Pandering Politicians: Ideological Changes from Primary to General Elections

William Godel, Evgenii Nikitin and Ye Wang

New York University

Abstract

Voters rely, at least in part, on politicians to inform them of their ideological positions. But over the course of a two stage campaign, U.S. politicians face two distinct constituencies. Does this system encourage politicians to manipulate their messaging, pandering to their constituencies over the course of a campaign? This question has important implications for understanding voter and politician behavior, as well as the effects of primaries on electoral outcomes. We address it by estimating politicians' expressed ideology in a dynamic setting, covering the 2018 primary and general elections, by classifying politician tweets through a semi-labeled Reddit dataset. Using this measure, we find that Republican politicians change significantly in their expressed ideology after a primary, although the effect of the primary is not consistent across all groups. Conversely, we find little evidence that Democratic politicians moderate their expressed sentiment after winning a primary.

notes for me : Do parties cater to the middle? Hard to test. But change in constituencies, direction of change is known. Do politicians follow? Social media is flexible and public, allows politicians to change quickly add social media contribution "ideal setting"

My things to do descriptive for winners and losers mean differences quartile graphs

1 Introduction

When candidates change constituencies, do they change themselves as well? How politicians can or cannot shape voter perceptions is fundamental to explaining how democratic processes function. The popular theory that politicians will often moderate their message when they transition from the primary to general election rests on assumptions that potentially go to the heart of any theory of elections, and verifying those assumptions would go a long way towards providing the empirical contours theories of elections must explain. Fundamentally, addressing this question can help to provide insight into the relationship between voters and the candidates they elect. While there is an extensive research literature into the question of how candidates interact with voters, including in two stage elections, rarely has data been brought to bear in order to adjudicate between plausible alternatives in a dynamic setting.

Answering this question requires measuring the ideological positions politicians are sending. While researchers have been successful in measuring a politician's ideology or expressed ideology in a variety of ways, including legislative voting (Marshal and Peress 2018), press releases (Grimmer 2010), advertised votes (Cormac 2016), surveys (Ansolabehere, Snyder Jr, and Stewart, 2001, Montagnes and Rogowski, 2014), twitter followers (Barbera, 2015), and donor networks (Bonica, 2013, 2014), the comparability of alternative measures of ideology remains a challenge (Tausanovitch and Warshaw 2018) and genuinely dynamic measures of politician expressed ideology remain elusive. In this paper, we present a novel estimation technique to measure politician's expressed ideology dynamically which we then use to estimate the effects of winning a primary on politicians' expressed ideology.

Substantively, we expect politicians to either change, moderate or otherwise reemphasize their expressed ideology when the election switches from a primary election, whose median voter is typically considered more partisan, to the general election, where the median voter is more moderate. We would expect this effect to potentially be more pronounced in competitive elections (Mayhew 1974). At the same time, despite the plausibility of this hypothesis, it could easily be the case that it does not hold (Fiorina 1974, Gulati 2004). Primary election voters may lose enthusiasm for a candidate who may contradict themselves or otherwise appear more moderate than they were during the primary, which could dampen turnout. Additionally, general election voters may not find a candidate's ideological moderation credible if it is significantly different than what they expressed during the primary. Given these plausible alternatives, it remains an open question if politicians do in fact pander to their different constituencies in two stage elections.

We attempt to answer this question by utilizing a machine learning classification system to label politician tweets on a left-right dimension. Firstly, we address the challenge of ideological classification with a novel source of pre-labeled data – Reddit. Reddit is an online discussion community that has strict subgroups devoted to the discussion of specific topics. Our classification algorithm is a two stage process whereby tweets are first classified by their similarity to various Reddit topics (the first level model), and then a second algorithm (the second level model) assigns them a value on the left-right spectrum. This two stage classification system allows us to utilize an extremely large and diverse pre-labeled dataset, while also placing candidates on the same left-right spectrum as their DW-NOMINATE score, which we use to train the second level of the model.

Using these labeled tweets, i.e., the tweets of incumbent politicians for whom DW-Nominate scores exist, we can then construct a dynamic estimate of a candidate's expressed ideology over the course of a campaign, including both primary and general elections. In order to identify ideological changes, we use both the staggering of primary dates as well as the tweets of U.S. Senators who are co-partisans of the candidates but who are not up for re-election themselves. By leveraging the tweets of both candidates and their co-partisans who are not of up for reelection we are able to construct novel estimates of changes in candidate expressed ideology dynamically, over the course of the entire campaign, and more specifically measure changes before and after the primary. Our findings are broadly in line with our predictions though somewhat surprising. We find that Republican politicians moderate their expressed sentiment significantly after winning a primary. This effect is heterogeneous across election types and is, unsurprisingly, more pronounced in competitive races. Conversely, we find little evidence that Democrats moderated their message, regardless of the election type or competitiveness.

2 Literature Review

Our question is fairly straightforward, do politicians moderate after the win a primary and enter a general election? This is not a novel hypothesis, it is present both in the research literature (Cormac 2016) and in popular accounts of elections. For example, during the GOP primary for the 2012 election, candidate Mitt Romney expressed support for strongly pro-life legislation, “I support the Hyde Amendment. ...I will reinstate the Mexico City Policy. ...I will advocate for and support a Pain-Capable Unborn Child Protection Act.” This fairly conservative policy position did not survive the candidates general election campaign. Later, after securing the partys nomination, Romneys stance on abortion evolved to become “Theres no legislation with regards to abortion that I'm familiar with that would become part of my agenda” (Estes, 2018).

