Prepared for presentation at the Midwest Political Science Association (MPSA) Meeting 2019

The Rhetorical Agenda: What Twitter Tells Us About Congressional Attention

Libby Hemphill, University of Michigan

Annelise Russell, University of Kentucky

Angela Schöpke, University of Michigan

Abstract

Understanding how Members of Congress (MCs) distribute their political attention is key to a number of areas of political science research including agenda setting, framing, and issue evolution. Tweets illuminate what lawmakers are paying attention to by aggregating information from newsletters, press releases, and floor debates to provide a birds-eye view of a lawmaker's diverse agenda. In order to leverage this data efficiently, we trained a supervised machine learning classifier to label tweets according to the Comparative Agenda Project's Policy Codebook and used the results to examine the differential attention that policy topics receive from MCs. The classifier achieved an F1 score of 0.79 and a Cohen's kappa with human labelers of 0.78, suggesting good performance. Using this classifier, we labeled 1,485,834 original MC tweets (Retweets were excluded) and conducted a multinomial logistic regression to understand what influenced the policy areas MCs Tweeted about. Our model reveals differences in political attention along party, chamber, and gender lines and their interactions. Our approach allows us to study MCs' political attention in near real-time and to uncover both intra- and inter-group differences.

Introduction

Lawmakers' public statements often garner as much attention as their policy proposals in Congress, if not more. Members of Congress (MCs) use press releases, television, and now social media strategically to communicate their policy priorities and preferences. Traditional studies of policy agenda setting have largely relied on legislative outputs (i.e., bills, co-sponsorships) to understand policy attention, but non-legislative actions, such as a senator's Twitter communication, can be a useful venue for agenda-setting research. Public statements can signal a lawmaker's policy intentions even before legislative action takes place (cites TKTK), and increasingly, Twitter is a common site for these statements to appear (cites TKTK). Scholars have traditionally measured an actor's policy agenda by institutional actions, such as the amendment process, roll call votes or committee assignments, but non-legislative tools are growing resources for policy attention studies (Carson, Engstrom, and Roberts 2007; Rocca and Gordon 2010; Shepsle 1979; Shepsle and Weingast 1987).

New media platforms such as Twitter provide a methodological advantage for studying how individual lawmakers decide what to pay attention to and how they strategically advertise policy priorities on a regular basis; tweets are relatively cheap and easy to post and to access, and they can occur at any point in the legislative cycle. MCs are adapting their daily routines to incorporate social media applications like Twitter, and the platform provides a mechanism for strategically maximizing time and effort toward their public-facing policy agenda (Lassen and Brown 2011). Twitter is a low-resource mechanism for agenda setting that does not require institutional action or media attention. In 140-characters, a member can support a policy, take credit for its success, and signal its advantages to his constituents.

We introduce a computational model to measure the policy agendas expressed by individual MCs in their regular communication on Twitter. We then test individual and institutional influences on how legislators explain their work to constituents, journalists, and partisans. By using a public and widely adopted communication tool like Twitter, we aggregate information found in newsletters, press releases, and floor debates to provide a birds-eye view of a lawmaker's diverse agenda. Members of Congress have sent millions of tweets in the last couple years, and in order to leverage this data efficiently, we trained a supervised machine learning classifier to categorize lawmaker tweets according to the U.S. Policy Agenda Project's Policy Codebook. We used the results to examine the differential attention that policy topics receive from MCs. Just as some issues garner a disproportionate amount of attention in roll calls or legislative hearings, individual politicians also skew their attention online. We catalogue these policy patterns with a classifier that achieved an F1 score of 0.79 and a Cohen's kappa with human labelers of 0.78, suggesting good performance for assessing the policy content among lawmaker tweets. Using this classifier, we labeled 1,485,834 original MC tweets (Retweets were excluded) and conducted a multinomial logistic regression to understand what influenced the policy areas MCs tweeted about. Our model reveals differences in political attention along party, chamber, and gender lines. Our approach allows us to study MCs' political attention in near real-time and to uncover both intra- and inter-group differences that not only highlight how MCs use social media but also reveal MCs' public agenda-setting setting behavior.

Agenda Setting

Elite actors have individual agendas, each with unique distribution of preference intensities (Rocca, Sanchez, and Morin 2011), and media communications reveal how lawmakers balance their attention among underlying goals and responsibilities. The complexity we expect across the system-wide policy agenda is similarly evident in a member's individual agenda where they strategically communicate policy information given the political environment and their own interests.

Attention is the relative amount of time an individual or institution spends on a given issue. An elected official's processing power is curbed by a "bottleneck of attention" where short-term memory allows us to attend to only limited elements of the environment at any given time (Jones and Baumgartner 2005). This scarcity of attention forces us to deal with problems one at a time, and limited processing restricts what information is available for an actor to relay on social media. How an actor communicates this information is the result of a skewed or selective process that results in some interests being ignored while others are prioritized (Boydstun 2013; Downs 1972).

Traditional agenda setting studies typically assess the policy agendas of legislative bodies or institutions, but the scarcity of agenda space, primacy of attention, and skewed distribution of policy issues are also present among individual lawmakers. Traditional agenda setting is the organizational analogue to attention allocation at the individual level (Jones and Baumgartner 2005). Individual agendas are also a comprised of a broad and complex set of policymaker priorities. The bills a senator introduces may shed light on his policy priorities, but to understand the complexity behind individual decision-making we need a non-legislative tool. Policy agenda setting is contingent on many factors — e.g., political climate, political feasibility, personal and constituent priorities — and lawmakers' agendas on Twitter allow for that contextualization.

Tweeting Political Agendas

Political communication is a useful tool for assessing lawmakers' policy agendas, and social media has emerged as a viable venue such that politicians no longer solely rely on newspapers and newsletters to frame their political messages. Twitter offers a window into politicians' decision making and provides new and previously unknown details of how politicians prioritize their agenda in pursuit of re-election. Lawmakers are integrating Twitter among their traditional media activities due to the platform's many advantages, including its low-cost publicity, user discretion, and multi-directional communication potential (Straus, Glassman, Shogan, and Smelcer 2013).

Twitter enables direct communication with both elite and mass publics with minimal opportunity costs. Actors have increased control over their own communications strategies with minimal time and resources expensed, and thus have the ability to better target communications to their base of followers. This low-cost effort has the outsized potential to increase the interactions between elite and mass publics by publicly broadcasting their agenda and creating an accessible record of government action (Bruns and Highfield 2012). Twitter is a broadcasting device for politicians (Gainous and Wagner 2014; Golbeck et al. 2018; Hemphill, Otterbacher, and Shapiro 2013), so being able to take advantage of its outreach capabilities is especially important to politicians and their staffers (Chi and Yang 2011). This public communication domain offers policymakers a relatively unfiltered credit claiming opportunity (Mayhew 1974) to highlight accomplishments and advertise a political brand. Press releases and C-SPAN coverage produce a public record, but neither is as readily-accessible for public observation or analysis as are 140-count messages in a scrolling feed that thousands follow.

On Twitter, the lawmakers have direct control of their message — bypassing media context and institutional constraints. Bypassing the traditional media filter means messages are not balanced with opposing viewpoints or prioritized according to the media agenda. Traditional media sources attract publicity and an audience but they come at the cost — either in discretion or lawmaker effort to appeal to journalists and editors. While traditional media messages may be an index of elite opinions or deferential to politicians (Bennett 1990), when information is published, how it is framed, and the context of an issue are all decisions out of the hands of elite actors. In newspapers or television broadcasts, the priorities of the politicians become integrated with the priorities of the news organization. Traditional media offer a regular, indirect measure of priorities whereas Twitter is more frequent and contextualized by only the political actor and his staff.

Twitter offers a type of new-media franking privilege where members can highlight their individual or party achievements and boost their image by deriding the opposition party. Social media platforms like Twitter, at their most core function, are a mechanism for conflict expansion, and actors have increased control over that expansion. Nothing attracts a crowd as quickly as a fight (Schattschneider 1960), and Twitter is one platform for politics to target that contagion effect.

Social media is not only a useful mechanism for political actors but it also is emerging as a medium for political communication studies. Different media allow for different modes of information production and consumption (Jungherr 2014), and the constraints of the technology underlying broadcast news are different than that of social media. The fact that social media can bypass traditional media institutions altogether requires a differentiation between how we study social and traditional media sources (Jungherr 2014; Shapiro and Hemphill 2017). A new digital logic requires research that takes advantage of Twitter's platform to aggregate attention and participate in dialogue not mediated by mass media.

