the same party are highly coordinated with each other, that would indicate an attempt to play to or maintain a party brand. If electoral cohorts are more similar to each other than the party overall, this would suggest fairly rapid shifts in branding. Rather than coding for conservative or liberal arguments, as is standard for party brand literature, I coded the number of mentions of specific policies, values, and campaign strategies. I use this to create a rank-ordering of message priorities for each candidate and use an adjusted borda count to aggregate party priorities. I then measure the distance between an individual's rank-ordered priorities and the party's aggregate priorities. This approach allows me to easily create a distribution of internal party coordination and identify both overall trends for party branding and coordination over specific types of messages. I use this approach to attempt to answer whether gubernatorial candidates coordinate on message, whether parties are different in their coordination, and on what topics candidates coordinate.

Methods

In this chapter, I approach the question of party coordination as a matter of priorities. On a social media site, there are no limitations on the number of times an account can tweet or on the topics an account can address. Therefore, candidates have the freedom to address as many or as few topics as they want. The topics that they do address should therefore be an accurate reflection of their platforms and personalities. Excluded topics are likely to be considered unimportant, or unlikely to help win support; they're simply not part of the candidate's branding. Because of this, I can represent the candidate's brand by quantifying the number of mentions of specific topics and ordering them from most-mentioned to least mentioned.

To get the party's brand, I simply aggregate the ranked priorities of same-party candidates. To do this, I use the Borda Count method, often used as a way to aggregate ranked-choice voting ballots (Borda, 1781). First, each candidate's priorities are sorted from most-mentioned to least-mentioned. Then, point values are assigned to each category, according to their position in the ranking. Due to the nature of the data and my purposes for it, categories with no mentions are common, so I use an alternate version of the Borda count that allows for 0-point values. Ties are broken alphabetically. Then, point totals for each category of tweet are summed across party members. The resulting point values, when sorted, estimate the priorities of each party at large. The benefit of this method for my purposes (as opposed to other RCV aggregation methods) is that it maintains a ranked order of aggregate preferences. This keeps the party's brand in the same format as the candidate's brand, and allows me to measure distance between the two.

To determine the distance between an individual candidate's messaging and the overall party brand, I rely on Kendall's tau, a measure of difference between ranked order preferences (Kendall, 1938). Each category is run through a series of pairwise comparisons with other categories. For each combination of categories, the individual candidate's ranking is compared to the party aggregate ranking. If the rankings match, the pair is concordant. If the rankings don't match, the pair is discordant. Kendall's tau is calculated by the number of concordant pairs minus the number of discordant pairs, divided by the total number of possible pairings. This results in a number between -1 (perfectly discordant) and +1 (perfectly concordant).

$$\tau = \frac{(\text{concordant pairs - discordant pairs})}{\binom{n}{2}}$$

To account for the possibility of parties attempting to change their branding in different election cycles, I repeated this process for each year of data individually (excluding 2008 and odd years, when not enough candidates were running to create a trustworthy party brand). This allows me to measure both how different each candidate is from others of their cohort and how different each cohort is from the overall party brand. I also examine coordination across policy, values, group ID messaging, and campaign tactics separately. All analyses below are run on the subset of gubernatorial candidates from 2010, 2012, 2014, 2016, and 2018 who tweeted at least once (N = 227).

RESULTS

How Coordinated Are Parties Overall?

I start with an examination of how coordinated parties are across all categories and years. The patterns indicated in Figure 3.1 demonstrate that parties are fairly coordinated on messaging. Furthermore, Democrats (m = .602, sd = .105) and Republicans (m = .611, sd = .083) are similar in their overall levels of coordination, with Republicans having a slightly narrower distribution. This indicates that there is some, though not perfect, consensus on how to use Twitter as a campaign platform. In chapter 2, I found that both parties favored campaign event announcements, policy messaging on the economy, taxes, healthcare, and education, and largely ignored other issues and strategies. Therefore, a moderate amount of coordination is unsurprising, though from those results, it would have been reasonable to expect a higher amount of coordination than indicated here.

Figure 3.2 expands on these trends by examining the same overall coordination with party messaging over time. While there are some outliers, especially among Democrats, the trend is relatively constant, with the mean coordination for most years hovering just above 0.6.

Are Cohorts More Coordinated?

Next, I examine whether candidates are more coordinated with others in their cohort than with the party at large. If coordination with one’s cohort is higher than with the party as a whole, this would indicate two things. First, that there are election-cycle changes in what candidates perceive will be effective Twitter messaging strategy. This could be due to changes in the national discussion or in the circumstances of the state’s economic success, or some other variable entirely. Second, and more important for this chapter, it would indicate that gubernatorial candidates across the country perceive the *same* changes to be beneficial. If there was less cohort coordination than party coordination, that would indicate that candidates were not adapting to election-year changes in messaging and relying on more traditional party branding.

Figures 3.3 & 3.4 demonstrate that both parties exhibit more coordination with their own cohort than with overall party messaging style. This means that not only do candidates perceive shifts in the electoral landscape between elections, but that parties respond to that shift as a group. Means and standard deviations are available in Table 1.

Figure 3.5 breaks cohort coordination down by year, and though there is more variation over time than overall party coordination, cohort coordination is still relatively stable over time.

What Are Parties Coordinated On?

Figures 3.6 and 3.7 show that Democrats are fairly coordinated on policy, campaign tactics, and group identification measures, but are much less coordinated on values, while Republicans are fairly coordinated on all four categories of tweet. To highlight the differences between the parties, Figures 3.7a through 3.7d isolate the Democrat-Republican differences for each category separately. Democrats are significantly more coordinated than Republicans in messaging around policy and campaign tactics, but the starkest difference is on values-based messaging. No Republican candidate falls below 0.25 on their tau score for values messaging, indicating a party that is strongly in agreement over these values. Meanwhile, Democrats are widely split on which values are important to discuss, with some candidates even demonstrating negative coordination (more discordant pairs than concordant pairs), indicating a nearly opposite set of priorities from the party at large when it comes to values-based language.

Figures 3.8 and 3.9 indicate how these trends have changed over time. Both Democrats and Republicans have become somewhat more coordinated over all four major categories, and the previous patterns hold constant across the entire time frame, with Republicans continuing to be more coordinated on values, while Democrats are more coordinated on campaign tactics and policy. These patterns remain the same when looking at cohort coordination (Figures 3.10, 3.11, 3.11a-3.11d, 3.12, 3.13).

Does Coordination Help, or Hurt?

Based on this understanding of party coordination, it may be interesting to examine whether coordination proved helpful to gubernatorial candidates over the past decade, in terms of actually winning races. Though a complete examination of this question is outside the scope of this data, I can provide answers to two preliminary questions. First, to what extent did party coordination relate to a candidate's probability of winning? And, to what extent did it affect their vote share? In order to address these questions, I shift to treating the data as sample data to allow for the use of regression analyses. The data includes two dependent variables relevant to these questions: a binary indicator of whether the candidate won or lost their election, and a continuous measure of vote share in the election results. Regression analyses were conducted including party and year as covariates.

Coordination's Effect on Victory

First, I examine the predicted probability of victory based on both party and cohort coordination. If there were no substantial relationship before adding covariates, no further analysis would be necessary. Figures 3.14 & 3.15 demonstrate that a substantively important relationship exists, at least in absence of common covariates. Republicans and Democrats both suffer a reduced chance of election victory the more coordinated they are with their overall party message. This effect disappears for Republicans, but not for Democrats, when shifting to a focus on cohort coordination instead of party coordination.

Before continuing, it's worth discussing what these results may mean. Due to the limited reach of Twitter, even with the added influence of traditional media drawing stories from the platform, I hesitate to imply any sort of causal influence of Twitter content to electoral results. To examine further, I use the total number of retweets each candidate received to subset the data by only the most engaging candidates. If the patterns observed above replicate across various levels of engagement, I can reasonably say that the above results are tapping into something unrelated to Twitter messaging. Figures 3.14a and 3.15a subset the data by only the top 50% of candidates on engagement (those whose tweets received over 250 retweets over the election cycle). The patterns remain the same, with Democrats facing a sharp decline in probability of victory the more coordinated they are with both the party and their cohorts, and Republicans receiving some penalty for party coordination that disappears when the cohort is examined separately.

Finally, since 250 retweets is still a relatively low number for an election cycle, I subset the data by only the top 25% of candidates on engagement (those who received over 1000 retweets over their election cycle). Figures 3.14b and 3.15b reflect this subset. The patterns are almost exactly the same for Democrats and Republicans, with the exception that Republican cohort coordination now shows a positive effect on electoral outcomes. This suggests that for Republicans only, Twitter content may actually have a causal effect on electoral outcomes, but for Democrats, Twitter content might just be a reflection of other factors external to the platform.

There are a couple of plausible explanations to consider. First, a candidate's Twitter content may reflect their overall campaign style – a focus on overall party message may reflect a lack of an individual identity to campaign on. This could have downstream effects in their canvassing efforts, advertising, etc. Second, the causal direction could flow from electoral results to Twitter content. Those who are in particularly difficult races that they are likely to lose may run with party-loyal messaging to curry favor from leadership for future patronage positions. There may be other factors external to Twitter that these results are tapping into, as well. These alternative explanations should be kept in mind as I continue examining these relationships.

The next step is to determine whether these effects remain significant when covariates are added. To that end, I run binomial regressions with election victory as the dependent variable. The first regression uses party coordination as the primary independent variable, and the second regression uses cohort coordination. Due to the relatively small N, I keep the number of covariates low. However, I include some of the most common covariates in political science research, including party (0 = Republican, 1 = Democrat), incumbency (0 = challenger, 1 = incumbent), the average same-party presidential vote share from the previous two presidential cycles, and a variable to indicate year (0 = 2010, 4 = 2018). Tables 2 & 3 report the results of these regressions.

Even after including some of the most common and powerful predictors of election success, message coordination with others of your own party still reaches traditional levels of significance. Coordination, whether with the party at large, or with your own cohort, is related to a lower chance of winning a statewide election. As expected, incumbency and the president's vote share are positively related to election victory.

Coordination's Effect on Vote Share

Looking at victory or loss alone is an incomplete picture. It's an outcome variable that doesn't allow for much nuance – there are a number of possible explanations for the patterns seen in the last section. Therefore, it's worthwhile to look at a candidate's vote share for a more intricate understanding of how election results are affected. Just as with the probability of victory, I begin by graphing the relationship between vote share and coordination. Figure 3.16 demonstrates the relationship between party coordination and vote share, while Figure 3.17 shows cohort coordination. Here, the picture is much less clear than in the previous section. Republicans appear to have an advantage across the spectrum of coordination, and cohort coordination appears to have a smaller penalty than party coordination.

The linear regression models presented in Tables 4 & 5 reinforce this interpretation. As with the binomial models, party coordination has a significantly negative effect on a candidate's vote share. However, message coordination with one's cohort, while negative in direction, is not significant ($p = 0.197$).

2018

The final investigation I pursue is how victory and vote share patterns change over time. Though the year of the election is a non-significant variable in all of the above regression models, here I present some data that may have different effects in 2018 than in the other years of the dataset. First, Figure 3.18 shows how much each year's ranked priorities differ from the overall party brand. In 2018, both Republicans and Democrats deviated from the overall party brand more than in any other year in the dataset.

Figures 3.19 and 3.20 examine the average party coordination and cohort coordination, respectively, of winners and losers in each year of the dataset. In 2018, election losers from both parties were more coordinated with the overall party brand than election winners from either party. However, Democratic-party winners, for the first time in the dataset, were more coordinated with members of their cohort than any of the other groups.

Finally, Figure 3.21 shows the relationship between a candidate's vote share and their party and cohort coordination in 2018. Higher levels of cohort coordination provided an advantage over higher levels of overall party coordination. However, this visualization adds needed context to previous findings by demonstrating that although cohort coordination was particularly advantageous for Democrats in 2018, it still did not bring their average vote share above 50%. Instead, these races were more competitive, but were still far from certain outcomes.

Taken together, these results indicate that in 2018, both parties experimented with somewhat different priorities in messaging. For Democrats, at least, the experiment seems to have worked. The change in messaging seems to have led to races becoming competitive that otherwise would have been much more difficult. For an example of the difference between 2018 and previous years, Figure 3.22 presents coordination patterns in 2014. Other years in the dataset look similar to 2014, with little differences in vote share between party and cohort coordination, and a nearly-universal Republican advantage. Only 2018 demonstrates a unique pattern, with Republicans losing vote share for coordination and Democrats gaining from coalescing around their new messaging strategy.

Conclusion

To summarize the findings of this chapter, I find that parties are fairly coordinated overall, and are slightly more coordinated with members of their own cohort than with the party overall, reflecting the changing messaging needs of each election cycle. Republicans and Democrats also demonstrate somewhat different patterns of coordination. Republicans demonstrate high levels of coordination around values-based language, while Democrats are highly uncoordinated on values. Democrats are mostly coordinated around policy and campaign tactics. When it comes to using this data to predict campaign success, coordination with larger party messaging largely hurts candidates. Finally, evidence suggests that 2018 is an aberration in messaging patterns. Both parties used somewhat different messaging priorities in 2018, and coordination around the new priorities had unusual effects on the electoral success of both parties.

