Analyzing Candidates' Ideological Messaging Throughout the Electoral System

Benjamin Wittenbrink

June 10, 2020

Abstract

Primary elections in the United States are widely believed to have a polarizing effect on elections. In particular, because primary elections tend to attract a loyal base of partisans, candidates are thought to take extreme positions in order to secure the party nomination. However, the political science literature contains surprisingly little empirical information on the evolution of candidate positions, largely due to poor data on the ideology of primary candidates. This paper addresses whether and how candidates strategically manipulate their ideological positioning in order to increase the chances of victory in the primary and general elections, as well as what factors and incentives account for these (or lack of) changes. Using congressional candidates? tweets from 2012 to 2018 to quantify the magnitude of differences in speech over the election cycle, I specify a multinomial model of speech and employ a bag-of-words approach to identify ideological rhetoric. My interest is in both documenting whether a shift in messaging occurs and exploring the nature and possible reasons for that shift. I will consider the effect of primary opponents? ideology, the divisiveness of the primary, and the general-election competitiveness of the district as well as the political ideology and favorability of party leaders. Initial results suggest that candidates do indeed converge toward the middle in competitive races.

Introduction

Primary elections in the United States are widely believed to have a polarizing effect on elections (Jacobson 2004, Burden 2004, Fiorina et al. 2006, Fiorina & Levendusky 2006). In particular, because primary elections tend to attract a loyal base of partisans (Brady, Han, & Pope, 2007, Brady et al. 2007, Burden 2001, 2004, Hall & Snyder 2015), candidates are thought to take extreme positions in order to secure the party nomination. However, the political science literature contains surprisingly little empirical information on the evolution of candidate positions, largely due to poor data on the ideology of primary candidates (Hall & Snyder, 2015). Assuming that candidates do strategically manipulate their ideological positioning in order to increase the chances of victory in the primary, a follow-up question emerges: upon attaining the party nomination, do candidates change their messaging to appeal to the general electorate? More generally, how does the ideological composition of campaign messaging change between the primary and the general election? What factors account for these (or the lack of) changes?

The proposed study addresses these central questions by using congressional candidates? tweets from 2012 to 2018 to quantify the magnitude of differences in speech over the election cycle. I focus on social media in particular, as it should better capture the rhetoric of candidates ? as opposed to policy proposals ? and as candidates likely use Twitter to appeal directly to voters. I specify a multinomial model of speech and employ a bag-of-words approach to identify ideological rhetoric. My interest is in both documenting whether a shift in messaging occurs and exploring the nature and possible reasons for that shift. I will consider the effect of primary opponents? ideology, the divisiveness of the primary, and the general-election competitiveness of the district as well as the political ideology and favorability of party leaders.

Literature Review

Conventional political punditry suggests that to capture their party?s nomination, candidates must run to the extremes in the primary to establish themselves as strong partisans and then, upon doing so, move to the middle to win over independent and more centrist general election voters. Nonetheless, conventional wisdom also suggests that candidates caught doing so are dogged by accusations of ?flip-flopping,? from which it may be difficult to recover.

Do either of these truisms have theoretical or empirical grounding? I proceed by considering a theoretical model of moderation, followed by a discussion of existing empirical work regarding the post-primary moderation hypothesis. I further outline possible scenarios where post-primary moderation might not take

place or where the expected changes are directionally ambiguous. Based on this consideration of the extant literature, I then propose a study to test my primary research questions.

Why might candidates moderate?

One main way in which ideological positioning and divergence have been studied is in the Hotelling-Downs model (Hotelling 1929, Downs 1957). The central finding of the model is the median voter theorem: both candidates maximize their chance of winning the election by converging to the position of the median voter. First, consider a unidimensional policy space along the real number line. For simplicity let this be given by $X = [0, 1]$ and further suppose there are two parties L and R. These voters are voters scattered along this interval according to a continuous density function f and median x^m such that their position on the interval represents their optimal bliss point. Assume that voters are uniformly distributed along X for simplicity. Voters possess single-peaked preferences with a maximum at their bliss point and their utility decreases monotonically with the distance from this point. In this simple Downsian model where parties only care about winning, parties will only have one equilibrium strategy, staking out positions at the median voter, i.e. $L = R = x^m$. In reality, the two major American political parties are significantly divergent in ideology and policy platforms, though there is evidence to suggest that the American two-party system results in more politically moderate candidates who are more concerned with independent voters than are candidates in multiparty systems (Burden 2001).

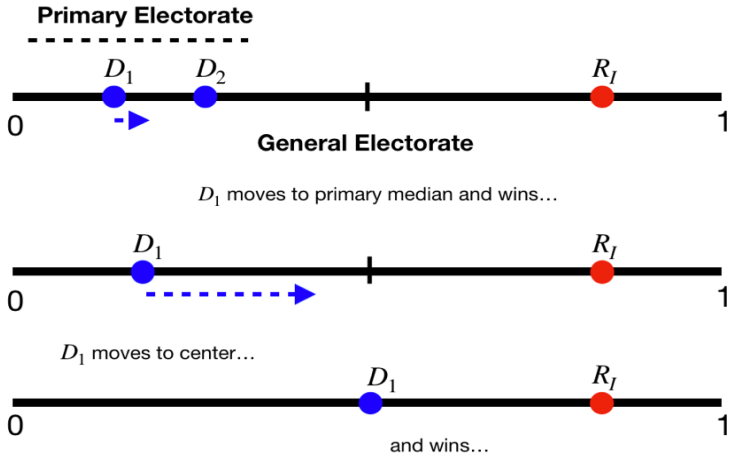
The Hotelling-Downs model can also be adapted to account for the effects of a primary race preceding the general election. The simplest application of the Downsian model to this two-stage example merely adds an independent preliminary ?primary? stage to the model. That is, any given candidate must first compete in the primary to secure the party?s nomination, and upon winning, then competes in the general electorate.

Cox (1990) identifies two sets of incentives that help to explain the candidate?s ideological positioning. Centrifugal forces push candidates toward the extremes of their party while centripetal forces pull candidates toward the center. In order to compete in the general election, candidates must successfully win over their primary voters. As primaries are dominated by partisan actors with strong ideological priors, centrifugal forces are paramount as candidates seek donations and support from the extreme end of the policy distribution. Upon winning the nomination, the candidate must then appeal to the median general voter and hence centripetal forces begin to dominate.

