Supplementary Information: Legislative Communication and Power: Measuring Leadership in the U.S. House of Representatives from Social Media Data

February 14, 2024

1 Introduction

In the following pages we provide technical details about the important steps in our paper's methodology: summary statistics and visualizations of our Twitter data; technical details and sensitivity analyses for our topic modeling; information useful for understanding the sensitivity of our PCA modeling decisions; summary statistics from our network modeling; and finally, details and sensitivity analysis of our dynamic analysis.

Upon publication, we will make our code and documentation available, along with a great deal of additional material that readers can use to examine our modeling decisions and the robustness of our results to those decisions, including detailed log files and estimation details. We will also provide our raw data, subject to Twitter's current policies about data sharing.

2 Twitter as a Congressional Communications Platform

2.1 The Data

In order to study leaders' influence in congressional communications, we collected the Twitter handles of 440 representatives from June 29th, 2019 to March 23, 2020 based on the official Twitter handles list[1] collected by C-SPAN.[2] Due to the limitations imposed by the public Twitter API, we were only able to get at most 3,200 of the latest tweets for each member of Congress at the time of first collection. However, we have captured all subsequent tweets since we began collecting data. In the end, we collected 1,252,505 tweets, including original posts and re-tweets. This data is high-frequency text data, which we exploit to study the dynamics of communication – using this granular data, we can find evidence of whether the party rank and file anticipate their leaders' communications on social media of vice versa. Further, this data can be used to detect communication relations in Congress, by using retweets to construct a network structure. This network structure can be used to uncover measures of influence using classical empirical methods.

For our sentiment-topic analysis, we focus on 315,000 tweets from all Democratic and Republican House of Representatives Twitter accounts in the 116th Congress in the calendar year 2019 and until March 24, 2020. Our analysis ends before the onset of the COVID-19 pandemic. For sentiment-topic analysis, we exclude retweets and quote tweets (since the legislator is not amplifying their own message, but potentially engaging with messages with which they disagree). Further, we do not want to create mechanical correlations between our sentiment-topic derived policy stance scores and leadership centrality scores, the latter of which are constructed from quote tweets and retweets.

For users' retweet posts, we simply identified the *source* users and *receiver* users. For users' original Twitter posts, we pre-processed each post text with the following procedures to prepare the data consistent with the theoretical model (Grimmer and Stewart, 2013; Denny and Spirling, 2018):

1. We excluded all non-English words; numerals, and non-standard symbols such as emojis.

[1] We did not include election or personal accounts in our datasets.

[2] https://twitter.com/cspan/lists/members-of-congress/members

2. We excluded names of all the legislators, and some common procedural words such as amendment and bill, as these words can mask the underlying policy stance of the tweet (such as when a bill relates to health care or taxes).

3. We lemmatized all remaining words with the WordNet Lemmatizer[3] in NLTK package Bird, Klein and Loper (2009);

4. We detect n-grams in the data, as we observe that politicians tend to used two or three syllable phases (for example, cap-and-trade, catch-and-release, and law and order).

2.2 The Distribution of Congressional Tweeting Behavior

In this section we provide summary statistics on the Twitter activity of the Members of the U.S. House of Representatives, during the time period covered in our study. Table SI 1 gives summary statistics for the entire dataset, by party. In Figure SI 1 we show the data on tweets by member in a histogram.

Table SI 1: Distribution of Tweeting Behavior: Entire Dataset

Party	Mean	Median	Minimum	Maximum	Standard Deviation
Democratic Party	894.45	797	43	3,200	520.46
Republican Party	528.31	457	11	2,732	417.05
All	727.17	597	11	3,200	509.33