Modeling Policy Agendas on Twitter

Studying political attention is essential to understanding how lawmakers distribute attention and and frame issues for voters and the public at large. Existing methods for studying the policy agendas of lawmakers primarily employ manual topic labeling, which is dependent on human effort and can be restrictive in terms of scope and scale. As members of Congress have expanded their use of social media for daily communications about policy problems, so too must the research methods that we use to understand how lawmakers engage various constituencies. To understand what effects those patterns of attention and how it differs among lawmakers, we seek alternative methods of policy agenda analysis.

To address the need for a comprehensive and consistent mechanism for measuring policy agendas on Twitter and at scale, we develop a computational model for estimating political attention. We leverage a sample of human-labeled congressional tweets to train a supervised machine learning classifier to label the policy topics in lawmakers' tweets. We test that classifier to evaluate the performance of our models against true labels, and expect that the trained classifier will serve as a viable alternative to manual coding techniques.

With a high-performing classifier, we can analyze what drives lawmakers' patterns of attention among a consistent set of policy topics on a much larger scale than possible by current content coding techniques. Senators' Twitter agendas are an ideal platform to address theoretically important questions about legislators' agenda-setting behavior and representation (Russell 2018), and the use of the machine learning classifier allows for for comprehensive analysis across a set of topics, from the Policy Agendas Project, that have been used over time to calculate policy attention across multiple policy outputs. By using all the tweets of lawmakers in Congress, the data allows us to use this coding scheme to test hypotheses common to inquiries of legislative activity and lawmaker homestyle. Because individuals develop unique styles of communication and legislative style (Bernhard and Sulkin 2018; Grimmer 2013), we expect their Twitter agendas and the issues they choose to prioritize for public messaging to reflect those patterns of communication. Policy attention is often dependent on political climate, issue emergence, policy frames, but at the individual level, we look at how lawmaker-specific characteristics and the political climate influence attention allocation to policy issues on Twitter. Research suggests that gender, party, and geographic differences are powerful explanations for legislators policy behavior (Atkinson and Windett 2018; Neiheisel and Niebler 2013), and we expect similar outcomes in their communication agendas (Gulati and Williams 2007).

Gender Effect

Twitter can be an effective tool for political minorities lacking the political power to shape the institutional agenda. The number of female lawmakers in Congress is increasing, but their small numbers relative to broader population may incentivize them to more readily seek out alternative agenda-setting spaces. Work by Evans and Clark (2016) suggests we should expect gender to have a direct effect on political candidates' social media messages. They find that women running for Congress discuss policy issues on Twitter at a higher rate, and those issues are often "women's issues." Stereotypes of female lawmakers as relationship-builders rather than policy experts may incentivize some women to be more active in policy communication on Twitter. Once in office, they may also adopt more "masculine" styles of communication that highlight policy preferences more often. Women may combat stereotypes by adopting broad, diverse agendas that allow them to develop a reputation and deter possible challengers (Atkinson and Windett 2018). Having an alternative agenda space on Twitter may enable female MCs to counter these stereotypes and compensate for perceptions that female lawmakers are less policy-capable. Based on this earlier work, we expect that women will discuss policy more often than their male counterparts and that their patterns of attention will differ.

Party Effect

Political parties have been characterized as a mechanism to serve and facilitate electoral goals (Mayhew 1974) and to maintain majority status (Aldrich 1995), and in order to do so members work collectively to maintain desired ends. A lawmaker's strategic action is therefore not only tied to individual actions, but alternatively constrained by the party and institution. Congress is simultaneously a representative assembly where members act as individuals responsible for meeting constituent demands and a lawmaking body where members act collectively, with increasing homogeneity among political parties (Davidson and Oleszek 1990; Hill and Williams 1993; Rohde 1991). There is a tension between the goals of individual members and that of the party or the collective (Damore and Hansford 1999), and that party pressure may lead elected officials to prioritize policy in light of party preferences. A member's decision to delegate between individual and collective goals is a trade-off in attributable costs and benefits (Hill and Williams 1993), and parties maintain ties with their members through the allocation of benefits —e.g., campaign funds during the next election cycle, open leadership positions or committee assignments. Party leaders have increased influence over the institutional agenda (Aldrich 1995; Cox and McCubbins 1993; Rohde 1991), but the extent to which leaders and the party influence the issues that individual lawmakers choose to address in their public agendas on Twitter has implications for both representation and the policy process. Based on this prior work, we expect to observe little within-party variation, but that the two parties will exhibit different patterns of political attention from one another.

Chamber Effect

Prior work found that House candidates provided less information about specific policy positions than did Senate candidates and suggests that media outlets that candidates control, such as their own websites or social media feeds, are useful vehicles for communicating issue stances in case those positions are reported by mainstream media (Gulati and Williams 2007). Research that examines chamber differences in Twitter behavior, explicitly, found that senators were, on average, less frequent tweeters than representatives (Hemphill, Otterbacher, and Shapiro 2013). Senators and representatives represent different constituencies--states and districts within states--and those constituencies likely require different political strategies. Prior research shows that the differing constituencies produce different patterns in federal spending policy and credit-claiming between the two chambers (Lee 2004). How distributive policies relate to other issues such as environment or civil rights and whether members of the house attend to more parochial concerns than their senate counterparts is not yet settled (Lee 2004; Parameswaran 2018). This prior literature on chamber differences suggests that the Senate will use Twitter less often to discuss policy and that the chambers will exhibit different patterns of attention.

Whether and how lawmakers use Twitter to communicate their policy agendas and how that communication differs among parties, chambers, and genders are open questions we address here. We first describe the construction of our classifier and then report and discuss the patterns of political attention we find after labeling the 115th US Congress' tweets.

Methods

Data

115th Congress Data

Using the Twitter Search API, we collected all tweets posted by official accounts linked to voting members of Congress during the 115th Congress which ran January 3, 2017 to January 3, 2019. We identified MCs' Twitter user names by combining lists of MC social media accounts from the UnitedStates project[1], George Washington Libraries[2], and the Sunlight Foundation[3]. Throughout 2017 and 2018, we periodically used the Twitter API to search for the user names in this composite list and retrieved the accounts' most recent tweets. We conducted our final search on January 3, 2019, shortly after the 115th Congress ended. In all, we collected 1,485,834 original tweets from 524 accounts. We included data for MCs who resigned (e.g., Ryan Zinke) and those who joined after special elections (e.g., Rep. Conor Lamb); we were also not able to confirm accounts for every state and district.

Metadata

We used UnitedStates project and Sunlight Foundation datasets to retrieve MC metadata information, including details about which state they represent, chamber, party, gender, and birthday. We determined age according to MCs' reported birthdays. We grouped these ages into age buckets 30-39, 40-49, 50-59, 60-69, 70-79, and 80-89. For each of six CMs ('gianforte', 'lindseygrahamsc', 'repblumenauer', 'repryanzinke', 'amashoffice', and 'senbillcassidy') that did not have birthday or state metadata available via UnitedStates project or Sunlight Foundation, we used data from their official websites to manually collect birthdays and state metadata.

Manually-Labeled Training Data

The original set of labeled Tweet data from Russell (2017, 2018) comprised 68,398 tweets. Of these tweets, 45,402 tweets were labeled with policy codes and 22,996 were labeled as not-policy tweets. We removed Retweets from this set to limit our classification to original tweets, resulting in a total set of 59,826 labeled tweets (39,704 policy tweets and 20,122 not policy tweets). Since not-policy tweets (topic "0") were overrepresented in our training dataset, in order to ensure that our supervised model could effectively learn patterns from policy-labeled tweets, we trained our best performing model (LR) on modified training sets from which we incrementally removed between 10% and 100% of randomly selected not-policy tweets (see Table 1). The best performing model according to F1 Score and Cohen's Kappa, is that for which we removed 90% of all not-policy tweets. This proportion of not-policy tweets reflects approximately the median proportion of tweets represented by any given topic, supporting a nearly-balanced training set. This Tweet set amounted to 41,716 tweets total (39,704 policy tweets and 2,012 not policy tweets).

Training and Evaluating a Supervised Classifier

Preprocessing

In preparing text for use in a machine learning model, the text is divided into sequences of characters called *tokens* that are then used for analysis. Often, tokens are words, but in some cases they are multi-word phrases or parts of words such as word stems. Classifiers then look for associations between tokens and classes. How the text is processed into tokens impacts how classifiers make decisions, and here we describe the decisions we made during preparing the text for classification (or *preprocessing*).

We left words intact and did not reduce them to stems or lemmas. In the case of tweets, we expect that misspellings may often indicate different semantic meanings among terms with the same stemmed root, and thus potential association of different spellings with different topics. Stemming in these instances can remove the nuance in potential semantic meaning achieved by misspellings (Schofield and Mimno 2016).