Tables

Table 1. Party & Cohort Coordination

	Democrats		Republicans	
	Mean	Std. Dev	Mean	Std. Dev
Party Coordination	.602	.105	.611	.083
Cohort Coordination	.638	.095	.634	.083

Table 2. Model 1: The Effect of Party Coordination on Victory

```
=============================================
                     Dependent variable:
                 ---------------------------
                           Winner
---------------------------------------------
Party Coord                -4.971**
                           (1.957)

Incumbent                   2.252***
                           (0.464)

Democrat                   -0.800**
                           (0.351)

Presidential Vote          11.289***
                           (2.077)

Year                        0.018
                           (0.110)

Constant                   -2.585*
                           (1.363)

---------------------------------------------
Observations                 227
Log Likelihood            -105.716
Akaike Inf. Crit.          223.431
=============================================
Note:            *p<0.1; **p<0.05; ***p<0.01
```

Table 3. Model 2: The Effect of Cohort Coordination on Victory

	Dependent variable:
	Winner
Cohort Coord	-3.340*
	(1.923)
Incumbent	2.269***
	(0.462)
Democrat	-0.750**
	(0.348)
Presidential Vote	11.078***
	(2.068)
Year	-0.008
	(0.108)
Constant	-3.340**
	(1.385)
Observations	227
Log Likelihood	-107.608
Akaike Inf. Crit.	227.215
Note:	*p<0.1; **p<0.05; ***p<0.01

Table 4. Model 3: The Effect of Party Coordination on Vote Share

	Dependent variable:
	Vote Share
Party Coordination	-10.336*
	(5.294)
Incumbent	6.267***
	(1.129)
Democrat	-4.829***
	(0.965)
Presidential Vote	44.566***
	(5.085)
Year	0.392
	(0.314)
Constant	32.545***
	(3.881)
Observations	227
Log Likelihood	-766.787
Akaike Inf. Crit.	1,545.574
Note:	*p<0.1; **p<0.05; ***p<0.01

Table 5. Model 4: The Effect of Cohort Coordination on Vote Share

	Dependent variable:
	Vote Share
Cohort Coordination	-7.080
	(5.471)
Incumbent	6.364***
	(1.133)
Democrat	-4.733***
	(0.969)
Presidential Vote	44.553***
	(5.156)
Year	0.332
	(0.314)
Constant	30.836***
	(4.002)
Observations	227
Log Likelihood	-767.871
Akaike Inf. Crit.	1,547.741
Note:	*p<0.1; **p<0.05; ***p<0.01

Figures

Figure 3.1 - Overall Party Coordination

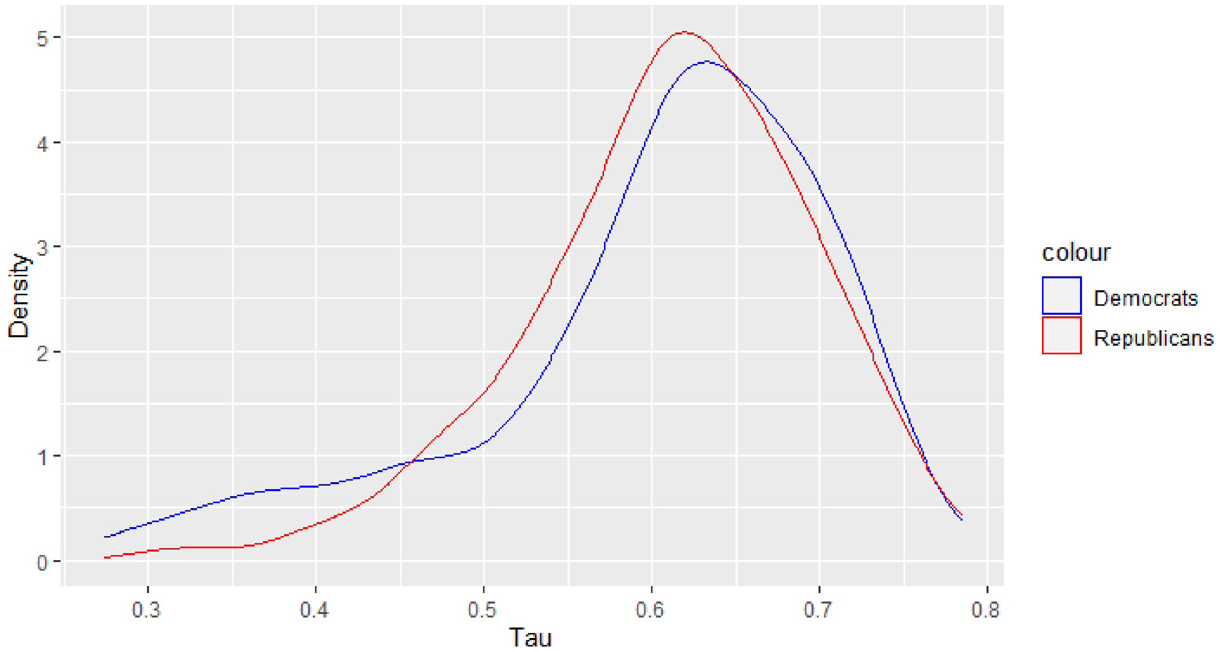


Figure 3.2 - Overall Coordination by Year

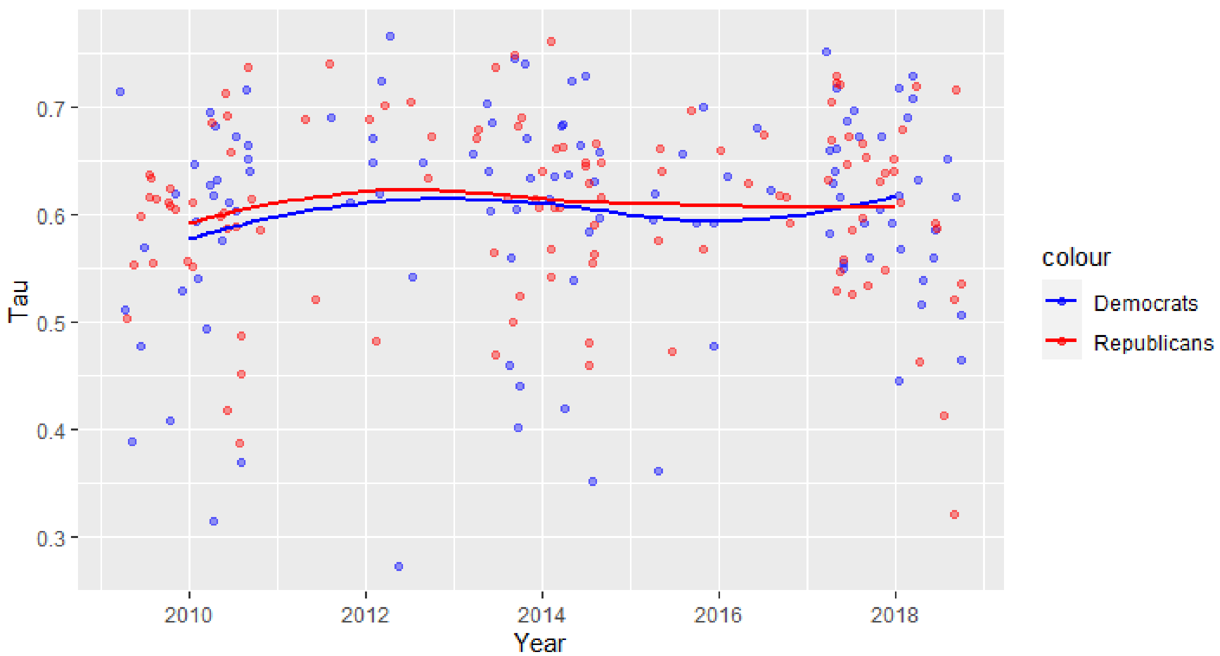


Figure 3.3 - Cohort vs. Overall Coordination - Democrats

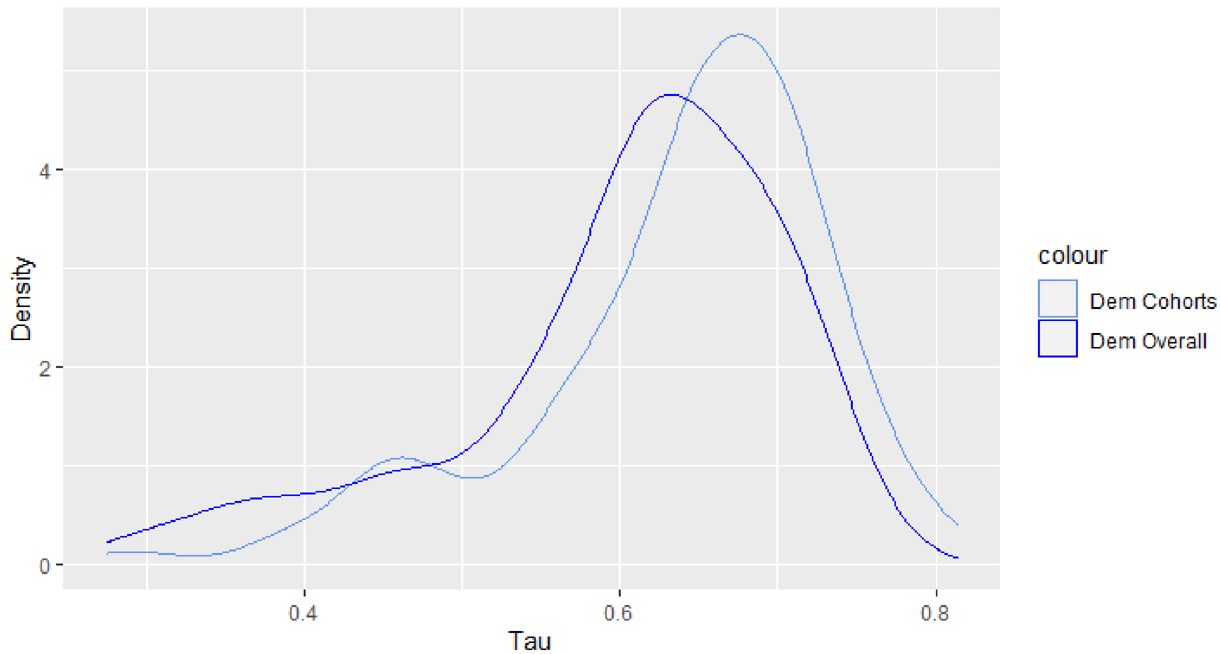
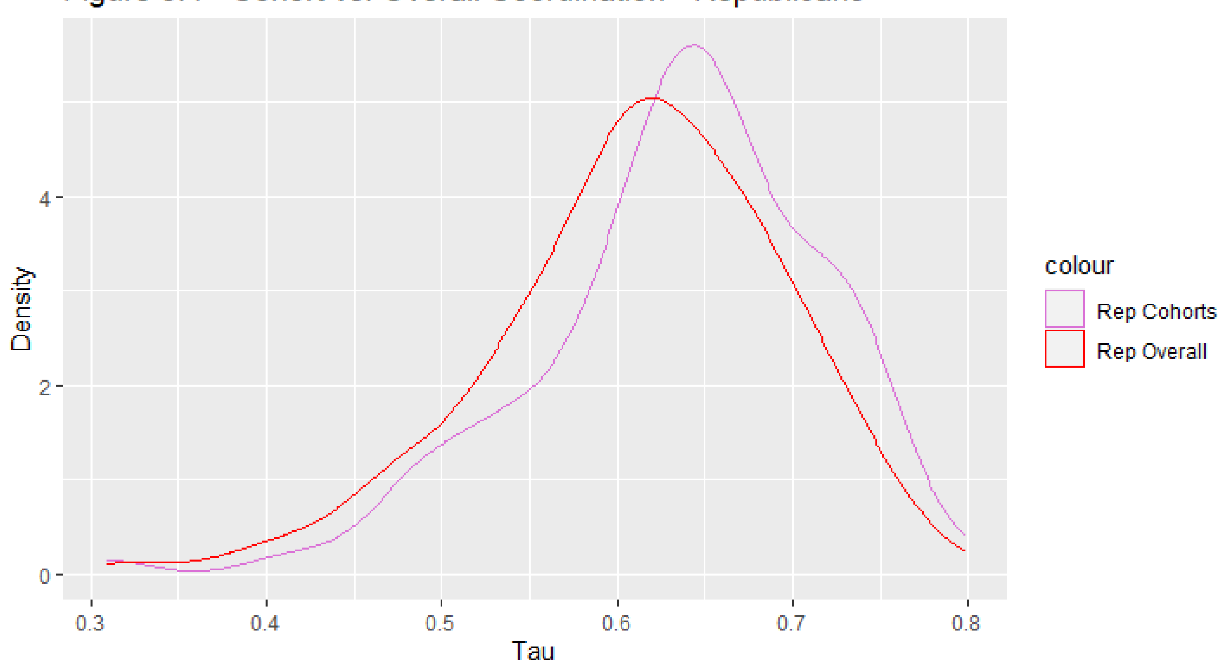


Figure 3.4 - Cohort vs. Overall Coordination - Republicans



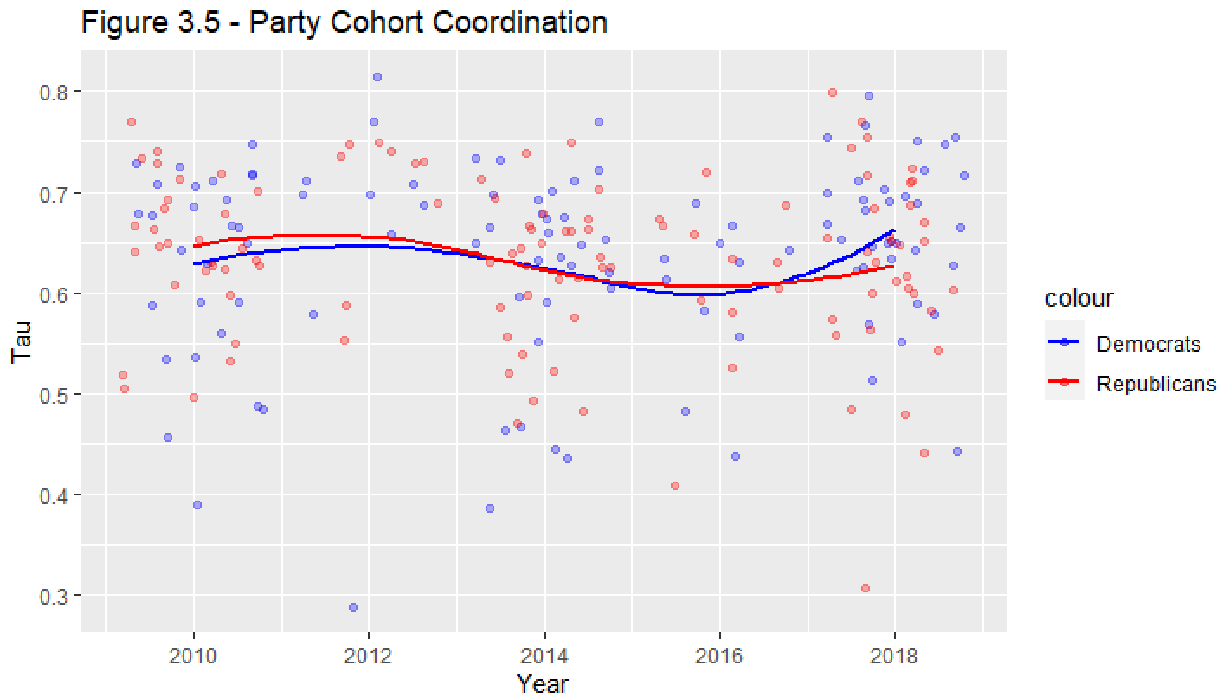


Figure 3.5 - Party Cohort Coordination
Tau
Year
colour
Democrats
Republicans
0.8
0.7
0.6
0.5
0.4
0.3
2010
2012
2014
2016
2018

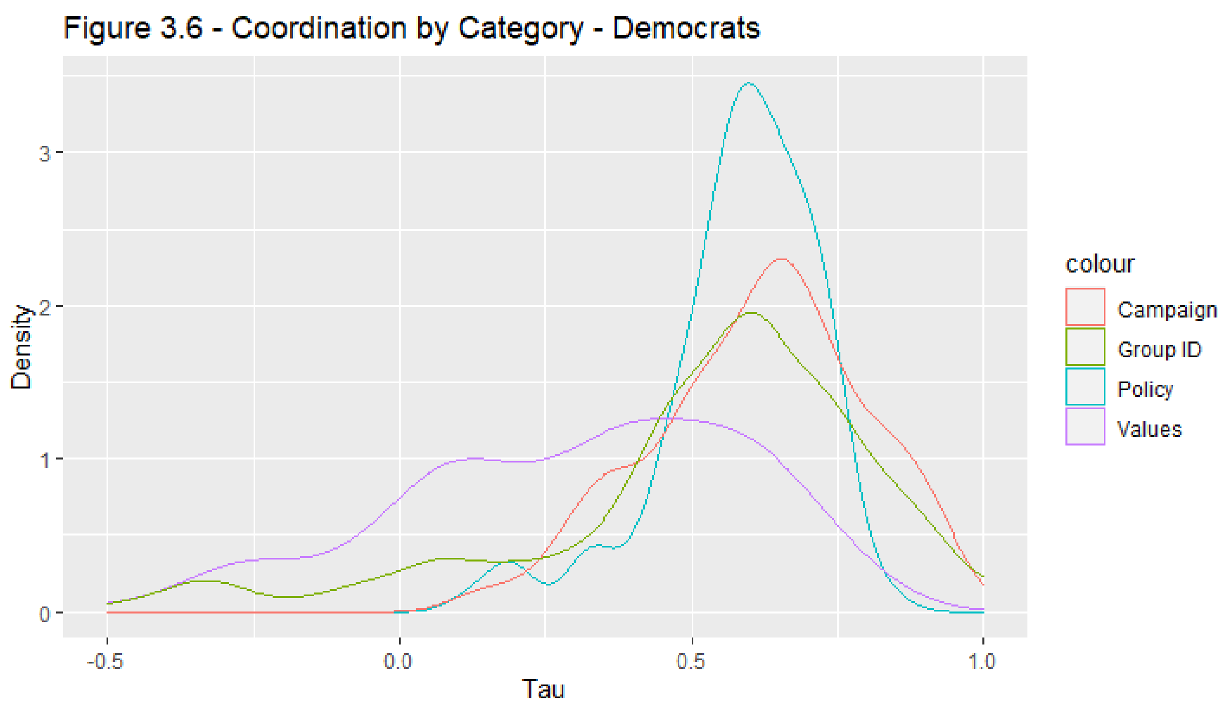


Figure 3.6 - Coordination by Category - Democrats
Density
Tau
colour
Campaign
Group ID
Policy
Values
3
2
1
0
-0.5
0.0
0.5
1.0