In accord with Cox?s predictions, the equilibrium strategy for the candidates in the Downsian model is to converge to the median primary voter x_p^m and then if the candidate wins, shift toward the median general

voter x_g^m. Under the assumption that party primary electorates are dominated by partisans, candidates should moderate between the primary and the general election, as illustrated below.

Figure 1:



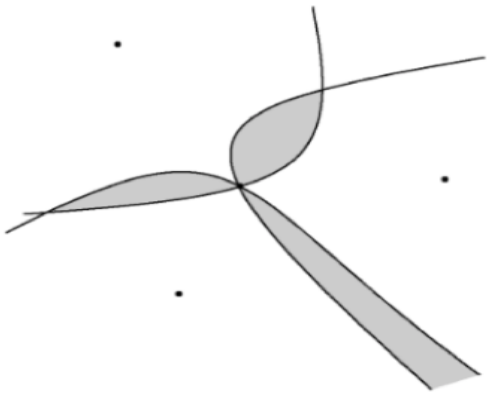
Importantly, a simple two-stage model such as this cannot account for potential interactive effects between the primary and general election stages. For example, a politician?s primary messaging may limit her messaging options for the general election because voters may punish her for having conflicting ideological messages. For example, the candidate may be perceived as dishonest or inauthentic. To account for such effects, Hummel (2010) and Agranov (2011) both develop formal theories of two-stage elections that include possible flip-flopping costs. In particular, Hummel finds that the magnitude of moderation is determined by the associated costs of flip-flopping and that candidates always adopt divergent policies. Agranov shows the existence of an equilibrium where the post-primary moderation hypothesis holds, and primary winning probabilities are inversely related with winning probabilities in the general. Importantly, Agranov finds that if the competitiveness of the general election is sufficiently high, the primary candidates have no incentive to put any weight on the moderate position. However, in elections with limited competition, candidates have an incentive to keep their true positions hidden in the primary. Similarly, Hummel finds that in uncompetitive elections the upsides of any moderation strategy are dominated by the risks of flip-flopping, and hence appearing dishonest.

Why might candidates not moderate?

Although the Downsian model predicts that candidates converge to the median voter, a robust empirical literature documents that in reality American politics is fairly polarized (Ansolabehere et al. 2001, Poole

and Rosenthal 1997). Hence, returning to the one-stage model, there are multiple potential explanations for ideological divergence within each of the two parties. First, median voter predictions are difficult to maintain in multi-dimensional policy spaces (Plott 1967). Assuming that voters are distributed along just three distinct policy spaces (e.g. economic issues, social issues, and foreign policy) renders a median voter equilibrium incredibly rare. For the case of three voters, the intuition behind nonexistence is illustrated below (Acharya: Course Notes, 2018):

Figure 2:



Indeed, the existence of this equilibrium only holds if the multidimensional space can be projected on a unidimensional space, ordering voters by type ? an assumption that seem far-fetched given the heterogeneity of voter preferences across issues. Another potential explanation for the observed divergence from the median voter theorem is that parties are also ideologically motivated (Calvert 1985, Wittman 1977). According to this account, parties do not know the location of the median voter x^m, which is uniform along $[0, 1]$, and are policy motivated with the payoffs $-\tau^2$ and $-(1-\tau)^2$ (for L, R respectively). Then, the unique solution exists when party L chooses $1/4$ and party R chooses $3/4$ (assuming $R \geq L$). Thus, when parties are also ideologically motivated, the equilibrium captures divergence.

Moreover, as noted by Burden (2004), candidates may choose not to adhere to the post-primary moderation hypothesis for a variety of other reasons: i) as mentioned above candidates may be ideologically motivated and are unwilling to compromise these convictions, ii) voters may actually punish candidates for flip-flopping on important issues, and iii) for incumbent candidates, they may have reputations that they need to protect or that may be difficult to transform. Similarly, Tomz and Van Houweling (2009, 2014) find empirical support that voters associate flip-flopping with negatively valenced trait characteristics, including dishonesty. Risk-

averse voters may prefer, therefore, a candidate who commits to a policy, rather than someone who alternates between multiple policies. Many recent presidential candidates, including John Kerry, Hillary Clinton, and Mitt Romney, have all been dogged by accusations of flip-flopping on policy issues.

Finally, there are circumstances that counteract moderation. Some candidates may run for ideological concerns, perhaps to get an issue on the table for discussion. Such candidates have little incentive to moderate, as winning the election is not their primary concern. In addition, many congressional districts are fairly partisan ? either solid blue or solid red ? and hence the primary electorate may be very similar to that of the general election. For instance, consider Texas? 13-th congressional district. The Cook Political Report estimates that the Partisan Voting Index for this district is R+33, essentially implying that this district is 33 percentage points more Republican than the country as a whole. In such a district, it seems a safe assumption that there should not be significant divergence between the median Republican primary voter and the median general election voter. Furthermore, in contexts where the race is not competitive and lacks intensity ? due to strong incumbency effects or unpersuadable voters ? there is little incentive for candidates to adopt extreme positions and then moderate, instead preferring to keep their position hidden (Agranov 2011).

Other, directionally ambiguous factors?

There are also circumstances that have a theoretically ambiguous relationship with respect to change in messaging between primary and general elections. First, when primaries are particularly competitive, and especially when there is great ideological divergence between candidates, the winning candidate may move in the direction of their main primary opponent in order to solidify support among their fellow partisans. Hacker (1965) suggests that divisive primaries can divide the party?s supporters, deterring some of them from supporting the victorious candidate. In this case, the candidate may attempt to unite the party by appealing to disaffected partisan voters. When the more extreme primary candidate wins the primary, this effect moves in the direction of moderation; however, if the more moderate candidate is nominated, this effect moves the general election candidate toward the extremes. This is not implausible: for instance, consider the 2016 Democratic Presidential Primary. Although Clinton, the more moderate candidate, won the nomination, the primary was by all accounts divisive and competitive, generating many news stories reporting on Sanders voters who refused to vote for ? or at least support enthusiastically ? Clinton. In response, Clinton adopted many of Sanders? more progressive policies (NYT 2016, Atlantic 2016). Hence, in this case instead of turning to the center and potentially alienating more voters from the candidate?s base, the nominee for the general election moves toward their opponent?s ideological point to consolidate support

from their base.

