12

A COMPUTATIONAL EXPLORATION OF THE EVOLUTION OF GOVERNMENTAL POLICY RESPONSES TO EPIDEMICS BEFORE AND DURING THE ERA OF COVID-19

Philip D. Waggoner

INTRODUCTION

Government in America is an extension of the people. Whether executive, legislative, or judicial, governmental actors at some level are put in place by, and thus act in response to, a voting population who desire responsiveness from their representatives. One of the primary methods of delivering on this representational arrangement is in the form of policymaking, where governmental actors offer policies aimed at correcting or addressing issues in society (Jones et al., 2009). While there are a host of avenues for policymakers to be made aware of pressing societal issues requiring policy action at some level (Waggoner, 2019), some issues are so apparent, there need not be an explicit cue from the public constituency. The COVID-19 epidemic is one of these types of issues, where the effects are so far reaching, governmental response through policymaking is ostensibly *expected*.

Yet, while the need for action may be overt, recent congressional debate of COVID-related legislation has demonstrated that the path to enacting policy responding to COVID-19 is not nearly as clear, simple, or even decorous. For example, congressional debate on the coronavirus economic stimulus bill with a price tag of $484 billion revealed harsh partisan mudslinging by both parties and in both chambers. Rep. Jayapal (D-WA) alleged Republicans offered a "bad bill," with the implication being they (Republicans) do not care as much as we (Democrats) do for families and workers (C-SPAN, 2020a). In a similar tone, but on the other side of the aisle and in the other chamber, Sen. Barrasso (R-WY) criticized the Democrats' approach to the issue, aided by a poster entitled "Pelosi is on Fantasy Island." His remarks included inflammatory terms that resembled the tone of the poster (C-SPAN, 2020b). These examples of harsh partisan debate highlight the fractured context that defines much of American public policymaking today. And though rooted in the representational responsibility of Congress to make policy in response to social and public health issues like COVID-19, the problem itself might be *apolitical*, but the elite response may not.

And to complicate matters, news coverage of highly salient issues like COVID-19 often fan the flames of political division, regardless of the direction of partisan-slant of the news outlet (Larcinese et al., 2011). News media are certainly imbued with a powerful role in the policymaking apparatus to responsibly report government policymaking back to the public. Yet, a grave limitation in this information transfer is a growing perception of biased and partisan-leaning news coverage in both directions (Perryman, 2019).

Thus, a multipart question emerges: First, is governmental policymaking on such widespread, apolitical issues characterized by political division? If so, for how long has this been the case? Put differently, is the brand of public policymaking we anecdotally see and hear about today a function of historical policymaking on similar types of issues? Or, is America in a unique era of division, where policymaking on far-reaching, nonpartisan issues is similarly tainted by partisan division? These questions are grounded in a deep literature finding elite partisan division in specific questions (Souva & Rohde, 2007), as well as in policymaking in general (Layman et al., 2010). And these elite partisan differences and their effects are not beholden to the realm of policymaking, but significantly influence mass behavior and public opinion (Berinsky, 2007; Druckman et al., 2013; Robison & Mullinix, 2016).

This study is aimed at exploring these questions from a broad exploratory lens, where patterns that naturally exist over time are able to emerge. Thus, using a suite of computational techniques, I am interested in exploring the evolution of governmental policymaking on epidemics.

In light of the aforementioned perception of bias that so often characterizes news coverage of such consequential, widespread issues and policy responses, I opt to look to the medium of policymaking itself: proposed bills. Specifically, I leverage an original data set of all U.S. congressional (Senate and House) bill metadata on COVID-19 and epidemics broadly defined from 1973 to 2020. These data are mined for cross-temporal comparison of congressional policymaking on epidemics.

My goal is to allow the policies offered by the elected representatives to speak for themselves, untarnished by news coverage or any perception of biased reporting. This approach will shed important light on two key points: first, whether congressional policymaking on related public health issues is an evolutionary process; and second, the contours of the landscape of policymaking in this "epidemics" issue space.

To explore the evolution of government policymaking responses to epidemics, there are two dimensions of variation of interest for present purposes: time and partisanship. For the time dimension, I address the evolutionary question, which allows for a deeper, contextualized understanding of the current policymaking climate in American politics in the era of COVID-19. The second dimension of partisanship is closely linked with the first. Namely, I am interested in exploring not only whether the types of bills introduced on addressing epidemics have changed over time, but especially in the modern, hyperpolarized era. The second part of the goal, then, is to detect whether and to what degree partisan differences appear.

Over the five stages of analysis detailed below, there were several striking patterns that emerged. Notably, the "what" of the policy substance remained relatively stable over time. That is, members of both parties tend to focus their policies on the epidemic in question, using terms related to the given epidemic. However, the "how" changes and grows steadily over time. In the earliest days of the study period, the tone of the policies was remarkably neutral. This trend faded away in favor of more pronounced sentiment over time, culminating in the starkest period of *negative* sentiment in the current COVID era. The trend was present for members of both parties and across both chambers.

Diving into the current COVID era explicitly, bigram network models showed that members of both parties tended to use terms that appeal to their bases in crafting bill descriptions. For example, Republicans invoked "China" and "small business," whereas Democrats invoked terms like "Medicaid" and "fair housing." An additional striking pattern is that Republicans are much more homogenous within their ranks as to the number and types of terms used. This is in comparison to Democrats, who use a much wider set of terms and cover many more topics in their policies. These patterns are in line with similar research demonstrating that the Democratic base is more fractious compared to the Republican base, which tends to be focused more on ideological purity and consistency (Grossmann & Hopkins, 2016).

EMPIRICAL STRATEGY

This exploration begins with the time dimension, and then is followed by the partisan dimension, and is organized by two time periods: *pre-COVID* (1973–2018) and *COVID* (2019–2020). There are five sections comprising the analysis: first, descriptive differences between the *COVID* and *pre-COVID* eras; second, topic models (the "what" question) and sentiment analysis (the "how" question) by decade; third, sentiment analysis by decade *and* party, bringing in the partisan dimension; fourth, deeper exploration into topic models for the *COVID* era only; and finally, bigram networks for the *COVID* era only.

Data and Preprocessing

The data used in this project include metadata on all bills related to (1) COVID-19 (spanning 2019 to 2020), and (2) epidemics broadly defined over a longer period, from 1973 to 2018 (i.e., policymaking in the *pre-COVID* era). These data were scraped from congress.gov and are also available in the C-SPAN Archives. The bill-level data includes several useful features: Congress number (e.g., 115th), year sponsored, descriptive bill title (different from and longer than short bill title), primary bill sponsor (name, district/state, and party affiliation), date of bill introduction, number of cosponsors, initial committee assignment, date of most recent action, and the most recent action (e.g., referred to another committee).