Figure 3.7 - Coordination by Category - Republicans

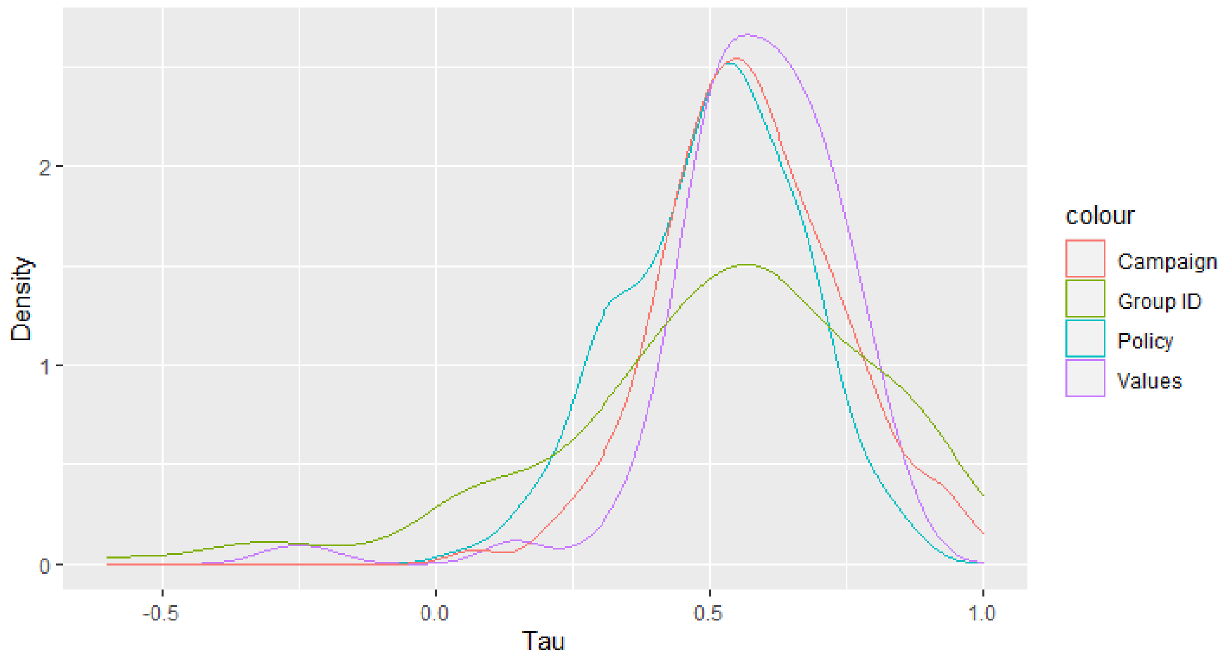


Figure 3.7a - Party Differences in Policy Coordination

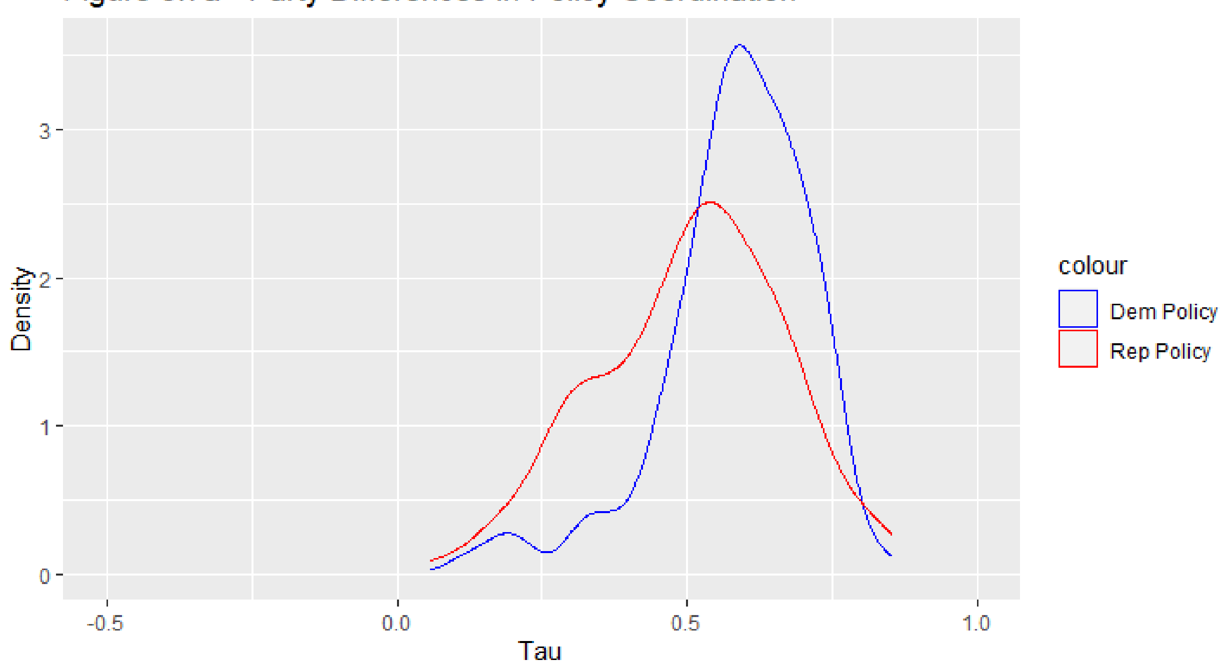


Figure 3.7b - Party Differences in Values Coordination

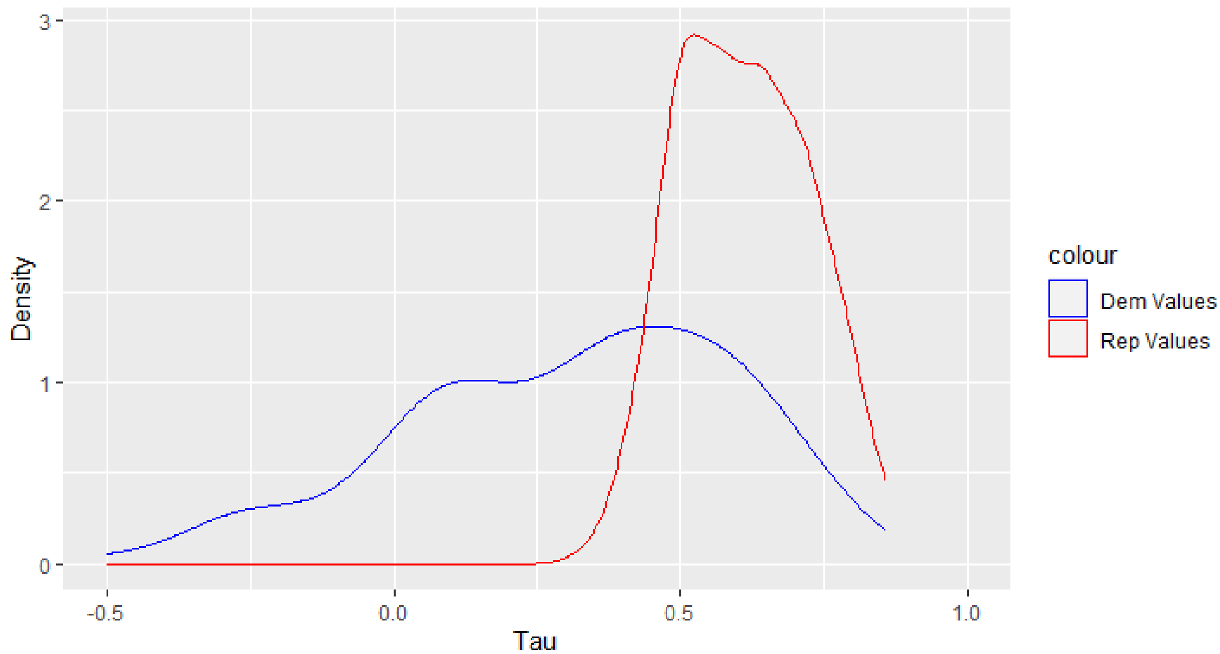


Figure 3.7c - Party Differences in Campaign Coordination

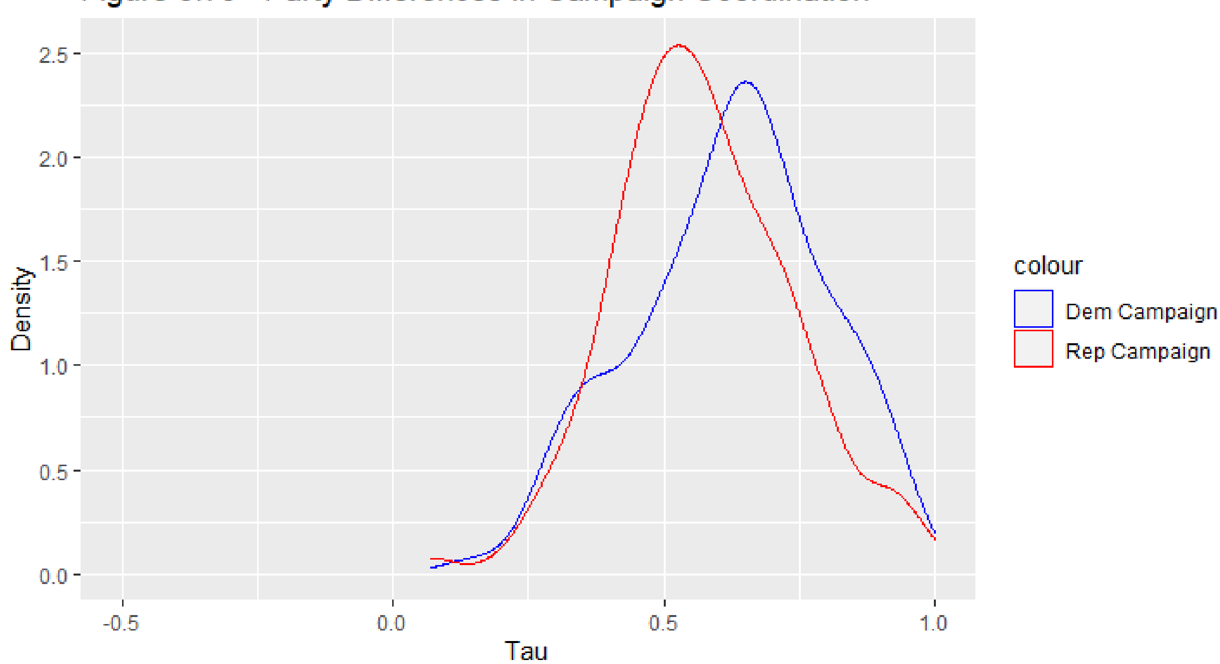


Figure 3.7d - Party Differences in Group Coordination

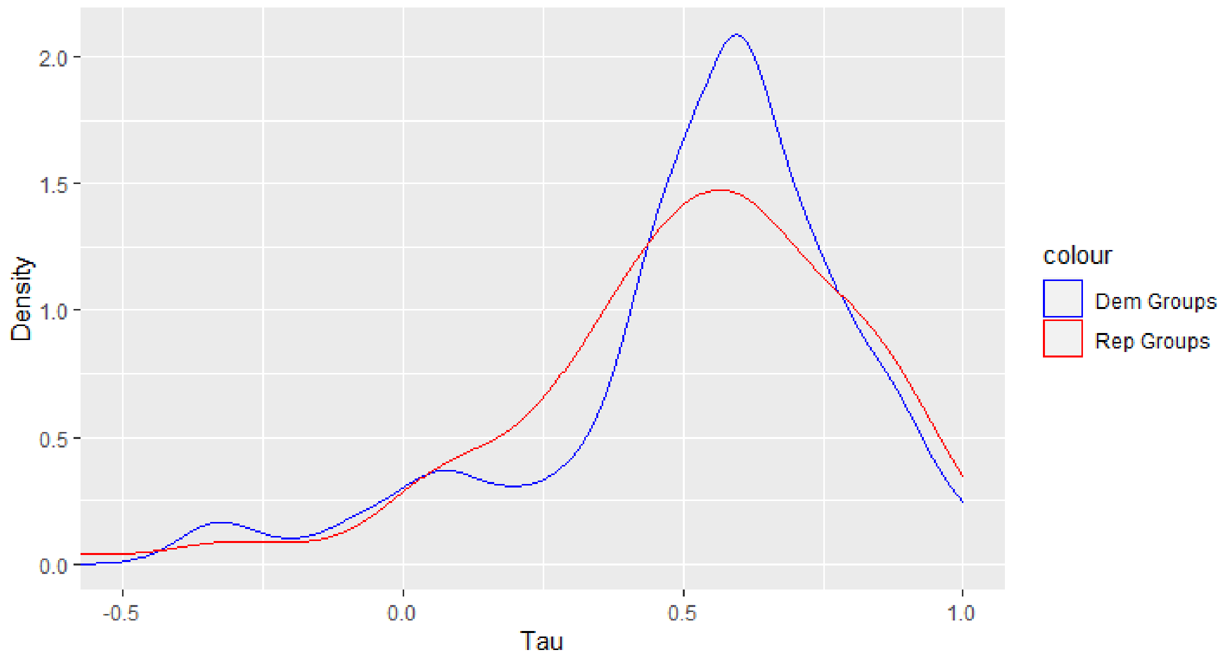


Figure 3.8 - Coordination by Category and Year - Democrats

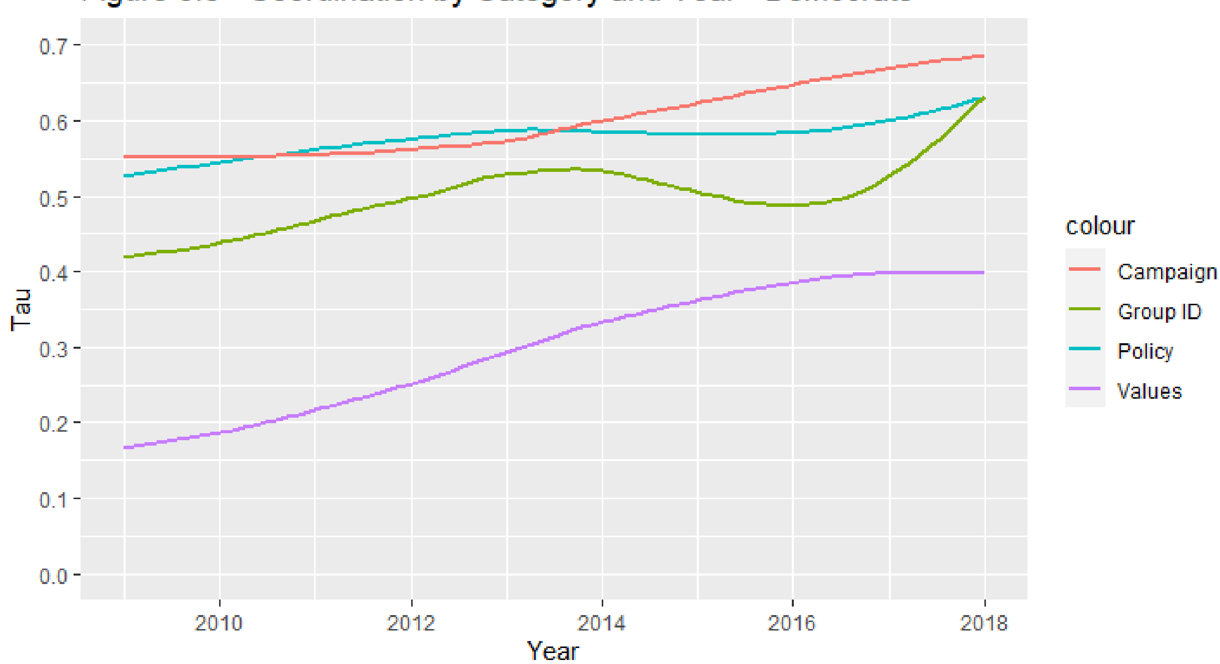