Second, an exogenous shock to the political or economic system may substantially alter the focus of campaigns. Feigenbaum and Hall (2015) demonstrate that localized economic shocks from trade with China cause legislators to vote in a more protectionist direction on trade but do not affect their voting on other issues. I suggest these effects are not limited to politicians? voting strategies, but likely extend to ideological messaging during campaigns. That is, candidates running for office may alter the political orientation of their campaign to capitalize on a significant external event. It is theoretically ambiguous whether flip-flopping costs are binding in the face of such a shock, and additionally, whether the candidate chooses to simply change their rhetoric on the specific issue (as in Feigenbaum and Hall) or on all the major campaign issues. Hence, depending on the nature of the exogenous shock, the candidate may respond by moving to the center or the extremes.

Third, internal party politics may affect a candidate?s ideology in the direction of or even away from party leadership. That is, Congressional candidates are likely not independent actors and thus may adjust the ideological orientation of their campaign to court the endorsement of party leaders in primaries. This is particularly salient when the candidate?s party leaders enjoy significant within-party favorability. For instance, in the 2018 midterm elections, FiveThirtyEight wrote ?an endorsement from Trump himself remains the most valuable prize in a Republican primary.? While this observation is certainly not causal, it suggests that candidates may attempt to synchronize their positions with party leaders to attain a valuable endorsement in the primary and then move to their preferred ideological point upon receiving the nomination. Alternatively, when candidate party leaders are particularly unpopular among out-of-party voters, candidates may seek to differentiate themselves from the party and run in the opposite direction of the party leaders.

Contribution

The proposed study has the potential to contribute to contemporary discussions about politicians? strategies ? particularly whether candidates ?run to the center? after winning the nomination. This question has important theoretical and policy implications, particularly in times of growing polarization among the electorate and Congress (Gentzkow, Shapiro, & Taddy 2019, Iyengar, Sood, & Lelkes 2012).

While there has been substantial theoretical work in the literature on the evolution of ideological positioning throughout the primary and general election, empirical work addressing this question has been limited. Hall and Snyder (2015) note that ?political science research often has surprisingly little to say about the overall electoral process,? instead tending to focus on the general or primary elections separately. The rare studies

that do analyze ideology throughout the entire electoral process tend to focus on incumbents, particularly due to the historical difficulty in measuring the ideological positions of losing primary candidates. For example, in a study of more than 2,000 candidates for the House, Rogowski (2012) reports that it was only possible to calculate the ideological positions for 190 candidates using responses for Project Votesmart?s National Political Awareness Test (from Hall & Snyder 2015).

The proposed study attempts to address this gap by employing text analysis of candidate messages (?tweets?) to capture the ideological positioning of congressional candidates in the primary and the general. I then test whether the deviations in ideological positioning are heterogeneous depending on particular features of the primary and general election. In particular, I consider features such as the primary opponents? ideology and the competitiveness of the primary, the competitiveness of the general, the political ideology and favorability of within-party leaders as well as localized exogenous shocks.

One of the main differences with the existing literature is the proposed measure of ideology. In general, one of the main obstacles to research on primary candidates is the lack of any objective measure of partisanship. As the literature has largely relied on roll-call voting scores (DW-NOMINATE) for congressional candidates, this obviously presents significant complications when extending the analysis to primary candidates ? the vast majority of whom will never serve in Congress.

As an alternative, studies have relied on donation data (as in McCarty, Poole & Rosenthal 2006, and Bonica 2013, 2014) from the FEC, estimating the political leaning of donors and then using the weighted average of donors to a candidate to calculate their ideology. While these donation-based estimates largely coincide with DW-Nominate ratings, they are not well suited for comparison between the primary and the general. First, strategic donations are a serious concern. The primary motivation for donors may be that of strategic interest rather than ideological conviction. People may donate to get access to a politician or based on electability concerns. In fact, Hall and Snyder (2015) conclude that donors act even more strategically than voters, ?wasting? very few donations on candidates outside of the top two in primaries. As it is not possible to distinguish between ideological and strategic donations, this issue has the potential to bias ideology estimates. Second, it stands to reason that donors are on average substantially wealthier than non-donors. This likely biases results in the direction of elites: for example, donation history will make campaigns look less populist than their rhetoric may suggest. Finally, structural differences between the general and primary significantly bias the result, as partisans may coalesce around their party nominee regardless of ideology and independents will choose a side. To provide some intuition for this, consider the example of a primary with two candidates, x_1 and x_2, and three donors/voters, y_1, y_2, and y_3. Let $x_1 = y_1, x_2 = y_2 = y_3$. Clearly, x_2

wins and has an observed ideological position of y_2. Suppose x_2 does not change their true position for the general election. Nonetheless, if y_1 donates to x_2, the ideological position of x_2 is now observed as $\frac{2y_2+y_1}{3}$ even though the x_2?s true position remained unchanged. A similar example can be constructed for independents, as independents likely do not participate in the primary election but then donate to a candidate in the general election, artificially biasing the candidate?s observed position closer to the independents bliss point (even if the candidate?s ideology actually remains unchanged). Hence, donation data do not provide a suitable metric for comparison from the primary to the general election.

Instead, the current study focuses on candidates? speech in social media, exploiting recent technical advantages in computing power and text analysis techniques. In particular, I suppose that ideological beliefs are manifested in candidates? rhetoric. For example, Republicans are much more likely to use the term ?death tax? while Democrats prefer the term ?estate tax? (Gentzkow & Shapiro 2010). Ideology may influence candidate speech through the language they choose to use to communicate and frame their positions ? as the example above suggests ? or through the topics they choose to discuss and omit.