From the bill data, a corpus was constructed based on the long, or "descriptive," bill titles. In some cases these titles, which act as brief summaries of the bills, are dozens of words in length. Thus, this choice was largely made for reasons of computational efficiency, such that if the full bill text were used, not only would the bill text offer a noisier signal as to the intent, and tone impacted by the legal jargon comprising congressional bill text, but also the massive size of the corpus would have led to an infeasible processing task for most personal computers. Substantively, long bill titles are carefully developed to give a summary of the full bill, thereby offering a signal of authors' intentions and goals behind writing the bill in the first place.

With the corpus of long bill titles constructed, I preprocessed and staged the data in line with traditional text mining techniques, including removing stop words (extraneous terms like the articles "the" and "a," but also domain-specific terms like "act" and "bill" that fail to add substantive meaning to the text), removing numbers and punctuation, stripping white space left behind from preprocessing, and performing various other cleaning tasks. The result is a corpus that is a bag of words, wherein word order is not important compared to the inclusion of words.

The full corpus was then staged as a document-term matrix (DTM), where documents (bills) are rows and individual terms are columns. Elements of the matrix are term frequencies. DTMs are required for fitting topic models. The other two techniques described below, sentiment analysis and bigram networks, do not require the data to be staged as a DTM, but rather require the corpus to be tokenized, or broken down into smaller chunks of text. For my purposes, I used two tokenizers for these stages respectively: *word* (single words) and *bigram* (two-word combinations).

Methods

Though deployed across five stages, there are three main text mining techniques used in this chapter: topic models, sentiment analysis, and bigram networks.

First, regarding topic models, there are a variety of ways of thinking about and modeling topic structure in text. But in general, most of these methods share the same goal: to uncover the latent structure of topics that define the "what" of a corpus—that is, the topics underlying bill long titles. Topic models of this sort are considered unsupervised, where there is no ground truth conditioning the modeling process, as well as a lack of an expected outcome from the run of the algorithm. Rather, the core assumption of topic models is that some structure of topics is *latent* and exists across the full document space. So, the task is to uncover these topics that likely characterize the space most efficiently. Importantly, as this is an unsupervised task, there is no set number of topics that formally defines the space; there are no labels. Rather, there is some configuration of topics that likely exist and precede production of the documents and words themselves. The goal is to recover this latent topic structure.

The topic model leveraged in this project is latent Dirichlet allocation (LDA) (Blei et al., 2003). In brief, LDA is an algorithm that starts with assuming a mixture of topics, *k*, which defines the document-feature space. Assuming the topics and topic memberships are Dirichlet distributed, the goal is to find the configuration of topics that represent the space the best. "Best" defined here is the unique combinations of words contributing to each topic. Each topic, then, is defined by a combination of words that frequently co-occur to some degree of proportion. For example, a topic relating to "America" might have the terms "United" and "States" associated with it to high degrees. Then, at the aggregate level, the optimal set of topics defining the corpus is a blend of topics that are individually compact, and well-separated from all other topics. This result would suggest not only that the topics are well-defined but that the corpus is clearly composed of a set of topics, as opposed to being a more opaque blend of topics.[1]

The next method used is sentiment analysis, or "sentiment scoring." This method measures the overall tone of a corpus based on the frequency of words that occur in the corpus as well as appear in a sentiment dictionary. A common use of sentiment analysis is to score some text as more "positive" or "negative" overall based on frequency of "positive" terms versus "negative" terms. Scoring is carried out based on the choice of tokenizer, which is the size of text into which the full corpus is broken down. For my purposes, all sentiment scores are based on a *word* tokenizer for a more granular look at the text. This is compared to many other possible tokenizers, such as scoring by sentences or even full paragraphs. The idea is that the algorithm uses a supplied dictionary of words that are scored as either "positive" (1) or "negative" (–1) and then scores words accordingly in the corpus that also appears in the dictionary. I use the Bing dictionary for all analyses that leverage sentiment analysis (Liu, 2012). For example, suppose a document includes the term "happy." If this term is also included in the sentiment dictionary and is scored as a "positive" (1) word, then this word gets a score of 1 in the text. The final step is to sum and average the scores to give a summary of the sentiment of a corpus, which in my case is either more negative or positive on balance. This is a simple, yet powerful approach to understand the tone and thus the "how" of a set of documents.

Lastly, I use bigram networks in the final stage. These networks are similar to topic models. Yet, instead of searching the space for an optimal configuration of topics that are defined by a set of words that frequently co-occur, bigram networks build a network representation of connections between the usage of terms (two, to be precise; hence *bi*gram). The nodes in the network represent the use of a term, and the edges represent the connections between the usage of multiple terms. Edges can be weighted to capture the frequencies of term co-occurrence, as I demonstrate below. These are extremely valuable for visualizing how terms that occur in a common space are linked to usage of other terms. This gives unique insight into the focus of the full document space.

ANALYSIS AND RESULTS

Descriptively Exploring the "Epidemics" Space

In the first stage of analysis, I present a high-level look at policymaking on epidemics across the two main periods: *COVID* (2019–2020) and *pre-COVID* (1973–2018). The purpose of this first stage is to offer a launching place to understand subsequent results exploring whether differences in policymaking exist over time. Importantly, in this first stage I am not yet looking at parties. Rather, I am setting the stage for exploring the first dimension of "time," which addresses the evolutionary question. Descriptive trends are presented in two word clouds in Figure 12.1, with the COVID era (a) and the pre-COVID era (b). Note that in light of limitations in diagnosing word clouds, bar plots of the top terms used at least 150 times are presented in Figures A.1 and A.2 in the Appendix to this chapter.

A few notable trends emerge. First, a rallying call is present in both eras, including terms like "emergency," "national," "supporting," "resolution," and so on. This is in line with naive expectations on government policymaking related to major epidemics, where the government is fulfilling its representational duty to respond to a crisis, while also signaling shows of strength and unity.

Further, it is interesting to note that in the pre-COVID era (Figure 12.1 [b]) the terms related to the epidemic in question are used. For example, terms like "hiv," "aids," and "drug" are used. This is in comparison to much less frequently used terms that might be associated with COVID-19, such as "COVID" or "coronavirus." Rather, the COVID-era plot (Figure 12.1 [a]) seems to focus more on relief-type legislation and response, which makes sense given the unprecedented widespread nature and impact of COVID-19.