[1] https://github.com/unitedstates/congress-legislators
[2] https://dataverse.harvard.edu/dataset.xhtml?persistentId=doi:10.7910/DVN/UIVHQR
[3] https://sunlightlabs.github.io/congress/ index.html\#legislator-spreadsheet

Given the prevalence of both English and Spanish language tweets, we removed English and Spanish stopwords using Python's Natural Language Toolkit (NLTK) English and Spanish stopword lists.

We employed NLTK's (Bird, Klein, and Loper 2009) TweetTokenizer with parameters set to render all text lower case, to strip all Twitter username handles, and to replace repeated character sequences of length three or greater with sequences of length three. We complemented initial tokenization with removal of punctuation (including emojis), URLs, words smaller than two letters, and words that contain numbers.

Vectorization is the process of turning texts into numerical vectors that indicate the presence or absence of various tokens (or features) in a given text. We tested each of three vectorization approaches: simple one-hot encoding approach using Scikit-learn's (Pedregosa et al. 2011) DictVectorizer, simple bag-of-words approach with Scikit-learn's CountVectorizer, and bag-of-words term frequency inverse document frequency approach with Scikit-learn's TfidfVectorizer. Of these vectorization approaches, we found the simple bag-of-words approach using unigrams to result in the best performing models. This means that we represented tweets as unordered collections (or bags) of tokens (or words) using vectors that indicate, for each word in the entire collection of tweets, which are present in a given tweet.

Model Selection

This project trained and tested each of four types of classification models: a random guessing baseline dummy model (D) using stratified samples that respect the training data's class distribution, a Naive Bayes (NB) model, a Logistic Regression (LR) model, and a Support Vector Machine (SVM) model. In each case, we used a 90-10 split for train-test data meaning that 90% of labeled tweets were used as training instances, and the models then predicted labels for the remaining 10%. We then compared the models' predictions with the human labels to evaluate their performance. After initial testing, for each of our top two performing models (LR and SVM), we subsequently evaluated whether the addition of Word2Vec (W2V) (Mikolov, Chen, Corrado, and Dean 2013) word embedding features or Linguistic Inquiry and Word Count (LIWC) (Pennebaker, Booth, and Francis 2015) features could improve their performance.

W2V models turn text corpora into numerical vectors (word embeddings), and groups these vectors of similar words together. Given sufficient data, these groupings can infer a word's meaning. For W2V feature embeddings, we attempted including each of W2V feature embeddings from Godin et al.'s pre-trained model (Godin, Vandersmissen, De Neve, and Van de Walle 2015), as well as from a W2V model we trained on our Tweet set.

LIWC dictionaries are collections of words categorized and subcategorized into semantic and syntactic categories (e.g. emotional tone, percentage of words in the text that are pronouns, affect, biological processes, leisure activities, swear words, etc.). We used the LIWC 2015 Dictionary composed of approximately "6,400 words, word stems, and select emoticons" grouped into approximately 90 categories and subcategories (Pennebaker, Booth, and Francis 2015). Using this dictionary, we extracted LIWC categorical groupings for each token (word) in our Tweet set.

We implemented each of D, NB, LR, SVM models with W2V and LIWC features using Scikit-learn with modifications to model configurations to improve predictive performance. Each of these models' performance is outlined in the "Supervised Model Performance" section, Table X.

Evaluation Measures

To evaluate the performance of our models, we calculated F1 scores for each model and Cohen's Kappa of predicted labels against true (i.e., human-coded) labels. F1 scores are essentially the weighted average of precision and recall and are a common performance measure for general classification models. There is no threshold or ladder of F1 performance, but higher is generally better.

Supervised Model Performance

Using a logistic regression (LR) classifier, we achieved the highest F1 score (0.79) in our ensemble. We compared our classifier to a dummy classifier using the same training data, and it achieved only F1 = 0.07. Given the difference between our classifier's score, the dummy, and the more complicated models and feature additions we evaluated (NB, SVM, W2V, and LIWC), we argue we have achieved high accuracy and that the LR classifier is the best available. We also compared our classifier's labels with the human labels and achieved a Cohen's kappa of 0.78, suggesting moderate agreement (McHugh 2012). Table 1 displays our models' results according to F1 and Cohen's Kappa scores.

Table 1: Classifier Performance Summary

Classifier	F1 Score	Cohen's kappa
D	0.07	-0.00
NB	0.70	0.69
LR	0.79	0.78
LR + pre-trained w2v features	0.78	0.76
LR + original w2v features	0.77	0.75
LR + LIWC	0.78	0.76
SVM	0.78	0.76
SVM + pre-trained w2v features	0.77	0.75
SVM + original w2v features	0.77	0.74
SVM + LIWC	0.78	0.76

Statistical Analyses

Our goal is to understand and explain how a member's party, chamber, and gender affect their political attention. We use the output of our classifier to measure that attention by assuming that a tweet's policy area class indicates attention to that area. Since the Comparative Agendas Project codebook includes 20 policy areas, and we used a multinomial logistic regression to approach this question. We chose policy area number 5, *Labor*, as our reference category and use the *nnet* package in R (Venables and Ripley 2002) to conduct these analysis. Labor is used as the base category given its moderate level of salience and inter-party appeal that spans from issues around workforce development to questions about fair pay and benefits.

Results

We summarize the variables used in our models in Table 2. To address whether there are differences among parties, chambers, and genders related to tweeting about policy generally, we first used logistic regression (LR) to predict the frequency of policy tweets on any topic. We fit models of the predictors independently, in combination, and with interaction terms. Using AIC comparisons and ANOVA, we found that the exhaustive model, that included all three independent variables and interactions among them, was the model of best fit when predicting frequency of policy-related tweets (see Table 3).

Table 2. Variables included in regression analyses

Variable	Definition
Dependent Variables	
policy_tweet	0 = not a policy-related tweet 1 = tweet is related to policy
policy_area	0 = tweet is not about policy 1-21 = major code from the Comparative Agendas Project Codebook that is most likely associated with the tweet
Independent Variables	
republican	0 = Democrat or Independent 1 = Republican
senate	0 = Representative 1 = Senator
man	0 = woman 1 = man

Table 3. Results of Logistic Regression predicting policy tweets

Term	Odds Ratio	SE
republican	0.656***	0.006
senate	1.610***	0.017
man	0.910***	0.006
republican:senate	0.691***	0.014
republican:man	1.137***	0.012
senate:man	0.927***	0.012
republican:senate:man	1.158***	0.026
(Intercept)	2.857***	0.014

Then, to identify patterns in the specific policies discussed, we fit six different multinomial logistic regression (MLR) models to determine the relationships between party, chamber, gender, and policy area. We fit each of the independent variables alone, then all three together, interacting party and chamber, and interacting all three terms. Using AIC comparisons and ANOVA, we found that the model that included all three independent variables and no interactions was the model of best fit. Table 4 shows the results of the best MLR[4]; it contains odds ratios and standard errors for each topic. The CAP codebook contains no code #11.

Overall we see more policy discussion in 2018 than in 2017 (see Figure 1). We see that policy discussion peaked on Twitter in April and May of 2018, increasing after the 2018 primaries, and decreasing after Congressional elections in November 2018. We also see that Democrats, Senators, and women tend to post policy tweets more frequently than Republicans, House Representatives, and men (see Table 4).

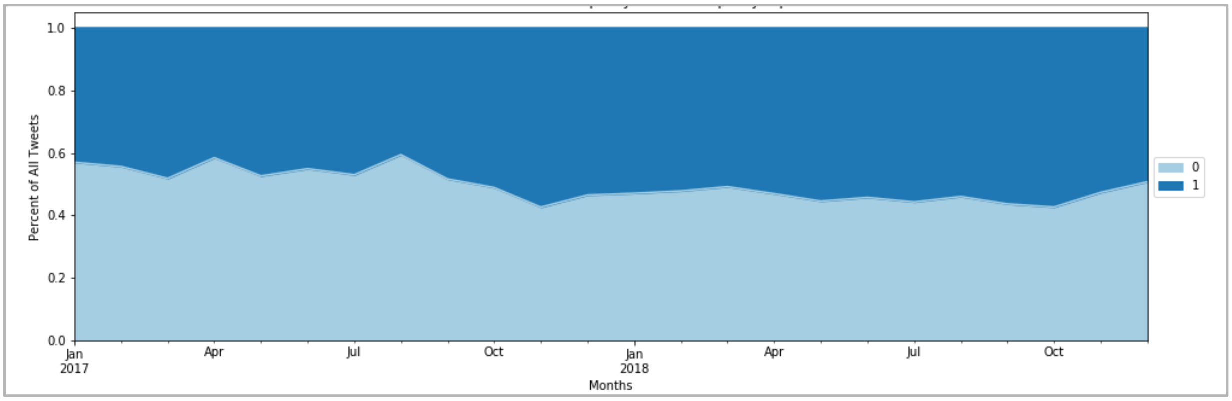

Figure 1. Proportion of tweets that were labeled “policy” or “not policy” over time. 0 = “not policy and 1 = “policy”

Table 4. Results of Multinomial Logistic Regressions. Odds ratios and standard errors are provided.									
CAP #	CAP Major Code	repub.	s.e.	senate	s.e.	man	s.e.	constant	s.e.
1	Macroeconomics	1.458***	0.012	0.872***	0.013	1.277***	0.014	2.034***	0.012
2	Civil Rights	0.633***	0.013	0.815***	0.014	0.860***	0.014	2.719***	0.012
3	Health	1.017	0.012	1.031**	0.012	0.993	0.013	3.749***	0.011
4	Agriculture	1.826***	0.019	1.338***	0.019	0.860***	0.021	0.340***	0.019
6	Education	0.882***	0.015	0.941***	0.015	0.893***	0.016	1.233***	0.014
7	Environment	0.494***	0.017	1.277***	0.016	1.506***	0.018	0.660***	0.016
8	Energy	1.972***	0.018	1.447***	0.018	1.255***	0.021	0.289***	0.019

[4] Complete results for all models are available in supplementary documents.