[3] https://www.nltk.org/_modules/nltk/stem/wordnet.html

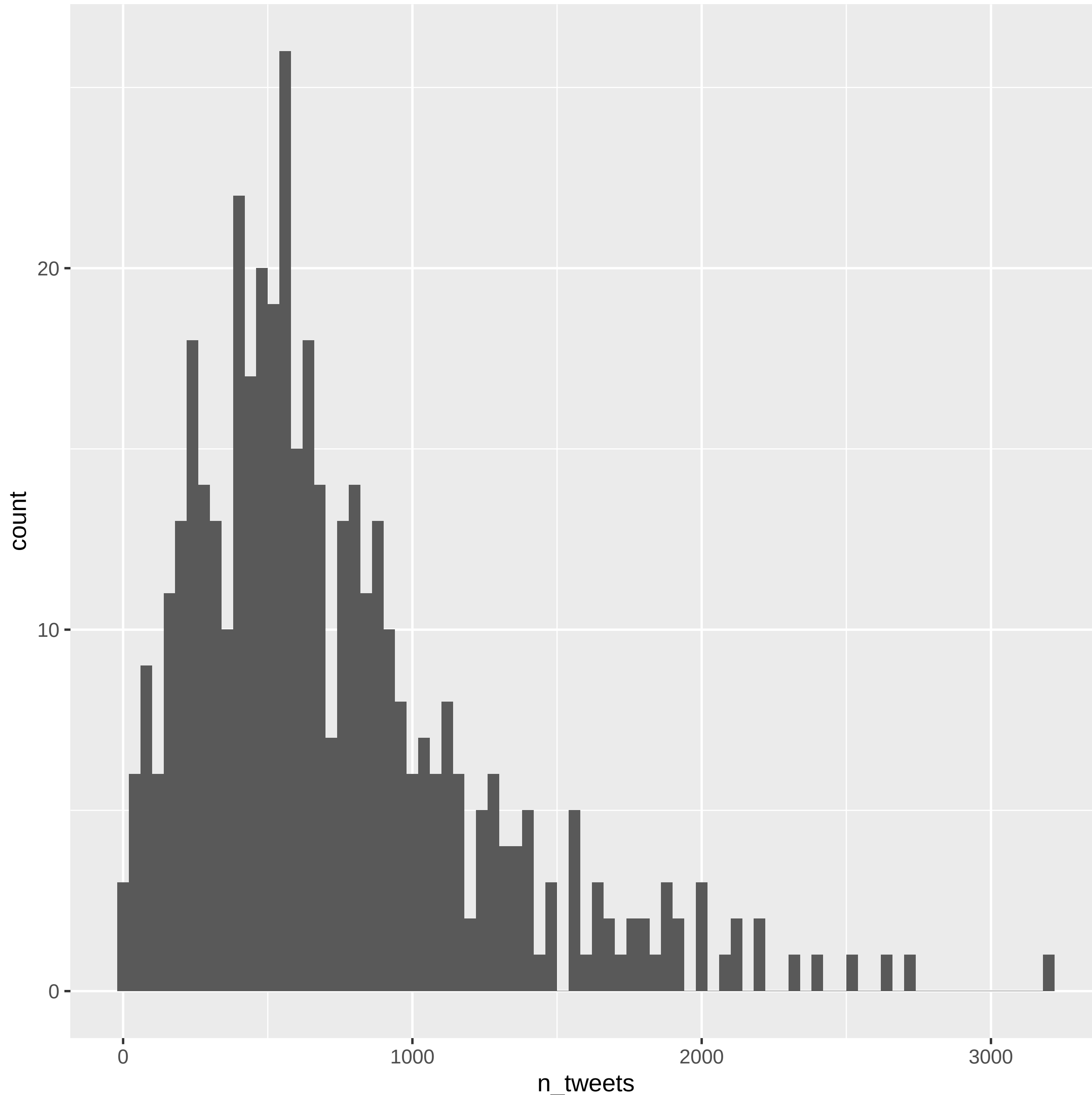


Figure SI 1: Distribution of Tweets by Member

3 Theory

3.1 Game Setting and Example

Here we summarize the model setting. In the next section of the Supplementary Information we provide intuition for how the model fits our setting using the 2019 government shutdown as an example. In this model, there are n party rank-and-file members who are deciding to advocate either policy stance A or B. The optimal policy choice depends on a state variable, θ. The state is the underlying political situation. It represents the party mood regarding an unexpected politically sensitive issue. Finally, members receive private signals m_i about the true state of the world, which are normally distributed.

In order to coordinate on a policy, a policy must have a sufficient threshold of support, p_A and p_B for policies A and B respectively. Conceptually, this is the informal level of consensus needed for the party to advocate a platform. Then, x is the number of party rank-and-file advocating policy A. Party members earn the following payoffs depending on their choice of policy stance and on the underlying state θ and support for policy A, x :

$$\begin{cases} u_A(\theta) = \exp\{\frac{\lambda\theta}{2}\} & \text{if } \frac{x}{n} > p_A\text{, adopt policy } A \\ u_B(\theta) = \exp\{-\frac{\lambda\theta}{2}\} & \text{if } p_B > \frac{x}{n}\text{, adopt policy } B \\ u_A = u_B = 0 & \text{if } p_A \geq \frac{x}{n} \geq p_B\text{, coordination failure} \end{cases} \tag{1}$$

Dewan and Myatt (2007) assume legislators play a threshold strategy and that they vote for policy stance A instead of the status quo, B, if and only if their private signal $m_i > m$ for some threshold m. They assume this private signal is distributed normally with mean θ and variance $\frac{1}{\psi}$. In the payoff structure, the sensitivity to the benefits of coordinating (electoral success, the continuation of good public policy) are captured by λ, the party's *need for direction.* This concept represents the importance of choosing the right messaging strategy and the gravity of choosing incorrectly. Conditional on state of the world θ, party rank-and-file advocate for A with probability $p = \Pr[m_i > m|\theta]$, which is distributed normally with standard normal CDF Φ by the distributional assumption on the signal m_i. The authors note that as n increases, $\frac{x}{n}$ approaches p by the Law of Large Numbers. The authors then note that assuming large n, policy A succeeds if $p > p_a$. Given the normality assumption on m_i, this condition is equivalent to $\theta > \theta_A$ where θ_A satisfies $p = \Phi[\sqrt{\psi}(\theta_A - m)]$. Similarly, the party adopts policy B if $\theta_B > \theta$ where θ_B satisfies $p = \Phi[\sqrt{\psi}(\theta_B - m)]$ This results in the following outcome structure:

$$\text{Outcome} \quad = \quad \begin{cases} \text{Coordinate on } A & \text{if } \theta > \theta_A \\ \text{Coordinate on } B & \text{if } \theta_B > \theta \\ \text{Coordination failure} & \text{if } \theta_A \geq \theta \geq \theta_B \end{cases} \tag{2}$$

$$\text{where} \qquad \begin{cases} \theta_A = m + \frac{\pi_A}{\sqrt{\psi}} \\ \theta_B = m + \frac{\pi_B}{\sqrt{\psi}} \end{cases} \tag{3}$$

where substitutions $\pi_A = \Phi^{-1}(p_A)$ and $pi_B = \Phi^{-1}(1 - p_B)$ have been made for clarity. The authors note that conceptually, π_A and pi_B measure the heights of the *barriers to coordination*.

Given this setting, the game sequence proceeds as follows:

1. Rank-and-file members receive a private signal $m_i|\theta$ for i in $1, ..., n$ that is conditioned on the true state of the world distributed with variance $\frac{1}{\psi}$, the *sense of direction*.

2. Leaders of the party decide to give a speech or not relaying their signal to the party.

3. Rank-and-file members adopt a policy stance they individually decide to advocate.

4. If the critical thresholds of rank-and-file members advocate for the same policy stance (π_A and π_B), the party successfully coordinates. These thresholds are called *barriers to coordination*. Otherwise, the party fails to coordinate.

5. Borrowing terminology from Dewan and Myatt (2007), rank-and-file members are willing to follow their leaders' signals based on a leadership index R:

$$R = \frac{\text{Barriers to Coordination} \times \text{Sense of Direction}}{\text{Need for Direction}} \tag{4}$$

6. The equilibrium strategies are characterized by R, which makes the concept of leadership precise in our context: When $R > 1$, rank-and-file members adopt the same signal as their leaders. For $R < 1$, rank-and-file members adopt a threshold that is biased towards the leaders' preferred threshold, increasing in R. That is, as R approaches 1, rank-and-file member play strategies biased in favor of their leaders' preferred strategies.