Topic Models and Sentiment Analysis Over Time

Building on the descriptive patterns discussed in the previous section and shown in Figures 12.1 (a) and (b), I now shift to probe the "what" and "how" questions explicitly. I start with constructing topic models by decade to explore the "what"

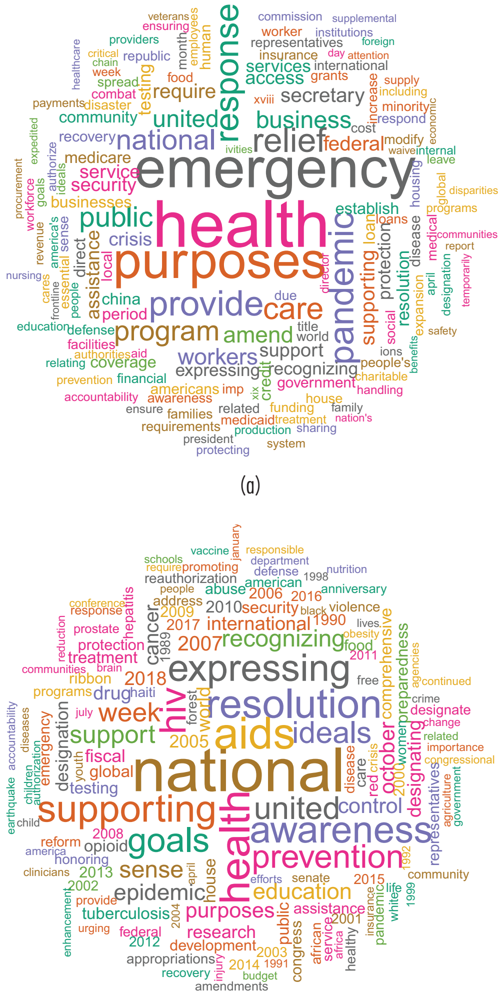


FIGURE 12.1 Word clouds of most frequently used terms: (a) COVID era; (b) pre-COVID era.

question pertaining to the topics that are present in the legislation across all periods in the data set. I then pivot to the "how" question by leveraging sentiment analysis, which will build on the "what" and give a clue as to the general tone of these bills on epidemics across the full study period.

Notably, with these two analytical approaches, I am interested in the evolutionary or "time" dimension discussed above. Topic models will help address related questions like Does evolution exist, or are bill topics relatively stable? and

Are foci of topic structures similar between eras or not? Sentiment analysis will also help address the evolutionary question, but in addition it will help address slightly different questions like What is the tone, and does it shift over time? and Do we see differences across chambers?

First, I present the results of the topic models. Recall that the goal of topic models is to find the optimal latent topic structure that likely defines a corpus. As this is an unsupervised problem, though there are many ways to think about optimality. For my purposes, I calculate and compare perplexity scores, which describe how well a model predicts some sample. Note, LDA models are generative, meaning they are interested in predicting distributions, which in my case is a mixture of topics in a single space. Calculating multiple perplexity scores varying the number of topics, *k*, in the mixture, I will pick the value of *k* for which perplexity is smallest, signaling that mixture of topics does the best job of predicting the full sample of terms. The optimal perplexity score, and value of *k*, varies across decade subsamples. These scores are presented in Figure 12.2. I then used the optimal k values for each of the respective topic models fit to each subsample of bills from each respective decade. The top words in each decade across each topic are presented in Figure 12.3.

A few notable trends are clear in the terms that characterize the different topics over time. First and foremost, in addition to the perplexity values in Figure 12.2, it is clear when zooming in on decades/periods that different topic structures define different periods. This is an important pattern as it provides a first clue that policymaking on epidemics is not a static endeavor. This initial signal would have been lost if a global topic model were fit on the full document space.

In the COVID era in plot (a) in Figure 12.3, four topics are addressing four distinct areas (a pattern that is corroborated by the clearly lowest value of perplexity at $k = 4$): topic 1 involves domestic relief for businesses, Medicaid, and general emergency response; topic 2 involves global security and health, seen by the three terms comprising the topic; topic 3 involves China and international affairs; and topic 4 involves workers, care, and assistance. These four topics not only make intuitive sense but they reflect the different ways in which legislators brand their policy proposals. Indeed, some tend to focus on marketing relief effort by focusing on domestic workers for example, whereas other bills tend to focus on the global aspects of the pandemic (topics 2 and 3).

Further (as shown in Figures 12.2 and 12.3), the 2000s decade (plot [c]) is more succinctly defined by only two topics, though less clearly separated compared to the COVID era (plot [a]) or the 1980s (plot [e]). This is seen in the appearance