9	Immigration	0.704***	0.016	0.710***	0.017	1.225***	0.017	0.928***	0.015
10	Transportation	0.983	0.019	1.122***	0.019	1.152***	0.021	0.382***	0.019
12	Law and Crime	0.751***	0.013	0.996	0.013	0.952***	0.014	2.479***	0.012
13	Social Welfare	0.541***	0.021	0.587***	0.023	0.833***	0.02	0.624***	0.017
14	Housing	0.644***	0.032	1.162***	0.031	1.028	0.032	0.147***	0.028
15	Domestic Commerce	1.266***	0.014	0.955***	0.015	1.112***	0.016	1.079***	0.014
16	Defense	1.696***	0.012	0.955***	0.012	1.181***	0.013	2.513***	0.012
17	Technology	1.145***	0.022	1.361***	0.022	1.048*	0.025	0.235***	0.022
18	Foreign Trade	1.343***	0.035	1.508***	0.034	1.146***	0.041	0.067***	0.037
19	International Affairs	1.385***	0.014	1.097***	0.014	1.358***	0.016	0.941***	0.014
20	Government Operations	1.224***	0.012	1.173***	0.013	1.247***	0.014	1.669***	0.012
21	Public Lands	1.146***	0.019	1.441***	0.019	1.004	0.021	0.362***	0.019

What patterns of political attention appear on Twitter?

We see that most policy topic areas receive little attention, and that low attention shows little variance over time (see Figure 2). Those that do receive more attention consistently, in periodic swells, or sporadically include topics 3 (health), 1 (macroeconomics), 12 (law and crime), 16 (defense), and 9 (immigration). *Health* received a peak in attention during the first half of 2017 and then leveled off over the remainder of the 115th Congress. *Macroeconomics* peaks in the fall of 2017. *Law and crime* received increased attention during the first half of 2018, which then diminishes for the rest of the year while remaining higher overall than 2017 rates of attention. *Defense* features a generally higher baseline than most topics, demonstrating some periodicity towards the end of 2017 and early 2018. For *Immigration*, we observe three noticeable peaks in the 4th quarter of 2017 and the 1st and 2nd quarters of 2018.

Figure 3 shows the percent of tweets represented by each policy area over time during the course of the 115th Congress. Figures 2 and 3 display the same data, and the stacking in Figure 3 reveals different trends and anomalies. We see that during the first three quarters of 2018, topics 14 (housing) and 18 (foreign trade) increasing slightly in their proportion of Tweet attention relative to other topics. During the third quarter of 2018, we see topic 13 (social welfare) experiencing a jump in Tweet attention. Finally, during the final quarter of 2018, we see topic 4 (agriculture) also experiencing a jump in Tweet attention.

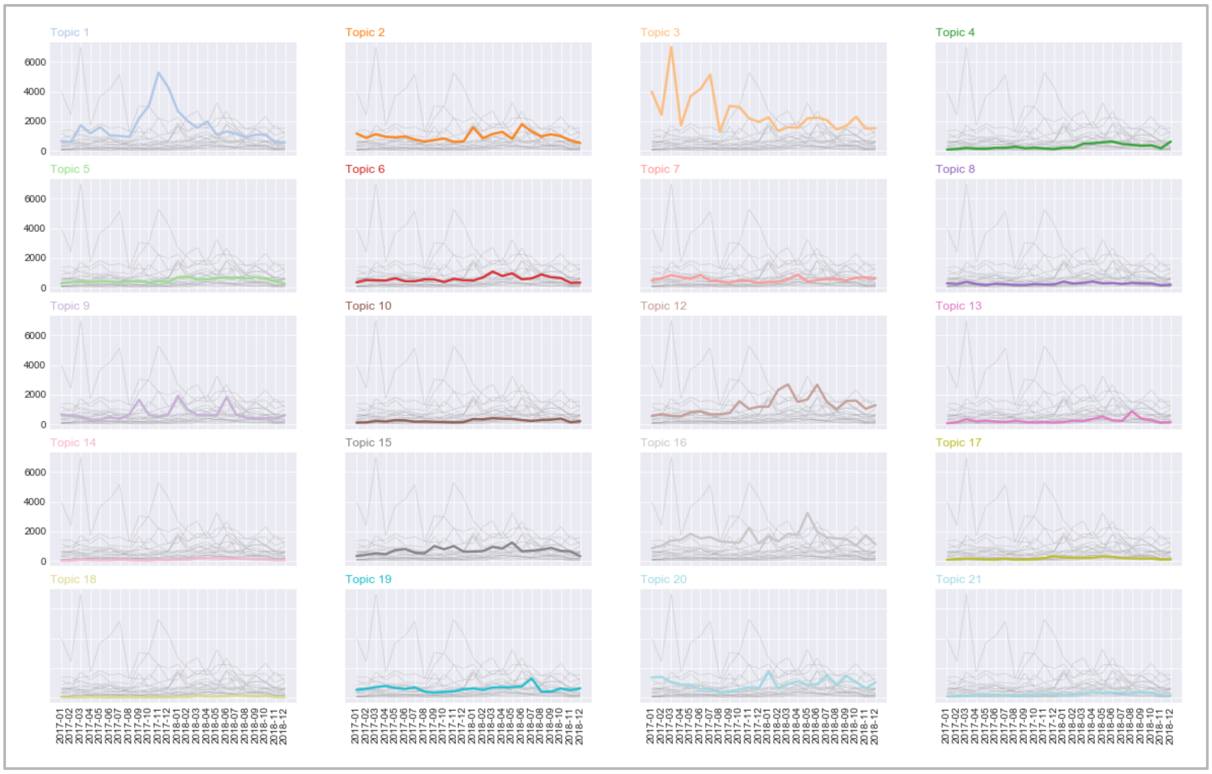

Figure 2. Frequency over time for each topic in CAP codebook. Results are presented with monthly frequencies.

Do the parties, chambers, and genders differ in the policies they discuss?

The results of the MLR (see Table 4) show that there are significant differences between parties, chambers, and genders for nearly all topics in the CAP codebook. We can look to the odds ratio for these variables' effects in order to understand the nature of these differences.

If we look to the effects of party, we see that the highest odds ratio exists for Republicans

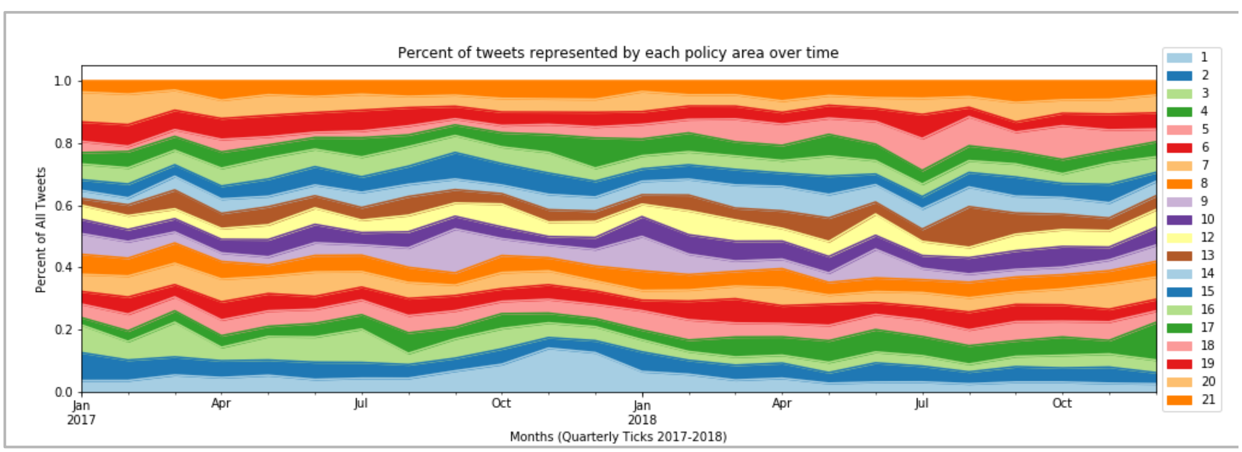

Figure 3. Percent of tweets in each policy area over time.

compared to Democrats with topic 8 (energy). The lowest odds ratio exists for Republicans compared to Democrats for topic 7 (environment). These odds ratios suggest that the Republican MCs are more likely to focus attention on *Energy* than are Democrats, and Democrat MCs are more likely to focus attention on *Environment* than are Republicans.