Figure 3.9 - Coordination by Category and Year - Republicans

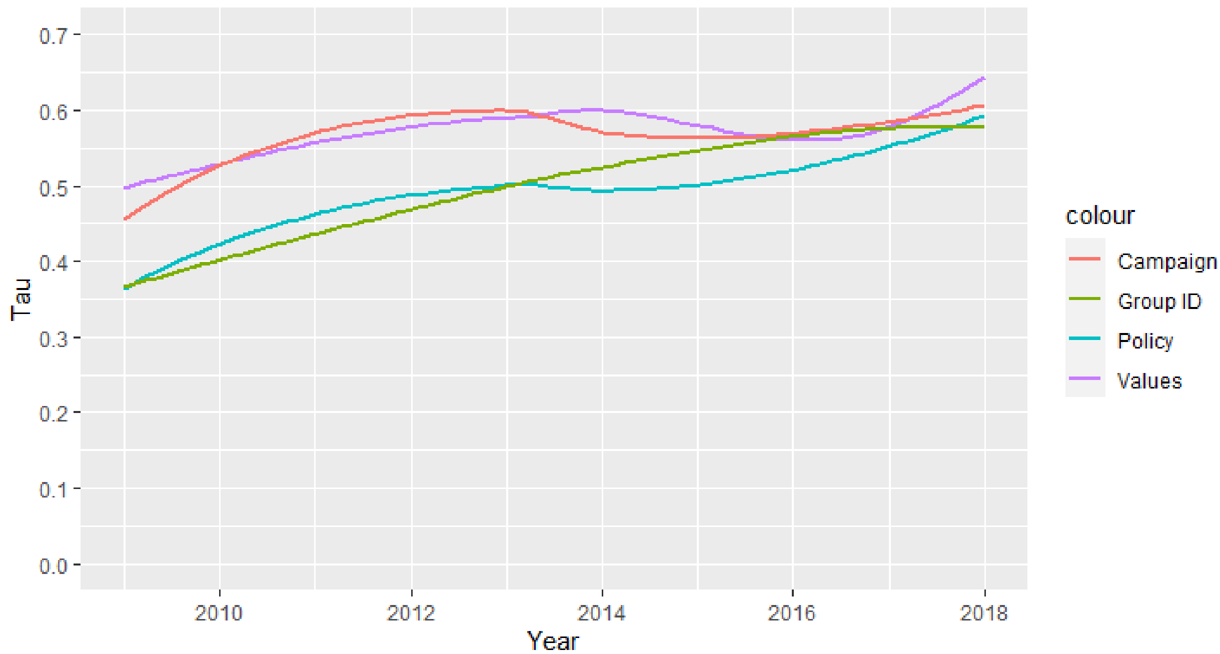


Figure 3.10 - Cohort Coordination - Democrats

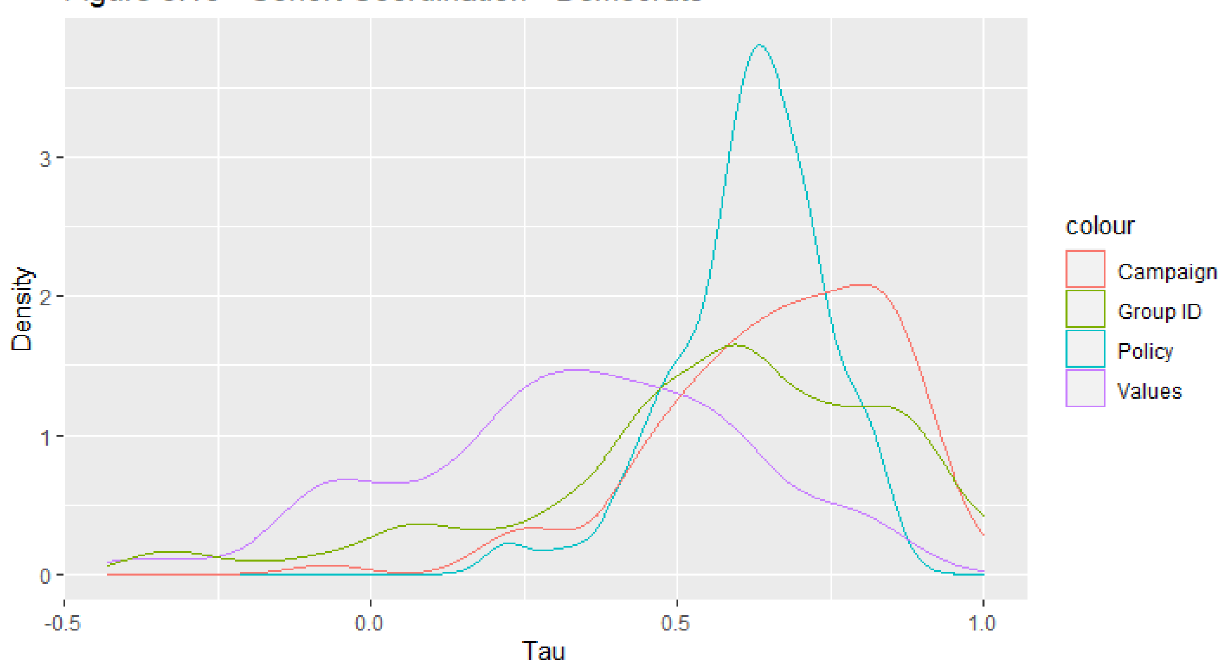


Figure 3.11. Cohort Coordination - Republicans

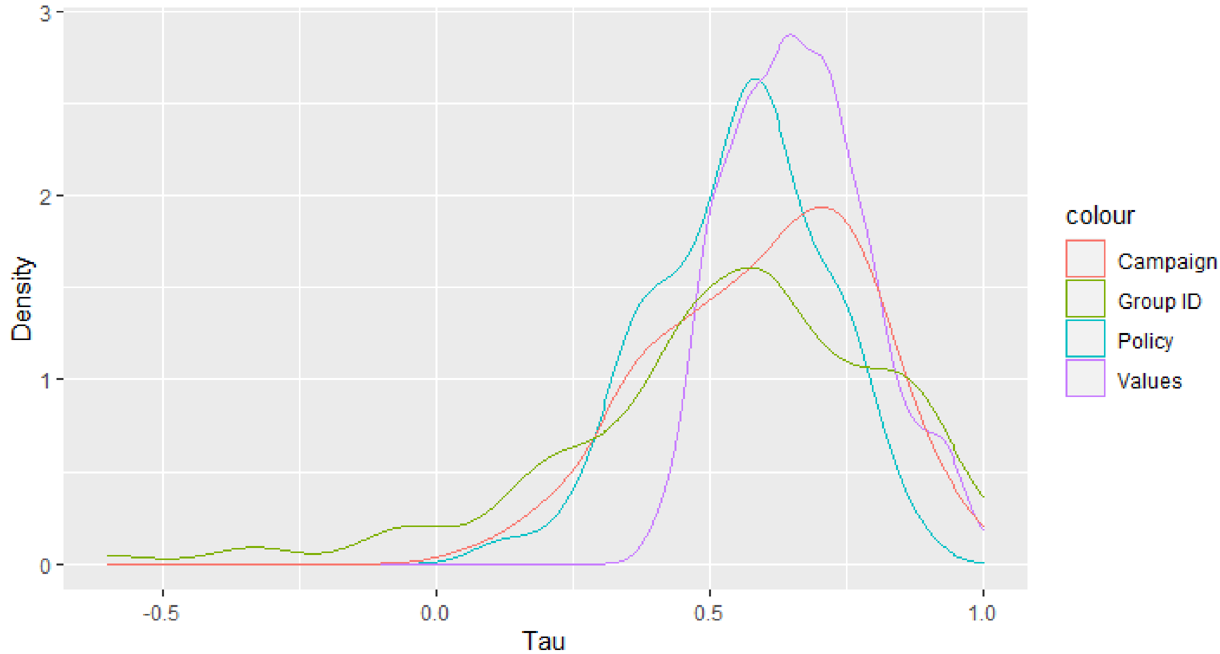
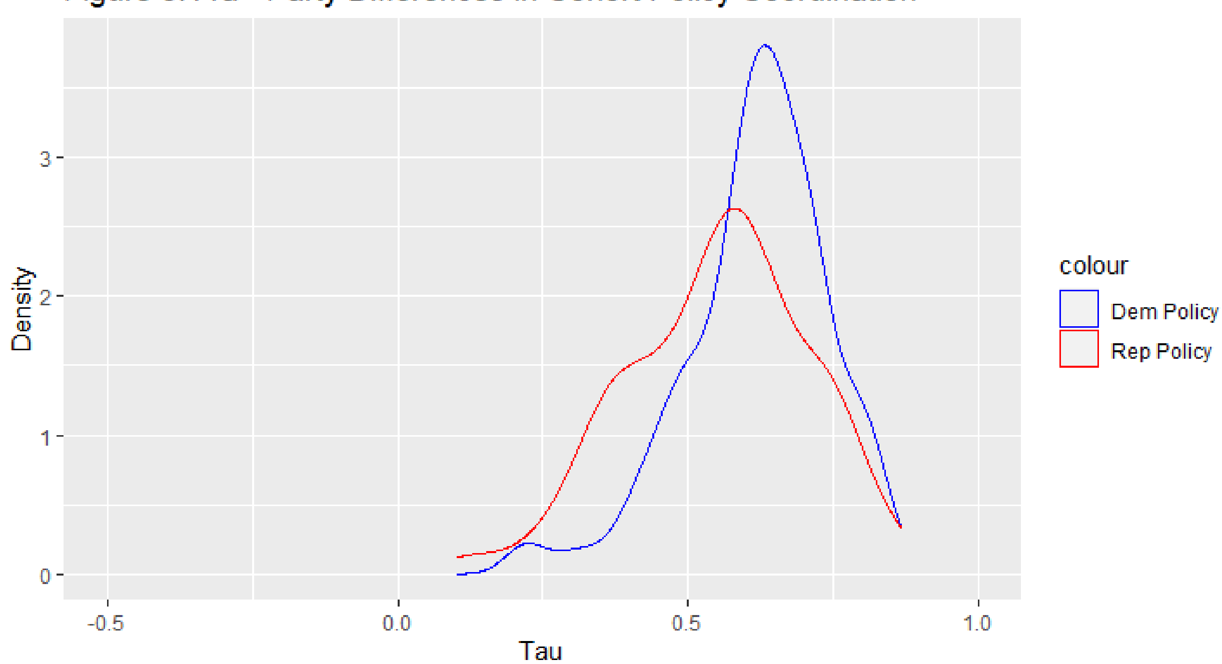


Figure 3.11a - Party Differences in Cohort Policy Coordination



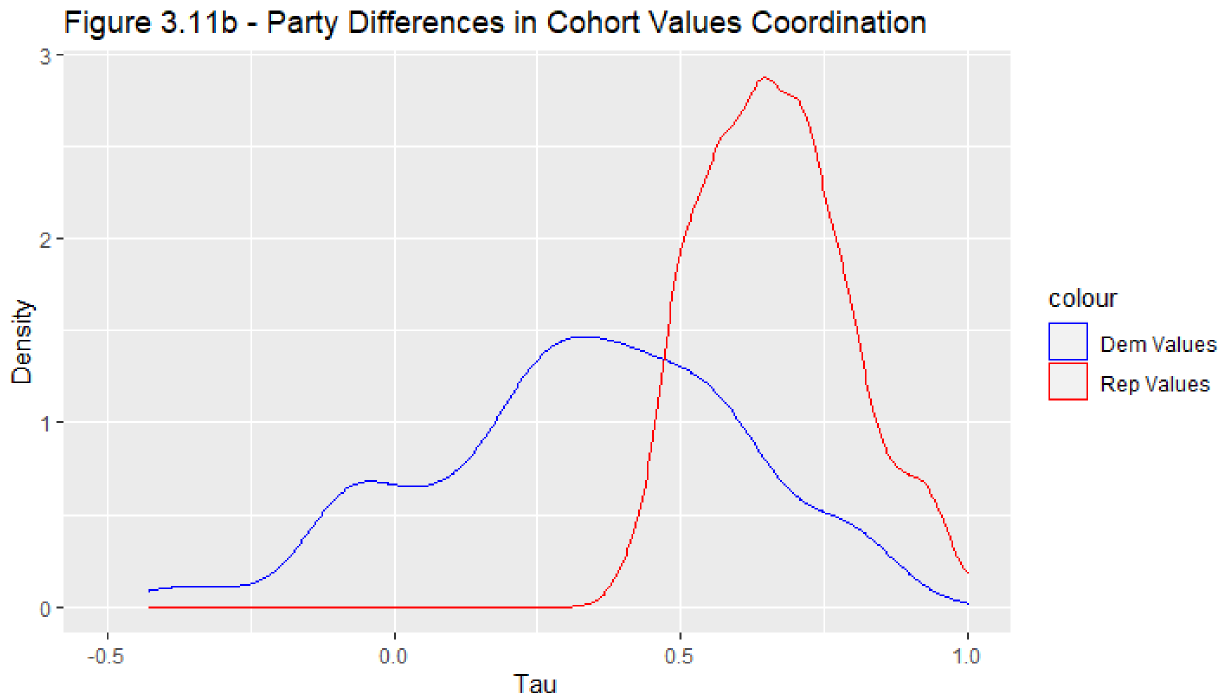


Figure 3.11b - Party Differences in Cohort Values Coordination
Density
3
2
1
0
-0.5
0.0
0.5
1.0
Tau
colour
Dem Values
Rep Values