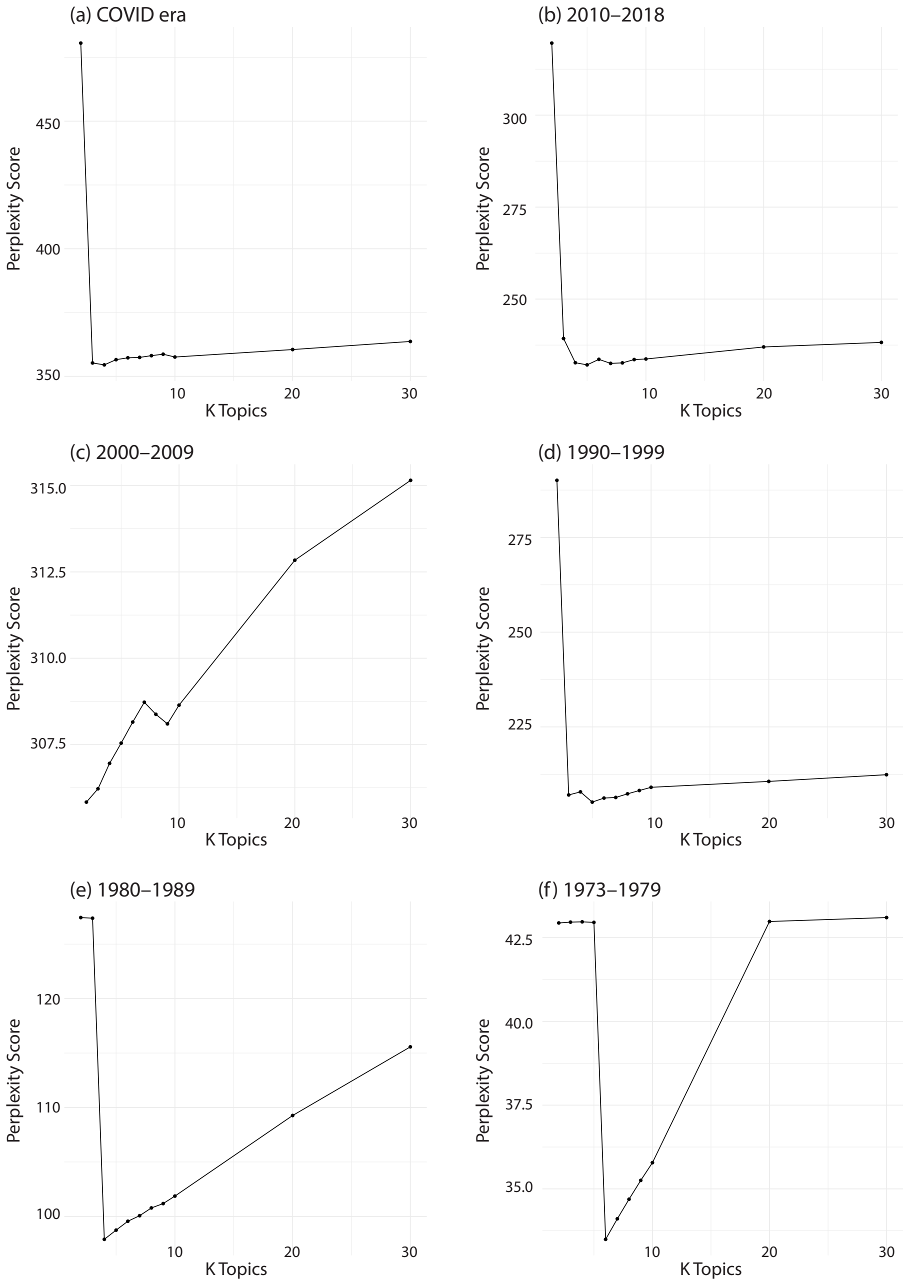
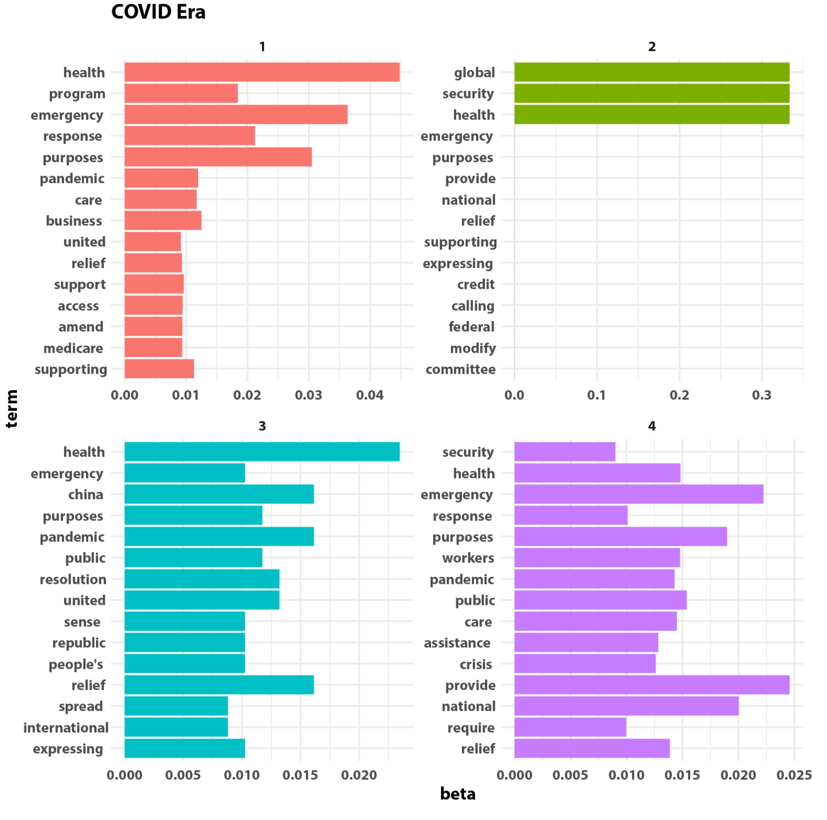


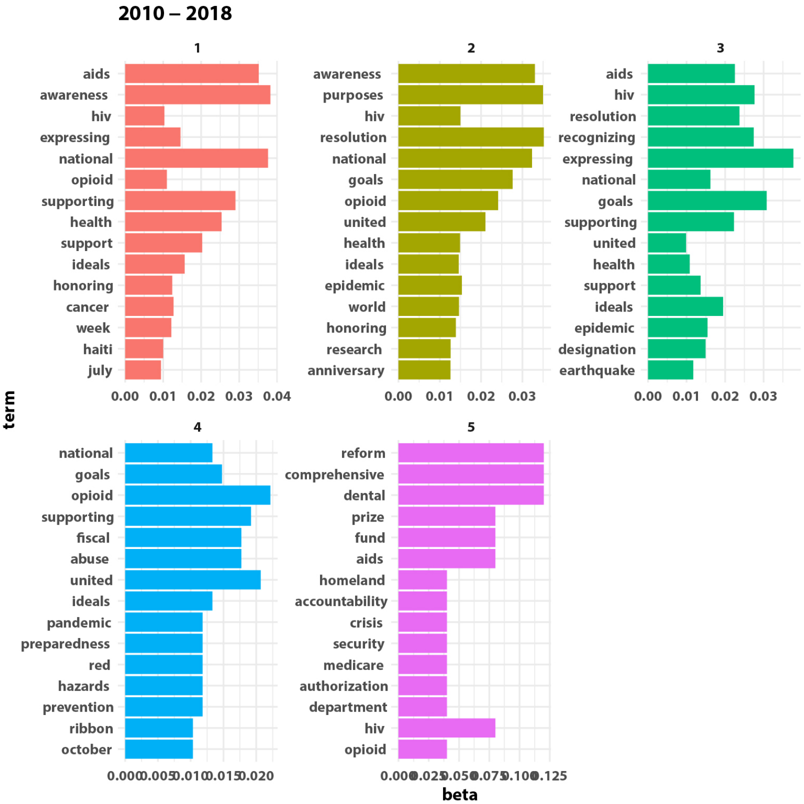
FIGURE 12.2 Perplexity across k topics by decade. (a) COVID era, (b) 2010s, (c) 2000s, (d) 1990s, (e) 1980s, (f) 1970s.

of several of the same terms in both topics. Substantively, this means that there is not a clear topic structure in legislation branded as addressing an "epidemic" in this decade. In the absence of a clear epidemic such as COVID-19, the casting of an epidemic (via use of the term) could be much more widely understood. For example, in plot (c) for the 2000s, topic 1 has terms like security and defense, whereas topic 2 has terms like education and health. Thus, while there may not be clear separation between types of epidemics and thus topics, it is still possible to pick up on temporal cues as to those issues considered as "epidemics" by policymakers at the time.



