Nationalized Elections, Localized Campaigns?

A Supervised Machine Learning Approach to Measuring Nationalized Political Rhetoric

Derek Holliday*

September 11, 2022

Recent research documents the "nationalization" of election results in the United States. Does this nationalization of electoral results extend to campaign rhetoric itself? I utilize quantitative text analysis and supervised machine learning techniques to determine this distinctiveness of gubernatorial rhetoric relative to candidates for national office, applying the analysis to a broad corpus of text. The results show a more nuanced picture of nationalized rhetoric; in some forums, such as televised debates, gubernatorial speech is easily distinguishable from the speech of presidential candidates. In less structured formats, gubernatorial speech is more likely to reference national topics. These results speak to the complex information environment voters may encounter in down-ballot contests and the subsequent challenges posed to political accountability.

*Ph.D. Candidate, UCLA, deholliday@g.ucla.edu

A growing literature in political science focuses on the "nationalization" of U.S. politics. Generally, this phenomenon refers to national political actors and issues influencing state and local political activity (Abramowitz and Webster 2016; Hopkins 2018; Sievert and McKee 2019). The most prominent results in this literature point to the increasing correlation between vote shares of presidential and down-ballot candidates of the same party. This nationalization of election results has potentially problematic effects on the quality of representation from the winners of such down-ballot contests. If voters are evaluating candidates not based on job demands specific to the elected office (which at the state and local level is in many cases orthogonal to the contours of national politics) but simply on the partisan affiliation of the candidates, holding an office-holder accountable becomes significantly more challenging.

A large portion of the nationalization literature links nationalized election results to a lack of information specific to down-ballot races (G. J. Martin and McCrain 2019; Moskowitz 2021). If voters have access to distinct down-ballot political information, they are more likely to make voting decisions less directly influenced by co-partisanship with national candidates. While these studies focus predominantly on the media environment, one source of political information remains understudied; the content of political campaigns themselves. A growing narrative has emerged within popular media that the content of down-ballot political campaigns has grown more similar to the campaigns of candidates for national office. By referencing the talking points of prominent national candidates, down-ballot candidates are perhaps able to easily identify themselves to voters who are already more familiar with the national candidates.

In this paper, I evaluate the extent to which the rhetoric of gubernatorial candidates reflects the topics referenced by their national counterparts. I consider a broad array of political speech, spanning televised electoral debates from 2000 to 2018, televised political advertisements from 2004 and 2008 election campaigns, and the Twitter activity of incumbent Governors and Members of Congress in 2018. I approach the measurement of nationalized rhetoric as a text classification problem by constructing a classification model based on topic representations of speech known to be of state or national providence. The results show a more complete and nuanced picture of gubernatorial campaign rhetoric; gubernatorial candidates overall tend to reference topic distinct from their national counterparts, but are more likely to "nationalize" their speech in televised ads and social media. These results have significant consequences for voter behavior; if voters are effectively exposed to gubernatorial campaign messaging, they are at least nominally able to make voting decisions that aren't as susceptible to nationalized pressures. However, voters are also susceptible to appeals that *are* more nationalized, and if these messages are the ones highlighted by media outlets, then nationalized voting behavior may follow.

The paper proceeds as follows; first, I briefly review existing literature on nationalized elections and their underlying mechanisms, following this discussion with anecdotal and theoretical extensions to campaign rhetoric. Second, I present my methodological approach to measuring the nationalization of political rhetoric. Third, I apply my approach to three corpuses: debates, advertisements, and Twitter. I conclude with a discussion of the results and implications for future research.

Nationalized elections

The correlation between Presidential and down-ballot vote shares of co-partisan candidates has steadily and steeply risen since the 1970s. The correlation between Presidential and gubernatorial candidates in midterm elections grew from a low of less than 0.2 in 1970 to a high of over 0.8 in 2010, with elections of presidents and governors occurring in the same year showing a similar but less dramatic rise of less than 0.3 in 1968 to around 0.7 in 2012 (Hopkins 2018). The rate at which the same party won both the Presidential and Senatorial contests in a given state rose from 52% in 1980 to 84% in 2012 (Sievert and McKee 2019), with similar trends in U.S. House elections (Jacobson 2015). Such nationalization extends to further down-ballot races as well, including State Supreme Court and Superintendent of Public Instruction

(Weinschenk et al. 2020; Weinschenk 2022).

A number of mechanisms have been theorized regarding the nationalization of American elections, which roughly fall under two (non-mutually exclusive) categories: identity and information. The identity mechanism explains nationalization as an extension of partisanship and polarization; that is, partisanship is an affective, expressive identity, so we should expect voters to vote for their preferred party (or against their non-preferred party) in any context (national, state, or local) as an expression of in-group solidarity and/or out-group antipathy (Huddy and Bankert 2017; Iyengar, Sood, and Lelkes 2012). For example, Abramowitz and Webster (2016) note the increase in straight-ticket party voting is associated with high levels of out-party antipathy, also known as negative partisanship. Webster (2020) finds a similar association between straight-ticket voting and anger. However, the extent to which voters are willing to make voting decisions purely on party identification may be bounded. Using a conjoint design, Mummolo, Peterson, and Westwood (2021) find voters punish excessive deviation from preferred positions on salient policies by co-partisan candidates. Costa (2021) also uses a conjoint design to find voters prefer candidates who provide substantive representation and constituency service over partisan affect. So, while partisanship contains information about ideology, it is not a substitute for ideology in the eyes of voters.

The information mechanism suggests decreasing access to meaningful information about state and local political contests drives increased nationalization. Without specific policy information to evaluate the candidates, voters simply default to the candidate of their preferred party. Hayes and Lawless (2018) note the steep decline in access to local news in the last decade, with a 10% reduction in issue coverage and a 33% reduction in the coverage of candidate traits in U.S. House of Representatives contests between 2010 and 2014. Since 2004, one in five newspapers has closed (Abernathy 2018). This decline in access to local news is associated with increase in nationalized news content and voting behavior (G. J. Martin and McCrain 2019; Moskowitz 2021). The more information voters have besides the party identification of the candidates, the more likely they are to make split-ticket decisions.

Both the identity and information hypotheses stress the importance of voters using available signals, including partisanship and policy stances, to make a judgment on candidate type. Even in the state context, nationalized signals are not devoid of information on candidate type. Platforms have homogenized and national and state parties are seen as more singular than separate (Caughey, Dunham, and Warshaw 2018; Hopkins 2018). The dimensions of state politics largely mirror the left-right contours of national politics (Caughey and Warshaw 2016; Shor and McCarty 2011).

Nationalized campaigns?

The strength of party identification, relative lack of prolonged media attention in down-ballot contests, and correlation between state and national political dimensions may incentivize candidates for subnational office to make nationalized rhetorical appeals during their campaigns to amplify their candidacy. The mechanisms discussed above largely focus on voter-level factors that influence the electorate's propensity to vote in a nationalized manner. While elite behavior certainly pushes the trend in a similar manner through cross-state homogenization of party platforms and the more rigorous sorting of partisans along ideological lines, the effects candidates themselves can have in any given election remains understudied.