In our case, we interpret the private signals m_i as a member's observation of the party's mood, which is derived from interpersonal conversation, social media stances from other party members, and party conference meetings and calls.[4] We interpret the leader's speech as the leadership of the parties tweeting out their talking points and messaging strategy to their members. We interpret the policy stances as the policy stances advocated on Twitter. In order to identify Dewan and Myatt (2007) we restrict the strategy space to what they consider a natural class of strategies, threshold strategies.

We interpret the policy stances on Twitter themselves as the the key strategic behavior. On Twitter, House party leadership and rank-and-file membership publicly and strategically communicate their policy stances. When R is high, we expect rank-and-file members to follow their leaders. When it is low, we expect rank-and-file members to be less likely to follow their leaders. Thus, the leadership index R suggests intuition for patterns of communication behavior we might expect. Using this intuition from this framework, we derive hypotheses regarding House party leadership behavior and the tendency of rank-and-file House members to follow their leaders.

[4] In order to link this theory to our empirical setting, we first note that House member Twitter accounts are managed both by staff and the legislator. We assume that the incentives of the congressional communication staff are aligned with the legislator they represent. Conversations with several House communication staffers suggest social media activity is coordinated at the office level under the direction of their principal.

4 Topic Analysis

In this section we discuss the details of the Joint Sentiment Topic model and our implementation. In the next section we provide technical details for the Joint Sentiment Topic model. The subsequent sections provides graphical material on the sensitivity of our results to modeling decisions.

4.1 The Joint Sentiment Topic Model

In this paper, we use a Joint Sentiment Topic (JST) model (Lin and He 2009) to obtain the topic diversity for members of the U.S. House of Representatives.Lin et. al (2012) describe their method as follows. Take a corpus of tweets C, which is a collection of D tweets $\{t_1, t_2, t_3, ..., t_D\}$. Each tweet itself is a collection of N_t words. Let the words in each tweet be denoted by $\{w_1, w_2, ..., w_{N_t}\}$. Now, each potential word in any tweet is indexed by a vocabulary, with V total terms $\{1, 2, 3, 4, ..., V\}$. Now, let J signify the total number of sentiment labels and L the total number of topics. Explicitly, the underlying data-generation process for the documents is summarized as follows:

1. For each sentiment label j in $\{1, 2, 3, ..., J\}$
   (a) For each topic k in $\{1, 2, 3, ..., L\}$ draw $\phi_{j,k}$ from $Dir(\lambda_j \times \beta^L_{j,k})$
2. For each tweet t, choose a distribution $\pi_t \sim Dir(\gamma)$
3. For each sentiment label j under tweet t, drawn a distribution $\theta_{j,k}$ $Dir(\alpha)$
4. For each word w_i in tweet t,
   (a) Draw sentiment j_i from Multinomial(π_t)
   (b) Draw topic label k_i from Multinomial(θ_{t,j_i}) which is conditioned on sampled sentiment j_i.
   (c) Draw word from per-corpus word distribution conditioned on sentiment label j_i and topic label k_i, i.e. choose a word from Multinomial(ϕ_{j_i,k_i}).

The hyperparameter α can be interpreted intuitively as the the prior observation counts for the number of times topic k associated with sentiment label j is sampled from a tweet. The hyperparameter β can be interpreted as the prior belief on the frequency at which words sampled from topic k are associated with sentiment label j, respectively, *ex ante*. Following this logic, λ can be treated as the prior belief on the number of times sentiment label j is sampled from a tweet before observing any tweets.

Observe that as β goes to 0, the model converges to a model of a single sentiment-topic. That is, one sentiment-topic label has probability 1, with all other labels being assigned 0. On the other hand, as β grows large, the limiting distribution is uniform over sentiment-topics. We expect that tweets, given their concise nature, are likely only to relate to very few topics at once, so we set these priors relatively small, following standard practice (such as in Lin and He (2009)).
[5]

[5]The model incorporates a prior over λ using a lexicon which suggests sentiment orientations for some 7000

4.1.1 Topic Selection

Table SI 2: Emblematic Tweets

handle	
@repdelbene	whats stake trillion dollar federal funding countless po
@reprokhanna	climate change isnt intergovernmental panel climate change report ef
@stevescalise	family small business dems held relief hostage day play politics tr
@repcuellar	homeland security questioned border patrol chief carla provost curre
@reparmstrongnd	finally chance ask ig horowitz question fbi investigation trump campaign fbi knew steele u

We select the number of topics based on the inflection point beyond which increases to coherence are small. Based on this criterion, we select 60 topics. To arrive at this number, we tuned the model starting from 5 topics and 10 topics increasing in increments of 10 up to 60 topics. Figure SI 2 shows that that topical coherence along an NPMI metric is maximized at 60 topics (which results in 180 Senti-Topics). Due to computational feasibility constraints, we can estimate at most 60 topics, but in addition to strong quantitative coherence, we show they have facial validity, as well.

Of course, there are many topic modeling approaches that could be used to reduce the dimensionality of the data. We use JST for two reasons. First, unlike the more widely known LDA model, JST estimates two layers, one for topics the other for sentiment. When it comes to political speech we believe this is superiod to methods that produce only estimates of the topic layer. Consider congressional debates on most important issues, like climate change. One side in a debate about some climate change policy will take the side endorsing the measure, and thus will engage in mostly positive rhetoric about the policy. The opposing side, though, will engage in mostly negative rhetoric about the policy. As we would like both negative and positive dimensions to be represented in our analysis, we use JST. Second, unlike topic modeling approaches like STM, the JST model imposes few restrictions on the sentiment-topic solution.

common words. For more details, see Lin and He (2009) and Lin et al. (2012). We use an R wrapper written around the authors' original `C++` code, found here: `https://github.com/linron84/JST` to estimate the model. We run the model for 1000 iterations after a burn-in of 1000. The model is computationally expensive, and it runs for about 9 hours prior before converging.