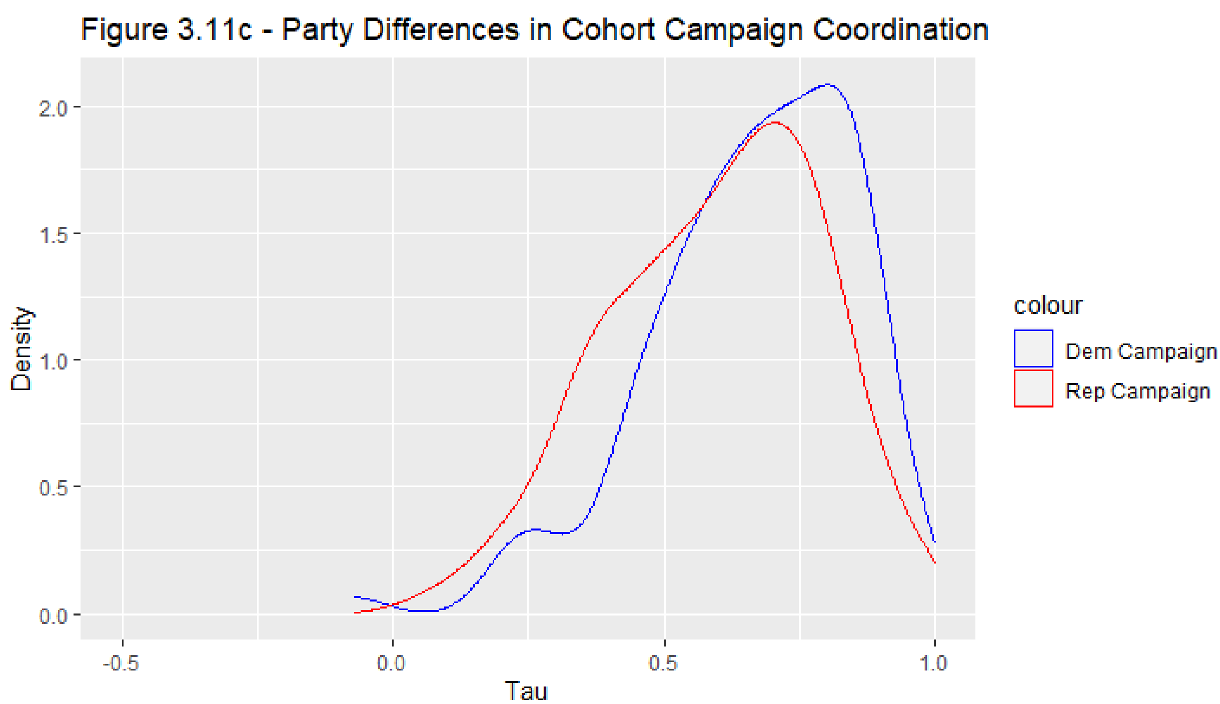


Figure 3.11c - Party Differences in Cohort Campaign Coordination
Density
2.0
1.5
1.0
0.5
0.0
-0.5
0.0
0.5
1.0
Tau
colour
Dem Campaign
Rep Campaign

Figure 3.11d - Party Differences in Cohort Group Coordination

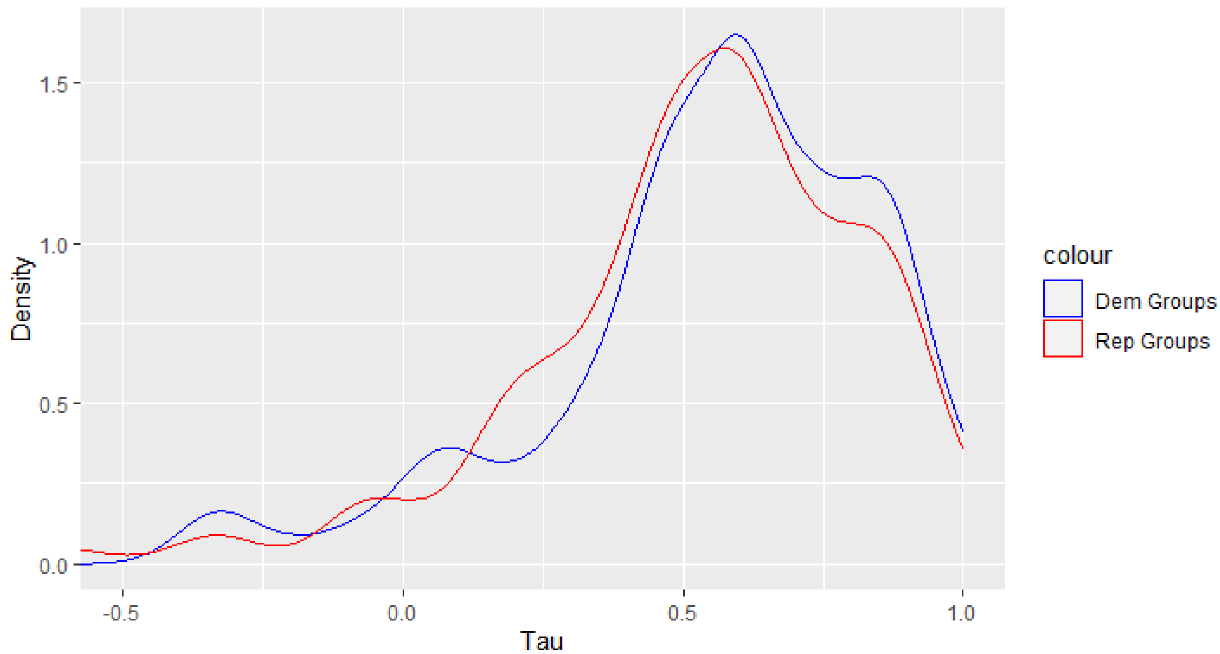


Figure 3.12. Cohort Coordination by Year - Democrats

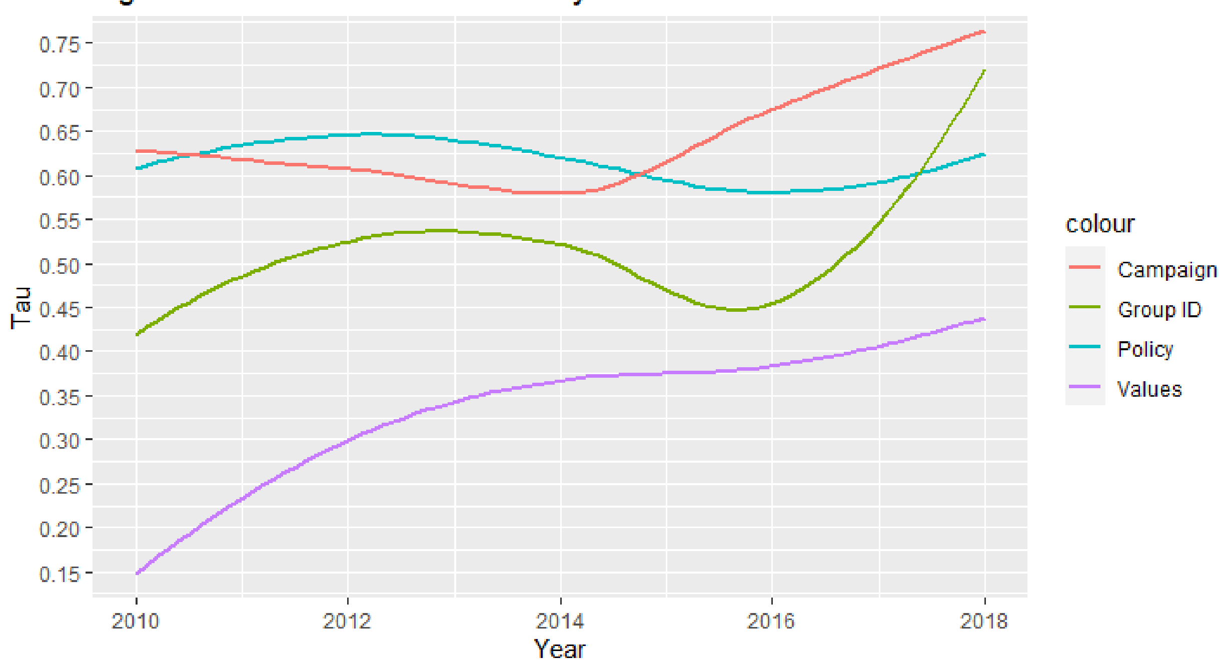


Figure 3.13 - Cohort Coordination by Year - Republicans

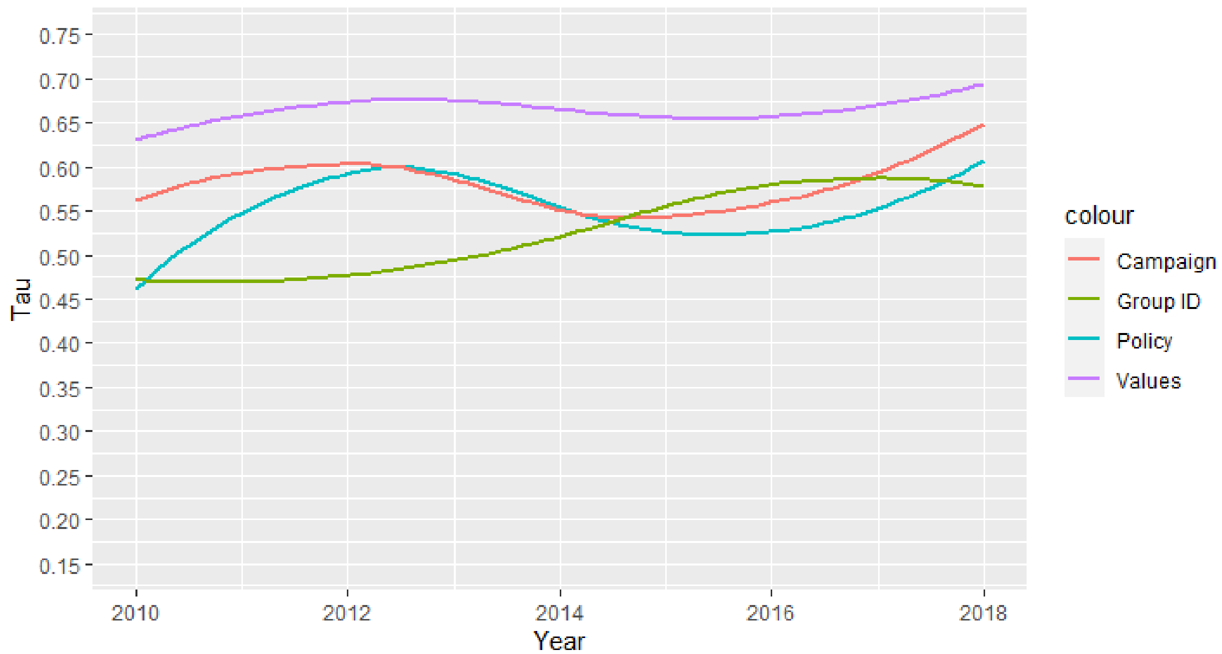
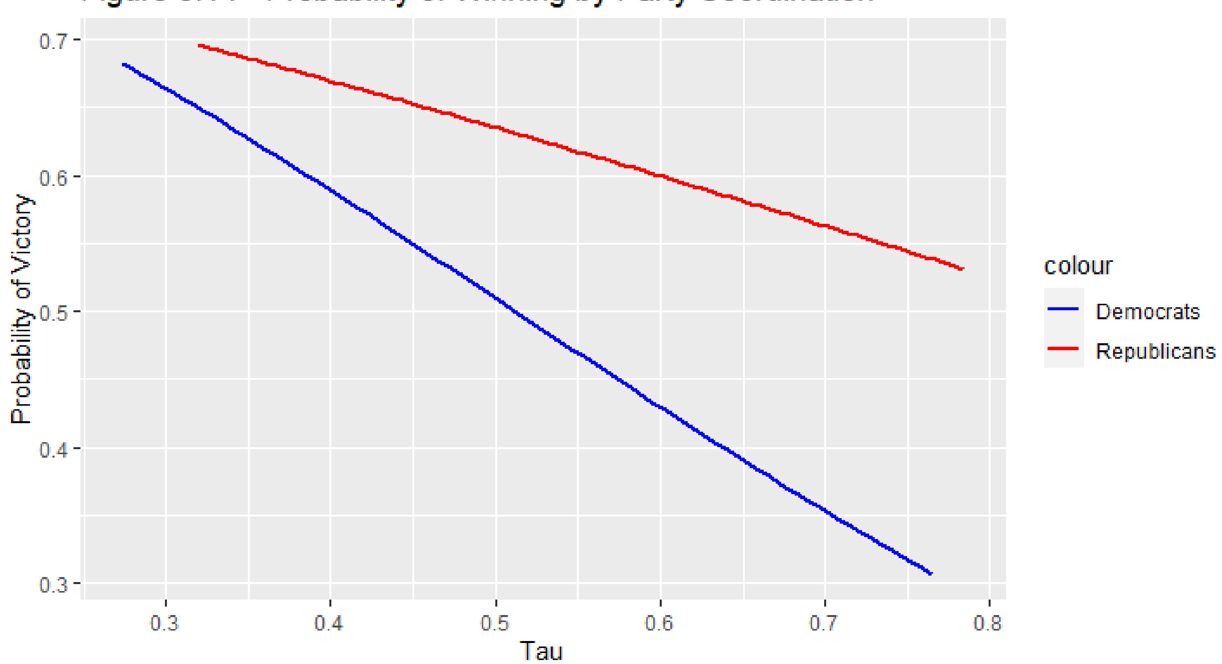


Figure 3.14 - Probability of Winning by Party Coordination



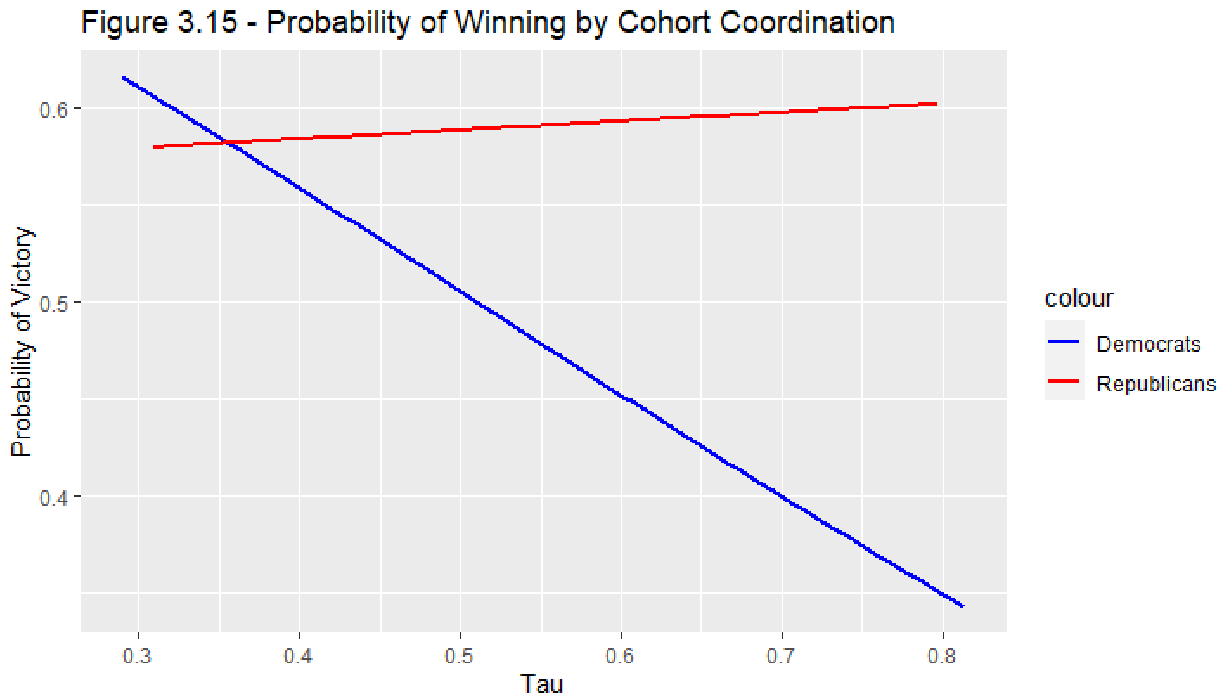


Figure 3.15 - Probability of Winning by Cohort Coordination
0.6
0.5
0.4
Probability of Victory
0.3
0.4
0.5
0.6
0.7
0.8
Tau
colour
Democrats
Republicans