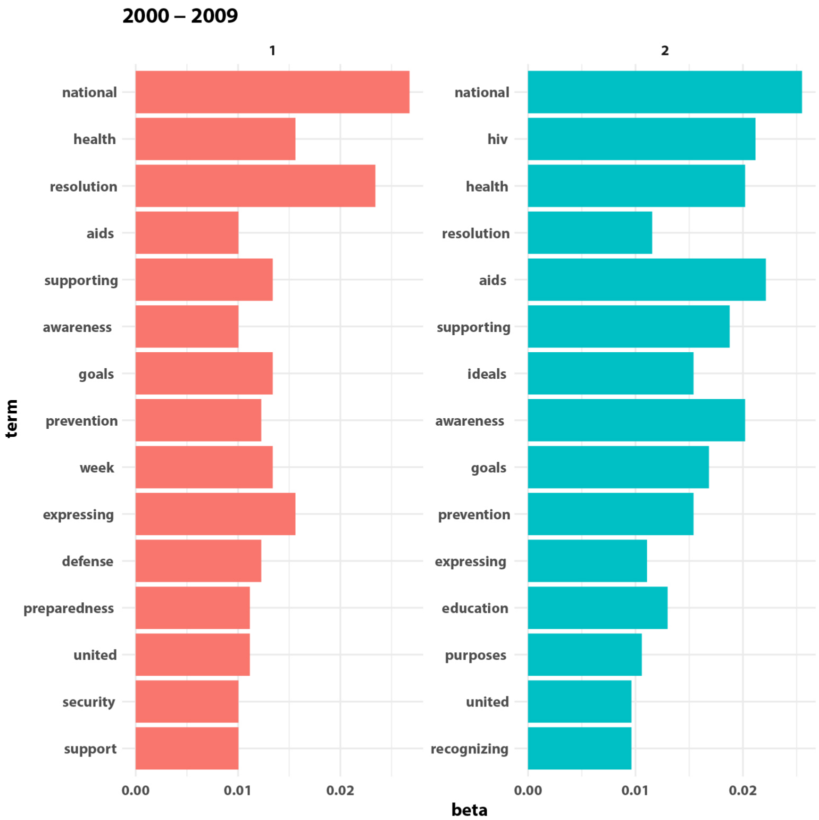
(a)

FIGURE 12.3 Topic model terms at optimal *k*.
(a) COVID era, (b) 2010s, (c) 2000s, (d) 1990s, (e) 1980s, (f) 1970s. (*Figure continued*)



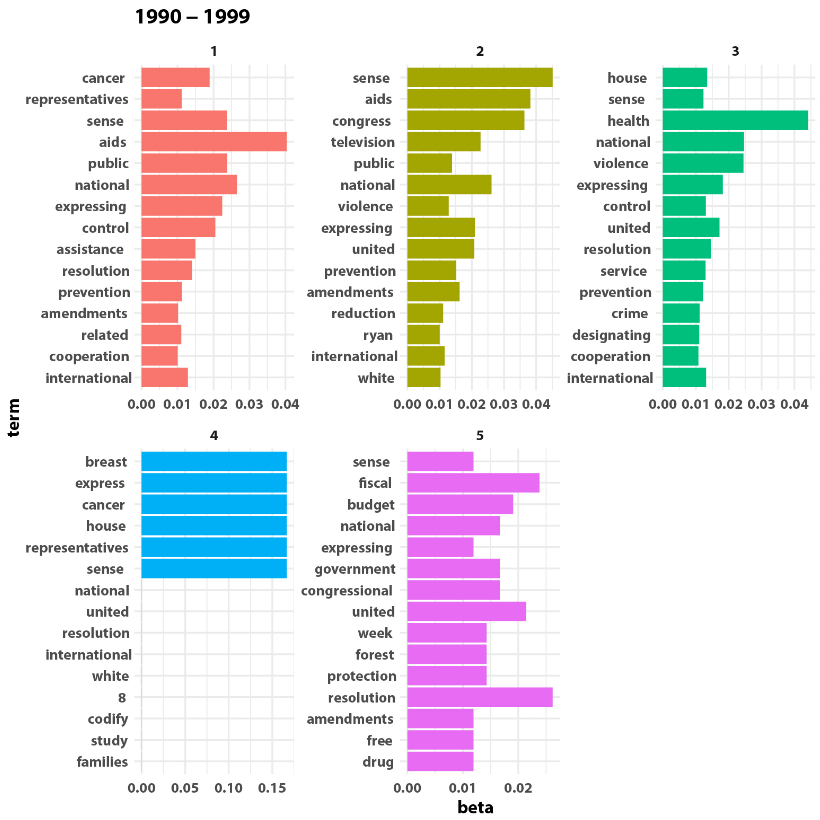
(b)

FIGURE 12.3 Topic model terms at optimal *k* (*continued*).
(a) COVID era, (b) 2010s, (c) 2000s, (d) 1990s, (e) 1980s, (f) 1970s. (*Figure continued*)



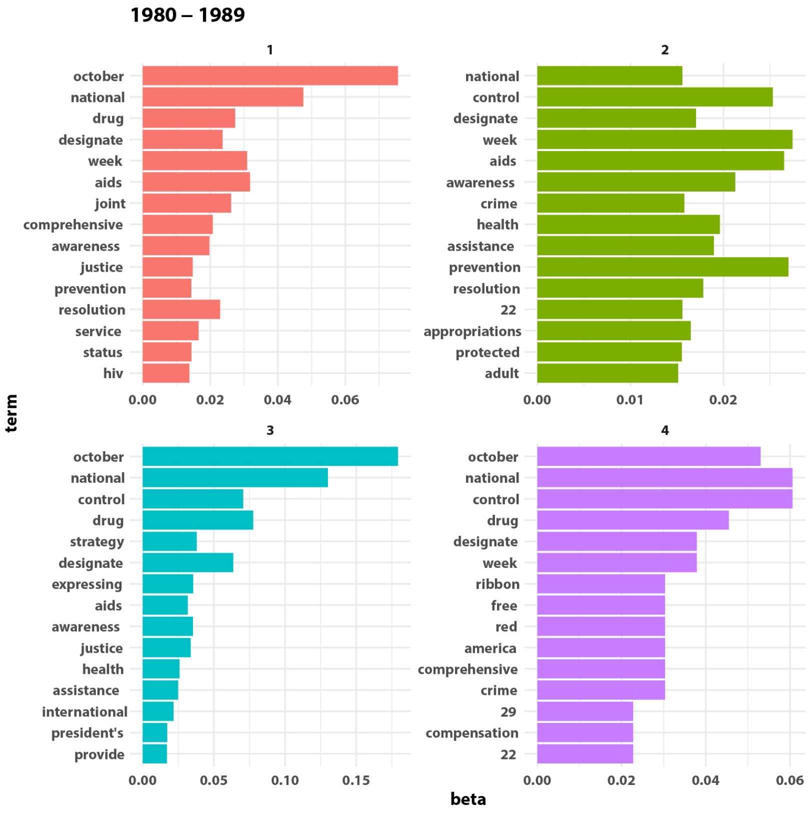
(c)