Media portrayals of gubernatorial campaigns have stressed their "nationalized" content. For example, national media outlets characterized the 2019 gubernatorial races in Kentucky, Louisiana, and Mississippi as being nationalized due to Donald Trump's personal involvement in some of the races and an emphasis on impeachment of Trump as a campaign issue (Manchester 2019; J. Martin 2019; Rojas and Alford 2019). Other outlets gave similar appraisals of other races, including Washington in 2016 ("Inslee… was happy to nationalize the governor's race, sounding at many events like he was running against Trump"), West Virginia in 2011 (where the Republican Governor's Association spend $3.5 million in ad buys in an attempt to link the democratic candidate to Obamacare), and Texas in 2010 ("Mr. Perry turned the race into a referendum on federal spending") (Brunner 2016; Catanese 2011; McKinley Jr. 2010). Some Governors have engaged in nationalized rhetoric themselves, with Governor Gavin Newsom of California characterizing supporters of the 2021 recall election as "a partisan, Republican coalition of national Republicans, anti-vaxxers, Q-Anon conspiracy theorists and anti-immigrant Trump supporters." At the very least, candidates for state office do not feel bound only address or espouse policies, individuals, and organizations exclusive to their own states.

Some empirical evidence exists to suggest political rhetoric has broadly nationalized in the same manner elections have. Das et al. (2022) analyze the tweets of incumbent Members of Congress, Governors, and mayors in 2018 utilizing a topic modeling approach to ascertain the level of semantic similarity between the different office holders. They find Members of Congress and Governors are almost indistinguishable in terms of their topical similarity, while mayors still seem to tweet about distinct topics. These findings give pause to the "all politics is national" hypothesis, at least at the local level, but still suggest gubernatorial rhetoric has nationalized parallel to the nationalization of electoral results. It is important to note, however, that the Twitter activity analyzed by Das et al. (2022) is not specific to campaigning, focuses on sitting incumbents, and may also contain content that is apolitical in nature. For example, New Jersey Governor Phil Murphy has, since February 2022, tweeted about his daily Wordle score, the Saint Peter's University men's basketball team, and changing the state bird of New Jersey to the middle finger for April Fools.

Both the identity and information mechanisms prove potentially useful in explaining this potential nationalization of campaign rhetoric. If voters are predominantly motivated by simple party identification, nationalizing one's campaign appeals in gubernatorial contests may boost signals of partisan type by linking candidates to more traditional, national-level policy positions. This makes candidates more "identifiably" Republican or Democrat. Alternatively, nationalizing campaign appeals may have the effect of diluting the pool of locale-specific information available to voters, instead focusing the information environment on national signals of partisan type.

While there are many plausible reasons to nationalize a gubernatorial campaign, there are equally plausible reasons to keep a campaign localized. The most obvious reason is voters may recognize a candidate running on nationalized appeals has no jurisdiction over the issue being discussed. Current research is divided on the extent to which voters hold politicians accountable for conditions under their jurisdiction; Arceneaux (2006) finds survey respondents tend to attribute credit/blame to offices which they (fairly accurately) assign functional responsibility to, whereas Brown (2010) finds partisanship moderates the attribution of functional responsibility and subsequent credit/blame. de Benedictis-Kessner and Warshaw (2020) find some evidence for both conclusions suing time series, cross-sectional models; voters routinely hold the president's party responsible for local economic conditions, but also hold governors accountable for such conditions. Therefore, the incentives for candidate to nationalize gubernatorial campaigns seems mixed.

Additionally, rhetorical context may influence the content of campaign appeals. Such differences may emerge from the perceived audiences of different medium and constraints on message length or content (Bossetta 2018; Owen 2014; Stier et al. 2018). From a nationalization perspective, we may expect messages broadcast through less geographically-defined medium to emphasize more national political themes, perhaps as a means to fund raise outside of one's jurisdiction (Reckhow et al. 2017). As the available space or time for the message decreases, we may also expect appeals to homogenize toward more familiar appeals to national partisanship, while longer-form messages can explore locale-specific details of certain issue dimensions. To fully understand the dynamics of nationalized rhetoric, we must therefore consider a broader array of rhetorical contexts.

Methodological approach

I approach the potential nationalization of political rhetoric during campaigns as a supervised text classification problem; the collection of words spoken or otherwise disseminated during a campaign can be categorized as having either national or state content. This approach allows me to measure the presence of nationalized rhetoric across a broad array of rhetorical contexts.

The workflow involves (1) defining training data where document "class" (national or state providence) is known, (2) quantitatively representing text using a topic-modelling approach, (3) fitting a classification model using the training data and quantitative text representation, and (4) predicting the class of test data using the trained classification model. I describe each step of the workflow in more detail below. The result of this workflow is a state/national classification rate of political campaign rhetoric at the state level, where a high national classification rate signifies campaign activities sharing similar characteristics to national-level rhetoric and where a high state classification rate signifies activities sharing similar characteristics to state-level rhetoric.

Training data

A classification approach requires a model to be fitted with data of known classes (state or national). This precludes the possibility of using words from campaign activities as part of the training process; by design, these activities are of ambiguous "class," being potentially more state or national in content than their providence would suggest. Additionally, the training data must be substantively representative of the classes I aim to predict. To this end, I synthesize a corpus of Presidential speeches (State of the Union addresses and opposition responses, inaugural addresses, official statements, and national party platforms) and gubernatorial speeches (State of the State addresses and budget addresses) representing national and state political content, respectively. The final training corpus contains 1,038 speeches and documents, 227 national and 811 state, spanning 2000-2018.

This training corpus is meant to distinguish between national and state political content via the policy discussions in each respective sphere. State of the Union/State speeches are particularly useful in this context, as they often involve explicit references to policy accomplishments and goals. However, this does not prevent certain words and phrases from existing in either the state or national contexts that are highly predictive of a particular class but devoid of policy content. For example, most gubernatorial State of the States addresses include the state names themselves and the names of residents for those states (such as "Californians" or "Hoosiers"). Including these words during the model fitting process would potentially allow for the model to "cheat" and accurately predict class not from the policy content of a speech but from these cheap signals of state providence. Therefore, during standard text pre-processing of the training data (stopwording, lemmatization, removal of very short or very rare words), I also remove all state names, names of state residents (including nicknames), references to the level of office (besides presidential), common audio transcription tags (laughter, applause), and common words without policy meaning (year, will, thank, etc.).

Quantitative text representation

To predict class, the text of any document needs to be quantitatively represented. In classic supervised learning approaches, the choice of quantitative text representation is driven by best out-of-sample prediction accuracy. In my application, however, the substantive meaning of the text representation is equally important, as I want to interpret the classification rate as a meaningful indicator of overall national content. Therefore, I represent all text in this paper as estimated topic proportions using a structural topic model (STM) approach.