This focus on social media interactions as opposed to congressional speeches or policy platforms is a key contribution of this proposed study. While there would also be significant merit in analyzing the policy platforms of candidates ? as platforms signal a candidate?s priorities and are considerably different across the ideological spectrum ? the proposed focus on social media ? Twitter in particular ? has significant advantages. Most importantly, focusing on tweets provides an opportunity to analyze the ideological rhetoric of candidates, which is likely more flexible than the policy positions outlined in policy platforms. That is, instead of becoming more centrist by changing their policy positions and likely suffering from the flip-flopping costs discussed above, through Twitter, candidates can alter their language and framing to sound more centrist. Hence, candidates can adopt more extreme policies that play well in the primary to rally support among partisans and attain the nomination, and then in the general alter their speech using more balanced rhetoric and bipartisan appeals to sound more moderate (Acree et al. 2019). For instance, a Democratic candidate with fairly left policies can employ appeals to moral values of authority and in-group loyalty to increase their conservative appeal while maintaining their credibility among party elites and activists due to their progressive policy platform (Enke 2018). Moreover, candidates may stress different parts of their agenda depending on the audience. For example, a Republican primary candidate may propose to ban abortion and build a wall along the Mexican border in the primary, but upon advancing to the general election, pivot to emphasizing plans to cut middle-class taxes and spur economic growth. Finally, candidates may even be able to exploit informational asymmetries of general election voters to make their policies appear more centrist than they really are (Iyengar & Simon 2000).

Finally, Twitter data is advantageous for several other reasons as well. First, social media presents the opportunity to observe candidates in real dialogue with their potential constituents. That is, policy platforms and speeches may rely on a nuanced understanding of policy and may only be studied carefully by policy wonks and highly educated or interested constituents. In contrast, social media provides a channel for candidates to directly engage a broader constituency in a more accessible manner. Second, focusing on tweets provides a more continuous estimation of the candidate?s evolving ideological positioning than would an analysis of policy platforms. If a candidate changes their policy platform and not just their rhetoric, they will likely remove all references to the prior platform. In contrast, given the high volume of tweets, candidates are unlikely to delete individual tweets from months prior. While still feasible, using policy platforms as data would pose significant technical challenges to the researcher: one could use internet archives to log changes of the content of the candidate?s website, yet this is not perfect as a website may not have been archived at the exact date of change, providing a fuzzy estimate of timing. Finally, from a purely technical perspective, focusing on tweets significantly increases the power of the study, as it provides a substantial increase in the n observed for each candidate.

Predictions

If candidates act in accordance with the Downsian model, we should expect the post-primary moderation hypothesis to hold. In congruence with Hummer (2010) and Agranov (2011), this effect should be particularly pronounced in competitive districts, where independent voters are likely to be most crucial to winning. This effect should also be stronger in districts with an open seat, as the candidates will be relatively unknown and thus not beholden to past promises or votes.

Alternatively, in particularly divisive and ideologically diverse primaries, we may expect the nominated candidate to deploy the ?rally-the-base? strategy and move toward the direction of her competitor to solidify in-party support. The direction, i.e. to the center or to the extreme, is ambiguous, dependent on the candidate and her opponents? ideological positions. Moreover, we might expect candidates to react to localized exogenous economic and political shocks (similar to Feigenbaum & Hall 2015). Additionally, we may expect candidates to move in the direction of or away from party leaders depending on within-party and out-of-party favorability. In particular, we might expect candidates in parties where leaders enjoy strong within-party favorability to align themselves ideologically with these leaders to court an endorsement. Upon receiving the nomination, the candidate may move in the direction of their desired ideological point, and hence the direction is theoretically ambiguous. On the other hand, if party leaders maintain significantly

unfavorability among non-party members, the candidate may seek to differentiate herself from the party and run toward the center, particularly in competitive elections.

Finally, we should not expect to observe any change between the primary and the general election if a candidate runs solely for ideological reasons. Similarly, in races that are exceptionally uncompetitive due to strong partisan identity there is no incentive to change positions in the general election (Agranov 2011). Moreover, incumbents are particularly sensitive to flip-flopping costs and hence have no incentive to moderate (or get more extreme).

Empirical Design

Data

My primary data source for the primary and general election candidates is the America Votes database. America Votes contains the aggregate level performance for each candidate in each state?s congressional primary and general election for the years 2012-2018. In order to collect a candidate?s rhetoric, I pull tweets for each candidate using the Twitter API. Each tweet contains the time and location that it was sent. In the case that there is not a unique or highly probable match between a Twitter account and the candidate, I will search Ballotpedia.org (or another campaign-related website) to identify if either have a Twitter account listed. The prevalence of Twitter among primary candidates is the main time constraint for this study, and I will adjust it accordingly if it appears that usage of Twitter is very low for candidates in earlier years. I don?t anticipate this to be a problem in 2016 or 2018.

Further, in order to construct the model, I will test two methods. First, I will use the text of the United States Congressional Record for the 112th through 114th Congress, obtained from HeinOnline. Hence, this spans the time frame from 2011 to 2017, matching my chosen primaries well. I plan to follow the procedure outlined by Gentzkow, Shapiro & Taddy 2019, parsing the Record into individual-level speeches, identifying the speaker, and then combining this with speaker-level data, including state, chamber, and gender. However, it is possible that representatives speak differently on the congressional floor than they tweet. For example, a congresswoman may tailor her congressional speeches to C-SPAN viewers and party operatives while she may use different speech when using Twitter to appeal to her constituents. To account for this potential problem, I will also pull tweets for every legislator for the first year of each Congress (112th -114th). I restrict the sample to the first year in order to avoid including campaign rhetoric in the model. This design is also not perfect, however, as it may capture pandering that representatives engage in on Twitter, leading

to a biased mapping of rhetoric to DW-NOMINATE scores. There are justifications to use both methods, and thus I plan to test both.

While the text documents and tweets are in a format that is easy for human comprehension, they must be substantially cleaned for text analysis algorithms that do not consider grammatical/sentence structure. Thus, I proceed with standard text-analysis pre-processing steps (Gentzkow, Kelly, and Taddy 2019). This includes removing hyphens and apostrophes and replacing all other punctuation with spaces, as well as any procedural speech, e.g. ?Madam speaker?. I then remove extremely common words from a list of ?stop words? (such as ?the?, ?a?, etc.) and apply the Porter stemming algorithm to reduce words of a common stem (Porter 1980). This list can be obtained from Gentzkow, Shapiro, & Taddy (2019). I then convert the text into vectors of frequency counts of two-word phrases (bigrams). Depending on the computational demand and the results of focusing on bigrams, I may also consider three-word phrases. Moreover, to ensure computational efficiency, I will discard any bigrams that do not appear in more than 15 speeches or are not used by more than 10 different congresspeople in a given Congress (Cook, 2019).