FIGURE 12.3 Topic model terms at optimal *k* (*continued*).
(a) COVID era, (b) 2010s, (c) 2000s, (d) 1990s, (e) 1980s, (f) 1970s. (*Figure continued*)



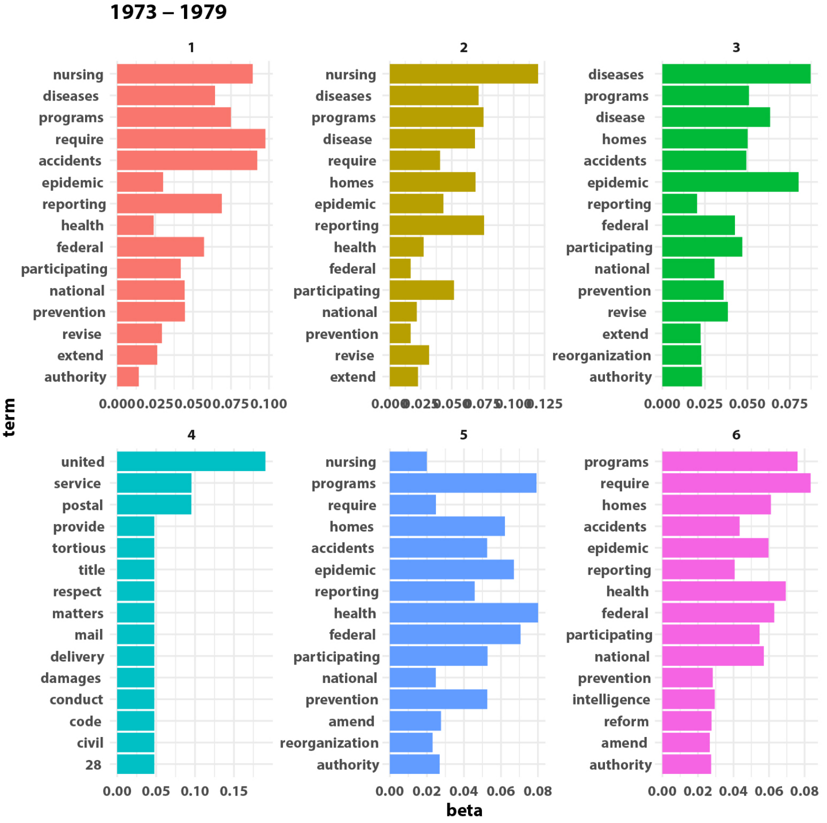
(d)

FIGURE 12.3 Topic model terms at optimal *k* (*continued*).
(a) COVID era, (b) 2010s, (c) 2000s, (d) 1990s, (e) 1980s, (f) 1970s. (*Figure continued*)



(e)

FIGURE 12.3 Topic model terms at optimal *k* (*continued*).
(a) COVID era, (b) 2010s, (c) 2000s, (d) 1990s, (e) 1980s, (f) 1970s. (*Figure continued*)



(f)

FIGURE 12.3 Topic model terms at optimal *k* (*continued*).
(a) COVID era, (b) 2010s, (c) 2000s, (d) 1990s, (e) 1980s, (f) 1970s.

To the evolutionary question at this point, it seems as though the content of proposed legislation tends to vary expressly with the epidemic in question, regardless of the specific epidemic and however broadly or narrowly defined that epidemic may be. Thus, to the "what" question on the topics comprising the introduced legislation over time, it appears as though policymaking is not evolutionary in the sense that trends in preceding time periods overtly spill over to affect topics in subsequent time periods. In other words, the 1990s do not seem *dependent* on the 1980s in the branding and definition of policy responses to epidemics. Rather, the epidemics of the decade are seemingly responded to with policy accordingly.

Building on these relatively stable patterns pertaining to the content of the policy proposals, I pivot now to explore the "how" question to add to the depth of the evolutionary question. More specifically, I am interested in understanding whether the tone in which policy responses to epidemics is evolutionary such that tone type (positive/negative) as well as intensity (proportion of positive/negative tone of the overall document space) builds over time, remains relatively stable, or decreases over time.

To accomplish this task I conducted a sentiment analysis by Congress, the results of which are presented in Figure 12.4.

Before discussing patterns, it is important to note that the sentiment scores presented in Figure 12.4 are scaled (divided by its standard deviation) but not mean-centered. This choice was made to account for variance in overall sponsorship rates over time, as well as across chambers, where the Senate typically introduces fewer bills than the House given the smaller size of membership. Further, the scores are disaggregated by chamber (yet not by party at this point), with dark gray for the U.S. House and light gray for the U.S. Senate. To read Figure 12.4, values below the 0.0 cut point suggest greater negative sentiment in the given chamber's sponsored bills on epidemics for the given Congress (two-year period). Values above the 0.0 dashed line point to greater positive tone for the given chamber and Congress.

The pattern indeed appears evolutionary, where in the earliest days of the study period (1970s–1980s), the tone of bills is largely neutral, with relatively small dips below and rises above the 0.0 cut point. Indeed, in some Congresses there were no sentiment scores registered, implying ultimate neutrality in tone. The intensity grows over time, picking up in the 1990s and culminating in the largest *negative* dip in both chambers in the current COVID era (the bottom right plot in Figure 12.4).

This pattern in bill sentiment is notably different from the patterns from the topic models, where different topic structures define different decades and different terms made up the topics by decade as well. Rather, regarding the "how" question pertaining to the tone of the bills on epidemics, we see a steadily building *intensity* in tone, both positive and negative, across both chambers over time.