Concerning the effects of chamber, we see that the highest odds ratio exists for Senators compared to House Representatives with topic 18 (foreign trade), and the lowest odds ratio for topic 13 (social welfare). These odds ratios suggest that Senators are more likely to focus attention on *Foreign Trade* than are House Representatives, and House Representatives are more likely to focus attention on *Social Welfare* than are Senators.

Finally, concerning the effects of gender, we see that topic 7 (environment) receives the highest odds ratio for men compared to women, and that topic 13 (social welfare) receives the lowest odds ratio. These odds ratios suggest that men are more likely to focus attention on *Environment* than women, and women are more likely to focus attention on *Social Welfare* than men.

Discussion

We presented a supervised machine learning model that is able to detect political topics in tweets and assign them to categories in a widely-used codebook for measuring political attention. This model enabled us to (1) observe patterns in Congress's political attention through the 115th Congress and (2) identify differences in political attention among lawmakers' parties, chambers, and genders. We found that the proportion of lawmakers' tweets that address policy issues stayed relatively stable throughout the congress, ranging between 41%-57% of tweets, but that the parties, chambers, and genders produced different patterns of issues within those policy tweets.

Our results show that Democrats, Representatives, and women are generally more likely to post policy-related tweets. However, the significance of the interaction terms in our LR indicate that these general patterns do not always hold. Rather, the effects of party, chamber, and gender depend on one another. We then examined policy tweets alone and found main effects for party, chamber, and gender on the relative attention topics receive.

Next, we discuss the topics *Health* and *Law and Crime* in more detail to illustrate the patterns we observed and begin to understand why those patterns emerged. We chose these two topics because they allow us to investigate the mechanisms that drive the attention patterns we observed. The two mechanisms we discuss are active legislative debates and public events.

Attention given to the topic of *health* was likely driven by active legislative debates. *Health* received the most attention during the first half of 2017 and then leveled off. We suggest that this one-time peak in attention was related to an intensive Congressional debate about repealing and replacing the Affordable Care Act (ACA) that peaked during the same period. 155,638 tweets were labeled as topic 3. Of these, 32,997 or approximately 21% address the ACA and its debates explicitly (i.e. #ACA, #repealANDreplace, Obamacare, aca). Figure 4 shows the total number of *Health* tweets in relation to ACA-related phrases. We can see that the ACA-related tweets reflect spikes in activity that mirror general spikes in health-related conversation. Though these moments in the tweet stream and legislative debates may not explain the entirety of the *health* spike during the first half of 2017 they suggest a correlation between legislative debates in Congress and the type of content that MCs tweet about that is worthy of further investigation.

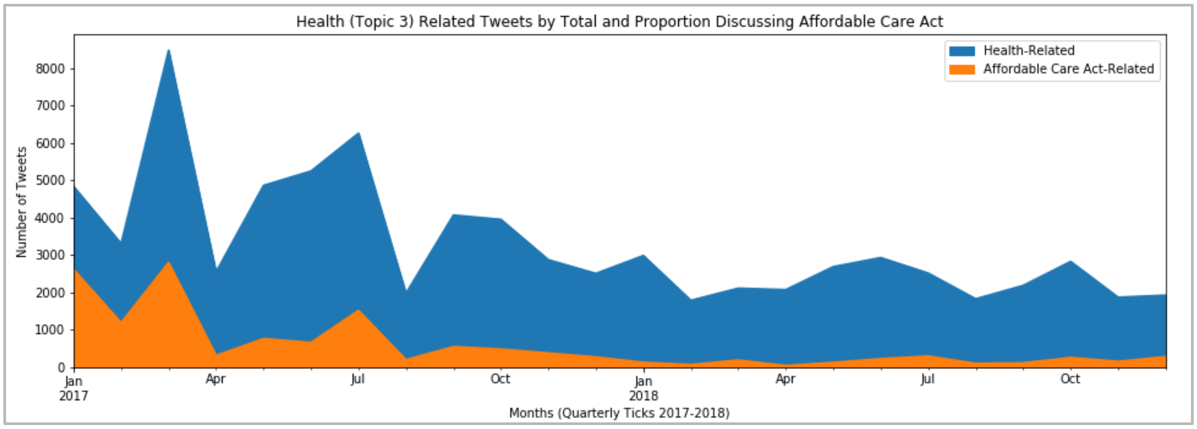

Figure 4. Frequency of tweets labeled *health* during the 115th Congress. The orange area marks tweets that explicitly mentioned the Affordable Care Act and/or related repeal efforts.

Public events are another likely driver of political attention, and topic 12, *law and crime*, illustrate the impact public events can have on Congress's attention. With *law and crime*, we observed an increase in attention during the first half of 2018 that then declines during the rest of the year (see figure X). Even this decrease results in higher attention to *law and crime* than throughout 2017. but remains higher overall than 2017 rates of attention.

Unlike *Health*, whose attention patterns have clear relationships to legislative action, *law and crime*'s attention patterns do not have immediate relationships to legislative action. To understand what may drive these patterns, we examined the specific tokens associated with *law and crime* to understand what subtopics exist in the broader discussions and plotted the frequencies of the top words associated with the topic in Figure X as well. We see, for example, that the word "gun" does not appear in the top 10 words mentioned among MC tweets until 2017 Q4, at which time it becomes the most frequently mentioned word. It continues increasing in frequency of use through 2018 Q1 and remains the most frequently used term through that quarter. In 2018 Q2, the word "children" replaces "gun" as the top term, pushing "gun" to second place. Several significant gun-related events occured in the U.S. during this period including a mass shooting as a concert in Las Vegas in October 2017 and a mass shooting at a high school in Parkland, Florida in February 2018. It is likely that these events are correlated with increased attention to keyword "gun" and then to "children". These findings suggest a correlation between public events and MCs' policy attention on Twitter.

The patterns among individual words under *law and crime* also suggest that our model is able to detect and reveal different framing strategies within broader political topics. Figures 6 and 7, show the frequencies of the *law and crime* topic and the top words that appear within it for each of the two major parties. We see that the word "gun", for example, does not appear at all in the top 10 words mentioned among Republicans during the 115th Congress. In contrast, at the time in which "gun" and "violence" spike in frequency of appearance among Democrats (2017 Q4, 2018 Q1, and 2018 Q2), we see the word "law" spiking among Republicans. Table 5 displays a comprehensive summary of the 10 most frequent words used in *law and crime* tweets during each quarter and whether they are common between parties. Since it is a summary of words in each quarter and not all words overlapped, the list includes more than 10 total words. We can see that there are a number of words that are used by only one party. For example, Republicans use the terms "border", "bipartisan", "communities", and "enforcement" where Democrats do not. In contrast, Democrats use the terms "assault", "ban", "children", and "policy" where Republicans do not. From this preliminary analysis, we begin to see distinct differences in