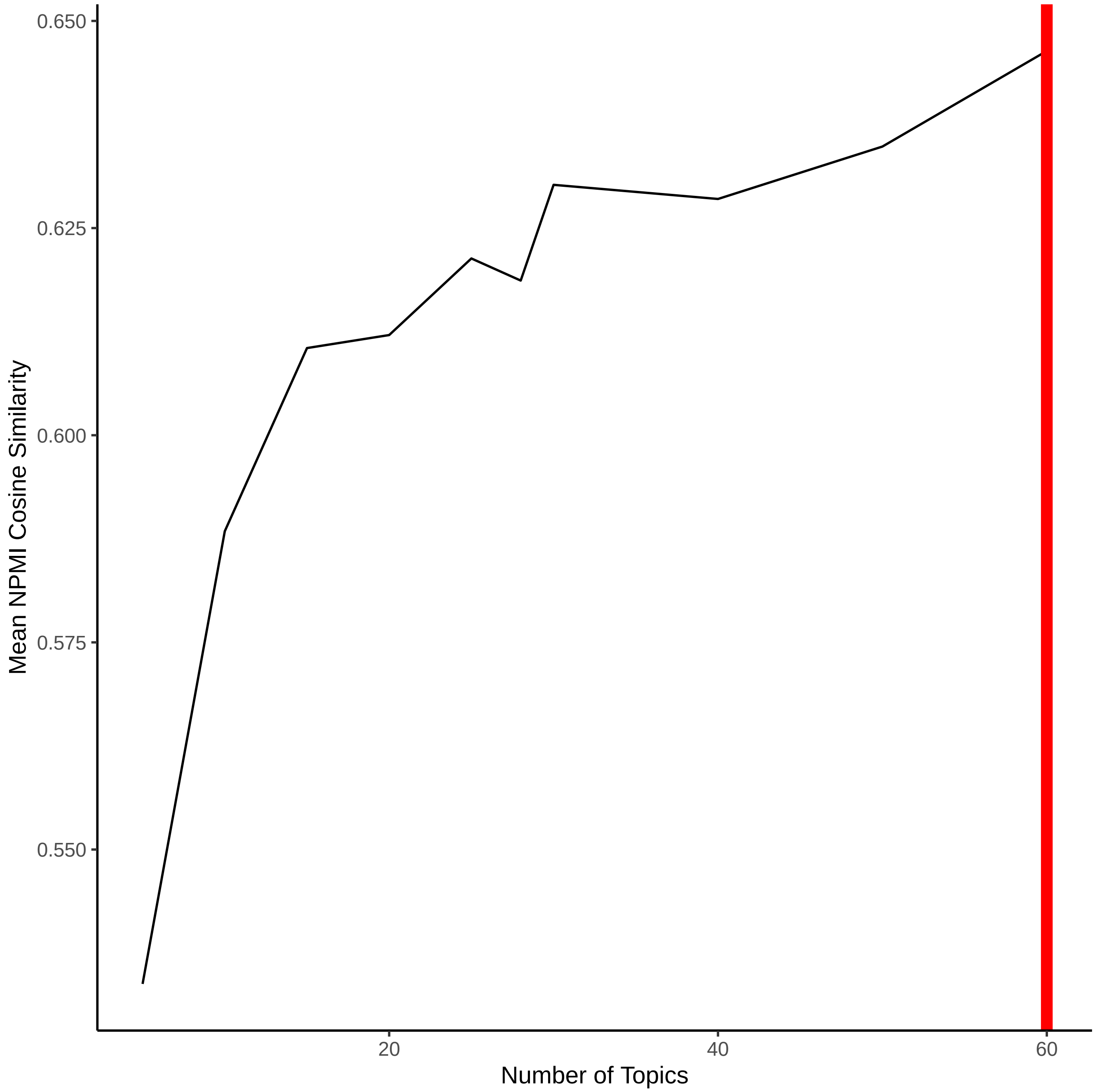


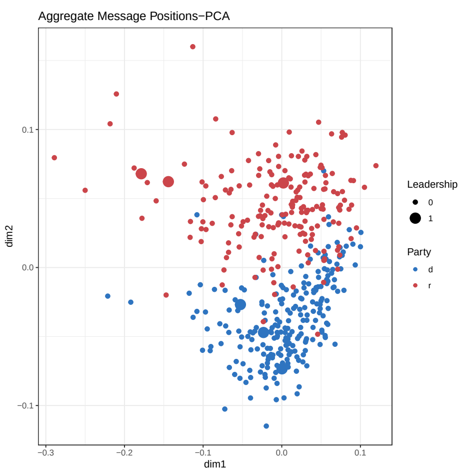
Figure SI 2: Coherence Score by Number of Topics

5 PCA Analysis and Summary

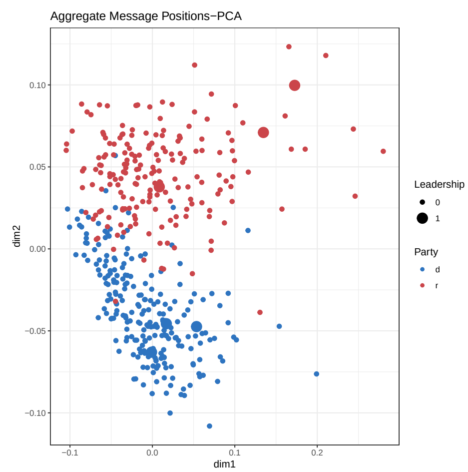
First, we show the topics which we identify as member-led in tables SI 3 and SI 4. These tables report the percent contribution to the overall variation in the data when we decompose the topic data using Principle Components Analysis.

5.1 Sensitivity to Topic Number - Democratic Party

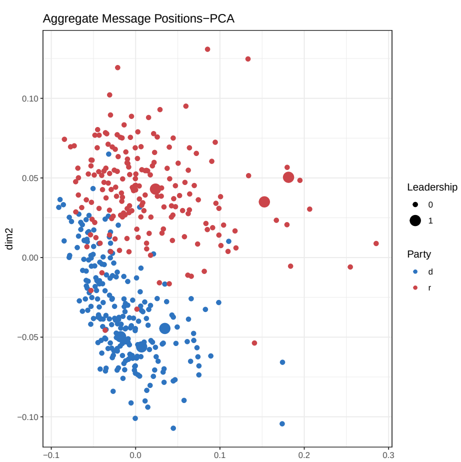
The following graphs show that the policy positioning which form the basis of the Need for Direction classification scheme are robust to changes in number of topics. The relative positioning and separation in the topical space is invariant to choice of topic number.