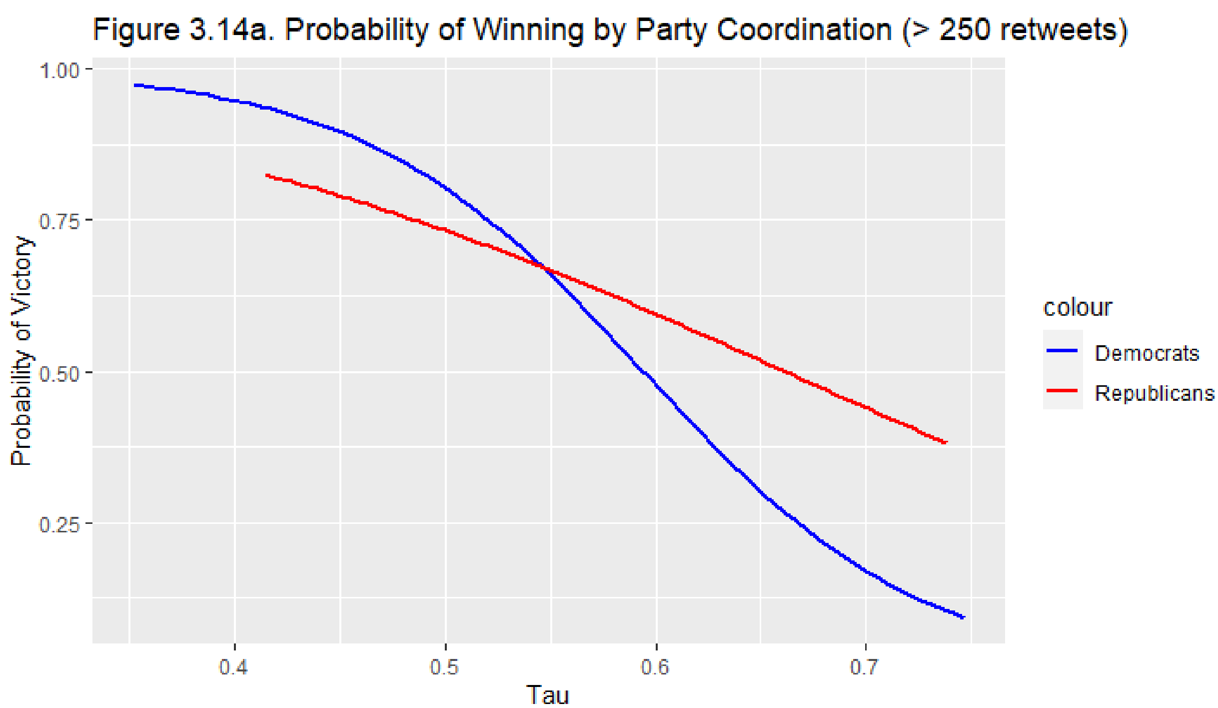


Figure 3.14a. Probability of Winning by Party Coordination (> 250 retweets)
1.00
0.75
0.50
0.25
Probability of Victory
0.4
0.5
0.6
0.7
Tau
colour
Democrats
Republicans

Figure 3.15a. Probability of Winning by Cohort Coordination (>250 retweets)

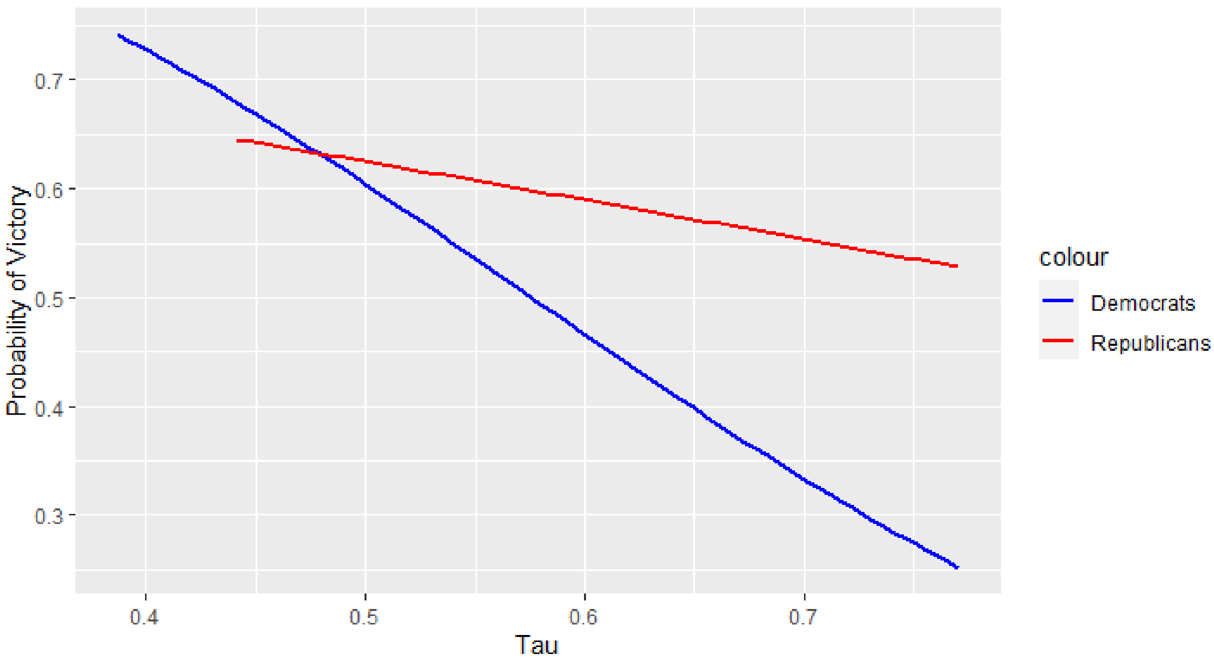


Figure 14b. Probability of Winning by Party Coordination (>1000 retweets)

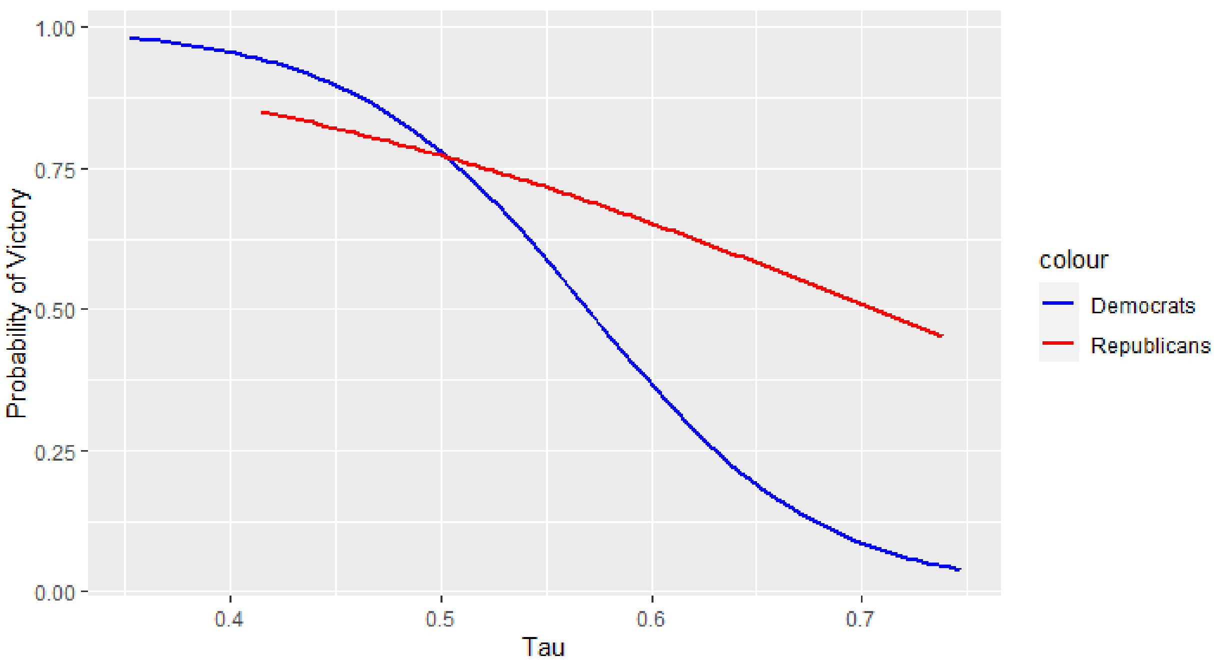


Figure 15b. Probability of Winning by Cohort Coordination (>1000 retweets)