STM treats texts as "bags of words." Like the latent Dirichlet allocation (LDA) approach, STM assigns words to topics and topics to documents probabalistically (Roberts, Stewart, and Airoldi 2016). The output gives word probabilities associated with each topic and the topic proportions for each document. STM builds upon LDA by allowing topic probabilities to vary according to researcher-specified covariates, allowing for resulting topics to more closely approximate the theorized data-generating process. In this application, I allow topics to vary as a function of providence (state or national) and year (binned in two-year intervals). Functionally, this means every document x_i is represented by a length k vector of topic proportions $\theta_{1 \ldots k}$, which are then used as the features of the classification model.[1]

The application of the STM process in this paper can be thought of as a text-as-data manifestation of other dimension-reduction techniques (such as principal components analysis) in machine learning. Methodologically, the process alleviates problems resulting from high-dimensional data such as data sparsity, computational complexity, and overfitting. Substantively, STM provides a more interpretable output than raw term frequencies and more closely captures the theoretical thrust of the nationalization hypothesis; certain collections of words are more "national" in nature than others. This is significantly more theoretically meaningful than any single word being an indicator of state or national providence.

Classification model

I train an assortment of classification models using the STM-generated topic proportions of the training data: logistic regression, naive Bayes, regularized logistic regression (lasso), support vector machine (SVM), and boosted gradient descent (XGBoost). Before fitting the models, I randomly hold out 20 percent of the training data as a validation set to evaluate out-of-sample model performance. Models requiring hyperparameter tuning are first evaluated for performance using area under the ROC curve (AUC) with 10-fold cross-validation. Furthermore, because of the class imbalance within the training data (significantly more state documents than national documents), I use a bootstrap-based synthetic oversampling technique called ROSE (Random Over-Sampling Examples) to prevent models from always predicting the majority class. The performance of each model for predicting the class of the documents in the

[1] k is a researcher-defined hyperparameter determining the number of documents. The results presented in this paper use $k = 40$, which was determined to have the best balance of semantic coherence, held-out likelihood, and minimization of residuals. See online appendix for details and results with alternative k specifications.

Table 1: Classification Model Performance on Heldout Documents

	National Documents		State Documents			
Model	Correct	Incorrect	Correct	Incorrect	Accuracy	AUC
Logistic Regression (Nonpenalized)	39	1	167	1	0.990	0.999
Naive Bayes	36	4	162	6	0.952	0.984
Penalized Logistic Regression (Lasso)	39	1	166	2	0.986	0.999
Boosted Gradient Descent (XGBoost)	28	12	168	0	0.942	0.994
Support Vector Machine	40	1	167	1	0.995	0.999

validation set are shown in Table 1.

All the models perform remarkably well, indicating there is sufficient textual differentiation between state and national rhetoric to perform classification in this manner. This also alleviates a potential concern that the state documents of the training data may themselves contain nationalized rhetoric. While such rhetoric may exist, the models are able to successfully determine which topics are most associated with state and national origin. I use the unpenalized logistic regression model as the final prediction model for the remainder of this paper because it has the joint-highest AUC and easily interpretable coefficients, but see the online appendix for prediction results from the other models.

Prediction results

The final step of the workflow is to fit the trained model to text from political campaigns to predict the class of each document. I consider three different mediums through which political candidates communicate with voters: televised debates, TV advertisements, and social media (Twitter). There are both theoretical and technical reasons to believe classification performance should vary by medium, which I will discuss in greater depth below.

Televised debates

I first consider the potential nationalization of rhetoric during televised political debates. Specifically, I analyze an original corpus of 397 electoral debates (86 presidential and 311 gubernatorial) between 2000 and 2018 retrieved from closed-captioned transcripts from the C-SPAN video archives, which were originally broadcast either directly on C-SPAN or through local public affiliates.[2]

Research on gubernatorial debates is rare, but the few studies that have been conducted conclude candidates largely focus on policy positions rather than character (Benoit, Brazeal, and Airne 2007) and viewers of debates are often able to correctly identify the eventual winner of the contest (Benjamin and Shapiro 2009). Research on the effects of presidential debates largely conclude such events have some short-term effect on candidate preference (Hillygus and Jackman 2003) and issue knowledge/salience (Benoit, Hansen, and Verser 2003). Practically, the debate context helps control for candidate-level confounders such as ideology, campaign resources, or campaign activity level that may bias results in a different context (such as television advertisements or social media).

While there is no particular reason why the messaging content of debates would deviate substantially from other mediums, there are reasons to believe there would be a high hurdle to find evidence of nationalization. The length of debates (typically at least an hour long) allows for greater depth and breadth of discussion across policy issues, making state-specific content perhaps more likely to appear in gubernatorial contests. The moderator of the debate (typically a member of the local media) may push candidates to give positions on more local issues of interest to that media market. The candidates themselves may believe the audience of such debates to be fairly well-informed, makes less detailed or policy-oriented appeals less effective. Still, media coverage of the debates the next day may focus on the headline-grabbing nationalized appeals made during the debates, and candidates are easily able to answer the

[2] Transcripts were retrieved through a combination of headless web browsing and scraping. Transcripts for non-closed captioned videos are not available.

questions they want to answer instead of the questions that are asked of them.

Figure 1 shows the results of applying the trained classification model to the C-SPAN debates corpus. The left panel shows the confusion matrix of the classification model, the upper right panel shows the average predicted probabilities of presidential and gubernatorial debates of being of national class over time, and the lower right panel shows the predicted class counts for just gubernatorial debates over time.

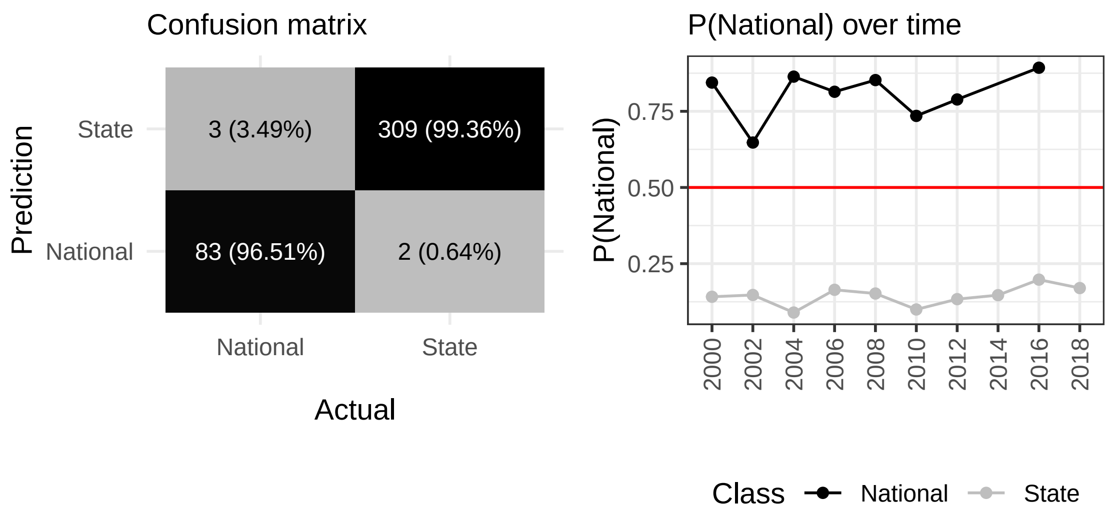


Figure 1: C-SPAN debate predictions, pooled and over time

These results give a consistent picture of rhetoric in debates; candidates predominately discuss topics germane to their jurisdictions. In purely statistical terms, collections of words more indicative of state (national) content are significantly more common in gubernatorial (presidential) debates, and the prevalence of those topics has not changed significantly over time.