At this point, a few key trends are clear. First, the topics of the proposed bills do not substantively deviate from the epidemics at hand (e.g., topics tend to focus on whatever the given epidemic is), implying little change in the types of policy being offered. Yet, when considering the tone or "how" of the bills, there

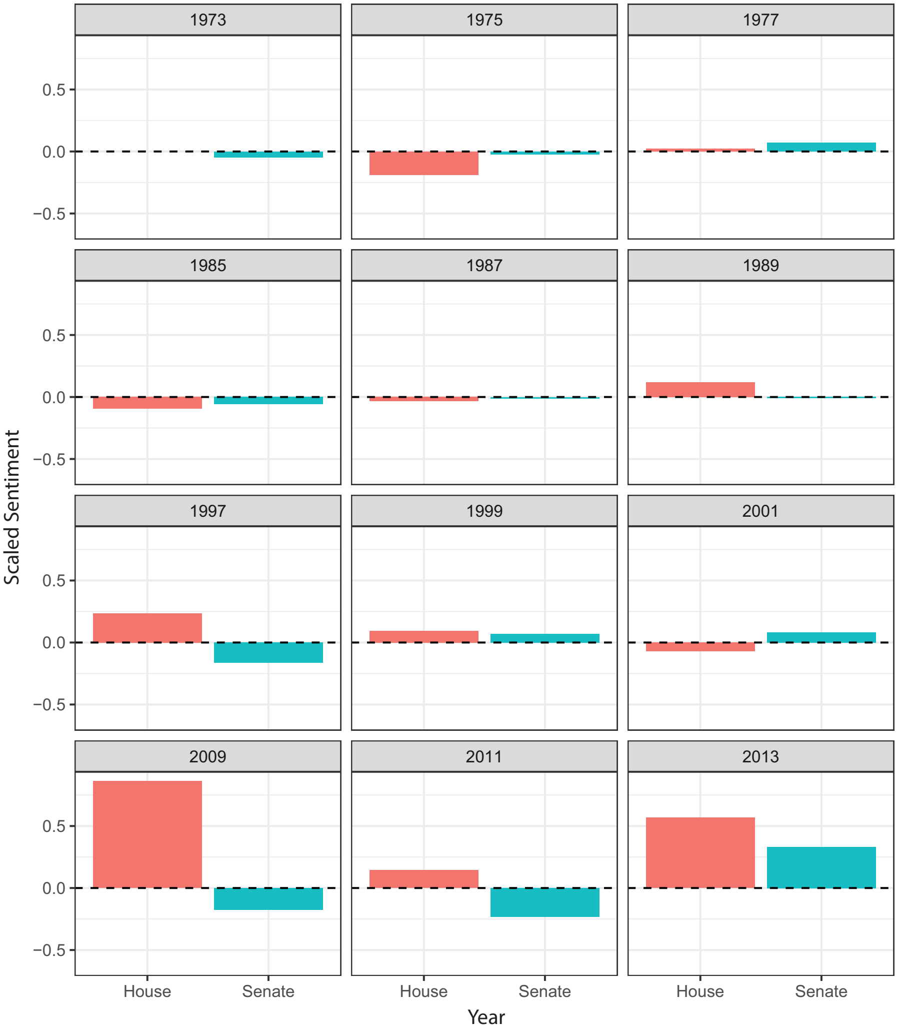

1973
1975
1977
1985
1987
1989
1997
1999
2001
2009
2011
2013
0.5
0.0
−0.5
Scaled Sentiment
House
Senate
Year

FIGURE 12.4 Scaled sentiment scores over time, by chamber.

appears to be an evolution in the tone of the policymaking. In the earliest days, the sentiment was largely neutral or absent entirely, with the intensity of tone of epidemic-related legislation increasing in more recent years. There was a spike in overly positive tone in the mid-2000s, and then a bottoming out of tone for both chambers in the COVID era (2019–2020).

In sum, the results from these two stages of analysis indicate that there seem to be evolutionary dynamics in *how* policy on epidemics is branded, but not necessarily in *that which* the policy is addressing.

Partisan Differences in Bill Sentiment

At this point, I pivot to address the other dimension of partisanship. I begin with this third stage in the analysis on sentiment analysis again, but this time disaggregated by the party of the sponsor instead of the chamber as in the previous stage. The results for the party-focused sentiment analysis are presented in Figure 12.5 (pp. 312–313).

Figure 12.5 is read the same as Figure 12.4, where scaled sentiment below the cut point on the y-axis suggests a generally negative tone in sponsored legislation compared to scaled sentiment scores above the 0.0 cut point, suggesting a generally positive tone in the proposed legislation addressing epidemics. In Figure 12.5, though, color varies by the party of the bill sponsor, with dark gray for Republicans, black for Democrats, and light gray for Independents.

A strikingly similar pattern exists at the party level as it did previously in Figure 12.4 at the chamber level, where tone intensity, both positive and negative, increases steadily over time. Also as in Figure 12.4, in Figure 12.5 there is a prominent drop in tone positivity (or an increase in negative tone) in the current COVID era. This suggests that there is likely an evolution to tone in proposed legislation along a partisan dimension as well. Both parties seem to be following a similar pattern. Yet is this enough to support the anecdotal motivation at the outset that policymaking on this apolitical issue of pandemics is characterized by divided partisan politics? Perhaps as a clue, but not in a systematic way. Indeed, the tone swings widely, but these patterns are not beholden to a single party, nor are they substantively political in nature, where one party might be more negative or positive than the other party. I come back to this evolutionary pattern in tone and limitations relating to partisan division in the discussion section at the end of this chapter.

Topic Structure of the COVID Era

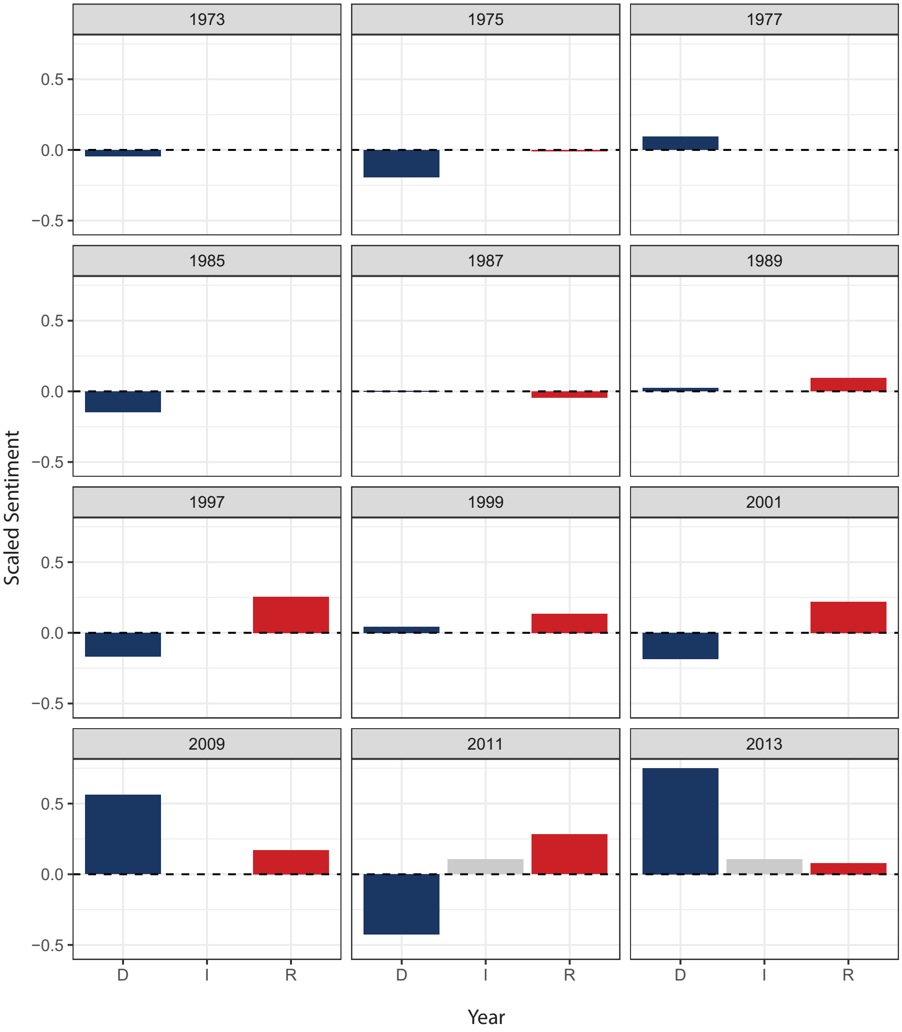
Given the contextual clarity around a long series of legislating on epidemics in American politics, I now shift focus to the COVID era to better understand the nature of policymaking in response to the massive epidemic with which the country is currently grappling. In this stage, I continue to probe the partisan dimension, but only in the COVID era.