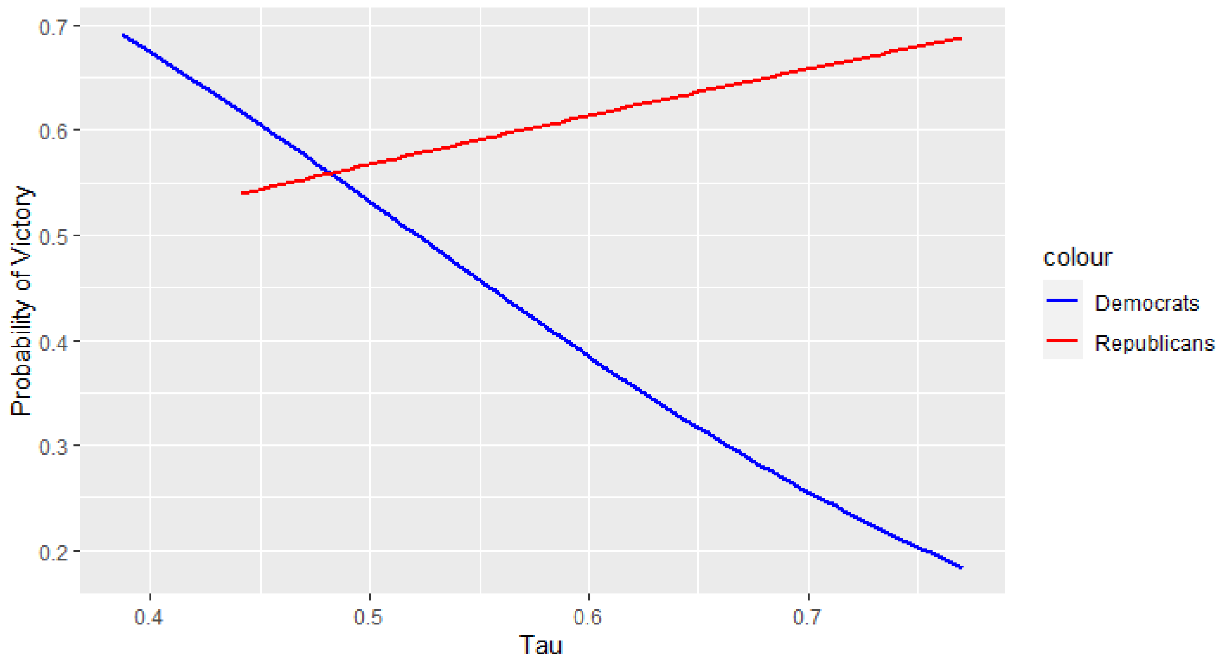


Figure 3.16 - Vote Share by Party Coordination

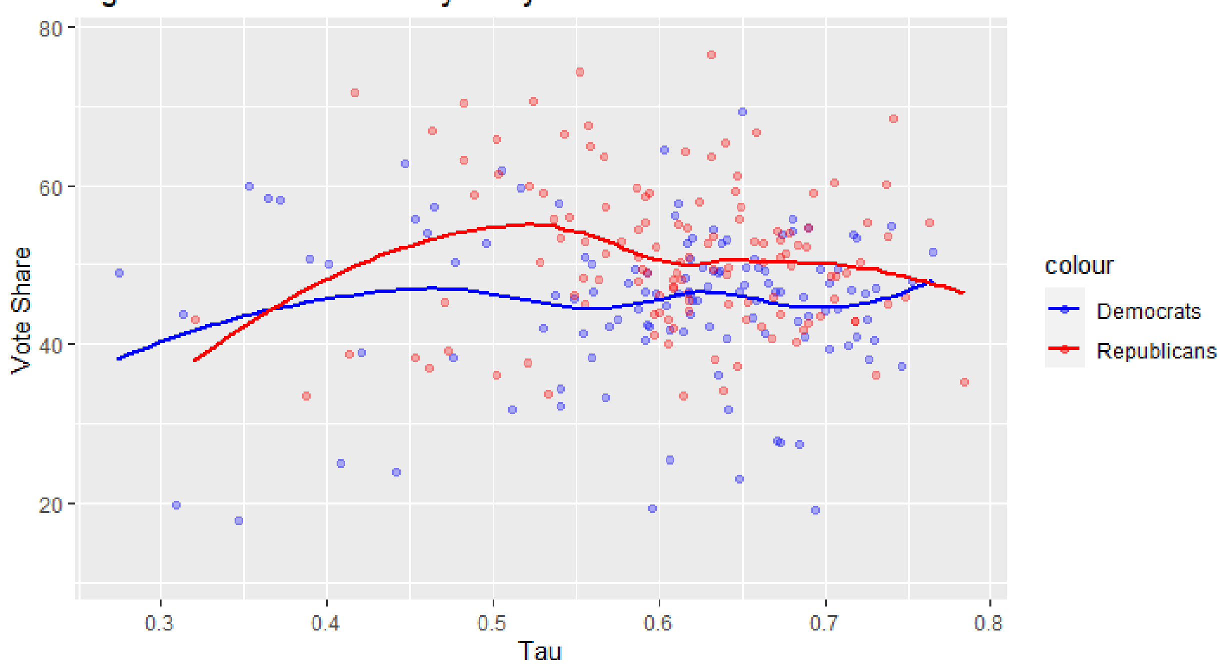


Figure 3.17 - Vote Share by Cohort Coordination

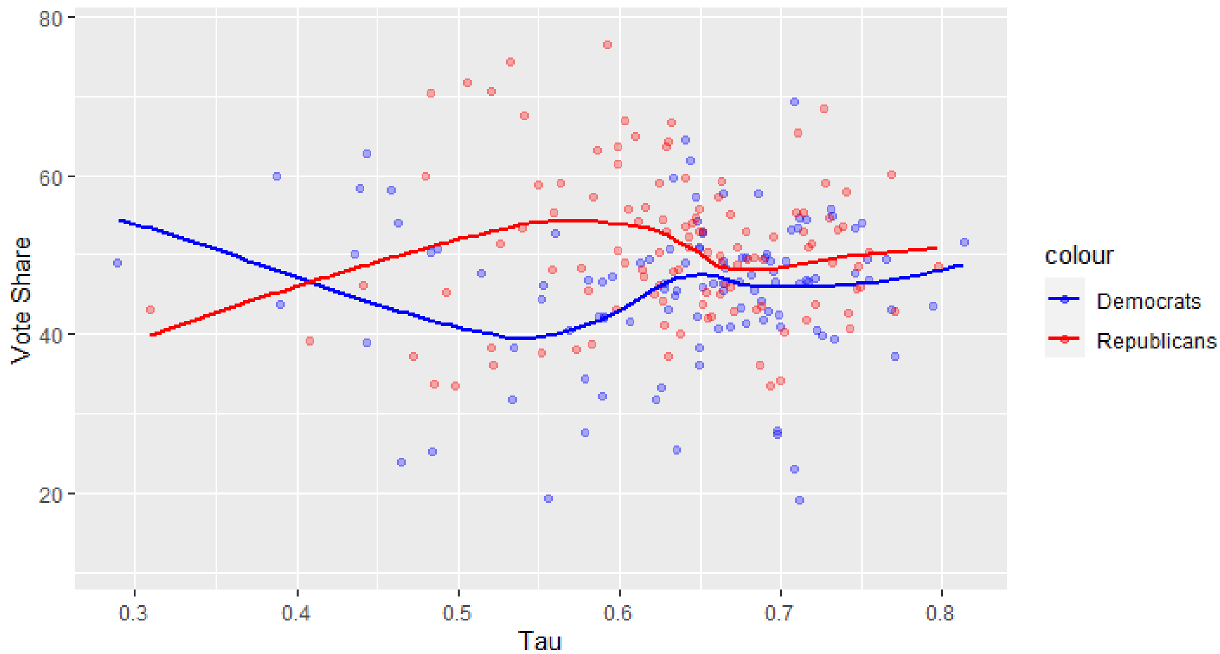


Figure 3.18 - Cohort coordination to Overall Message

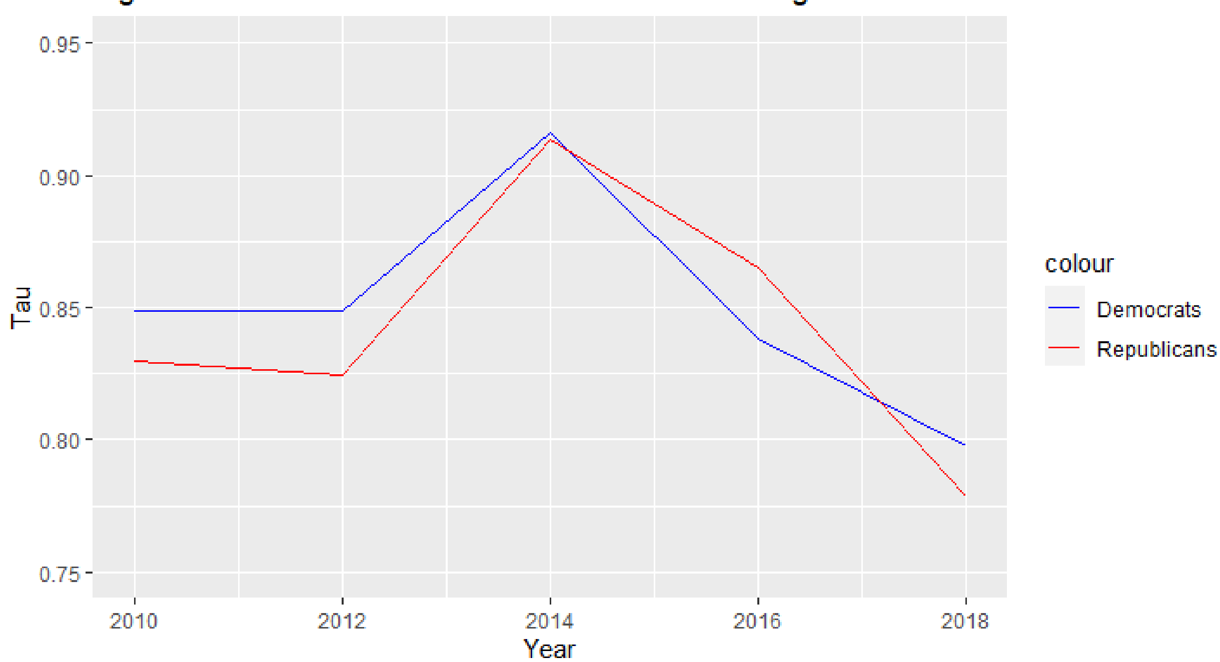


Figure 3.19 - Average Party Coordination by Year

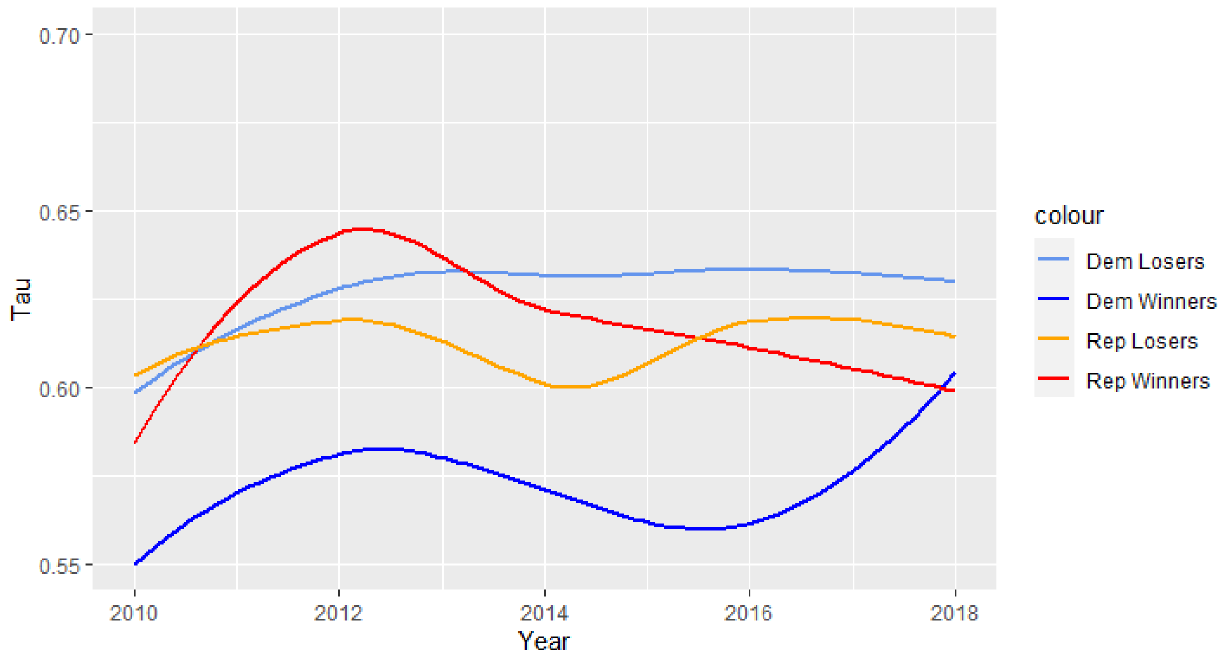


Figure 3.20 - Average Cohort Coordination by Year

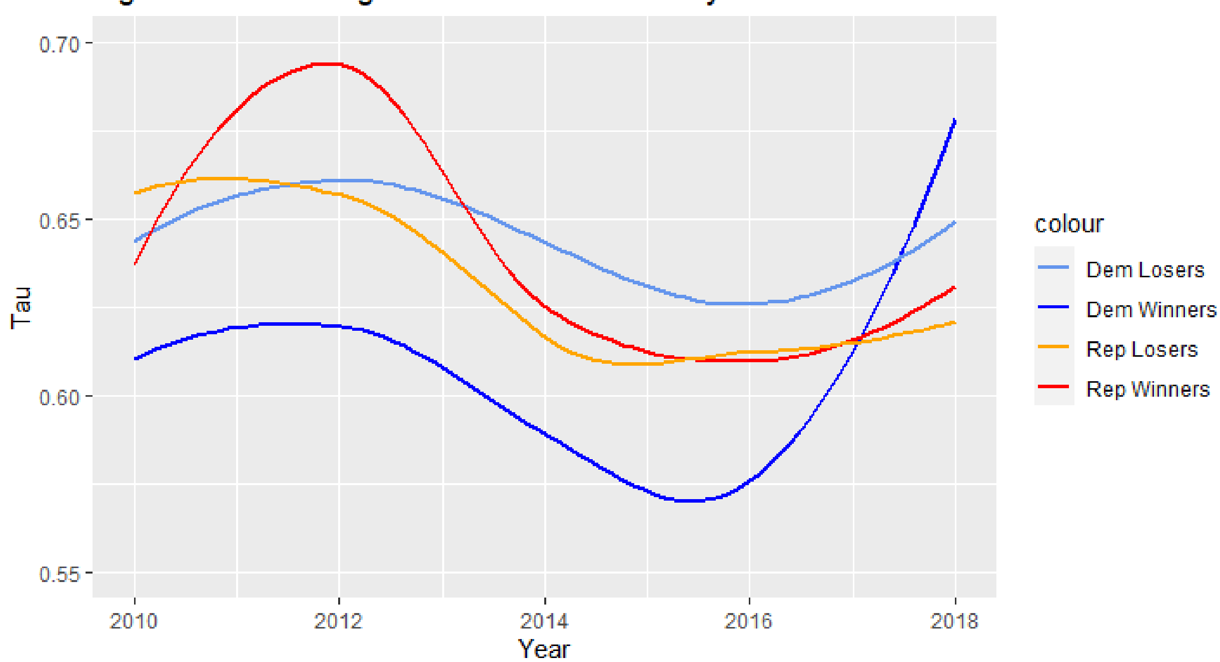


Figure 3.21 - Vote Share by Coordination in 2018

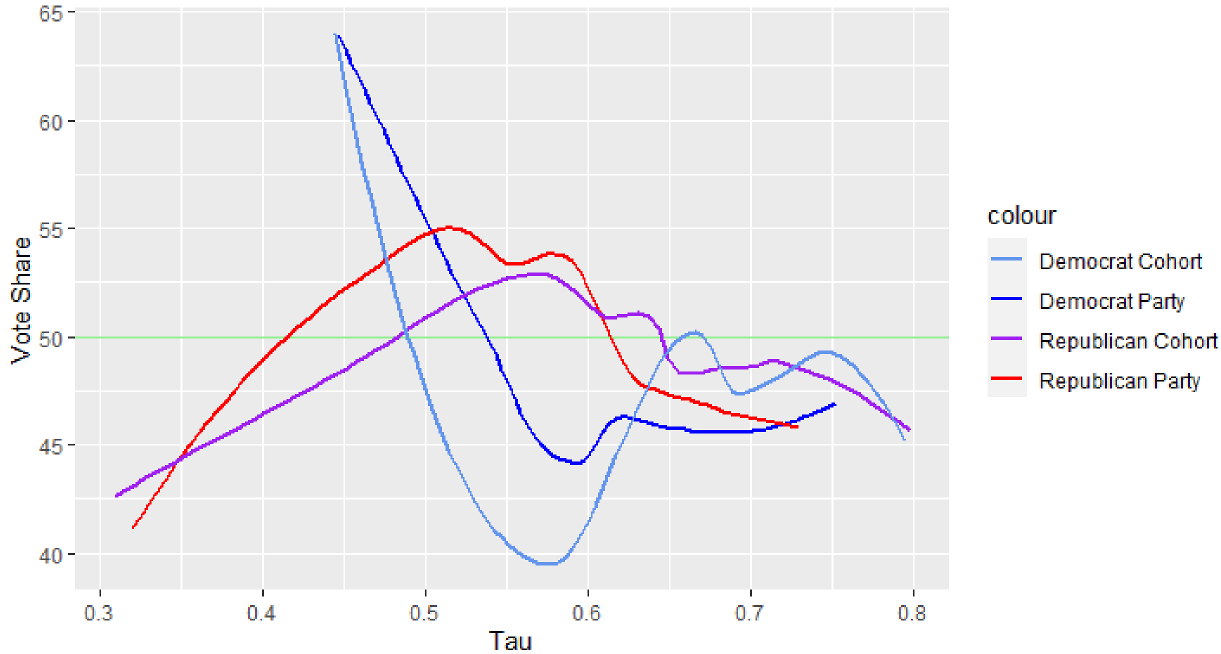
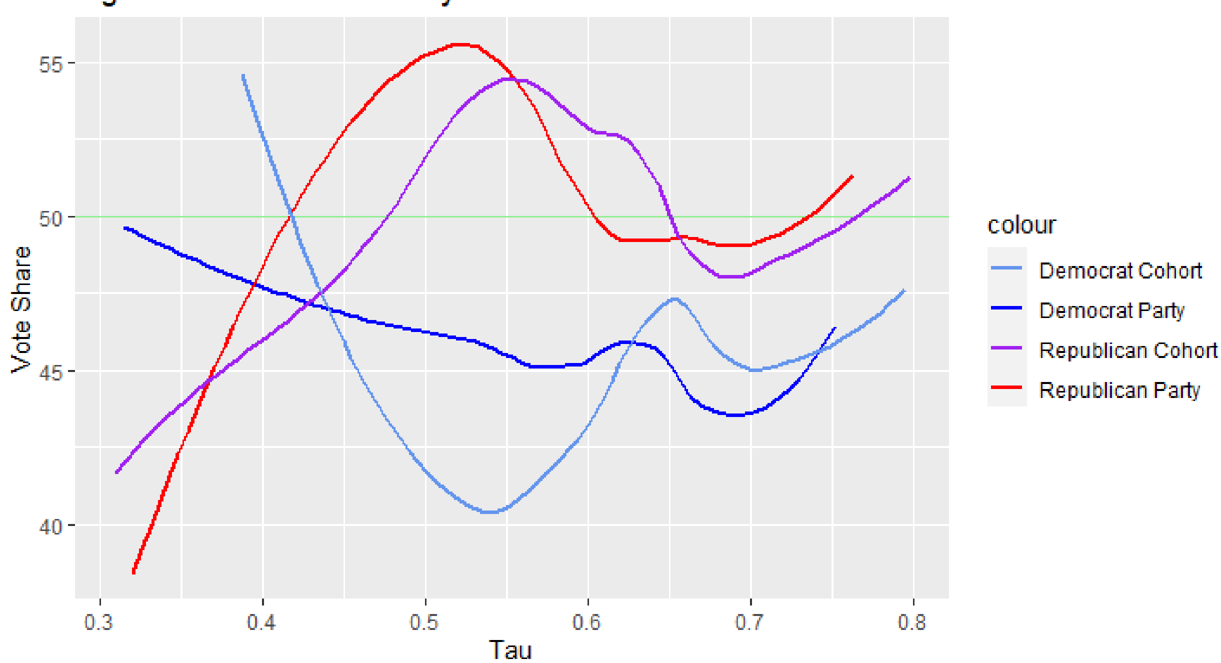


Figure 3.22 - Vote Share by Coordination in 2014



References

Abramowitz, A. I., & Saunders, K. L. (2008). Is polarization a myth? *The Journal of Politics, 70*(2), 542-555.

Borda, J. d. (1781). 'M'emoire sur les les e'lections au scrutiny', M'emoires de l'Acad'emie Royale des Sciences. *English translation by a. DEGRAZIA, Isis (1953), 136*.

Brown, A. R. (2017). The challenge of local party brand differentiation. *The Social Science Journal, 54*(1), 51-55.

Bruns, D., Langner, T., & Fischer, A. (2017). The origins of brand love: A typology of starting points. *Marketing: ZFP–Journal of Research and Management, 39*(2), 38-48.