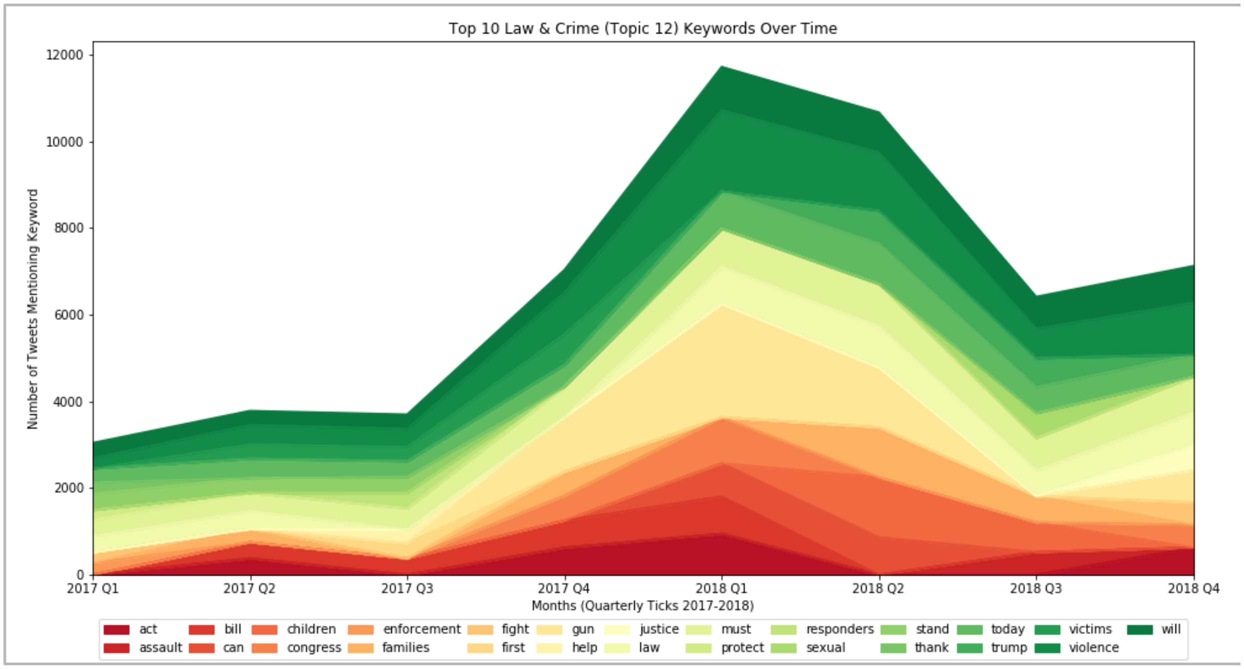

Figure 5. Frequency of tweets labeled *law and crime* during the 115th Congress. Each shaded area indicated the proportion of tweets that contained a specific word associated with the topic.

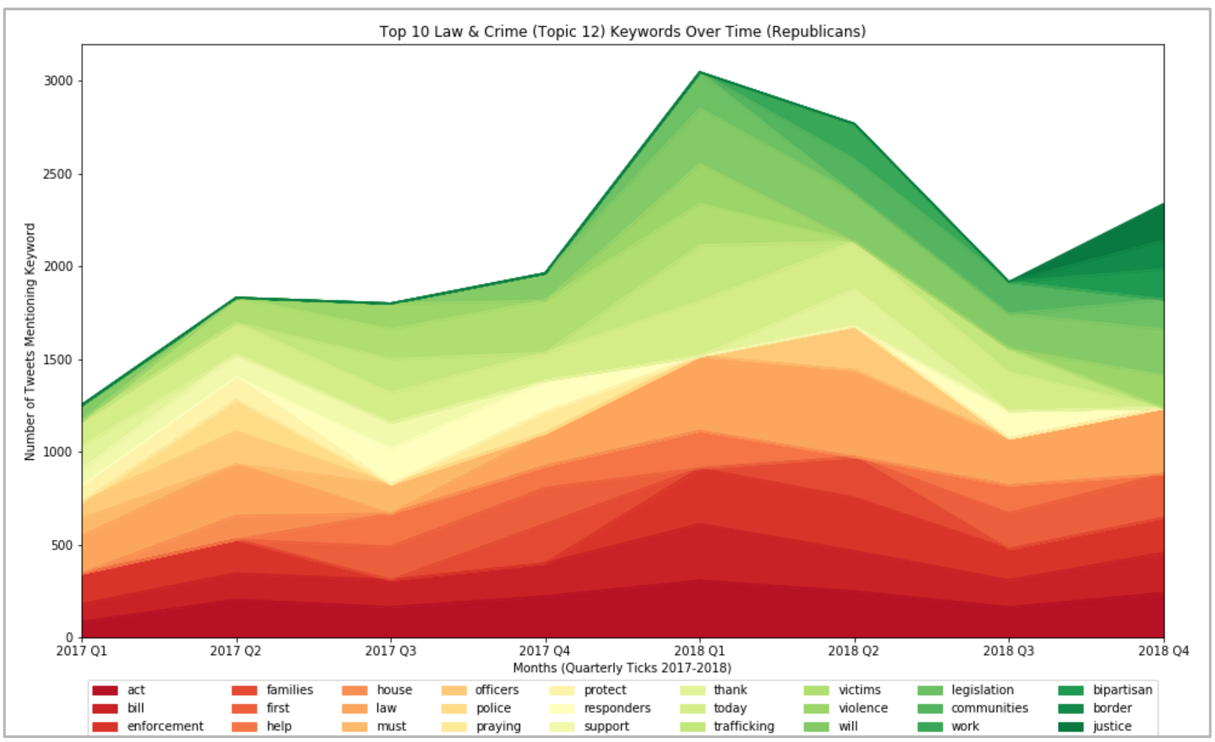

Figure 6. Frequency of tweets labeled *law and crime* and posted by Republicans during the 115th Congress. Each shaded area indicated the proportion of tweets that contained a specific word associated with the topic.

the *way* that Republicans and Democrats speak about the same policy topic on Twitter. This finding suggests that our model is useful for measuring political attention and for identifying political frames.

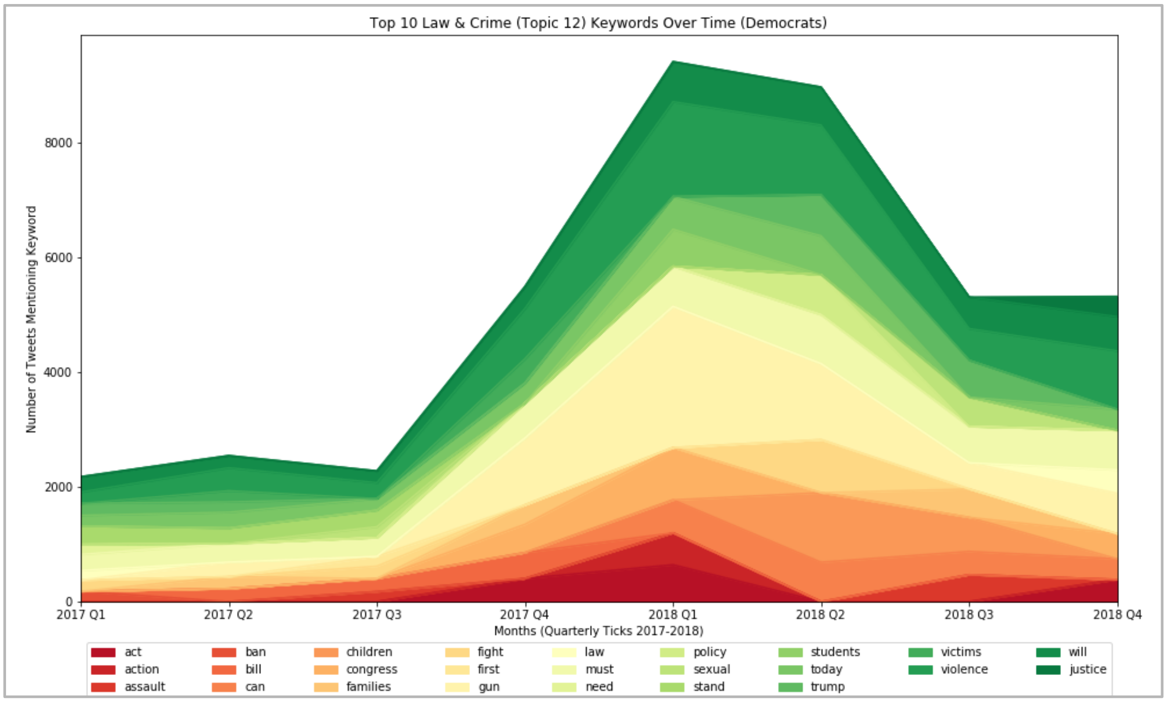

Figure 7. Frequency of tweets labeled *law and crime* and posted by Democrats during the 115th Congress. Each shaded area indicated the proportion of tweets that contained a specific word associated with the topic.

Table 5. 10 most frequent words used during each quarter and whether they are common between parties. Since the top 10 words during each quarter were evaluated and not all words overlapped, this table includes more than 10 words. Those fields marked in red with a “-” symbol indicate that the given word was not included in the party’s top 10 words during any quarter. A green field marked with an “x” symbol indicates that the given word was included in the party’s top 10 words during at least one quarter.