It is important to note that these results, particularly the predicted probability of national origin, are not meant to indicate the exact proportion of content in a debate that reflects state or national topics. The debates themselves *are* quantitatively represented as different proportions of topics, but the weights with which those topic proportions are translated into predictions of class are not uniform. An alternative method of quantitatively representing texts may break debates into more granular pieces (such as sentences), classify each individual sentence as either state or national, then tally the number of sentences in each predicted class. The advantage of the method I use in this paper is the overall sense of content for each document; in general, are the topics discussed in the document more consistent with documents of known origin?

For example, consider the 2002 New Mexican gubernatorial debate between Republican John Sanchez, Democrat Bill Richardson, and Green Party candidate David Bacon. The classification model gave this debate a predicted national probability of 16.9%, but this of course masks the full heterogeneity of topics covered during the debate, which touched on national topics such as NAFTA and the war in Iraq. Figure 2 gives a more detailed view of the debate as a treemap of estimated topic proportions, with the size of each tile (labeled with the topic number) representing the size of the proportion and the color representing the model estimate of its relative “state-ness” or “national-ness” (and coefficients that weren’t significant predictors of either). Here, topics that lean “national” make up a larger proportion of the total debate than the 16.9% total suggested by the predicted probability, but the predictive weight of the state topics lowers shifts the prediction to the state side.

TV advertisements

Next, I consider the potential nationalization of rhetoric in televised campaign advertisements. This medium is perhaps the modal form of campaigning in the eyes of constituents and the most commonly studied campaign messaging medium in political science. It does, however, present some unique challenges to the classification methodology utilized in this paper.

Most obviously, the quasi built-in controls for candidate ideology and campaign resources in the debate context are absent from televised ads. Candidates with larger war chests may be able to air more ads referencing a broader range of topics, whereas more cash-restricted

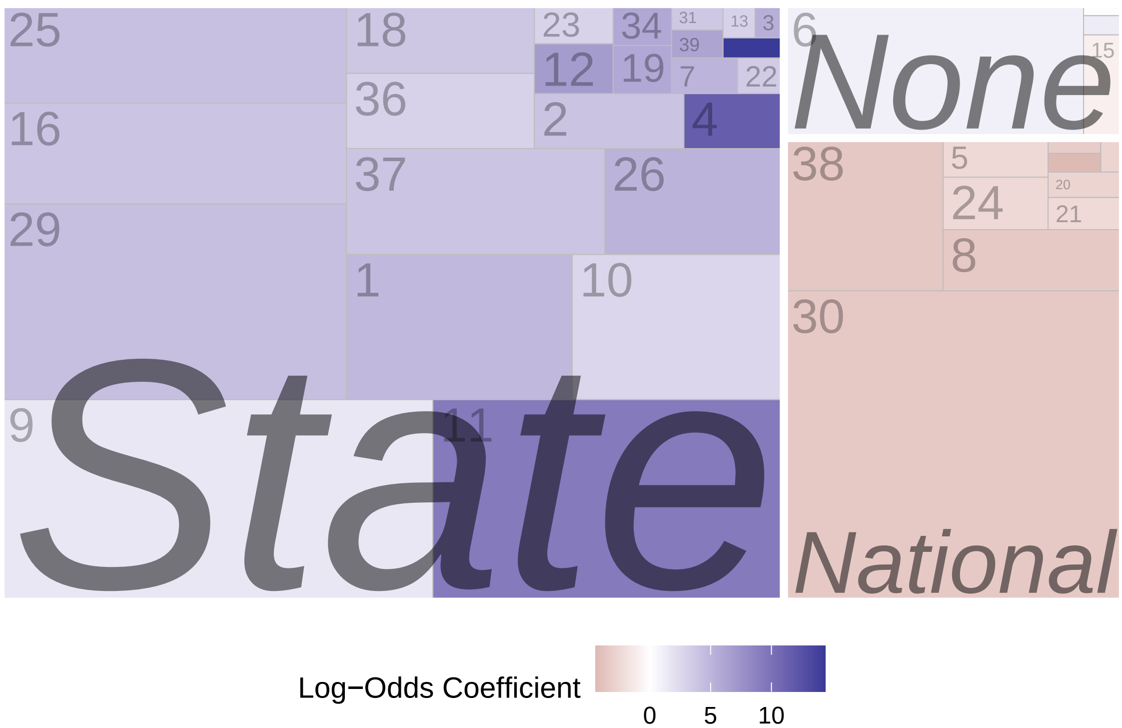


Figure 2: Topic proportions of 2002 New Mexico gubernatorial debate

candidates could be forced to focus their message around just a few talking points. A related problem is the unique content of ads overall; they are much shorter than debates, tend to be more negative, and, while they still speak predominantly about policy, they do often incorporate references to general candidate character or background. Furthermore, candidates increasingly have the ability to target advertisements to particular audiences for particular purposes. Certain messages may be broadcast to swing voters as persuasive content, while other messages may be broadcast to candidates' bases to turn out the vote. This is all to say that the content of advertisements is likely substantially different from the content of debates, which has consequences for the potential for nationalized campaigning strategies.

For this paper, I analyze 2,334 televised advertisements from presidential (1,528 ads) and gubernatorial (806) campaigns in 2004 and 2008. These ads are provided by the Wisconsin media project, with transcripts scraped from PDF storyboards.[3] These ads include those run in both the primaries and general elections and by both candidates and interest groups.

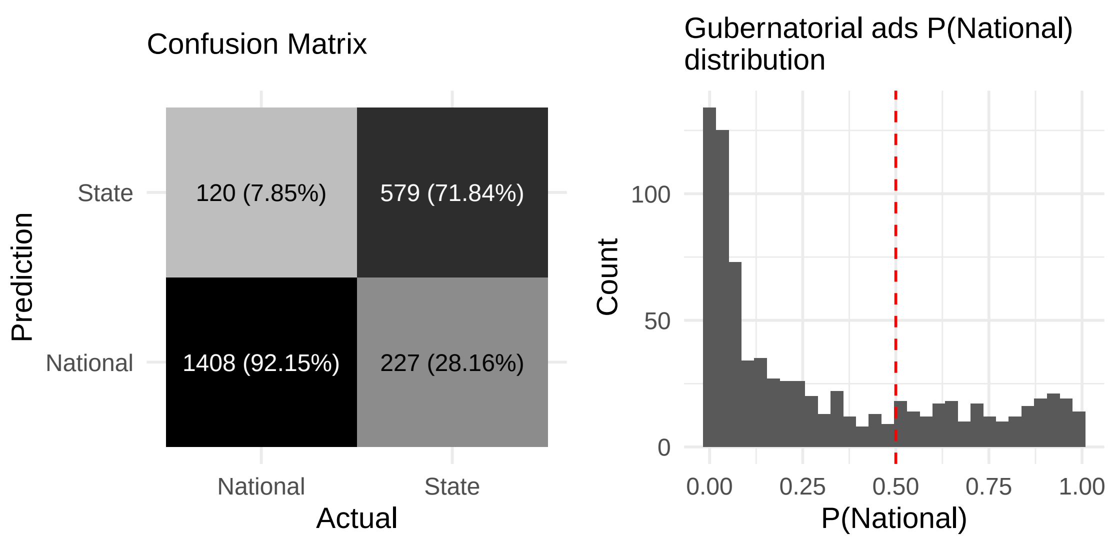


Figure 3: Televised advertisements predictions

[3]Only the data from 2004 and 2008 are presented in PDF storyboards with embedded text. While storyboards are available in 2000 and 2002, the text can only be extracted via more advanced techniques such as optical character recognition, which is a potential source of error. Years after 2008 do not included textual data.