This question speaks to multiple research literatures. Firstly, it directly addresses questions with respect to candidate messaging, particularly during elections. Cormac (2016) showed that politicians advertise more moderate voting records during elections, while other researchers emphasize different types of messages depending on the competitiveness of the election, constituency and the politician (Grimmer). Additionally, Ash, Morelli, and Van Weelden (2017) show that legislators are more likely to address divisive issues in their congressional speeches before elections. Additionally, Grose, Malhotra, and Parks Van Houweling (2015) finds that US legislators will reply to mail from ordinary voters and explain their congressional votes. Broockman (2013) and Broockman (2014) and Kalla and Broockman (2016) emphasize the role of ethnicity and campaign donations in deciding the responsiveness of legislators, respectively. Barberá et al. (2014) uses the Granger causality model to argue that posts of politicians are influenced by trends on Twitter. Our paper contributes to this literature by measure how, in one dimension, politicians communicate during campaigns on Twitter, as well as how this messaging changes dynamically over the course of the campaign.

Our research also contributes to the larger literature on candidate responsiveness and how candidates respond to elections. With respect to behavior, e.g, legislative voting, executive actions or sentencing (by elected judges), there is significant evidence that the proximity of elections alone can change politician behavior. For example, researchers have found that politicians respond to elections via legislative behavior (Canes-Wrone et al. 2002, Marshall and Peress 2018), sentencing Gordon and Huber (2007) in the case of judges, and legislative responsiveness (Griffin 2006). Our novel technique allows us to measure politician measuring dynamically over the course of a two stage election, allowing us to potentially draw inferences about how much candidates respond in their messaging when their elective constituency changes.

This paper also speaks to the literature on the effects of two stage elections, particularly research concerning what types of candidates are selected and how political outcomes may be shaped by having two stage elections. There is significant empirical evidence that primaries may shape both the types of politicians selected (Brady, Han, and Pope (2007), (Hall, 2015), (Adams and Merrill, 2008), (Hall and Thompson, 2018), (David, Hahrie, and Jeremy, 2011)) and affect candidate messaging ((James and Samuel, 2008)). There have also been significant efforts to model the two stage election process(landa....). Typically, ideological messaging is taken as fixed value over the course of a campaign. One exception is Banda and Carsey (2015), in which the authors propose a two-stage election model to explain the behavior of candidates in primaries. They find that candidates usually converge to the platform of opponents in the primary, but diverge from or ignore the platform of opponents in the general election. This paper speaks directly to this literature by testing if candidates do in fact switch their ideology between primary and general election campaigns.

Party asymmetry ?

Methodologically, our method for estimating expressed sentiment builds on a technique originally introduced by Nikitin (2018). This method utilizes a semi-labeled Reddit dataset to extract more information from a given source of text than would otherwise appear to be available from a direct machine learning mapping from text to labels. Consequently we offer a novel method for measuring ideology dynamically, in addition to current approaches that utilize bridge votes (Marshall and Peress 2018) or survey responses (Ansolabehere, Snyder Jr, and Stewart, 2001, Montagnes and Rogowski, 2014), measures which are too coarse to measure changes over the course of a single campaign.

Additionally, this method addresses an issue outlined by Tausanovitch and Warshaw (2016), who show that alternative measures of ideology have week intra-party correlation with DW-Nominate scores. Because our method maps tweets to DW-Nominate scores our measure achieves not only superior performance with respect to intra-party correlation, but also explicitly attempts to map to a well known and understood measure of ideology (DW-Nominate) rather than alternative measures, such those introduced by (Barberá et al., 2014), Bonica (2014), Shor and Rogowski (2015), Ramey (2016), and Joesten and Stone (2014).

3 Data & Methods

In this section, we describe the data we collected and our method for converting politician tweets into a measure of expressed sentiment on a unidimensional ideological spectrum. We also discuss the motivation for and implementation of the method.

3.1 Twitter Data

The data for analysis comes from Twitter from the 2018 midterm elections. We attempted to collected data for all primary winners for the Senate, House of Representatives and Governorships. 2018 presented a novel opportunity for data collection as Twitter began, for the first time, to add a formal label to political candidates twitter accounts. For example, candidate for Senator in Ohio Sherrod Brown had ”US Senate candidate, OH” affixed to his twitter account. This label was created by Twitter, in conjunction with Ballotpedia, and was opt out for candidates, meaning candidates automatically had these label affixed to their twitter pages unless they asked for them to removed. Using a list of all candidates for our elections of interest, we were able to identify 747 twitter accounts formally associated by Twitter with candidates. We then manually identified the primary twitter accounts for the remaining candidates for whom our automated search was unsuccessful [1]. Finally, we collected the tweets of incumbent Senators by utilizing the list of twitter accounts for sitting public officials from Propublica, and downloaded tweets only for those Senators not up for Reelection.

This provided us with a data set that contained in total 967 politicians and a total of 722,001 tweets. Of these, 900 were engaged in an election that year with the remaining 67 being incumbent Senators not up for reelection. Of the candidates, we had 486 Democrats and 414 Republicans in the dataset. Among those not up for reelection there were 24 Democrats and 43 Republicans. By election type, our dataset contained 72 candidates for governor, 762 candidates for the House, and 66 candidates for the Senate (all including incumbents). We limited our analysis exclusively to Republicans and Democrats[2].