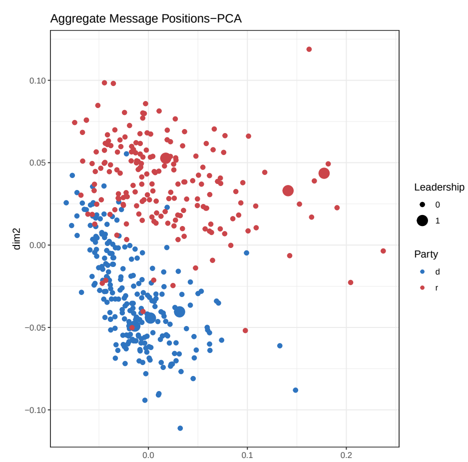
(a) 25 Topics



(b) 28 Topics



(c) 30 Topics



(d) 40 Topics

Figure SI 3: PCA Embeddings for Policy Stances, Varying by Topic Number

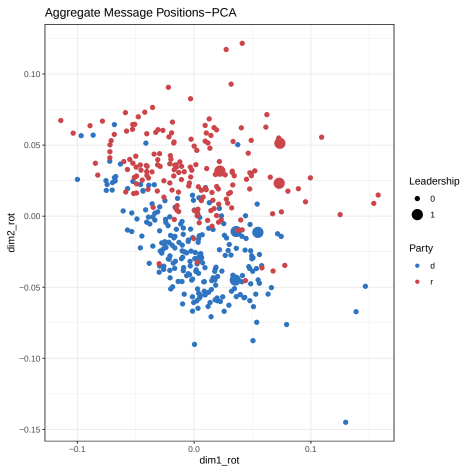
Table SI 3: PCA Topic Contributions - Member Driven 115th

Topic	Contribution
Lowest Unemployment Rate - Positive	1.26
Guests at Capitol Hill-Neutral	1.24
LGBT Equality-Negative	1.23
Fight for Civil Rights-Negative	1.15
Retweeting a Controversial Statement-Negative	1.11
Climate Change-Positive	1.11
Partisan Attacks on Trump/Biden-Negative	1.01
Important Meetings-Negative	0.99
Trump Admin Undermines Country - Negative	0.97
Budgetary Legislation - Negative	0.86
Committee Hearings-Positive	0.86
Hurricane Relief-Negative	0.86
Trump Climate Policy-Negative	0.85
Health Care Expansion - Neutral	0.84
Foreign Election Interference-Negative	0.82
Women's Pay - Positive	0.82
Supreme Court Nominations-Negative	0.73
Thoughts and Prayers - Negative	0.71
Floor Speeches-Negative	0.68
Student Loan Relief-Positive	0.67

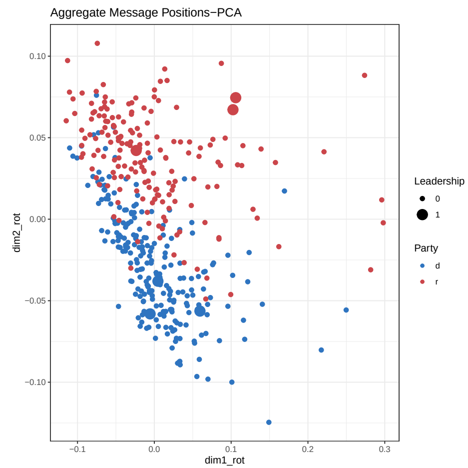
Table SI 4: PCA Topic Contributions - 116th Member Driven

Topic	Contribution
Family Seperations-Negative	1.18
Pro-Life Policy - Negative	1.14
China/Hong Kong Protests-Negative	1.11
Republican Senate Legislation-Negative	1.10
Prevent Gun Violence-Negative	1.08
Trump Admin Undermines Country - Negative	0.96
Fight for Civil Rights-Negative	0.93
Meuller Investigation - Negative	0.88
Trump Asuylum Policy	0.86
Enjoyable Visit - Positive	0.74
LGBT Equality-Negative	0.71
Social Security/Postal Service - Neutral	0.70
Health Care Expansion - Neutral	0.70
Trump Climate Policy-Negative	0.68
Mitch Mcconnel's Senate-Negative	0.67
Partisan Votes - Negative	0.60
Voting Rights - Positive	0.56
Law Enforcement - Positive	0.55
Honoring Cultural History-Negative	0.55
Protect Health Insurance - Neutral	0.54