Figure 3 shows the results of applying the trained classification model to these televised advertisements. While advertising content still predominately consists of topics germane to candidates' jurisdictions, there is a higher proportion (26.9%) of gubernatorial ads that are classified as being national. The left panel shows a similar classification rate of presidential rhetoric as in the debate context, which indicates the televised ad medium is not necessarily biased toward more state-like content. The right panel of figure 3 shows the distribution of the predicted probabilities of national classification for just the gubernatorial ads. Unsurprisingly, most of the predictions are strongly in the state directly, but a large number are classified as very national in content.

To give an example of one such gubernatorial ad, the Alliance for North Carolina ran an attack ad on Pat McCrory in October 2008 that the classification model assigned an 87% probability of being national in content. The brief transcript reads:

> The big developers, energy companies, and the banking industry just love Pat McCrory and George Bush. Why? Because McCrory and Bush have the same economic philosophy. Less regulation and less oversight to help these companies make even more profit. The result, economic collapse and a Wall Street Bailout. Who ends up paying? You the middle-class. Pat McCrory, stop supporting Bush economics and start supporting more regulation and oversight of big business.

This ad clearly attempts to link McCrory to Bush policies with fairly little state-specific information, instead using terms that would be equally applicable in any other state ("Wall Street bailout" and "middle-class"). While this is a compelling example, a majority of advertisements are still classified as being predominantly state content.

Twitter

Finally, I analyze a more modern form of campaign rhetoric; social media. Specifically, I analyze the Twitter of Members of Congress and Governors in office during 2018 using the

Das et al. (2022) corpus of tweets. This corpus contains 952,425 tweets from sitting Members of Congress and 101,546 tweets from incumbent governors.

This corpus is unique in this paper for many reasons. First, the tweets are not specific to the campaign timeframe, and therefore don't explicitly count as "campaigning." Second, the tweets only account for incumbents and do not include the tweets of their challengers. Third, the "national" comparison in this context is Members of Congress, not communication from Presidents. This is important because Members of Congress operate at the national stage while being beholden to district even more localized than their gubernatorial counterparts, so we might expect communication to be split between national and state topics. Finally, and perhaps most importantly, Twitter represents a fundamentally different avenue through which politicians communicate with supporters. Twitter can and is used to campaign, but can also be used for ostensibly non-political activity, like cheering on a local basketball team or engaging in more general political hobbyism. Outside of terms of service violations, there are really no restrictions on what can or can't be said on Twitter by politicians.

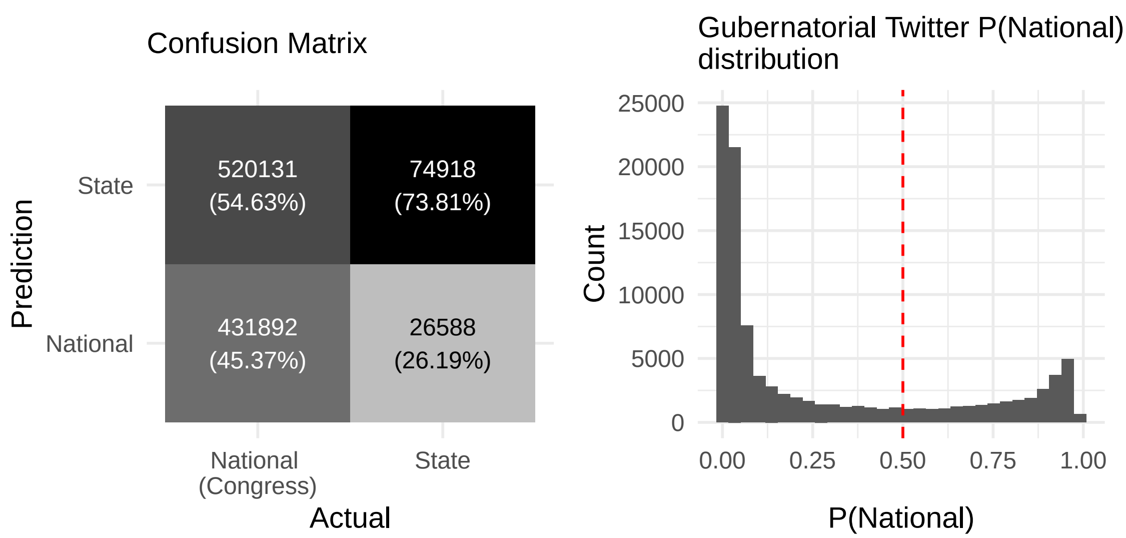


Figure 4: Twitter predictions

Figure 4 shows the results of the Twitter analysis. Similar to rhetoric in televised ads, governors still communicate more on topics related to state politics, but engage significantly in national political topics as well. Members of Congress are more evenly split between national and state topics, which is likely a function of their accountability to district-level pressures. The right panel of Figure 4 shows the distribution of predicted probabilities for national classification of gubernatorial tweets. Most tweets show strong "state-ness," but there does seem to be a slight bimodal distribution, with a large portion of tweets being classified as heavily national.

The results of figure 4 share similarities with those discussed by Das et al. (2022), but are difficult to directly compare. Das and coauthors' primary analysis uses a topic similarity approach, wherein 100 different topic models are trained and the resulting distributions of topic proportions can be compared using a distance metric. While the authors focus on the difference between national and mayoral Twitter, they do present results comparing Governors and Members of Congress. They find the median topic distance between Governors and Members of Congress is about 14% greater than the median distance between Governors and other Governors, indicating a small but not insignificant difference in the topics discussed by the two sets of politicians. Similarly, I find the "national" and "state" classification rates between Governors and Members of Congress differ by about 19%. Again, these numbers cannot be directly compared, but the general trend suggests some similarity but not complete overlap between the two levels of government.

Discussion

The preceding results indicate an overall picture of gubernatorial campaign rhetoric that gives pause to the "all politics is national" hypothesis. Across all mediums, a comfortable majority of communications were classified as primarily consisting of state topics. This approached almost 100% in debates, but closer to 75% in televised advertisements and social media posts on Twitter. The lower state classification rates in the latter two mediums suggest gubernatorial candidates do engage in some degree of nationalization when they have the flexibility to do so. The debate context is fairly constrained, so when those constraints are lifted and the field of possible topics expands beyond topics presented by a debate moderator, we would expect to see somewhat higher rates of nationalization.