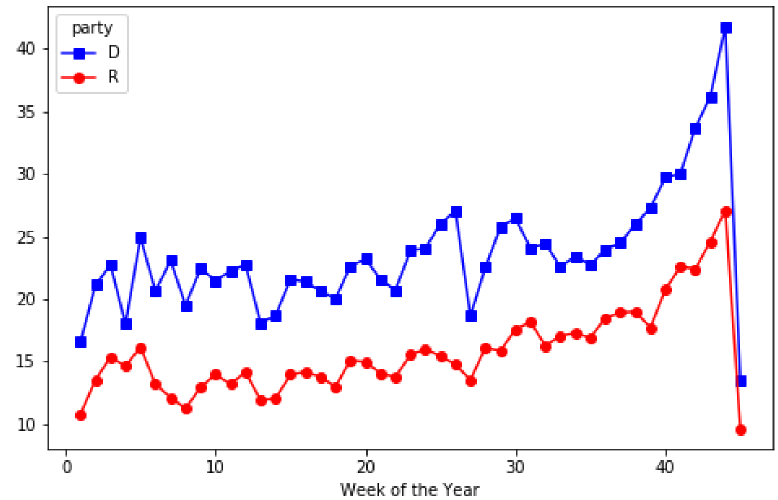
Our dynamic estimates of ideology (explained below) are at the weekly level and are based on all the tweets of a given politician in a given week. Below is a graph of the average number of tweets of individuals by party.

From the first figure, it is clear that Democrats are tweeting at a significantly higher rate than Republicans. Despite this, nonetheless Republicans appear to be tweeting on average

[1] For candidates with multiple accounts, if they had a formal twitter flag identifying them as a candidate we used the account that was flagged. If they did not, in the case where both accounts were active, we selected the account that had more followers. If one or more accounts was no longer being actively updated, we chose the more recently used account.

[2] For independents, we dropped all independents for elections other than the Senate. For the Senate, we labeled Senator Bernie Sanders a Democrat because he participated in a Democratic primary. Conversely, we dropped Senator Angus King because he did not participate in either a Democratic or Republican primary. Additionally we dropped the Mississippi special election as it contained a jungle primary, as well as all races from Louisiana for the same reason.

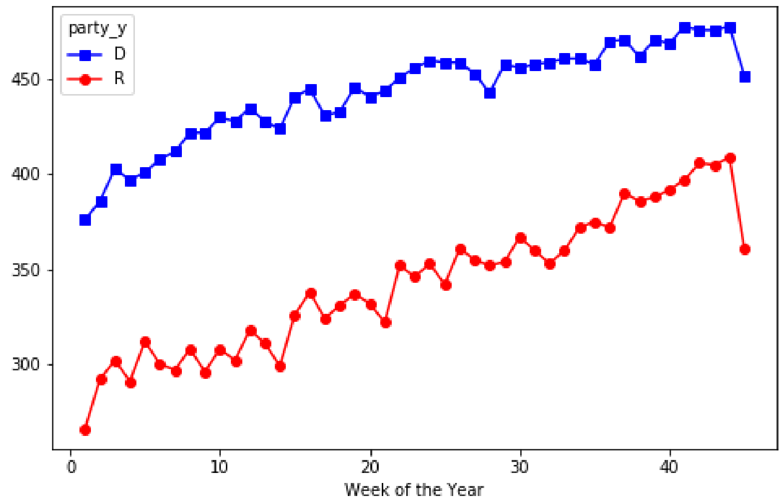
Figure 1: Average Number of Tweets per Candidate, per Week, by Party



about fifteen times per week for most of the campaign, with Democrats tweeting above 20 tweets per week on average.

Finally, the figure below shows the number of candidate observations by party, by week. This figure shows that on average about 350 Republicans are tweeting at least once per week (out of a total of 414) and roughly 450 democrats are tweeting once per week (out of a total of 486). Consequently, of those who are tweeting, they are tweeting fairly often both within and across weeks.

Figure 2: Average Number of Candidate Observations per Week by Party



3.2 Reddit Embeddings

The pipeline is as follows. First, we collected comments posted by users of the website Reddit.com. Reddit is a social media discussion platform composed of thousands of topic-specific forums known as **subreddits**. Users post content to a specific subreddit, sharing it within that specific community, typically containing their commentary or opinion, sometimes with a link to an external website. Within these forums users can upvote and downvote posts, which are then tallied, moving more upvoted comments to the top of the page and downvoted posts to the bottom in proportion to the number of votes. From the way these votes are utilized by users, highly upvoted comments are those that enjoy significant support among the community and downvoted comments the converse. For this project, we collected comments from 142 subreddits posted from January 2015 to December 2017. Our selection of subreddits was driven by the particular task at hand, so a large proportion of the subreddits that were used were related to politics and political issues.

We collected 10,000 (or fewer if that many were unavailable) of the most upvoted comments with a non-negative rating for each subreddit-month pair. As a result, we have roughly

48 million Reddit comments. For the sake of computational efficiency, we randomly sampled 70,000 comments (or fewer if unavailable) from each subreddit, which leaves us with approximately 7 million comments. Following that, we tokenize each comment using the following preprocessing rules: we remove all URLs, subreddit names, usernames, quotes of the other comments, punctuation, and stop words, and lowercase all tokens. These tokenized sentences are used as input for a fastText model (Joulin et al., 2016), which is trained to predict which subreddit each comment was posted in.

Figure 3 shows the architecture of the fastText model. Each word in the vocabulary is initialized with a random vector representation (word embedding). These representations are stored in an embedding matrix (i.e., a table look up layer) of size $K \times n$, where K is the number of unique words in vocabulary and n is the dimensionality of the model (in this case 100). A vector representation of the document is then calculated as a component-wise mean of the token vectors ("make" + "america" + "great" + "again"). This sentence vector constitutes a hidden layer d of dimensionality n. Vector d can be interpreted as a vector representation of the current document. This vector is then multiplied by model weights matrix U (also initialized randomly) in order to obtain an output activation for each class (e.g., */r/The_Donald, /r/hillaryclinton, /r/StarWars*). Matrix U has dimensionality $m \times n$, where m is the number of possible classes (3 in the current example). Finally, the Softmax activation function is then applied to the output activations to make sure that final class activations sum up to 1. These activations can be interpreted as predicted probabilities for each class. In this example, this final vector would have the probability that a given comment belong to each of the possible classes. Parameters of this model that are learned during training phase are (a) an embedding matrix and (b) a weight matrix U.