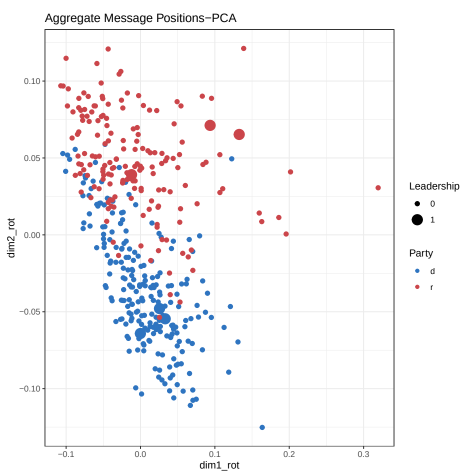
5.2 Dynamic Policy Stance Analysis



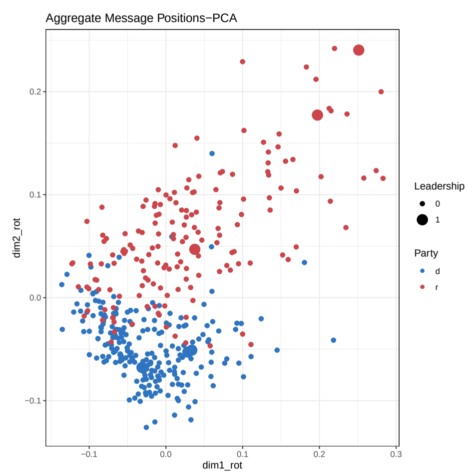
(a) Week 8 - 2019



(b) Week 20 - 2019



(c) Week 32 - 2019



(d) Week 2-2020

Figure SI 4: Changes in Time of Policy Stances

6 Time Series and Vector Autoregression

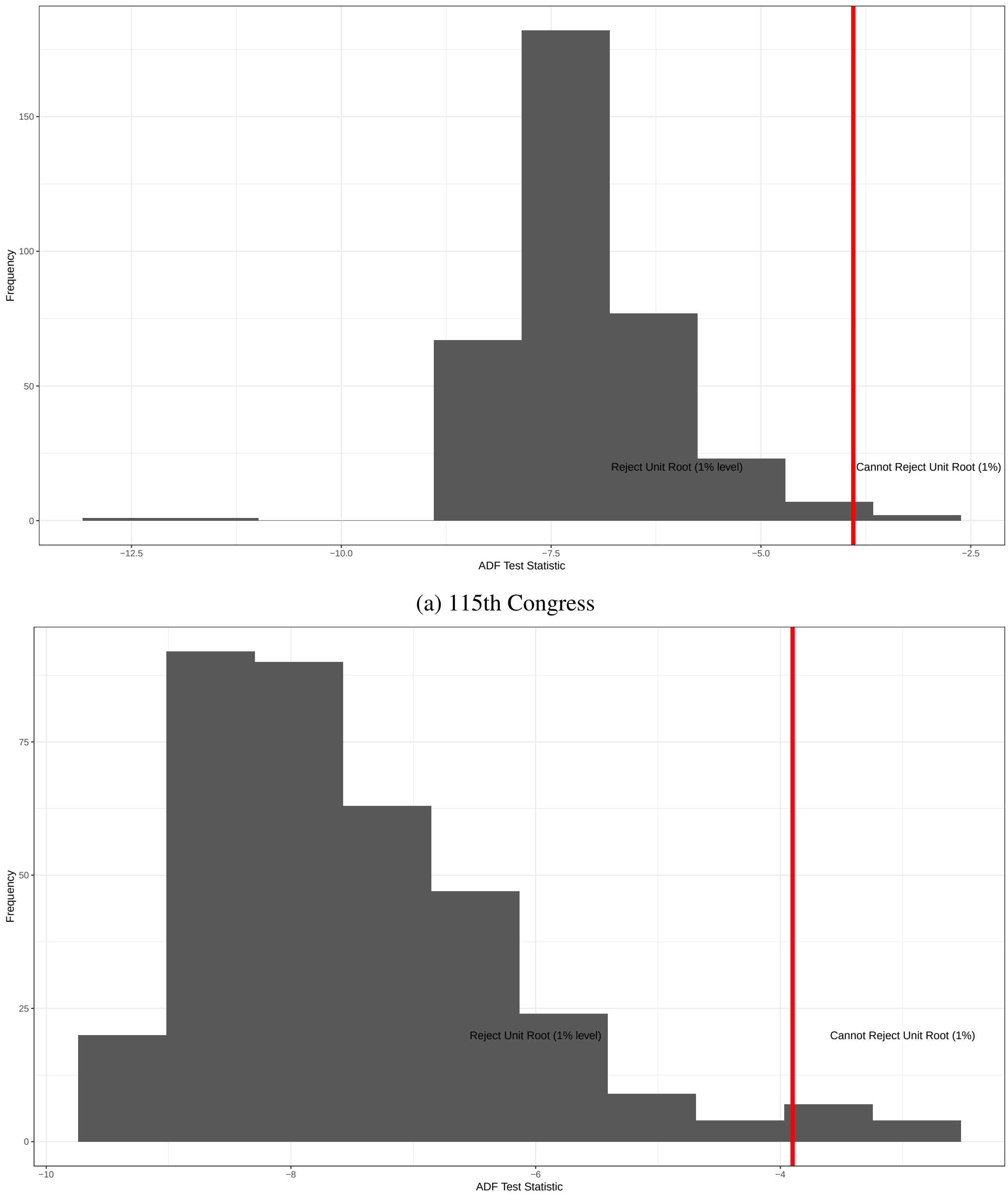
This section provides details for our dynamic analysis, in particular our vector autoregression methodology.

First, we sample some key time series to show the stationarity assumption – which is key to the validity of the VAR – holds across a variety of topics. We also show the full histogram of Augmented Dickey Fuller statistics, which tests for non-stationarity. The vast majority of our time series are consistent with the stationarity assumption, rejecting the unit root at the 1% level for over 95% of topics for the Democratic and Republican Parties in both the 115th and 116th Congresses.

Finally, we show a robustness check and that institutional leadership influence is substantively large. In Table SI 5 shows that institutional leaders exert on average more influence than the most followed accounts in each party and the leadership of the other party. On average, leaders exert double the influence as leaders from the other party on their members, as well nearly double the influence as the most followed accounts from within the same party. This latter finding highlights the relative strength of institutional leadership within the party caucus relative to the influence of members of the party who are popular with the public social media.