Term	Rep	Dem	Term	Rep	Dem	Term	Rep	Dem	Term	Rep	Dem
act	x	x	enforcement	x	-	need	-	x	thank	x	-
action	-	x	families	x	x	officers	x	-	today	x	x
assault	-	x	fight	-	x	police	x	-	trafficking	x	-
ban	-	x	first	x	x	policy	-	x	trump	-	x

bill	x	x	gun	-	x	praying	x	-	victims	x	x
bipartisan	x	-	help	x	-	protect	x	-	violence	x	x
border	x	-	house	x	-	responders	x	-	will	x	x
can	-	x	justice	x	x	sexual	-	x	work	x	-
children	-	x	law	x	x	stand	-	x			
communities	x	-	legislation	x	-	students	-	x			
congress	-	x	must	x	x	support	x	-			

The patterns we observed in the *health* and *law and crime* topics provide initial insights into the mechanisms that drive the patterns of political attention MCs exhibit on Twitter. For these topics, legislative action and public events were influential, and future work could conduct similar analyses for other categories to understand what influences the temporal dynamics of attention, especially given the relatively small variation the general patterns exhibit. We turn now to additional avenues for future research that leverages our model and its unique contributions.

Future work

We experimented with a number of approaches to improving the performance of the model but recommend that future work explore other potential improvements. For instance, including topic vectors or hashtag co-occurrence features may improve these models. In addition to model improvements, we suggest future work should test correlations between legislative action and political attention. For instance, a future study could leverage work on legislative topic categories (Purpura and Hillard 2006) to label both legislation and tweets and compare their distributions. In a similar vein, future work could perform a similar comparison between traditional media attention and Twitter attention to understand relationships between public events and MC attention. Our analysis focused on party, chamber, and gender, but other characteristics of MCs such as their region, the competitiveness of their elections, and their state's and district's local issues likely also impact their attention. Future work could include measures of these and other characteristics to develop a more nuanced understanding of what drives MCs' political attention as expressed online. Finally, based on our preliminary analyses, we demonstrated that our model is capable of capturing differences in both attention and framing, and future work should examine the potential for this frame detection to facilitate framing and messaging research.

Conclusion

Understanding how Members of Congress (MCs) distribute their political attention is key to a number of areas of political science research including agenda setting, framing, and issue evolution. We demonstrated that it is possible to exploit MCs' Twitter behavior to study their political attention and found notable differences in attention between parties, chambers, and genders. We have outlined directions for future research that test hypotheses concerning correlation between legislative debate, national events, and MCs' topic attention on Twitter suggested by our model. Our model enables researchers to efficiently label social media according to established policy area categories and to use those labels to study changes and patterns of political attention.

Appendix

Table A1. Topic Distribution and Top Term Features Associated with Each Policy Code

CAP #	CAP Label	Freq.	Prop.	Associated terms
1	Macro-economics	115,258	10.9%	budgetconference, unemployment, fiscalcliff, budgetdeal, renewui, manufacturing, budget, taxreform, sequestration, fiscal, debt, debtceiling, debtcrisis, sequester, dontdoublemyrate
2	Civil rights	80,958	7.7%	enda, passenda, nsa, marriageequality, surveillance, paycheckfairness, vra, abortion, lgbt, stopthebans, txlege, marchonwashington, marriage, snowden, talkpay
3	Health	155,638	14.7%	defundobamacare, obamacare, healthcare, obamacares, medicare, mentalhealth, nih, makedclisten, obamacareinthreewords, ocare, cancer, autism, stoprxdrugabusewv, flood, nsa
4	Agriculture	18,380	1.7%	farmbill, gmo, freedomtofish, fisheries, farm, sugar, agricultural, catfish, fishermen, crop, monsanto, fishing, nominee, stamp
5	Labor	41,067	3.9%	fmla, minimumwage, laborday, wia, raisethewage, familyact, righttowork, minimum, miners, pensions, cojobs, nlrb, jobs, obamacare
6	Education	43,748	4.1%	talkhighered, studentloan, tuition, studentloans, dontdoublemyrate, prekforall, investinkids, headstart, erate, restoreta, educational, stem, tribal, walmart
7	Environment	31,176	2.9%	actonclimate, climate, leahysummit, chemsafetyact, climatechange, chesbay, oilspill, keeptahoeblue, pollution, riograndedelnorte, tsca, carbontax, americarecyclesday, brownfields, nsa
8	Energy	22,401	2.1%	energyefficiency, hydropower, reca, helium, coal, energyindependence, biofuels, tva, americanenergy, keystonepipeline, cleanenergy, whitehouses, shovel, nebraska
9	Immigration	35,416	3.3%	immigration, immigrationreform, cirmarkup, cir, immigrants, momento, amnesty, immigrant, dreamers, hoevencorker, daca, econ, farmbill, nsa, acted
10	Transportation	17,903	1.7%	skagitbridge, obamaflightdelays, bridgeact, faa, airport, thud, transportation, maritime, highway, airports, harbormaintenancetax, rail, amndmnt, alleviate

12	Law and crime	88,135	8.3%	vawa, gun, voicesofvictims, guncontrol, gunsense, gunviolence, guns, passmjia, mjia, adoption, msa, sexualassault, nra, firearms, amndmnt
13	Social welfare	16,172	1.5%	snap, foodstamps, hungry, freecellphones, poverty, nutrition, seniors, older, socialsecurity, nationalservice, protectseniors, disappointment, welfare, hunger
14	Housing	5,514	0.5%	gsereform, fha, housing, fhfa, homeless, homelessness, liheap, mortgage, gse, revitalization, walkable, homeowners, affordablehousing, ndhfa
15	Domestic commerce	52,471	5.0%	fixflood, safechemicalsact, detroitbankruptcy, sandy, sandyrecovery, flood, fixfloodinsurancenow, patent, smallbiz, fema, smallbusiness, tourism, sandyaid, startupact, appears
16	Defense	148,929	14.1%	stolenvalor, drones, veterans, drone, ndaa, backlog, veteransday, endthevabacklog, missiletonowhere, nomination, assault, obamacare, immigration, energy
17	Technology	11,700	1.1%	marketplacefairness, nonettax, broadband, cyber, internetsalestax, cable, mfa, marketplacefairnessact, fcc, cybersecurity, internet, nasa, bolden, commissioners
18	Foreign trade	3,964	0.4%	trade, export, drywall, exports, olympic, overseas, concerning, automakers, currency, event, plants, connecticut, buyamerican, arms
19	International affairs	58,220	5.5%	benghazi, syria, egypt, waterstrategy, alqaeda, foreignrelations, israel, libya, ukraine, standwithisrael, nuclear, immigration, obamacare, nsa
20	Government operations	92,833	8.8%	nomination, irsscandal, postal, filibusterreform, irs, nominations, inform, nominee, gopshutdown, endgridlock, perez, confirmation, judges, nuclearoption, nominees
21	Public lands	17,807	1.7%	commissiononnativechildren, indiancountry, native, wildfires, monument, indian, wrda, wildfire, park, fundourparks, forest, parks, grazing, trashed

References

Aldrich, John H. 1995. *Why Parties?: The Origin and Transformation of Political Parties in America*. University of Chicago Press. https://market.android.com/details?id=book-DUah_LP8qdUC.

Atkinson, Mary Layton, and Jason Harold Windett. 2018. "Gender Stereotypes and the Policy Priorities of Women in Congress." *Political Behavior*. https://doi.org/10.1007/s11109-018-9471-7.

Bennett, W. Lance. 1990. "Toward a Theory of Press-State Relations in the United States." *Journal of Communication* 40(2): 103–127. http://dx.doi.org/10.1111/j.1460-2466.1990.tb02265.x.

Bernhard, William, and Tracy Sulkin. 2018. *Legislative Style*. University of Chicago Press. https://market.android.com/details?id=book-YRNDDwAAQBAJ.

Bird, Steven, Ewan Klein, and Edward Loper. 2009. *Natural Language Processing with Python: Analyzing Text with the Natural Language Toolkit*. "O'Reilly Media, Inc." https://market.android.com/details?id=book-KGIbfiiP1i4C.

Boydstun, Amber E. 2013. *Making the News: Politics, the Media, and Agenda Setting*. University of Chicago Press. https://market.android.com/details?id=book-SSE3AAAAQBAJ.

Bruns, Axel, and Tim Highfield. 2012. "Blogs, Twitter, and breaking news: The produsage of citizen journalism." In *Produsing Theory in a Digital World : The Intersection of Audiences and Production in Contemporary Theory*, ed. Rebecca Ann Lind. New York: Peter Lang Publishing, Inc., p. 15–32. http://snurb.info/files/2012/Blogs,%20Twitter,%20and%20Breaking%20News.pdf.

Carson, Jamie L., Erik J. Engstrom, and Jason M. Roberts. 2007. "Candidate Quality, the Personal Vote, and the Incumbency Advantage in Congress." *The American political science review* 101(2): 289–301. https://www.cambridge.org/core/journals/american-political-science-review/article/candidate-quality-the-personal-vote-and-the-incumbency-advantage-in-congress/6189B5B91F87C4FBE046CE95D96AEE8B (Accessed March 23, 2019).