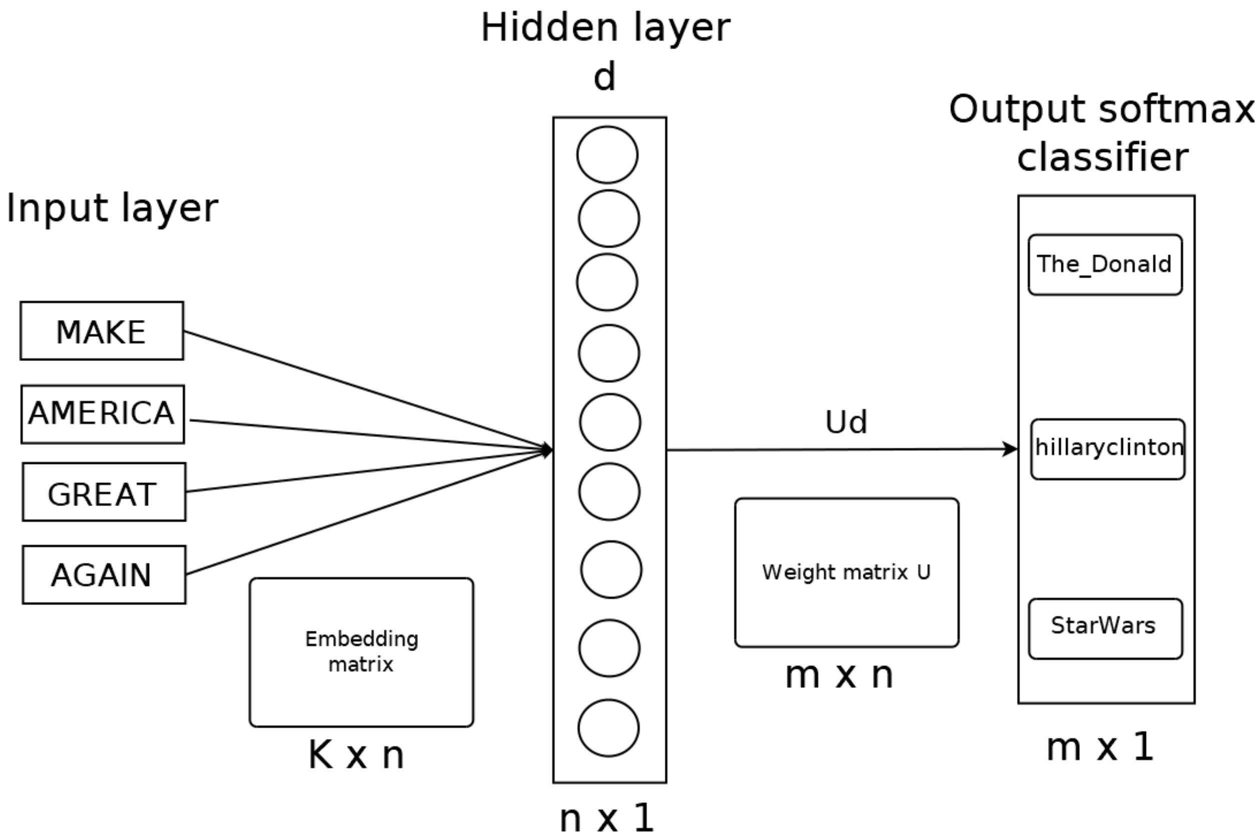


Figure 3: Architecture of fastText model

We trained three separate fast text models that very only by the output labels assigned to each comment. The first model used as output labels the actual subreddits themselves (a total of 142). For the second model we constructed our own labels by combining similar subreddits, e.g., feminist subreddits, conservative subreddits, Christian subreddits, etc. For the third model we train on only three labels that we again created from combining different subreddits. These three labels were *Liberal*, *Conservative* and *Nonpolitical*.

For each of these three models, we train a 100-dimensional model for 5 epochs with bigrams of tokens as features, using 10% of the dataset to evaluate performance of the model. When this is completed we have three separate fasttext models, each producing one dimensional vectors. We combine these three vectors by concatenation, which we then apply the standard scaler from scikit learn () to, for use in the second level model. This final vector from the first level model we will refer to as a Reddit embeddings.

3.3 Second-level model

In order to train the second level model, we then calculate Reddit embeddings for the collection of all tweets posted by Representatives of the 115th Congress from January 2016 to November 1st 2017. The tweets were preprocessed in the same way as Reddit comments with one additional procedure – hashtags and twitter handles were split into separate tokens (e.g., #womenmarch $\implies$ "women march", @realDonaldTrump $\implies$ "real donald trump"). Consequently, each tweet by a member of Congress produces a Reddit embedding. We then created a weighted average of these Reddit embeddings using how political a given tweet is as weights. We obtain this value by subtracting from one the probability that the tweet is nonpolitical (which is derived in the first level model). Once this procedure is complete, we have a weighted average of all a politician's Reddit embeddings during the training period, based on their tweets.

Following this, we use these embeddings to train second-level models to predict DW NOMINATE scores for the 115th Congress (downloaded from `http://k7moa.com/Weekly_Constant_Space_DW-NOMINATE_Scores.htm`). These second level models take as input these averaged Reddit embeddings and map them to that legislators DW NOMINATE score. DW NOMINATE is a statistical method that uses congressional roll call votes to generate unidimensional left-right ideology estimates for each politician who has ever served a term in Congress (Poole and Rosenthal, 1984). DW NOMINATE has been shown to be a reliable and consistent measure of politicians' ideal points. When fitting this model, we are assuming that on average over a long period of time, Twitter accounts of politicians should represent or at least map to their true ideological positions.