Hence, for some intuition, the sentence ?I do not approve of death taxes.? would yield the bigram representations ?approve death? and ?death tax?. One significant limitation of this methodology is that the contradictory sentence ?I approve of death taxes? would yield the exact same bigram representations. Nonetheless, there is evidence that partisans use ideologically-loaded terminology ? for example, Democrats use the phrase ?estate tax? whereas Republicans say ?death tax? ? so the scope of this problem is ambiguous (Gentzkow & Shaprio 2010). After all of these steps I am able to attain the main input to our model, a matrix for each time period whose rows correspond to speakers and whose columns correspond to bigrams. An element of this matrix gives the count of the number of times a phrase j has been spoken by speaker i in time period t.

A similar procedure must be applied to the Twitter data for the candidates, though the structure is slightly different. In particular, I must first separate the observed tweets by date: tweets in between the beginning of each candidate?s campaign (measured at the announcement date) and the day following the primary election are coded as ?primary? while tweets after the primary election up to the general election are coded as ?general?. After this initial restructuring, I collapse all the tweets in each group into documents by week and run the pre-processing steps as described above. As there are no longer ?speeches?, infrequent bigrams are removed according to the ?term frequency-inverse document frequency? (tf-idf) metric where tf is the count of occurrences in a document and idf is the log of one over the share of documents containing the phrase (Gentzkow, Kelly, and Taddy 2019).

Model

I proceed by taking advantage of the observed word count vectors from Congress to build a model of speech to predict the partisanship of each candidate in both the primary and the general. In particular, I choose to model polarization using the Multinomial Inverse Regression (MNIR) from Taddy (2013). MNIR has many desirable properties. Importantly, it allows for dimension reduction, significantly increasing computational speed, while maintaining estimates of sentiment. The multinomial model is also a natural model for speech, as it supposes that people choose phrases and combinations of words instead of assuming the independence of individual words (Cook, 2019).

For each speaker i I observe the vector of counts c_{it} of bigrams where t refers to the session of Congress. Define the total amount of speech for each speaker to be

$$m_{it} \coloneqq \sum_{j \in \mathcal{J}} c_{ijt}$$

where $\mathcal{J}$ is the set of unique bigrams spoken by i. Then, I assume that the phrase counts for each speaker comes from the multinomial distribution:

$$c_{it} \sim \text{MN}\left(m_{it}, q_{it}(x_{it})\right)$$

where x_{it} is a set of speaker characteristics including the DW-NOMINATE score, and $q_{it}(\cdot)$ is the probability of speaking each phrase given by

$$q_{ijt}(x_{it}) = \frac{\exp\left(\eta_{ijt}\right)}{\sum_{j \in \mathcal{J}} \exp\left(\eta_{sj}\right)}$$

$$\eta_{ijt} = \alpha_{jt} + \gamma_{jt} x_{it} + \varphi_{jt} \mathbb{1}_D$$

where $\mathbb{1}_D$ is an indicator function for whether the speaker i is a Democrat. Here, the coefficients γ_{jt} and φ_{jt} can be interpreted as the effects on utility of the characteristics x_{it} and party affiliation $\mathbb{1}_D$ respectively on the propensity to use the phrase j in session t.

It is important to acknowledge the necessary assumptions of this model. Since the model assigns a unique text-generating function to each speaker, I make the assumption that each speaker i's speech is independent of all other speakers in the data. Moreover, the model implies that speaking a phrase j is only a function of party affiliation and speaker characteristics and also that each phrase is independent of all other phrases used by the speaker. While these assumptions may be fairly restrictve, the model has performed well in previous studies attempting to extract meaning from political speeches (Groseclose and Milyo 2005, Taddy 2013, and Gentzkow, Shaprio, and Taddy 2019).

Estimation of the multinomial regression above is computationally difficult. In order to ease computation, I will approximate the likelihood of the counts with the likelihood of a Poisson model (as in Gentzkow, Shapiro, & Taddy 2019) and make use of a penalized objective function with an L_1 penalty in order to impose sparsity on the loadings. The usage of the sharp L_1 penalty implies that many of the loadings on the bigrams can be expected to be zero, which should reduce overfitting.

As detailed in Taddy (2013) in order to get a lower dimension a sufficient reduction score as detailed in Taddy 2013 is given by

$$z_{it} = \frac{(\boldsymbol{\gamma} + \boldsymbol{\varphi})c_{it}}{m_{it}}$$

where $x_{it}, \mathbb{1}_D \perp c_{it}, m_{it} \mid z_{it}$ and $\boldsymbol{\gamma}$ and $\boldsymbol{\varphi}$ are the coefficient vectors at the bigram level using mentions across all speakers. This reduction is computationally efficient, as it allows us to directly model the party affiliation of text against the SR projection values.

Now that I have constructed the multinomial distribution and reduced the dimension of the speech vectors, I can use the observed speech counts for the representatives to project their position on the DW-NOMINATE scale. To do so, I fit the simple forward linear regression model $y_{it} \sim z_{it}$ where y_{it} represents the DW-NOMINATE Score.

Once I have run this regression, I can calculate the dimension-reduced score z_{it} for each candidate in each time period from their observed counts and use the coefficients from above to predict the corresponding DW-NOMINATE scores. Thus, upon aggregating the results, I attain the projected average DW-NOMINATE score for each primary candidate and for each geneeral candidate. Importantly, the difference between these two scores yields the quantity of interest: each general election candidate's change in ideological positioning from the primary. Formally, let $\mathcal{P}$ be the set of all primary candidates and $\mathcal{G}$ be the subset of all general election candidates. With slight abuse of notation, then for each candidate $i \in \mathcal{P}$ I observe y^p_{it} and for each $j \in \mathcal{G}$ I observe y^g_{jt}. Thus I define the following quantities:

$$\text{Ideology Change} \coloneqq y^g_{jt} - y^p_{jt}$$

$$\text{Extremity Change} \coloneqq |y^g_{jt}| - |y^p_{jt}|$$

for $j = i$. These will be my main dependent variables in the subsequent analyses.