Table SI 5: IRFs Robustness

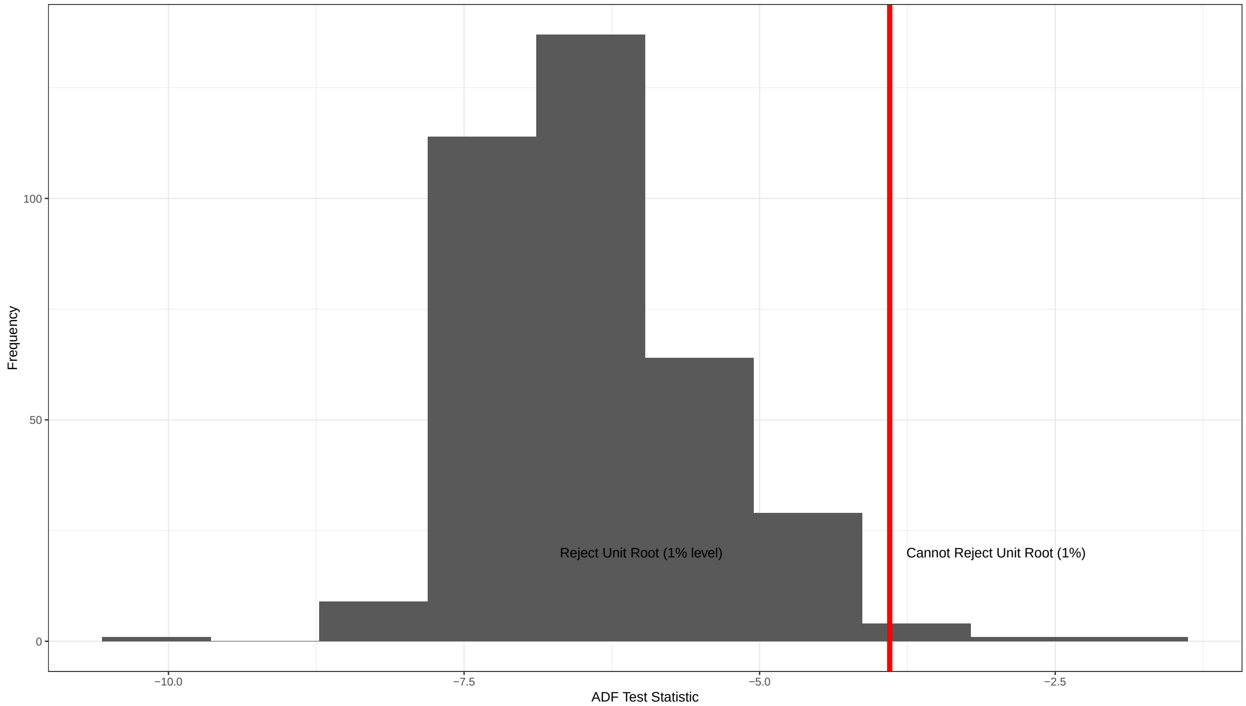
Leaders	Most Followed	Cross-Party Leaders	Total Percent Contribution	Congress	Party
0.202	0.160	0.042	57.245	115	Democratic
0.297	0.063	0.134	57.245	115	Republican
0.184	0.050	0.135	64.409	116	Democratic
0.408	0.364	0.257	64.409	116	Republican



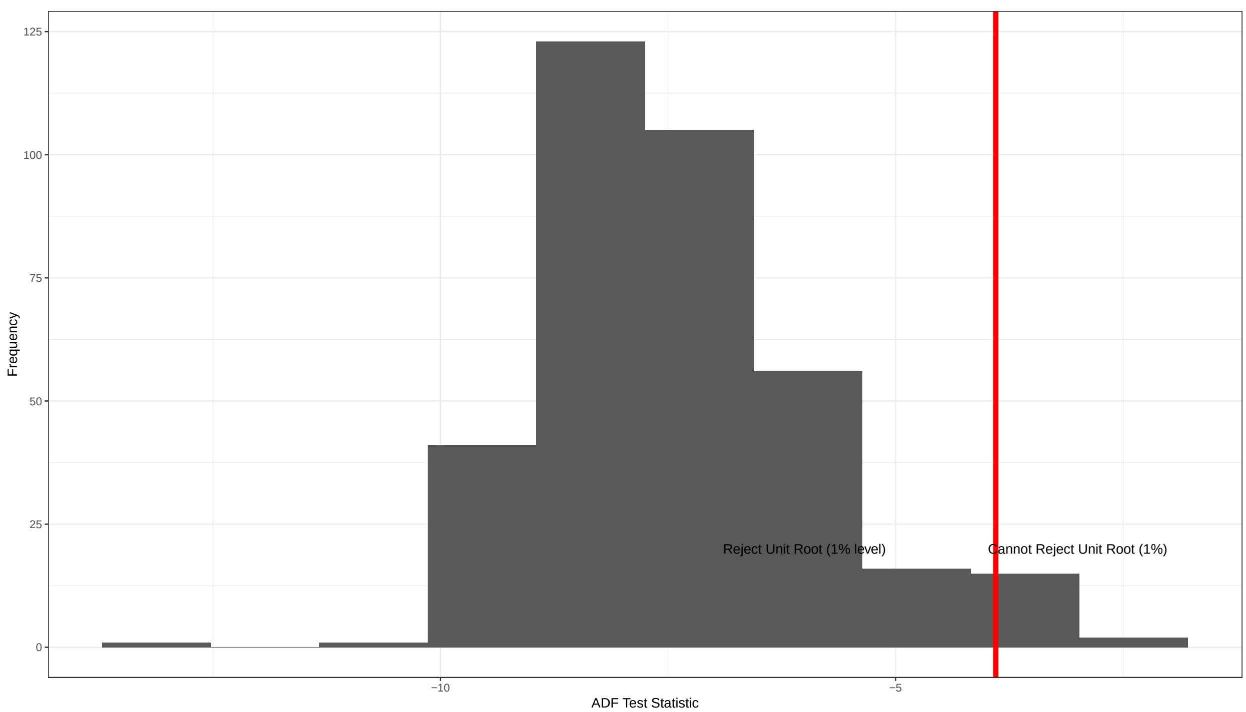
(a) 115th Congress

(b) 116th Congress

Figure SI 5: ADF Unit Root Test Statistics for all Topics: **Republican Party**. This Figure shows the distribution of tests ADF statistics for unit roots. All statistics to the left of the line represent topics for which we reject the null of a unit root at the 1% level, implying the stationarity assumption is satisfied