We tested several different second level models:

- Linear regression with one predictor ("expert model") – we used our conservativeness score as a simple benchmark. In this case it was used in a linear regression with an intercept that allowed it rescale it to the DW-NOMINATE scale.

- Elastic net – A least squares linear regression with ℓ_1 and ℓ_2 regularization
- word2vec
- two different bag of words

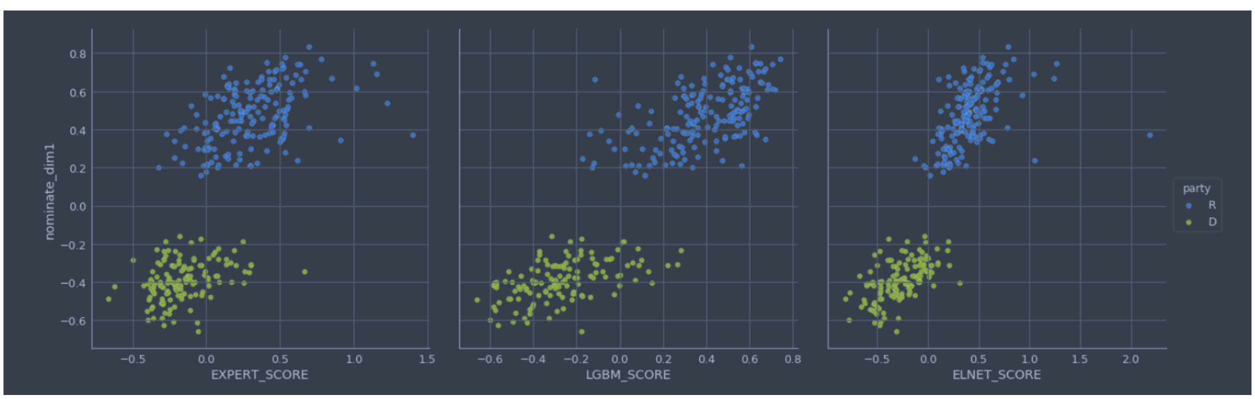
Models were tuned using 5-fold cross-validation, and their performance was assessed on 20% holdout set. From the table below it is clear that the more sophisticated models (Elastic Net and LightGMB) had far superior performance to the unsupervised Expert model.

Model	**Correlation**	**MSE**
Expert	0.74	0.109
LightGBM	0.88	0.034
Elastic Net	0.88	0.041

Table 1: Model Performance on Holdout Set

Figure 4 is a graph of the predictions versus the actual DW-NOMINATE scores.

Figure 4: Predictions of DM NOMINATE vs. Actual Score



3.4 Benchmarks

A two-level model is obviously more complicated than a one level model. This increased complexity is justified by the superior predictive capacity of these models relative to a direct mapping of tweets to DW NOMINATE scores. The motivation behind this decision is quite simple. First, by using Reddit we utilize much more text data containing much relevant ideological information than would otherwise be available, which boosts predictive performance and makes our models more generalizable. Second, our models also become much more interpretable since we can examine which subreddits have the greatest effects on predictions. This allowed us to confirm if the subreddits we would expect, in particular political subreddits, were contributing the most as features in the second level of the model, which in fact they were.

In order to demonstrate the benefit of using this setup empirically, we used three additional sets of features as benchmarks:

- Bag-of-Words (raw counts of words in tweets for each Representative)
- word2vec embeddings (Mikolov et al., 2013) – word2vec model was trained on all tweets, and then embedding for each politician is just an element-wise average of all word embeddings
- pretrained fastText embeddings (Bojanowski et al., 2016) – embedding for each politician was calculated as an element-wise average of all pre-trained word embeddings

We trained support vector regression (SVR) on these three sets of features. Table 2 shows results on the test set for the best model found by cross-validation. As we can see, supervised Reddit models significantly outperform all three benchmark models.

3.5 Analysis

After the first level models and our proffered second level models were trained and tuned we then applied the model to our target data. These tweets were preprocessed identically to how representative tweets were processed. For each individual we created a weekly estimated of their expressed ideology based on a weighted average of their tweets for that week.

The figures below show our estimates of Democratic and Republican average predicted DW-NOMINATE score over the course of a campaign, with the second figure showing the

Model	Correlation	MSE
Reddit-Expert	0.74	0.109
Reddit-LightGBM	0.88	0.034
Reddit-Elastic Net	0.88	0.041
Bag-of-Words	0.81	0.074
Word2Vec	0.82	0.049
Pretrained fastText	0.79	0.059

Table 2: Model Performance on Holdout Set

changes in expressed sentiment relative to before and after their primary. These graphs are taken from our two-level model and are prior to performing our estimation strategy.