Chi, Feng, and Nathan Yang. 2011. "Twitter adoption in Congress." *Review of Network Economics* 10(1). http://www.bepress.com/rne/vol10/iss1/3.

Cox, Gary W., and Mathew D. McCubbins. 1993. *Legislative Leviathan: Party Government in the House (California series on social choice and political economy; 23)*. University of California Press.

Damore, David F., and Thomas G. Hansford. 1999. "The Allocation of Party Controlled Campaign Resources in the House of Representatives, 1989-1996." *Political research quarterly* 52(2): 371–385. https://doi.org/10.1177/106591299905200206.

Davidson, R., and W. Oleszek. 1990. *Congress and its members*. Congressional Quarterly.

Downs, Anthony. 1972. "Up and down with ecology-The' issue-attention' cycle"." *The Public interest* 28(Summer): 38–50.

Evans, Heather K., and Jennifer Hayes Clark. 2016. "'You Tweet Like a Girl!': How Female Candidates Campaign on Twitter." *American Politics Research* 44(2): 326–352. https://doi.org/10.1177/1532673X15597747.

Gainous, Jason, and Kevin M. Wagner. 2014. *Tweeting to Power: The Social Media Revolution in American Politics*. OUP USA. https://market.android.com/details?id=book-KYdeAQAAQBAJ.

Godin, Fréderic, Baptist Vandersmissen, Wesley De Neve, and Rik Van de Walle. 2015. "Multimedia Lab @ ACL WNUT NER Shared Task: Named Entity Recognition for Twitter Microposts using Distributed Word Representations." In *Proceedings of the Workshop on Noisy User-generated Text*, , p. 146–153. http://www.aclweb.org/anthology/W15-4322.

Golbeck, Jennifer et al. 2018. "Congressional twitter use revisited on the platform's 10-year

anniversary.” *Journal of the Association for Information Science and Technology* 69(8): 1067–1070. http://doi.wiley.com/10.1002/asi.24022.
Grimmer, Justin. 2013. *Representational Style in Congress: What Legislators Say and Why It Matters*. Cambridge University Press. https://market.android.com/details?id=book-O552AgAAQBAJ.
Gulati, Girish J., and Christine B. Williams. 2007. “Closing the gap, raising the bar: Candidate web site communication in the 2006 campaigns for congress.” *Social science computer review* 25(4): 443–465. http://ssc.sagepub.com/content/25/4/443.abstract.
Hemphill, Libby, Jahna Otterbacher, and Matthew Shapiro. 2013. “What’s congress doing on twitter?” In *Proceedings of the 2013 conference on Computer supported cooperative work*, CSCW ’13, New York, NY, USA: ACM, p. 877–886. https://dl.acm.org/citation.cfm?doid=2441776.2441876 (Accessed March 16, 2018).
Hill, Jeffrey S., and Kenneth C. Williams. 1993. “The decline of private bills: Resource allocation, credit claiming, and the decision to delegate.” *American journal of political science*: 1008–1031.
Jones, Bryan D., and Frank R. Baumgartner. 2005. *The Politics of Attention: How Government Prioritizes Problems*. University of Chicago Press. https://market.android.com/details?id=book-HPYdDVu_ghMC.
Jungherr, Andreas. 2014. “The Logic of Political Coverage on Twitter: Temporal Dynamics and Content.” *The Journal of communication* 64(2): 239–259. http://onlinelibrary.wiley.com/doi/10.1111/jcom.12087/abstract (Accessed May 27, 2015).
Lassen, David S., and Adam R. Brown. 2011. “Twitter: The Electoral Connection?” *Social science computer review* 29(4): 419–436. https://doi.org/10.1177/0894439310382749.
Lee, Frances E. 2004. “Bicameralism and Geographic Politics: Allocating Funds in the House and Senate.” *Legislative Studies Quarterly* 29(2): 185–213. http://doi.wiley.com/10.3162/036298004X201140.
Mayhew, David R. 1974. *Congress: The Electoral Connection*. Yale University Press. https://market.android.com/details?id=book-j17QomTrD1EC.
McHugh, Mary L. 2012. “Interrater reliability: the kappa statistic.” *Biochemia medica: casopis Hrvatskoga drustva medicinskih biokemicara / HDMB* 22(3): 276–282. https://www.ncbi.nlm.nih.gov/pubmed/23092060.
Mikolov, Tomas, Kai Chen, Greg Corrado, and Jeffrey Dean. 2013. arXiv [cs.CL] “Efficient Estimation of Word Representations in Vector Space.” http://arxiv.org/abs/1301.3781.
Neiheisel, Jacob R., and Sarah Niebler. 2013. “The Use of Party Brand Labels in Congressional Election Campaigns.” *Legislative Studies Quarterly* 38(3): 377–403. http://dx.doi.org/10.1111/lsq.12019.
Parameswaran, Giri. 2018. “Bargaining and Bicameralism.” *Legislative Studies Quarterly* 43(1): 101–139. http://doi.wiley.com/10.1111/lsq.12179.
Pedregosa, Fabian et al. 2011. “Scikit-learn: Machine Learning in Python.” *Journal of machine learning research: JMLR* 12(Oct): 2825–2830. http://www.jmlr.org/papers/v12/pedregosa11a.html (Accessed May 27, 2016).
Pennebaker, J. W., J. Booth, and M. E. Francis. 2015. *LIWC2015*. https://liwc.wpengine.com/.
Purpura, Stephen, and Dustin Hillard. 2006. “Automated classification of congressional legislation.” *Proceedings of the 2006 national conference on Digital government research - dg.o ’06*. http://dx.doi.org/10.1145/1146598.1146660.
Rocca, Michael S., and Stacy B. Gordon. 2010. “The Position-taking Value of Bill Sponsorship in Congress.” *Political research quarterly* 63(2): 387–397. https://doi.org/10.1177/1065912908330347.
Rocca, Michael S., Gabriel R. Sanchez, and Jason L. Morin. 2011. “The Institutional Mobility of Minority Members of Congress.” *Political research quarterly* 64(4): 897–909. https://doi.org/10.1177/1065912910379225.

Rohde, David W. 1991. “Parties and Leaders in the Postreform House.” *University of Chicago Press*. http://www.press.uchicago.edu/ucp/books/book/chicago/P/bo3638205.html (Accessed June 7, 2017).

Russell, Annelise. 2018. “The Politics of Prioritization: Senators’ Attention in 140 Characters.” *Forum* 16(2): 331–356. http://www.degruyter.com/view/j/for.2018.16.issue-2/for-2018-0020/for-2018-0020.xml.

Russell, Annelise. 2017. “U.S. Senators on Twitter: Asymmetric Party Rhetoric in 140 Characters.” *American Politics Research*: 1532673X17715619. http://dx.doi.org/10.1177/1532673X17715619.

Schattschneider, Elmer Eric. 1960. *The Semisovereign People: A Realist’s View of Democracy in America*. New York: Holt, Rinehart and Winston.

Schofield, Alexandra, and David Mimno. 2016. “Comparing Apples to Apple: The Effects of Stemmers on Topic Models.” *Transactions of the Association for Computational Linguistics* 4(0): 287–300. https://www.transacl.org/ojs/index.php/tacl/article/view/868 (Accessed December 13, 2018).

Shapiro, Matthew A., and Libby Hemphill. 2017. “Politicians and the Policy Agenda: Does Use of Twitter by the U.S. Congress Direct New York Times Content?” *Policy & Internet* 9(1): 109–132. http://dx.doi.org/10.1002/poi3.120.

Shepsle, Kenneth A. 1979. “Institutional Arrangements and Equilibrium in Multidimensional Voting Models.” *American journal of political science* 23(1): 27–59. http://www.jstor.org/stable/2110770.

Shepsle, Kenneth A., and Barry R. Weingast. 1987. “The Institutional Foundations of Committee Power.” *The American political science review* 81(1): 85–104. https://www.cambridge.org/core/journals/american-political-science-review/article/institutional-foundations-of-committee-power/BB794F94572ECAC56F8BDEA3150CCE60 (Accessed March 23, 2019).

Straus, Jacob R., Matthew Eric Glassman, Colleen J. Shogan, and Susan Navarro Smelcer. 2013. “Communicating in 140 Characters or Less: Congressional Adoption of Twitter in the 111th Congress.” *PS, political science & politics* 46(1): 60–66. https://www.cambridge.org/core/services/aop-cambridge-core/content/view/6D65BAB70C1D58D133B70C14F3AEF628/S1049096512001242a.pdf/div-class-title-communicating-in-140-characters-or-less-congressional-adoption-of-twitter-in-the-111th-congress-div.pdf (Accessed August 13, 2018).

Venables, W. N., and B. D. Ripley. 2002. *Modern Applied Statistics with S*. Springer. https://market.android.com/details?id=book-X_KQcDpSB8MC.