Butler, D. M., & Powell, E. N. (2014). Understanding the party brand: experimental evidence on the role of valence. *The Journal of Politics, 76*(2), 492-505.

Canes-Wrone, B., Brady, D. W., & Cogan, J. F. (2002). Out of step, out of office: Electoral accountability and House members' voting. *American Political Science Review*, 127-140.

Clarke, A. J. (2020). Party Sub-Brands and American Party Factions. *American Journal of Political Science, 64*(3), 452-470.

Cox, G. W., Kousser, T., & McCubbins, M. D. (2010). Party power or preferences? Quasi-experimental evidence from American state legislatures. *The Journal of Politics, 72*(3), 799-811.

Cox, G. W., & McCubbins, M. D. (2005). *Setting the agenda: Responsible party government in the US House of Representatives*: Cambridge University Press.

Fenno, R. F. (1978). *Home Style: House Members in Their Districts*: Little, Brown.

Grimmer, J., Messing, S., & Westwood, S. J. (2012). How words and money cultivate a personal vote: The effect of legislator credit claiming on constituent credit allocation. *American Political Science Review, 106*(4), 703-719.

Grossmann, M., & Hopkins, D. A. (2016). *Asymmetric politics: Ideological Republicans and group interest Democrats*: Oxford University Press.

Huddy, L., Mason, L., & Aarøe, L. (2015). Expressive partisanship: Campaign involvement, political emotion, and partisan identity. *American Political Science Review, 109*(1), 1-17.

Kendall, M. G. (1938). A New Measure of Rank Correlation. *Biometrika, 30*(1/2), 81-93. doi:10.2307/2332226

Koger, G., & Lebo, M. J. (2017). *Strategic Party Government: Why Winning Trumps Ideology*: University of Chicago Press.

Krehbiel, K. (1996). Institutional and partisan sources of gridlock: A theory of divided and unified government. *Journal of Theoretical Politics, 8*(1), 7-40.

Lau, R., & Redlawsk, D. P. (2001). Advantages and Disadvantages of Cognitive Heuristics in Political Decision Making. *American Journal of Political Science, 45*, 951.

Lupu, N. (2016). *Party brands in crisis: Partisanship, brand dilution, and the breakdown of political parties in Latin America*: Cambridge University Press.

Mann, T. E., & Ornstein, N. J. (2016). *It's even worse than it looks: How the American constitutional system collided with the new politics of extremism*: Basic Books.

Pope, J. C., & Woon, J. (2009). Measuring changes in American party reputations, 1939—2004. *Political Research Quarterly, 62*(4), 653-661.

Smith, G., & French, A. (2009). The political brand: A consumer perspective. *Marketing theory, 9*(2), 209-226.

Snyder Jr, J. M., & Groseclose, T. (2000). Estimating party influence in congressional roll-call voting. *American Journal of Political Science*, 193-211.

Snyder Jr, J. M., & Ting, M. M. (2002). An informational rationale for political parties. *American Journal of Political Science*, 90-110.

Thomas, E. F., & McGarty, C. A. (2009). The role of efficacy and moral outrage norms in creating the potential for international development activism through group-based interaction. *British Journal of Social Psychology, 48*(1), 115-134.

Bibliography

Abramowitz, A. I. and K. L. Saunders (2008). "Is polarization a myth?" The Journal of Politics **70**(2): 542-555.

Auter, Z. J. and J. A. Fine (2016). "Negative campaigning in the social media age: Attack advertising on Facebook." Political Behavior **38**(4): 999-1020.

Ballantine, P. W. and B. A. Martin (2005). "Forming parasocial relationships in online communities." ACR North American Advances.

Benoit, W. L. (2004). "Election outcome and topic of political campaign attacks." Southern Communication Journal **69**(4): 348-355.

Benoit, W. L. (2004). "Political party affiliation and presidential campaign discourse." Communication Quarterly **52**(2): 81-97.

Bond, R. M., et al. (2012). "A 61-million-person experiment in social influence and political mobilization." Nature **489**(7415): 295-298.

Borda, J. d. (1781). "'M'emoire sur les les e'lections au scrutiny', M'emoires de l'Acad'emie Royale des Sciences." English translation by a. DEGRAZIA, Isis (1953) **136**.

Bouchet-Valat, M. and M. M. Bouchet-Valat (2015). Package 'SnowballC'.

Brown, A. R. (2017). "The challenge of local party brand differentiation." The Social Science Journal **54**(1): 51-55.

Bruns, D., et al. (2017). "The origins of brand love: A typology of starting points." Marketing: ZFP–Journal of Research and Management **39**(2): 38-48.

Butler, D. M. and E. N. Powell (2014). "Understanding the party brand: experimental evidence on the role of valence." The Journal of Politics **76**(2): 492-505.

Canes-Wrone, B., et al. (2002). "Out of step, out of office: Electoral accountability and House members' voting." American Political Science Review: 127-140.

Carsey, T. M. (2009). Campaign dynamics: The race for governor, University of Michigan Press.

Carsey, T. M., et al. (2011). "Strategic candidates, campaign dynamics, and campaign advertising in gubernatorial races." State Politics & Policy Quarterly **11**(3): 269-298.

Clarke, A. J. (2020). "Party Sub-Brands and American Party Factions." American Journal of Political Science **64**(3): 452-470.

Clarke, A. J. (2020). "Party Sub-Brands and American Party Factions." American Journal of Political Science **64**(3): 452-470.

Cox, G. W., et al. (2010). "Party power or preferences? Quasi-experimental evidence from American state legislatures." The Journal of Politics **72**(3): 799-811.

Cox, G. W. and M. D. McCubbins (2005). Setting the agenda: Responsible party government in the US House of Representatives, Cambridge University Press.

Dalager, J. K. (1996). "Voters, issues, and elections: Are the candidates' messages getting through?" The Journal of Politics **58**(2): 486-515.

Evans, H. K., et al. (2014). "Twitter style: An analysis of how house candidates used Twitter in their 2012 campaigns." PS, political science & politics **47**(2): 454.

Fenno, R. F. (1978). Home Style: House Members in Their Districts, Little, Brown.

Flowers, J. F., et al. (2003). "The media, the campaign, and the message." American Journal of Political Science **47**(2): 259-273.

Gainous, J. and K. M. Wagner (2013). Tweeting to power: The social media revolution in American politics, Oxford University Press.

Geer, J. G. and L. Vavreck (2014). "Negativity, information, and candidate position-taking." Political Communication **31**(2): 218-236.

Gentry, J., et al. (2016). "Package 'twitteR'." Cran. r-project.

Glassman, M. (2010). Social networking and constituent communication: Member use of Twitter during a two-week period in the 111th Congress, DIANE Publishing.

Glassman, M. E., et al. (2013). Social networking and constituent communications: Members' use of Twitter and Facebook during a two-month period in the 112th Congress. Washington DC: Congressional Research Service, CRS Report for Congress.

Golbeck, J., et al. (2010). "Twitter use by the US Congress." Journal of the American society for information science and technology **61**(8): 1612-1621.

Graham, T., et al. (2013). "Between broadcasting political messages and interacting with voters: The use of Twitter during the 2010 UK general election campaign." Information, communication & society **16**(5): 692-716.

Grimmer, J., et al. (2012). "How words and money cultivate a personal vote: The effect of legislator credit claiming on constituent credit allocation." American Political Science Review: 703-719.

Grossmann, M. and D. A. Hopkins (2016). Asymmetric politics: Ideological Republicans and group interest Democrats, Oxford University Press.

Haselmayer, M., et al. (2017). "Partisan bias in message selection: Media gatekeeping of party press releases." Political Communication **34**(3): 367-384.

Hemphill, L., et al. (2013). "Framing in social media: How the US Congress uses Twitter hashtags to frame political issues." Available at SSRN 2317335.

Hemphill, L., et al. (2016). "# Polar Scores: Measuring partisanship using social media content." Journal of Information Technology & Politics **13**(4): 365-377.

Hemphill, L., et al. (2020). "What Drives US Congressional Members' Policy Attention on Twitter?" Policy & internet.

Hemphill, L. and A. M. Schöpke-Gonzalez (2020). Two Computational Models for Analyzing Political Attention in Social Media. Proceedings of the International AAAI Conference on Web and Social Media.

Hetherington, M. and J. Weiler (2018). Prius or pickup?: How the answers to four simple questions explain America's great divide, Houghton Mifflin.

Hill, S. J., et al. (2013). "How quickly we forget: The duration of persuasion effects from mass communication." Political Communication **30**(4): 521-547.

Holbrook, T. M. and S. D. McClurg (2005). "The mobilization of core supporters: Campaigns, turnout, and electoral composition in United States presidential elections." American Journal of Political Science **49**(4): 689-703.

Huddy, L., et al. (2015). "Expressive partisanship: Campaign involvement, political emotion, and partisan identity." American Political Science Review: 1-17.

Imai, K. and K. Khanna (2016). "Improving ecological inference by predicting individual ethnicity from voter registration records." Political Analysis: 263-272.

Jerit, J. (2008). "Issue framing and engagement: Rhetorical strategy in public policy debates." Political Behavior **30**(1): 1-24.

Johnston, R., et al. (1992). Letting the people decide: Dynamics of a Canadian election, Stanford University Press.

Kalla, J. L. and D. E. Broockman (2018). "The minimal persuasive effects of campaign contact in general elections: Evidence from 49 field experiments." American Political Science Review **112**(1): 148-166.

Kendall, M. G. (1938). "A New Measure of Rank Correlation." Biometrika **30**(1/2): 81-93.

Koch, J. W. (2008). "Campaign advertisements' impact on voter certainty and knowledge of house candidates' ideological positions." Political Research Quarterly **61**(4): 609-621.

Koger, G. and M. J. Lebo (2017). Strategic Party Government: Why Winning Trumps Ideology, University of Chicago Press.

Krehbiel, K. (1996). "Institutional and partisan sources of gridlock: A theory of divided and unified government." Journal of Theoretical Politics **8**(1): 7-40.

Krupnikov, Y. (2011). "When does negativity demobilize? Tracing the conditional effect of negative campaigning on voter turnout." American Journal of Political Science **55**(4): 797-813.

LaMarre, H. L. and Y. Suzuki-Lambrecht (2013). "Tweeting democracy? Examining Twitter as an online public relations strategy for congressional campaigns'." Public relations review **39**(4): 360-368.

Lantz, B. (2013). Machine learning with R, Packt publishing ltd.

Lau, R. and D. P. Redlawsk (2001). "Advantages and Disadvantages of Cognitive Heuristics in Political Decision Making." American Journal of Political Science **45**: 951.

Lodge, M., et al. (1995). "The responsive voter: Campaign information and the dynamics of candidate evaluation." American Political Science Review **89**(2): 309-326.

Lupu, N. (2016). Party brands in crisis: Partisanship, brand dilution, and the breakdown of political parties in Latin America, Cambridge University Press.

Mann, T. E. and N. J. Ornstein (2016). It's even worse than it looks: How the American constitutional system collided with the new politics of extremism, Basic Books.

Mason, L. (2018). Uncivil agreement: How politics became our identity, University of Chicago Press.

McClurg, S. D. (2006). "The electoral relevance of political talk: Examining disagreement and expertise effects in social networks on political participation." American Journal of Political Science **50**(3): 737-754.

McMillen, D. (2016). Moving Out, Moving In: The National Archives' Important Role When the Presidency Changes Hands. Prologue Magazine. **48**.

Meyer, T. M., et al. (2020). "Who gets into the papers? Party campaign messages and the media." British Journal of Political Science **50**(1): 281-302.

Nickerson, D. W. (2008). "Is voting contagious? Evidence from two field experiments." American Political Science Review: 49-57.

Petrocik, J. R. (1996). "Issue ownership in presidential elections, with a 1980 case study." American Journal of Political Science: 825-850.

Pew Research Center (2019). January 2019 Core Trends Survey. Washington, D.C.

Pietryka, M. T. and D. A. DeBats (2017). "It's not just what you have, but who you know: Networks, social proximity to elites, and voting in state and local elections." American Political Science Review: 360-378.

Pope, J. C. and J. Woon (2009). "Measuring changes in American party reputations, 1939—2004." Political Research Quarterly **62**(4): 653-661.

Ridout, T. N. and G. R. Smith (2008). "Free advertising: How the media amplify campaign messages." Political Research Quarterly **61**(4): 598-608.

Rosenstone, S. J. and J. M. Hansen (1993). Mobilization, participation, and democracy in America, Longman Publishing Group.

Rosenstone, S. J. and J. M. Hansen (2001). "Solving the puzzle of participation in electoral politics." Controversies in voting behavior: 69-82.

Russell, A. (2018). "US senators on twitter: Asymmetric party rhetoric in 140 characters." American Politics Research **46**(4): 695-723.

Schulman, K. (2016). "The Digital Transition: How the Presidential Transition Works in the Social Media Age." from https://obamawhitehouse.archives.gov/blog/2016/10/31/digital-transition-how-presidential-transition-works-social-media-age.

Shapiro, M. A. and L. Hemphill (2017). "Politicians and the policy agenda: Does use of Twitter by the US Congress direct New York Times content?" Policy & internet **9**(1): 109-132.

Shogan, C. J. (2010). "Blackberries, tweets, and YouTube: Technology and the future of communicating with Congress." PS: Political Science and Politics: 231-233.

Sides, J. and A. Karch (2008). "Messages that mobilize? Issue publics and the content of campaign advertising." The Journal of Politics **70**(2): 466-476.

Silge, J., Robinson, D. (2016). ""tidytext: Text Mining and Analysis Using Tidy Data Principles in R." _JOSS_ **1**(3).

Smith, G. and A. French (2009). "The political brand: A consumer perspective." Marketing theory **9**(2): 209-226.

Snyder Jr, J. M. and T. Groseclose (2000). "Estimating party influence in congressional roll-call voting." American Journal of Political Science: 193-211.

Snyder Jr, J. M. and M. M. Ting (2002). "An informational rationale for political parties." American Journal of Political Science: 90-110.

Spiliotes, C. J. and L. Vavreck (2002). "Campaign advertising: Partisan convergence or divergence?" The Journal of Politics **64**(1): 249-261.

Straus, J. R., et al. (2013). "Communicating in 140 characters or less: Congressional adoption of Twitter in the 111th Congress." PS, political science & politics **46**(1): 60.

Straus, J. R., et al. (2016). "Congressional social media communications: Evaluating Senate Twitter usage." Online Information Review.

Stromer-Galley, J., et al. (2018). "Tweeting the attack: Predicting gubernatorial candidate attack messaging and its spread." International Journal of Communication **12**: 22.

Therriault, A. (2015). "Whose issue is it anyway? A new look at the meaning and measurement of issue ownership." British Journal of Political Science **45**(4): 929-938.

Thomas, E. F. and C. A. McGarty (2009). "The role of efficacy and moral outrage norms in creating the potential for international development activism through group-based interaction." British Journal of Social Psychology **48**(1): 115-134.

Vavreck, L. (2009). "The message matters." The Economy and Presidential Campaigns, Princeton, Princeton, NJ.

Vavreck, L. (2014). "Want a better forecast? Measure the campaign not just the economy." PS, political science & politics **47**(2): 345.

Verzani, J. (2014). "GWidgets2 [Electronic resource]." CRAN-Package gWidgets2.

Yiannakas, G.

Yu-Wei, C. D. C. (2015). Machine learning with R cookbook, Packt Publishing Ltd.