Figure 5: Average Expressed Ideology Overtime by Party

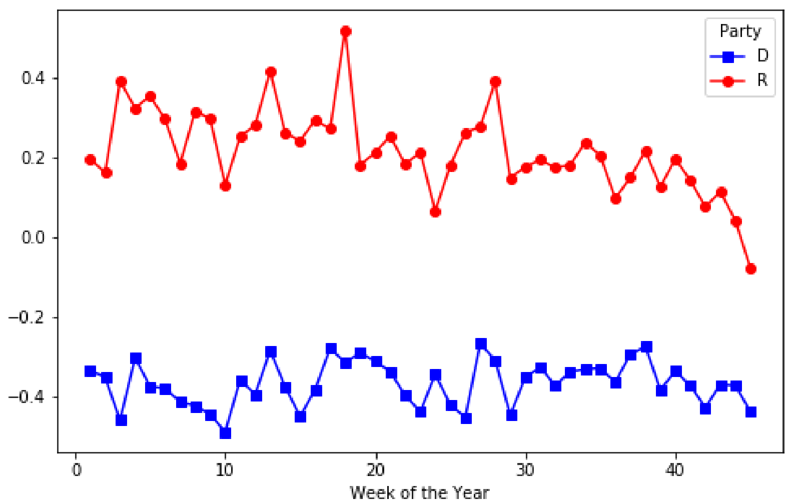
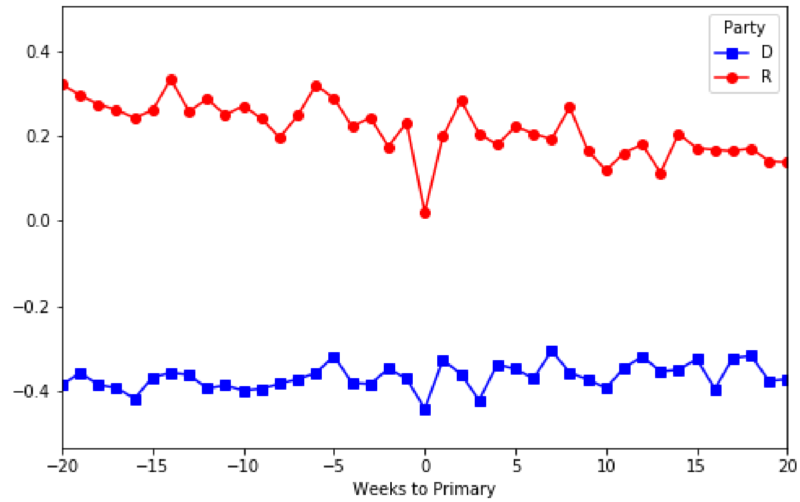


Figure 6: Average Expressed Ideology Weeks to Primary by Party



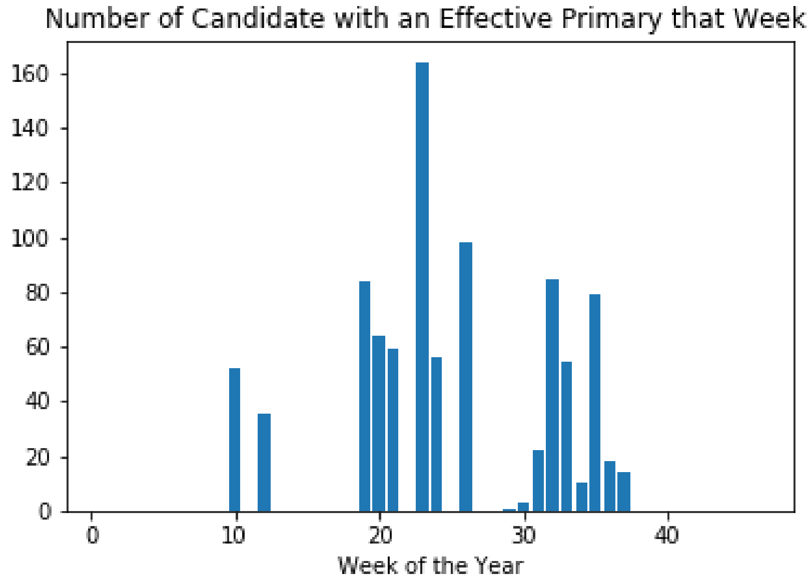
4 Estimation Strategy

The parameter of interest in this project is the impact of the primary's end on politicians' expressed sentiment. Our theoretical framework predicts that only politicians faced with electoral pressure will be affected after the time period. We consider these politicians as the treated group and others as the control group. Using terminologies in causal inference, our estimand is the treatment effect of the treated (ATT).

Our dataset has a clear panel structure with both unit and period dimensions. But as mentioned before, our measurement of the outcome is not perfect and may contains noise. Our basic estimation strategy is a fixed effects models where our treatment is a dichotomous variable for before and after treatment. We interact his treatment with party in to order estimate the effects of winning a primary by party.

An important component of our analysis is that primaries are spread out over the course of the year, rather than all occurring at once. Consequently, overtime changes that are due to a change electorate should be separable from a general time trend. The figure below shows the number of candidates experiencing a primary in a given week.