While this paper has focused largely on the classification of single pieces of communication in isolation, it is possible that the real engine of information nationalization is the media environment reporting on, circulating, and commenting on the communications. Media plays a major role in how voters engage with campaign materials. It is possible that while most campaign messaging from gubernatorial candidates focuses on state topics, the few communications that *are* nationalized are circulated more widely by the media. During the gubernatorial race in Kentucky in 2019, for example, coverage from national outlets like the New York Times largely focused on the Republican incumbent's (Matt Bevin) allegiance with and affinity for Donald Trump. In the debate between the Bevin and his Democratic challenger Andy Beshear, the topic of impeachment *do* arise, but it was constrained to a single question. For the most part, the rest of the debate revolved around issues germane to Kentucky politics.

Further work must be done to both extend the corpus of text analyzed for nationalized content and determine if there exists a relationship between nationalized rhetoric and nationalized results. While I address a wide variety of mediums here, campaigns devote resources to many more. These include radio, flyers, websites, and other social media outlets such as Facebook and TikTok. Linking nationalized rhetoric to the nationalization of results is beyond the scope of this paper, but future work should consider the media markets in which differentially nationalized content is utilized by campaigns. Finally, future work should consider "downstream" nationalization occurring through media coverage of the campaigns. While the initial campaign rhetoric may be fairly germane to state topics, state-level media outlets may focus significantly more on the nationalized aspects of those campaigns.

The results of this paper have consequences for how we understand voter interaction with campaigns. These campaign activities *do* offer a source of information to voters that is functionally distinct from national politics. How voters then process this information is of subsequent importance to better understand the information environment voters must navigate within a nationalized context.

References

Abernathy, Penelope Muse. 2018. *The Expanding News Desert*. Center for Innovation; Sustainability in Local Media, School of Media; Journalism, University of North Carolina at Chapel Hill.

Abramowitz, Alan I., and Steven Webster. 2016. "The Rise of Negative Partisanship and the Nationalization of U.S. Elections in the 21st Century." *Electoral Studies* 41: 12–22.

Arceneaux, Kevin. 2006. "The Federal Face of Voting: Are Elected Officials Held Accountable for the Functions Relevant to Their Office?" *Political Psychology* 27(5): 731–54.

Benjamin, Daniel, and Jesse Shapiro. 2009. "Thin-Slice Forecasts of Gubernatorial Elections." *The Review of Economics and Statistics* 91(3): 523–36.

Benoit, William L., LeAnn M. Brazeal, and David Airne. 2007. "A Functional Analysis of Televised U.S. Senate and Gubernatorial Campaign Debates." *Argumentation and Advocacy* 44(2): 75–89.

Benoit, William L., Glenn J. Hansen, and Rebecca M. Verser. 2003. "A Meta-Analysis of the Effects of Viewing U.S. Presidential Debates." *Communication Monographs* 70(4): 335–50.

Bossetta, Michael. 2018. "The Digital Architectures of Social Media: Comparing Political Campaigning on Facebook, Twitter, Instagram, and Snapchat in the 2016 U.S. Election." *Journalism & Mass Communication Quarterly* 95(2): 471–96.

Brown, Adam R. 2010. "Are Governors Responsible for the State Economy? Partisanship, Blame, and Divided Federalism." *The Journal of Politics* 72(3): 605–15.

Brunner, Jim. 2016. "Gov. Jay Inslee Defeats Republican Challenger Bill Bryant." *The Seattle Times*.

Catanese, David. 2011. "Tomblin Wins W.Va. Gov Race." *Politico*.

Caughey, Devin, James Dunham, and Christopher Warshaw. 2018. "The Ideological Nationalization of Partisan Subconstituencies in the American States." *Public Choice* 176(1): 133–51.

Caughey, Devin, and Christopher Warshaw. 2016. "The Dynamics of State Policy Liberalism, 1936–2014." *American Journal of Political Science* 60(4): 899–913.

Costa, Mia. 2021. "Ideology, Not Affect: What Americans Want from Political Representation." *American Journal of Political Science* 65(2): 342–58.

Das, Sanmay, Betsy Sinclair, Steven W. Webster, and Hao Yan. 2022. "All (Mayoral) Politics Is Local?" *The Journal of Politics* 84(2): 1021–34.

de Benedictis-Kessner, Justin, and Christopher Warshaw. 2020. "Accountability for the Local Economy at All Levels of Government in United States Elections." *American Political Science Review* 114(3): 660–76.

Hayes, Danny, and Jennifer L. Lawless. 2018. "The Decline of Local News and Its Effects: New Evidence from Longitudinal Data." *The Journal of Politics* 80(1): 332–36.

Hillygus, D. Sunshine, and Simon Jackman. 2003. "Voter Decision Making in Election 2000: Campaign Effects, Partisan Activation, and the Clinton Legacy." *American Journal of Political Science* 47(4): 583–96.

Hopkins, Daniel J. 2018. *The Increasingly United States: How and Why American Political Behavior Nationalized*. University of Chicago Press.

Huddy, Leonie, and Alexa Bankert. 2017. "Political Partisanship as a Social Identity." *Oxford Research Encyclopedia of Politics*.

Iyengar, Shanto, Gaurav Sood, and Yphtach Lelkes. 2012. "Affect, Not Ideology: A Social Identity Perspective on Polarization." *Public Opinion Quarterly* 76(3): 405–31.

Jacobson, Gary C. 2015. "It's Nothing Personal: The Decline of the Incumbency Advantage in US House Elections." *The Journal of Politics* 77(3): 861–73.

Manchester, Julia. 2019. "GOP Seeks to Nationalize Gubernatorial Elections." *The Hill*.

Martin, Gregory J., and Joshua McCrain. 2019. "Local News and National Politics." *American Political Science Review* 113(2): 372–84.

Martin, Jonathan. 2019. "Kentucky Governor's Race Tests Impact of Impeachment in States." *The New York Times*.

McKinley Jr., James C. 2010. "Perry Re-Elected in Texas Governor Race." *The New York Times*.

Moskowitz, Daniel J. 2021. "Local News, Information, and the Nationalization of U.S. Elections." *American Political Science Review* 115(1): 114–29.

Mummolo, Jonathan, Erik Peterson, and Sean Westwood. 2021. "The Limits of Partisan Loyalty." *Political Behavior* 43(3): 949–72.

Owen, Diana. 2014. "New Media and Political Campaigns."

Reckhow, Sarah, Jeffrey R. Henig, Rebecca Jacobsen, and Jamie Alter Litt. 2017. ""Outsiders with Deep Pockets": The Nationalization of Local School Board Elections." *Urban Affairs Review* 53(5): 783–811.

Roberts, Margaret E., Brandon M. Stewart, and Edoardo M. Airoldi. 2016. "A Model of Text for Experimentation in the Social Sciences." *Journal of the American Statistical Association* 111(515): 988–1003.

Rojas, Rick, and Jeremy Alford. 2019. "In Louisiana, a Narrow Win for John Bel Edwards and a Hard Loss for Trump." *The New York Times*.