Preliminary Results

I have collected Twitter data from two sources. First, I have gathered all Tweets from sitting Senators and Representatives during the 116th Congress, which will be used for the baseline model. Second, I have

gathered all Tweets from the two-party candidates in the 2018 congressional general election. These are the individuals whose ideology we wish to predict. Eventually, if the data permits, I hope to collect data on the primary candidates as well.

Using a mutual information test, I calculate the 10,000 most partisan phrases in the baseline data. Calculating the phrase counts for each sitting Congressperson, I then fit the baseline multinomial logistic model. Analogously, calculating the phrase counts for each of the election candidates, I am able to estimate ideological predictions for each candidate over time. Currently, I am able to do so at the monthly level; however, hopefully by expanding the feature set, I will eventually be able to calculate these estimates at the weekly level.

Below, I plot the smoothed polynomial fits of the ideological predictions in the 40 most competitive House elections.

Zooming in to take a closer look at four:

Figure 1: Competitive House Races

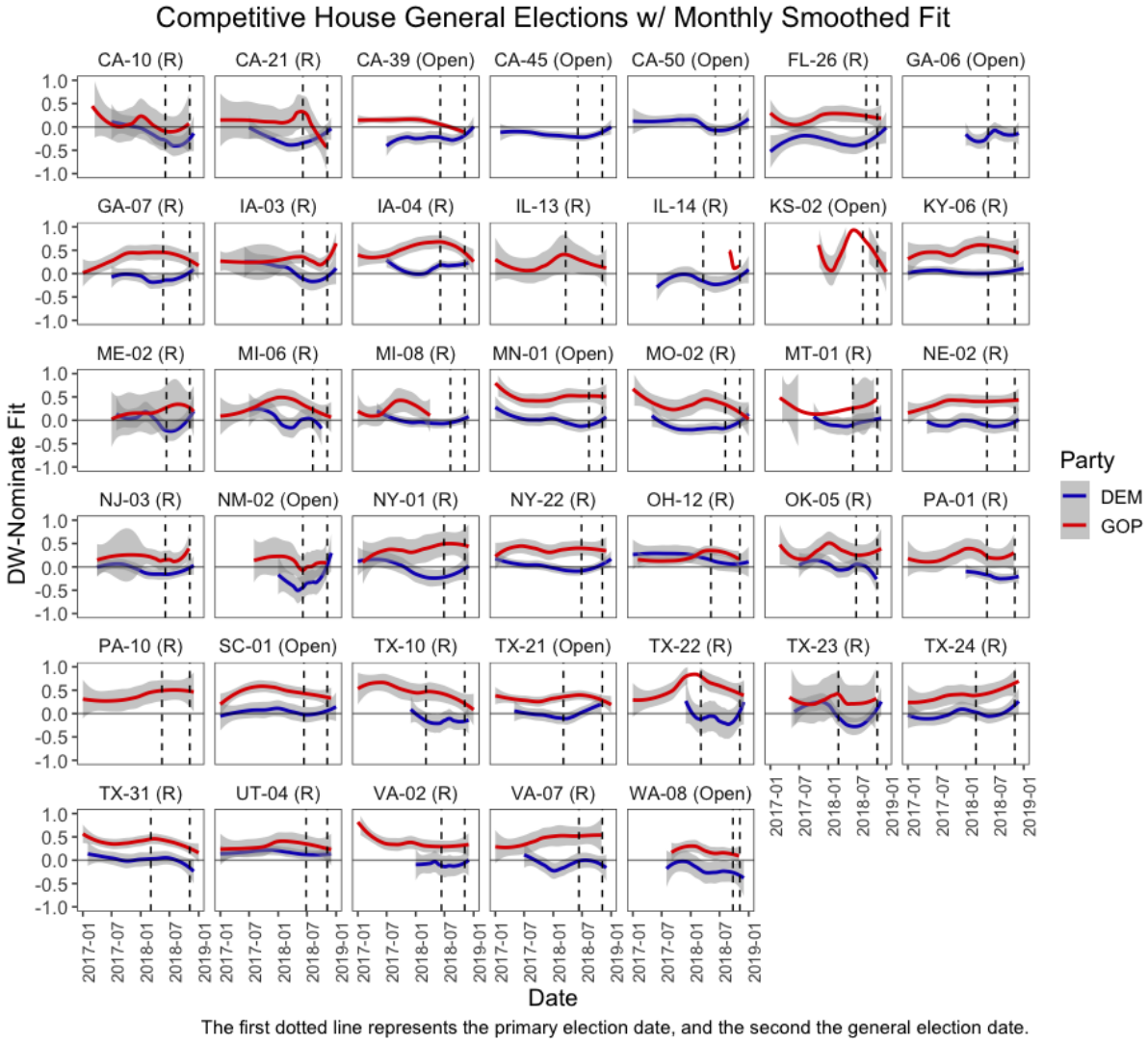
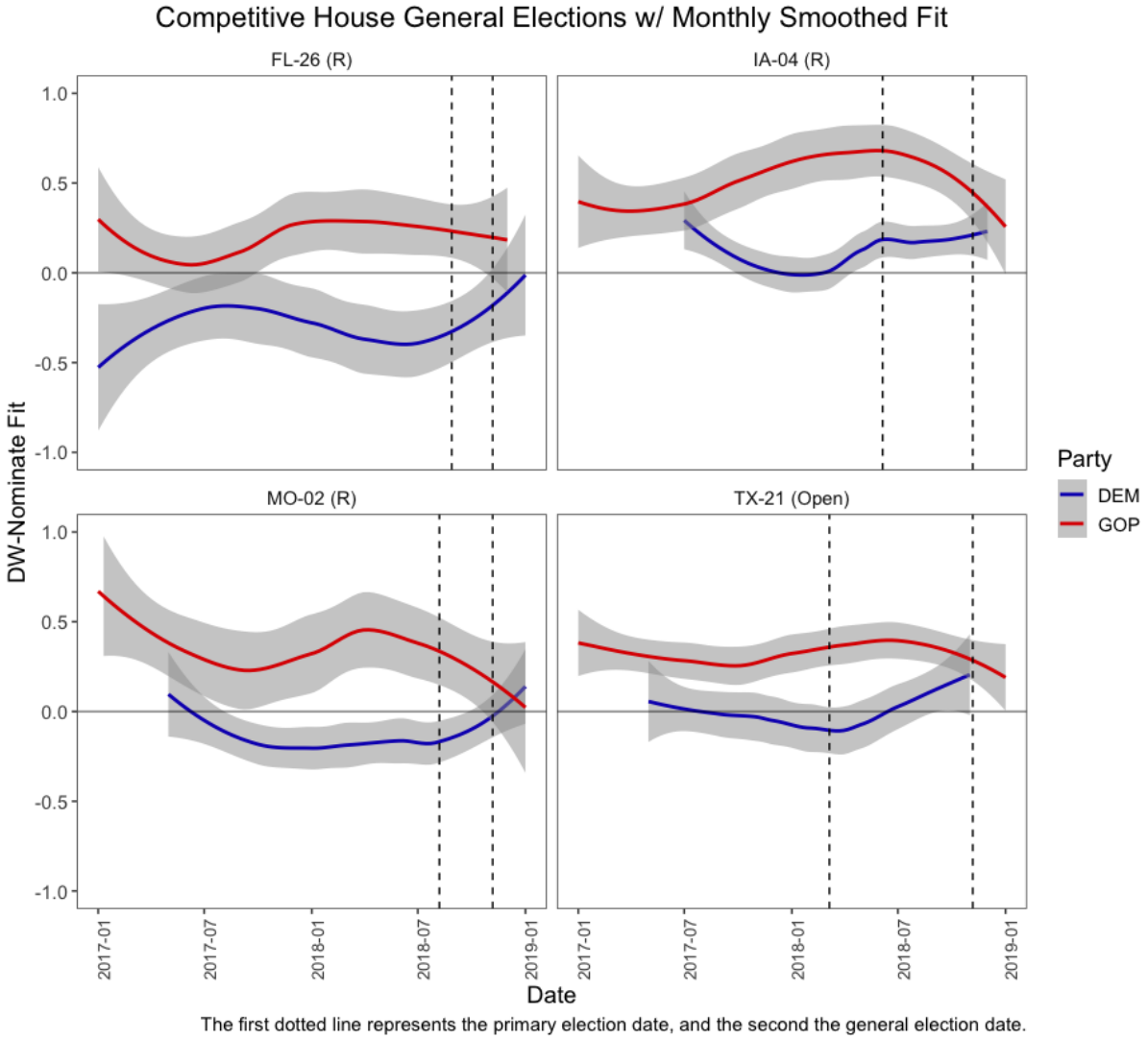