Figure 7: Number of Candidates Experiencing a Primary in a Given Week



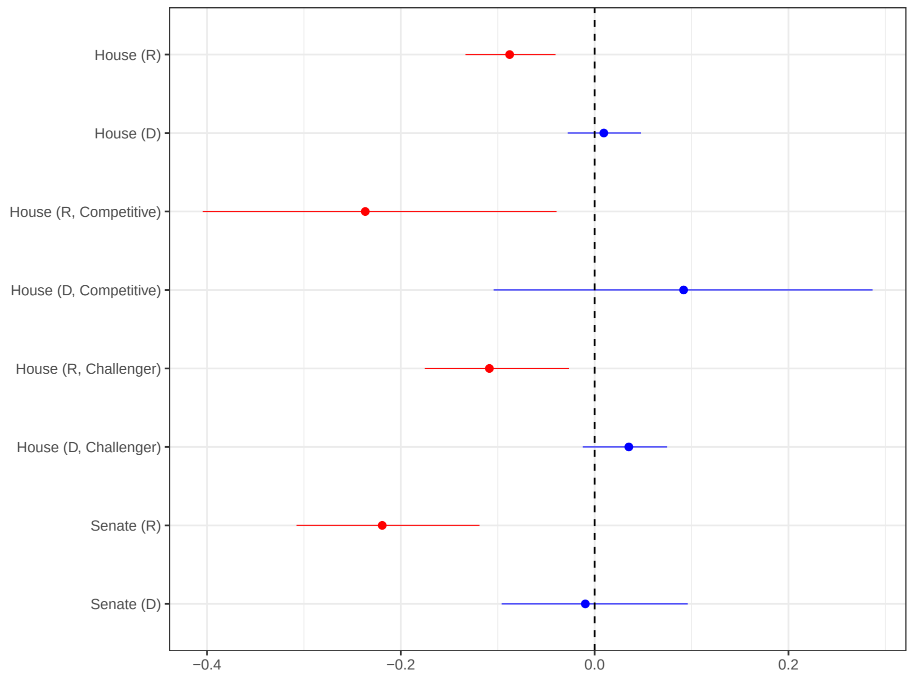
5 Results & Discussion

We performed the above estimation strategy, grouping observations by party and incumbency. We also performed the same analysis using all three different second level models so that any effects could have their robustness confirmed by alternative models. In the graphs below, greater values are more conservative and lesser values are more liberal[3]. Our findings are somewhat mixed. Most noticeably, it is clear from figure **??** that Democratic incumbents become more conservative after the primary and this finding is significant at the 5% level for both of our more sophisticated second level models.

From the coefficient chart, we find that Republican politicians across all groups moderate significantly after winning the primary. This effect appears to be significantly more pronounced for competitive races as well (those races where the margin of victory was forecast to be less than 20 percentage points). It also appears to be similarly pronounced for Republicans running for Senate seats. Conversely, Democratic candidates appear to show

[3]In the appendix we show how the treatment effects change with time periods.

Figure 8: Effect of the End of Primary on Expressed Ideology



virtually no change in expressed sentiment, and this null effect holds across groups and the competitiveness of the race.

Overall our analysis suggests that there is flexibility in the expressed sentiment of politicians over the course of an election campaign, particularly with regards to the period before and after a primary. At the same time, there is also an asymmetry across parties with respect to this change in expressed sentiment. Given the scale of the effects estimated it is unlikely that we would estimate consistent estimates of Republican moderation and find virtually no effects for Democrats at all simply due to measurement noise. This strongly suggests there is an asymmetrical effect associated with the end of a primary between the candidates from the two parties.

6 Conclusion

This paper attempts to find evidence that politicians pander to their constituencies; if, when the constituency changes, so does the politician, or at least what they say. In order to do this, we trained a two level model to predict a candidate's unidimensional left - right ideology using only their tweets. We then applied this model to the tweets of all candidates running for a governorship, House or Senate seat in 2018, and used their co-partisans in the Senate (those not up for reelection) as controls.

Our findings were mixed. We found general evidence that politicians did in fact moderate after the primary, but also some intriguing evidence that this effect was asymmetric between the parties, with Republicans moderating after the end of the primary, while Democratic politicians remained ideologically constant. Given the nature of Twitter, in particular the discretion it gives politicians with respect to what they tweet and address, it is likely that in other venues where politicians have greater discretion with respect to who will receive their message or how formal that message is (i.e., in print or not) this change in expressed sentiment may be even more pronounced.

Despite this, our findings could be challenged by the limitations of our analysis. In particular, twitter may not be representative of the larger expressed ideology or sentiment of a candidate. If candidates use twitter in a unique way, for example as a medium for the dissemination of innocuous information that is not representative of their larger campaign, then any findings derived from twitter may not generalize to the candidates themselves. In that case, our analysis can only speak to the ideological flexibility politicians exhibit on twitter rather than in the general campaign.

This qualification may in fact strengthen our findings though. Twitter is an optional medium. Politicians are not obligated to use it, and certainly not obligated to use it in an engaging way. Tweets are also extremely public and more or less permanent. This is in stark contrast to other forms of political expression where politicians may be obligated to address issues (e.g., town halls or debates) and can address private audiences (e.g., private speeches or fund raisers). Given twitter is a much easier medium for politicians to avoid topics which they may know could alienate voters, and any sentiment they express will become a more or less permanent and searchable, it is perhaps all the more surprising that politicians nonetheless still exhibit a degree of flexibility that is sufficiently strong and systematic to identify.

Overall the evidence suggests that there are important questions worth pursing with respect changes in expressed sentiment by politicians. Perhaps most fundamentally, formal exploration of two stage election processes should incorporate possible multiple and heterogeneous signals by politicians in order to explain why politicians send multiple signals and, importantly, why voters may or may not believe them. If possible, future work should also explore the heterogeneity that may occur between incumbents and non-incumbents and differences between parties with respect to repeated signaling during a two stage campaign.

Future research should be able to bring more data to bear on these questions as well, helping to providing a rich set of stipulated facts with which formal theorists can work from. Finally, this analysis also suggests the viability of machine learning models for political classification. By working only with an outside source of Reddit data we were able to utilize extremely limited amounts of text to classify politicians' ideologically dynamically. These findings were also generally robust to our second level model and estimation strategy. This is a strong endorsement for ability of social media data and machine learning techniques to tackle heretofore difficult challenges.