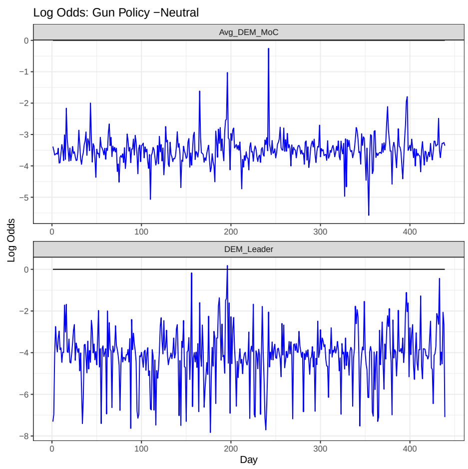


(a) 115th Congress

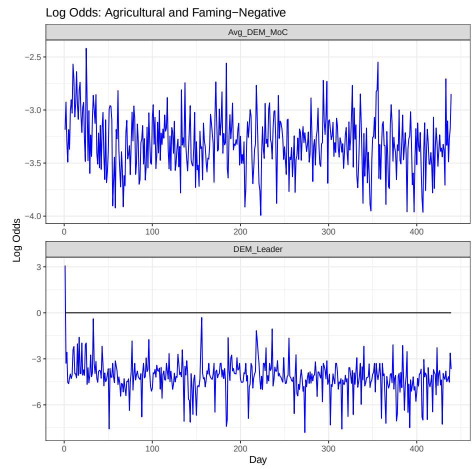


(b) 116th Congress

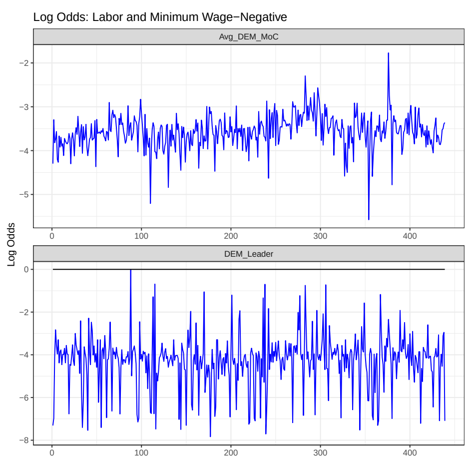
Figure SI 6: ADF Unit Root Test Statistics for all Topics: **Democratic Party**. This Figure shows the distribution of tests ADF statistics for unit roots. All statistics to the left of the line represent topics for which we reject the null of a unit root at the 1% level, implying the stationarity assumption is satisfied



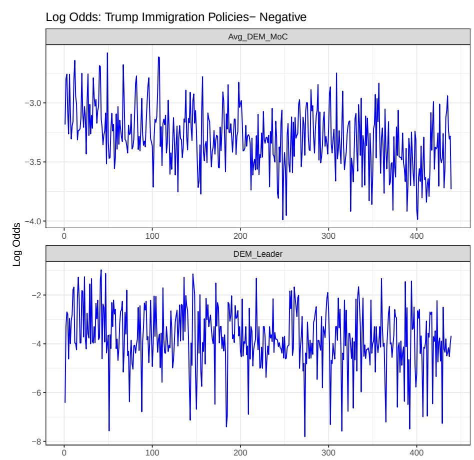
(a) Gun Policy



(b) Agriculture

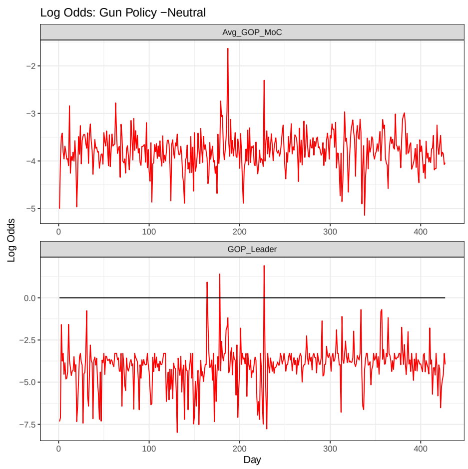


(c) Minimum Wage

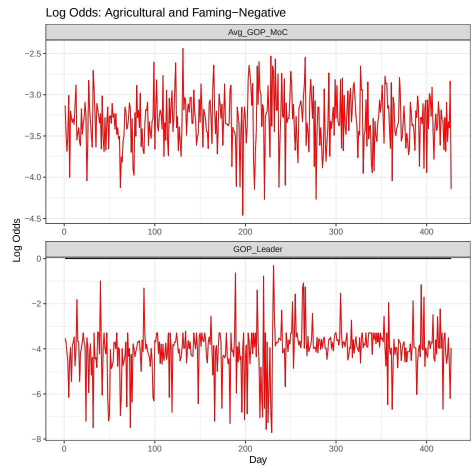


(d) Immigration Policies

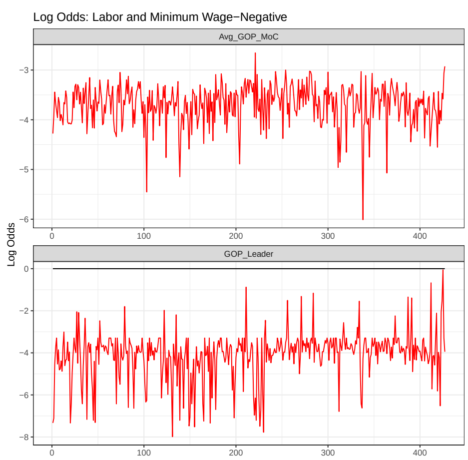
Figure SI 7: Stationarity in Log Odds of Daily Propensity of Discussion- Democratic Party



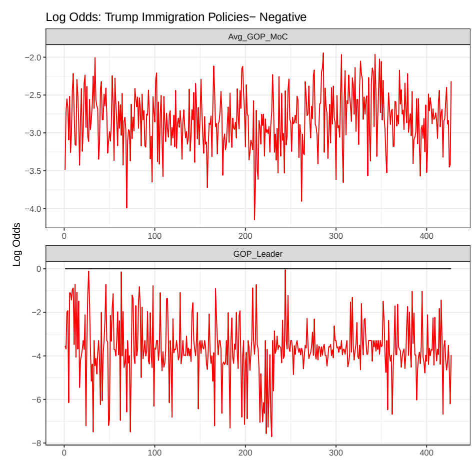
(a) Gun Policy



(b) Agriculture



(c) Minimum Wage



(d) Immigration Policies

Figure SI 8: Stationarity in Log Odds of Daily Propensity of Discussion- Republican Party

References

Bird, Steven, Ewan Klein and Edward Loper. 2009. *Natural language processing with Python: analyzing text with the natural language toolkit*. " O'Reilly Media, Inc.".

Denny, Matthew J. and Arthur Spirling. 2018. "Text Preprocessing for Unsupervised Learning: Why It Matters, When It Misleads, And What To Do About It." *Political Analysis* 26:168–189.

Grimmer, Justin and Brandon M. Stewart. 2013. "Text as Data: The Promise an Pitfalls of Automatic Content Analysis Methods for Political Texts." *Political Analysis* 21:267–297.

Lin, Chenghua and Yulan He. 2009. "Joint Sentiment/Topic Model for Sentiment Analysis." *Proceedings of ACM International Conference on Information and Knowledge Management (CIKM-2009)* .