Figure 2: Subset of Competitive House Races



Of course, these are smoothed polynomial fits, and the actual point estimates are much noisier. Nonetheless, these plots provide some suggestive evidence of overall convergence. To look at this more rigorously, I estimate the following model

$$\hat{y}_{i,t} = \beta x_{i,t} + \varphi_i + \epsilon_{i,t}$$

where $\hat{y}_{i,t}$ is the predicted ideology and $x_{i,t}$ represents the number of days until the general election for candidate i at time t. Finally, $\boldsymbol{\varphi}$ is a vector of candidate fixed effects, serving to center the estimates. In order to understand whether the candidates converge over the general election, we let $t = 0$ be the end of the primary.

I run this specification for Democrats and Republicans separately, as we expect the coefficient β to be of opposite sign given the nature of the DW-NOM scale. In particular, as we expect candidates to moderate over the course of the election, we should expect β to be positive for Republicans and to be negative for Democrats. We also run the specification restricting just to candidates in competitive elections and just to incumbents. We expect moderating forces to be particularly severe in competitive elections, as candidates are forced to win over more independents/ideological moderates. In contrast, we expect incumbents to be least likely to moderate, as they have known political histories and changing their rhetoric likely imposes flip flopping costs. Estimating these specifications, yields the following tables:

Table 1: GOP Candidate Ideological Fit Over Election

	Dependent variable:		
	Fitted DW-NOMINATE		
Sample	All	Competitive	Incumbent
Days Until General Election	0.00058***	0.00061**	0.00030
	(0.00016)	(0.00028)	(0.00021)
Candidate FE	X	X	X
Observations	1,383	277	801
R^2	0.60346	0.76479	0.62888

Notes: DW-NOM scores closer to 0 are more moderate, to 1 are further right. Thus positive coefficients imply that the candidate moderates, while negative coefficients suggest the candidate got more extreme. All observations are weighted in accord with the observation's proportion of bigram counts relative to that candidate's total.

* Significant at the 10% level.

** Significant at the 5% level.

*** Significant at the 1% level.

Table 2: DEM Candidate Ideological Fit Over Election

	Dependent variable:		
	Fitted DW-NOMINATE		
Sample	All	Competitive	Incumbent
Days Until General Election	-0.00052^{***}	-0.00077^{***}	-0.00080^{***}
	(0.00010)	(0.00016)	(0.00019)
Candidate FE	X	X	X
Observations	1,832	330	711
R^2	0.68913	0.73193	0.62020

Notes: DW-NOM scores closer to 0 are more moderate, to -1 are further left. Thus, negative coefficients imply that the candidate moderates, while positive coefficients suggest the candidate got more extreme. All observations are weighted in accord with the observation's proportion of bigram counts relative to that candidate's total.

* Significant at the 10% level.

** Significant at the 5% level.

*** Significant at the 1% level.

The average difference between the primary election and the general election is 160 days. This implies that candidates moderate on average by $160 \cdot \hat{\beta}$, i.e. Republican candidates moderate on average by 0.093 and Democratic candidates moderate on average by 0.083 DW-NOMINATE points from the end of the primary to the end of the general. This is substantial moderation! Moreover, the effect for candidates in competitive elections is larger for both Democrats and Republicans, as hypothesized. Interstingly, while the coefficient for incumbent Republicans is significantly smaller (though insignificant) as hypothesized, the coefficient for incumbent Democrats is even larger.