Shor, Boris, and Nolan McCarty. 2011. "The Ideological Mapping of American Legislatures." *American Political Science Review* 105(3): 530–51.

Sievert, Joel, and Seth C. McKee. 2019. "Nationalization in U.S. Senate and Gubernatorial Elections." *American Politics Research* 47(5): 1055–80.

Stier, Sebastian, Arnim Bleier, Haiko Lietz, and Markus Strohmaier. 2018. "Election Campaigning on Social Media: Politicians, Audiences, and the Mediation of Political Communication on Facebook and Twitter." *Political Communication* 35(1): 50–74.

Webster, Steven W. 2020. *American Rage: How Anger Shapes Our Politics*. Cambridge: Cambridge University Press.

Weinschenk, Aaron et al. 2020. "Have State Supreme Court Elections Nationalized?" *Justice System Journal* 41(4): 313–22.

Weinschenk, Aaron. 2022. "The Nationalization of School Superintendent Elections." *Social Science Quarterly* 103(3): 597–606.

Online Appendix for Nationalized Elections, Localized Campaigns?

A Supervised Machine Learning Approach to Measuring Nationalized Political Rhetoric

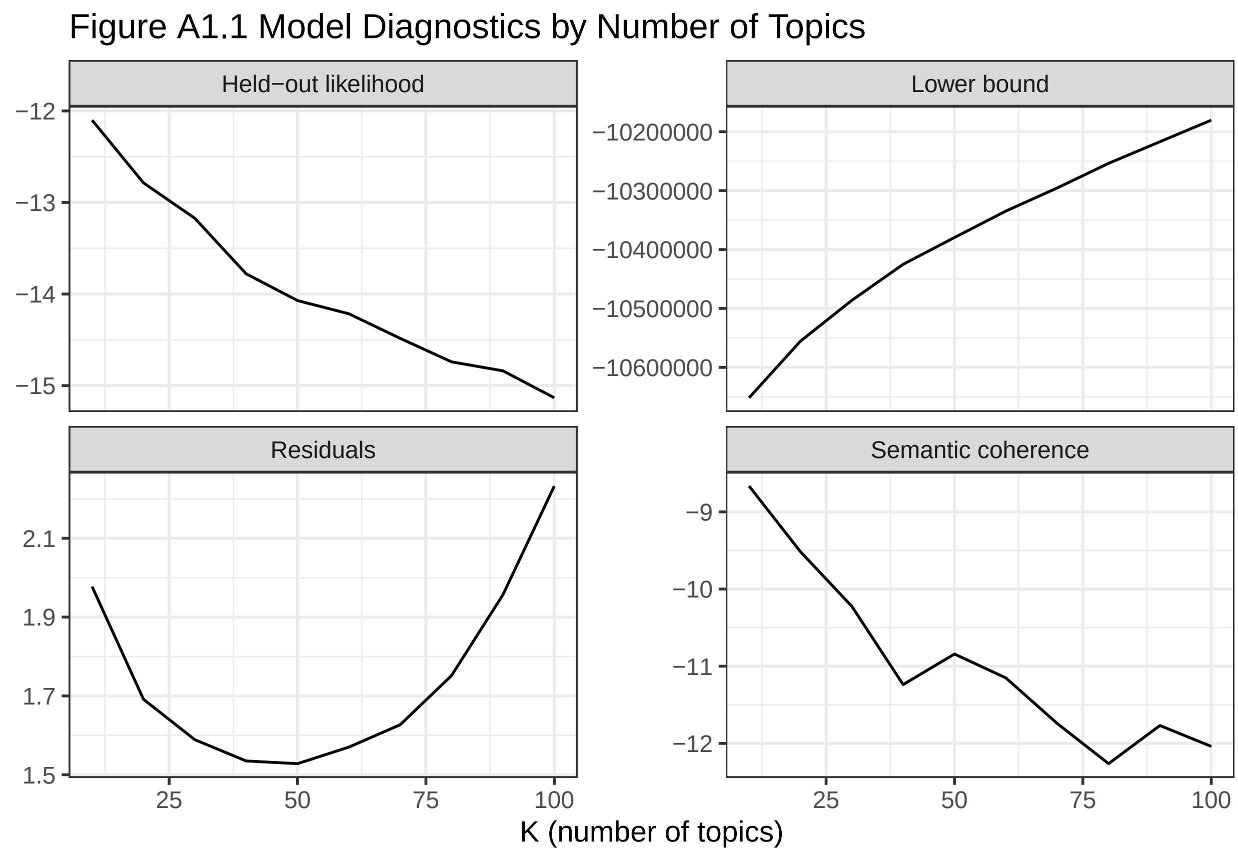
Derek Holliday*

2022-09-09

*Ph.D. Candidate, UCLA, deholliday@g.ucla.edu

Appendix A1: K Selection

One researcher-defined parameter in structural topic modeling is the number of topics k. There is no set rule for selecting k, but researchers should ideally balance a number of factors, including the likelihood, residuals, and semantic coherence. Increasing/decreasing k will mechanically tend to improve some factors but detract from others. Figure A1.1 shows a number of model diagnostics, including the held-out likelihood, lower bound of hte marginal log-likelihood, residuals, and mean semantic coherence. Semantic coherence gives a measure of word co-occurrence; when co-occurring words are also highly likely words within a topic, semantic coherence is high.



The diagnostics suggest a topic number between 40 and 50 would be a good balance of fit and coherence. I use $k = 40$ for the models presented in the paper so as to lean against potential overfitting with higher numbers of topics.

One may be concerned different topic numbers yield different results. To alleviate this concern, I perform the same analysis presented in Table 1 in the tables below, varying k. The table shows the accuracy of the models in predicting the held-out training documents. Across all values of k, the results remain remarkably similar.

More broadly, because I do not attempt to assign post-hoc meaning to the content of certain topics (for example, interpreting one topic as the "education" topic and another as the "military" topic), my particular approach to measuring nationalization is less vulnerable to k misspecification. The topics in my methodological approach are simply collections of words that co-occur more or less frequently in state or national contexts. How those topics are sliced is less of a concern than how predictive the topics are of state or national content as a sort of dimension-reduced representation of the text itself.

Table 1: Classification Model Performance on Heldout Documents - Varying K

K	Logistic Regression (Nonpenalized)	Naive Bayes	Penalized Logistic Regresssion (Lasso)	Boosted Gradient Descent (XGBoost)	Support Vector Machine
10	0.990	0.981	0.981	0.990	0.990
20	0.990	0.928	0.995	0.942	0.990
30	0.995	0.976	0.995	0.971	0.995
40	0.990	0.952	0.986	0.928	0.995
50	0.986	0.942	0.986	0.918	0.986
60	0.995	0.933	0.976	0.909	0.995
70	0.990	0.938	0.981	0.918	0.990
80	0.995	0.938	0.981	0.947	0.995
90	0.990	0.933	1.000	0.865	0.990
100	0.990	0.875	0.966	0.962	0.990

Appendix A2: Complete Model Results

While I use unpenalized logistic regression for the main model of the paper, other classification model approach yield fairly similar results. I present the results in the tables below. Generally, the worst-performing model is boosted gradient descent, which tends to overpredict the prevalence of state content.