7 Appendix

7.1 Subreddit List

This is the list of subreddits that were used in our model:

- 4chan
- 911truth
- AgainstHateSubreddits
- AmericanPolitics
- Anarchism
- Animals
- AntiPOZi
- AskFeminists
- AskThe_Donald
- AskTrumpSupporters
- Ask_Politics
- BlackPeopleTwitter
- Blackfellas
- Catholicism
- centrist

- Christians
- Classical_Liberals
- Communist
- Conservative
- ConservativesOnly
- CredibleDefense
- CringeAnarchy
- DarkEnlightenment
- DebateAltRight
- EnoughTrumpSpam
- Feminism
- Firearms
- Fitness
- Fuckthealtright
- GenderCritical
- GunsAreCool
- HillaryForAmerica
- HillaryForPrison
- ImGoingToHellForThis
- Impeach_Trump

- JordanPeterson
- Judaism
- KotakuInAction
- LGBTnews
- Liberal
- Libertarian
- MGTOW
- MachineLearning
- MensRights
- metacanada
- Music
- NeutralPolitics
- POLITIC
- PoliticalDiscussion
- PoliticalHumor
- PussyPass
- Republican
- SandersForPresident
- ShitRConservativeSays
- SocialDemocracy

- SocialJusticeInAction
- socialism
- TheRedPill
- The_Donald
- TrueChristian
- TrumpCriticizesTrump
- TumblrInAction
- TwoXChromosomes
- WhiteRights
- WikiLeaks
- ainbow
- alltheleft
- altright
- antiwar
- askaconservative
- atheism
- baseball
- blackladies
- books
- climateskeptics

- conservatives
- conspiracy
- dankmemes
- democracy
- democrats
- demsocialist
- enoughpetersonspam
- environment
- Equality
- esist
- eupoltics
- european
- feminisms
- food
- gadgets
- gaming
- guncontrol
- hillaryclinton
- islam
- labor

- lifehacks
- lgbt
- moderatepolitics
- movies
- neoliberal
- new_right
- news
- obama
- onguardforthee
- politics
- prochoice
- programming
- progressive
- progun
- prolife
- relationships
- republicanism
- republicans
- science
- sex

- sjwhate
- soccer
- socialjustice101
- sports
- technology
- television
- travel
- ukpolitics
- uncensorednews
- uspolitics
- whiteknighting
- women
- worldnews
- worldpolitics
- beer
- Advice
- funny
- BasicIncome
- Bad_Cop_No_Donut
- randpaul

- humanrights
- restorethefourth
- privacy
- Green
- Good_Cop_Free_Donut
- AskHistorians
- Cooking
- streetwear

References

Adams, James and Samuel Merrill (2008). "Candidate and party strategies in two-stage elections beginning with a primary". In: *American Journal of Political Science* 52.2, pp. 344–359.

Ash, Elliott, Massimo Morelli, and Richard Van Weelden (2017). "Elections and divisiveness: Theory and evidence". In: *The Journal of Politics* 79.4, pp. 1268–1285.

Banda, Kevin K and Thomas M Carsey (2015). "Two-stage elections, strategic candidates, and campaign agendas". In: *Electoral Studies* 40, pp. 221–230.

Barberá, Pablo et al. (2014). "Leaders or followers? Measuring political responsiveness in the US Congress using social media data". In: *Annual Meeting of the American Political Science Association.*

Bojanowski, Piotr et al. (2016). "Enriching Word Vectors with Subword Information". In: arXiv: `1607.04606`. URL: `http://arxiv.org/abs/1607.04606`.

Brady, David W, Hahrie Han, and Jeremy C Pope (2007). "Primary elections and candidate ideology: Out of step with the primary electorate?" In: *Legislative Studies Quarterly* 32.1, pp. 79–105.

Broockman, David E (2013). "Black politicians are more intrinsically motivated to advance blacks interests: A field experiment manipulating political incentives". In: *American Journal of Political Science* 57.3, pp. 521–536.

— (2014). "Distorted communication, unequal representation: constituents communicate less to representatives not of their race". In: *American Journal of Political Science* 58.2, pp. 307–321.

David, Brady, Han Hahrie, and Pope Jeremy (2011). "Primary Elections and Candidate Ideology: Out of Step with the Primary Electorate?" In: *Legislative Studies Quarterly* 32.1, pp. 79–105. eprint: `https://onlinelibrary.wiley.com/doi/pdf/10.3162/036298007X201994`. URL: `https://onlinelibrary.wiley.com/doi/abs/10.3162/036298007X201994`.

Estes, Adam (2018). *Newly Moderate Mitt Romney Says He Won't Pursue Abortion Laws.* URL: https://www.theatlantic.com/politics/archive/2012/10/newly-moderate-mitt-romney-says-he-wont-pursue-abortion-laws/322710/.

Gordon, Sanford C and Gregory A Huber (2007). "The effect of electoral competitiveness on incumbent behavior". In: *Quarterly Journal of Political Science* 2.2, pp. 107–139.

Grose, Christian R, Neil Malhotra, and Robert Parks Van Houweling (2015). "Explaining explanations: How legislators explain their policy positions and how citizens react". In: *American Journal of Political Science* 59.3, pp. 724–743.

Hall, Andrew B (2015). "What happens when extremists win primaries?" In: *American Political Science Review* 109.1, pp. 18–42.

Hall, Andrew B and Daniel M Thompson (2018). "Who Punishes Extremist Nominees? Candidate Ideology and Turning Out the Base in US Elections". In: *American Political Science Review*, pp. 1–16.

James, Adams and Merrill Samuel (2008). "Candidate and Party Strategies in TwoStage Elections Beginning with a Primary". In: *American Journal of Political Science* 52.2.

Joulin, Armand et al. (2016). "Bag of Tricks for Efficient Text Classification". In: arXiv: 1607.01759. URL: http://arxiv.org/abs/1607.01759.

Kalla, Joshua L and David E Broockman (2016). "Campaign contributions facilitate access to congressional officials: A randomized field experiment". In: *American Journal of Political Science* 60.3, pp. 545–558.

Mikolov, Tomas et al. (2013). "Distributed representations of words and phrases and their compositionality". In: *Advances in neural information processing systems*, pp. 3111–3119. arXiv: arXiv:1310.4546v1.

Poole, Keith T. and Howard Rosenthal (1984). "The Polarization of American Politics". In: *The Journal of Politics* 46.4, pp. 1061–1079. eprint: https://doi.org/10.2307/2131242. URL: https://doi.org/10.2307/2131242.