References

1. Acree, B., J. Gross, N. Smith, Y. Sim, and A. Boydstun. 2020. "Etch-a-Sketching: Evaluating the Post-Primary Rhetorical Moderation Hypothesis." *American Politics Research*, 48(1): 99-131.
2. Agranov, M. 2011. Flip-flopping, intense primaries and the selection of candidates. (Note: working paper rather than published 2016 version).
3. Ansolabehere, S., J. Snyder, Jr., and C. Stewart III. 2001. ?Candidate Positioning in U.S. House Elections,? *American Journal of Political Science* 45:136-159.
4. Blanchard, E.J., Bown, C.P. and Chor, D.. 2019. Did Trump's Trade War Impact the 2018 Election?. *NBER Working Paper No. w26434*
5. Brady, David W., Hahrie Han, and Jeremy C. Pope. 2007. "Primary elections and candidate ideology: Out of step with the primary electorate?." *Legislative Studies Quarterly* 32(1): 79-105.
6. Bonica, Adam. 2013. ?Ideology and Interests in the Political Marketplace.? *American Journal of Political Science* 57(2): 294-311.
7. Bonica, Adam. 2014. ?Mapping the Ideological Marketplace.? *American Journal of Political Science* 58(2): 367-387.
8. Burden, B. C. 2001. "The polarizing effects of congressional primaries." In *Congressional primaries and the politics of representation*, ed. P.F. Galderisi, M. Ezra, and M. Lyons. New York: Rowman & Littlefield Publishers.
9. Burden, Barry C. 2004. "Candidate positioning in US congressional elections." *British Journal of Political Science* 34(2): 211-227.
10. Calvert, R. 1985. The value of biased information: A rational choice model of political advice. *The Journal of Politics*, 47(2), pp.530-555.
11. Cook, C. 2019. "Measuring political ideology of presidential candidates." Unpublished project.
12. Cox, G. W. 1990. "Centripetal and centrifugal incentives in electoral systems." *American Journal of Political Science*, 34, 903-935.
13. Downs, A. 1957. *An economic theory of democracy.* New York, NY: Harper & Row.
14. Enke, ., 2018. Moral values and voting *National Bureau of Economic Research No. w24268*

15. Erikson, R., and G. Wright. 2000. ?Representation of Constituency Ideology in Congress.? In *Continuity and Change in House Elections*, ed. D.W. Brady, J.F. Cogan, and M.P. Fiorina. Stanford, CA: Stanford University Press.

16. Feigenbaum, J., and A. Hall. 2015. "How legislators respond to localized economic shocks: evidence from Chinese import competition." *The Journal of Politics* 77(4): 1012-1030.

17. Fiorina, M., S. Abrams, and J. Pope. 2006. *Culture war? The Myth of a Polarized America.* New York: Pearson Longman.

18. Fiorina, M., and M. Levendusky. 2006. ?Disconnected: The Political Class versus the People.? In *Red and Blue Nation? Characteristics and Causes of American Polarized Politics*, ed. P.S. Nivola and D. Brady. Washington, DC: Brookings Institution Press.

19. Gentzkow, M., Kelly, B. and M. Taddy. 2019. Text as data. *Journal of Economic Literature*, 57(3), pp.535-74.

20. Gentzkow, M., and J. Shapiro. 2010. ?What Drives Media Slant? Evidence from U.S. Daily Newspapers.? *Econometrica* 78(1): 35?71.

21. Gentzkow, M., Shapiro, J.M. and M. Taddy. 2019. Measuring Group Differences in High?Dimensional Choices: Method and Application to Congressional Speech. *Econometrica*, 87(4), pp.1307-1340.

22. Hacker, A., 1965. Does a ?Divisive? Primary Harm a Candidate's Election Chances?. American Political Science Review, 59(1), pp.105-110.

23. Hall, A. 2015. "What happens when extremists win primaries?." *American Political Science Review*, 109(1), pp.18-42.

24. Hall, A. and J. Snyder, Jr. 2015. "Candidate ideology and electoral success." Unpublished Manuscript, Stanford University.

25. Hall, A. and J. Snyder, Jr. 2015. "Information and Wasted Votes: A Study of US Primary Elections". *Quarterly Journal of Political Science*, 10(4): 433-459.

26. Hirano, S., J.M. Snyder Jr, S. Ansolabehere, and J.M. Hansen. 2010. "Primary elections and partisan polarization in the US Congress." *Quarterly Journal of Political Science.*

27. Hotelling, H. 1929. Stability in competition. *The Economic Journal*, 39(153), 41-57.

28. Hummel, P. 2010. Flip-flopping from primaries to general elections. *Journal of Public Economics*, 94, 1020-1027.

29. Iyengar, S. and Simon, A.F. 2000. "New perspectives and evidence on political communication and campaign effects." *Annual review of psychology*, 51(1), pp.149-169.

30. Iyengar, S., Sood, G. and Lelkes, Y. 2012. "Affect, not ideologya social identity perspective on polarization." *Public opinion quarterly*, 76(3), pp.405-431.

31. Jacobson, G. 2004. *The Politics of Congressional Elections*, 6th ed. New York: Pearson.

32. Margalit, Y., 2019. "Political responses to economic shocks." *Annual Review of Political Science.*

33. McCarty, N., K. Poole and H. Rosenthal. 2006. *Polarized America: The Dance of Ideology and Unequal Riches.* Cambridge, MA: MIT Press.

34. Plott, C. 1967. "A notion of equilibrium and its possibility under majority rule." *The American Economic Review*, 57(4), pp.787-806.

35. Porter, M. 2009. "The English (Porter2) Stemming Algorithm." Accessed at http://snowball.tartarus.org/algorithms/english/stemmer.html on December 6, 2019.

36. Rogowski, J. 2012. ?Primary Systems, Candidate Platforms, and Ideological Extremity.? Unpublished Manuscript.

37. Schaffner, B. 2012. *Politics, Parties, and Elections in America.* Seventh Edition. Boston: Wadsworth.

38. Taddy, M. 2013. ?Multinomial Inverse Regression for Text Analysis.? *Journal of the American Statistical Association* 108 (503): 755?70.

39. Tomz, M. and Van Houweling, R. 2009. "The electoral implications of candidate ambiguity." *American Political Science Review*, 103(1), pp.83-98.

40. Tomz, M. and Van Houweling, R., 2014. "Political repositioning: A conjoint analysis." Unpublished manuscript, Stanford University.

41. Wittman, D. 1977. "Candidates with policy preferences: A dynamic model." *Journal of economic Theory*, 14(1), pp.180-189.