Table 2: Classification Model Performance on C-SPAN Debates

	National Documents			State Documents		
Model	Correct	Incorrect	Accuracy	Correct	Incorrect	Accuracy
Logistic Regression (Nonpenalized)	83	3	0.965	309	2	0.994
Naive Bayes	83	3	0.965	259	52	0.833
Penalized Logistic Regression (Lasso)	84	2	0.977	304	7	0.977
Boosted Gradient Descent (XGBoost)	56	30	0.651	311	0	1.000
Support Vector Machine	84	2	0.977	306	5	0.984

Table 3: Classification Model Performance on TV Ads

	National Documents			State Documents		
Model	Correct	Incorrect	Accuracy	Correct	Incorrect	Accuracy
Logistic Regression (Nonpenalized)	1408	120	0.921	579	227	0.718
Naive Bayes	1196	332	0.783	631	175	0.783
Penalized Logistic Regression (Lasso)	1413	115	0.925	579	227	0.718
Boosted Gradient Descent (XGBoost)	803	725	0.526	786	20	0.975
Support Vector Machine	1414	114	0.925	575	231	0.713

Table 4: Classification Model Performance on Twitter

Model	National Documents			State Documents		
	Correct	Incorrect	Accuracy	Correct	Incorrect	Accuracy
Logistic Regression (Nonpenalized)	431892	520131	0.454	74918	26588	0.738
Naive Bayes	164250	787773	0.173	93404	8102	0.920
Penalized Logistic Regression (Lasso)	398934	553089	0.419	78205	23301	0.770
Boosted Gradient Descent (XGBoost)	32378	919645	0.034	100478	1028	0.990
Support Vector Machine	443220	508803	0.466	73934	27572	0.728

Appendix A3: Custom Stopwords

One assumption of my methodological approach is the classification model is distinguishing between meaningful state and national textual content. A concern related to this assumption would be the ability of the model to "cheat;" that is, classifying based on words that aren't necessarily meaningful indicators of state or national topics or policies. Such words could include state names, the names of state capitols, nicknames for state residents, or common phrases spoken during speeches in particular contexts. The topic modeling approach somewhat alleviates this concern; using collections of words instead of individual words as predictors softens the impact of any one outlier word. However, I also remove the following stopwords from the entire corpus of text:

[1] "alabama, alaska, arizona, arkansas, california, colorado, connecticut, delaware, florida, georgia, hawaii, hawai, hawai'i, idaho, illinois, indiana, iowa, kansas, kentucky, louisiana, maine, maryland, massachusetts, michigan, minnesota, mississippi, missouri, montana, nebraska, nevada, new, hampshire, jersey, mexico, york, north, carolina, dakota, ohio, oklahoma, oregon, pennsylvania, rhode, island, tennessee, texas, utah, vermont, virginia, washington, wisconsin, wyoming, house, district, representative, senate, senator, senatorial, governor, gubernatorial, governorship, question, crosstalk, applause, laughter, moderator, candidate, gentlemen, ladies, congressman, state, will, alaskan, thank, year, rebuttal, speaker, lieutenant, upstate, downstate, make, must, assemblyman, assemblywoman, assemblyperson, want, need, inaudible, commonwealth, aloha, isn, ve, didn, don, 1998, 1999, 2000, 2001, 2002, 2003, 2004, 2005, 2006, 2007, 2008, 2009, 2010, 2011, 2012, 2013, 2014, 2015, 2016, 2017, 2018, 2019, 2020, alabamians, alaskan, arizonan, arkansan, californian, coloradan, connecticuter, delawarean, washingtonian, floridian, georgian, hawaiian, idahoan, illinoisan, indianian, iowan, kansan, kentuckian, louisianian, mainer, marylander, massachusettsan, michiganian, minnesotan, mississippian, missourian, montanan, nebraskan, nevadan, hampshirites, jerseyan, mexican, yorker, ohioan, oklahoman, oregonian, pennsylvanian, islander, carolinian, dakotan, tennessean, texan, utahn, vermonter, virginian, washingtonian, west virginian, wisconsinite, wyomingite, alabama, arizonian, arkie, californiac, coloradoan, connecticotian, muskrat, floridan, malihini, idahoer, illinoian, indianer, hawkeye, kanser, kentucker, louisianan, mainiac, marylandian, michigander, mississipper, cornhusker, nevadian, hampshireman, jerseyite, tar, heel, nodak, buckeye, oklahomians, oregoner, pennamite, rhodian, south carolinan, volunteer, texian, utahan, toner, mountaineer, cheesehead, wyomingian, bamer, inuit, arkansawyer, californio, connecticutian, kamaaina, illinoyer, hoosier, iowegian, jayhawk, kentuckeyite, cajun, massachusite, michiganese, jerseyite, okie, sandlapper, big bender, tejano, badger, wyoman, aleut, nutmegger, islander, illini, michigine, haida, michiganite, inupiaq, wolverine, aleutian, inuk, alabamianss, alaskans, arizonans, arkansans, californians, coloradans, connecticuters, delawareans, washingtonians, floridians, georgians, hawaiians, idahoans, illinoisans, indianians, iowans, kansans, kentuckians, louisianians, mainers, marylanders, massachusettsans, michiganians, minnesotans, mississippians, missourians, montanans, nebraskans, nevadans, hampshiritess, jerseyans, mexicans, yorkers, ohioans, oklahomans, oregonians, pennsylvanians, islanders, carolinians, dakotans, tennesseans, texans, utahns, vermonters, virginians, washingtonians, west virginians, wisconsinites, wyomingites, alabamas, arizonians, arkies, californiacs, coloradoans, connecticotians, muskrats, floridans, malihinis, idahoers, illinoians, indianers, hawkeyes, kansers, kentuckers, louisianans, mainiacs, marylandians, michiganders, mississippers, cornhuskers, nevadians, hampshiremans, jerseyites, tars, heels, nodaks, buckeyes, oklahomianss, oregoners, pennamites, rhodians, south carolinans, volunteers, texians, utahans, toners, mountaineers, cheeseheads,

wyomingians, bamers, inuits, arkansawyers, californios, connecticutians, kamaainas, illinoyers, hoosiers, iowegians, jayhawks, kentuckeyites, cajuns, massachusites, michiganeses, jerseyites, okies, sandlappers, big benders, tejanos, badgers, wyomans, aleuts, nutmeggers, islanders, illinis, michigines, haidas, michiganites, inupiaqs, wolverines, aleutians, inuks, montgomery, juneau, phoenix, little rock, sacramento, denver, hartford, district of columbia, dover, tallahassee, atlanta, honolulu, boise, springfield, indianapolis, des moines, topeka, frankfort, baton rouge, augusta, annapolis, boston, lansing, st paul, jackson, jefferson city, helena, lincoln, carson city, concord, trenton, santa fe, albany, raleigh, bismarck, columbus, oklahoma city, salem, harrisburg, san juan, providence, columbia, pierre, nashville, austin, district of columbia, salt lake city, montpelier, richmond, olympia, charleston, madison, cheyenne”