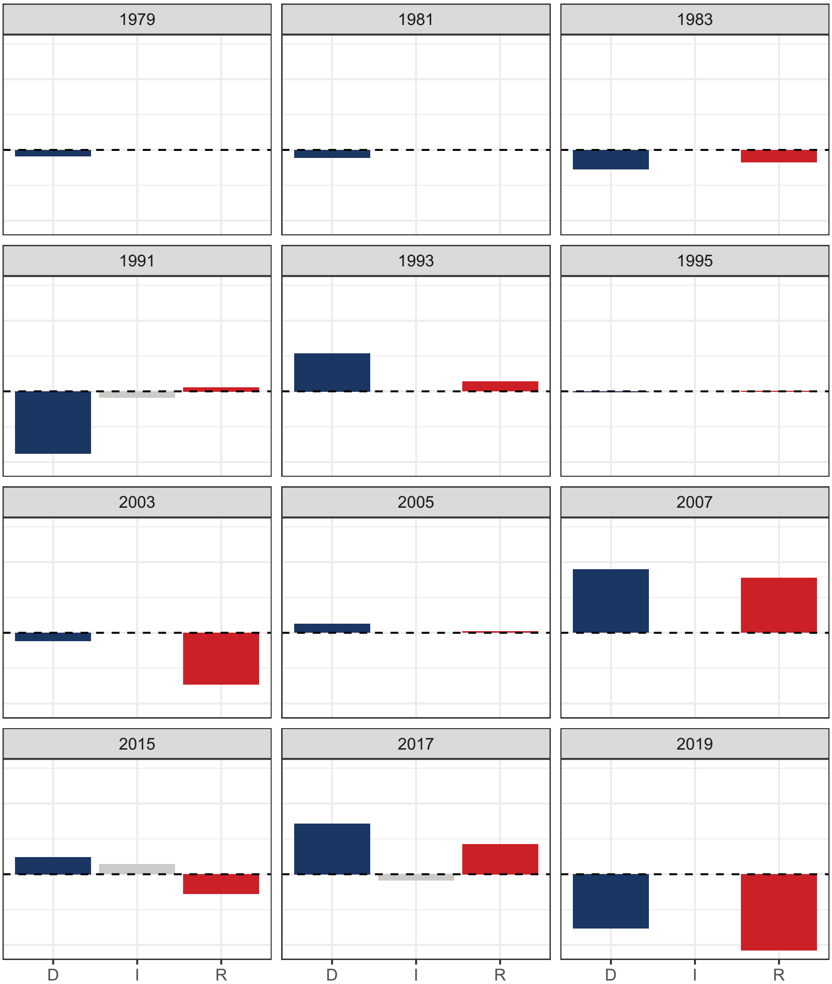
In light of the topic model results previously found, it may be reasonable to expect both parties to discuss COVID similarly. However, given the introductory anecdotal evidence on fractious approaches to policymaking on COVID-19, there is room to expect the parties to approach COVID from very different perspectives as well. These differences may be present in their proposed legislation. Thus, I return to the "what" question explored using topic models in the COVID era only.

Recall in the earlier topic models, I did not explicitly account for party. At this stage, while I will not account for party in the estimation of the model (e.g., using a structural topic model), I will instead proceed to fit a topic model with $k = 2$ and pull the results apart by party affiliation to understand whether latent partisan differences exist in topics. To do so, I start by examining the proportions of γ values by party affiliation. γ scores from topic models measure the probability a bill is associated with a given topic. Conditioning by party of the sponsor, I gain insight into the probabilities of bills *sponsored by different parties* being associated with one of the two topics. The results are shown in Table 12.1 (p. 314).

Most notably in Table 12.1, the probabilities of Democrats and Independents sponsoring bills related to topic 1 is higher than for topic 2, with $\gamma = 0.832$ and 0.528 for Democrats and Independents, respectively. This makes intuitive sense in that Independents in Congress nearly always caucus with Democrats. And adding to this, Republicans are more likely to sponsor bills related to topic 2 at $\gamma = 0.669$, compared to topic 1, with a value of 0.331. As such, there seems to be a clear partisan distinction in sponsored bills. Though stability in general topics was uncovered earlier, here I explicitly account for party of the sponsors, allowing for partisan differences in policymaking to emerge. But what terms define these topics? See Figure 12.6 (p. 314) for a bar plot of the topics with color conditioned on party.

In Figure 12.6 there is clear partisan difference in the branding of proposed legislation. For example, Democrats' bills include terms like "pandemic" and


1973
1975
1977
1985
1987
1989
1997
1999
2001
2009
2011
2013
0.5
0.0
−0.5
D
I
R
Scaled Sentiment
Year



Party
D
I
R

Note: Bars below the dashed line indicate greater overall negative sentiment, compared to bars above the dashed line indicating greater overall positive sentiment.

FIGURE 12.5 Scaled sentiment scores over time, by party.

TABLE 12.1 *γ by Party and Topic*

Party	Topic	Probability (γ)
Democrat	1	0.832
Independent	1	0.528
Republican	1	0.331
Democrat	2	0.168
Independent	2	0.472
Republican	2	0.669

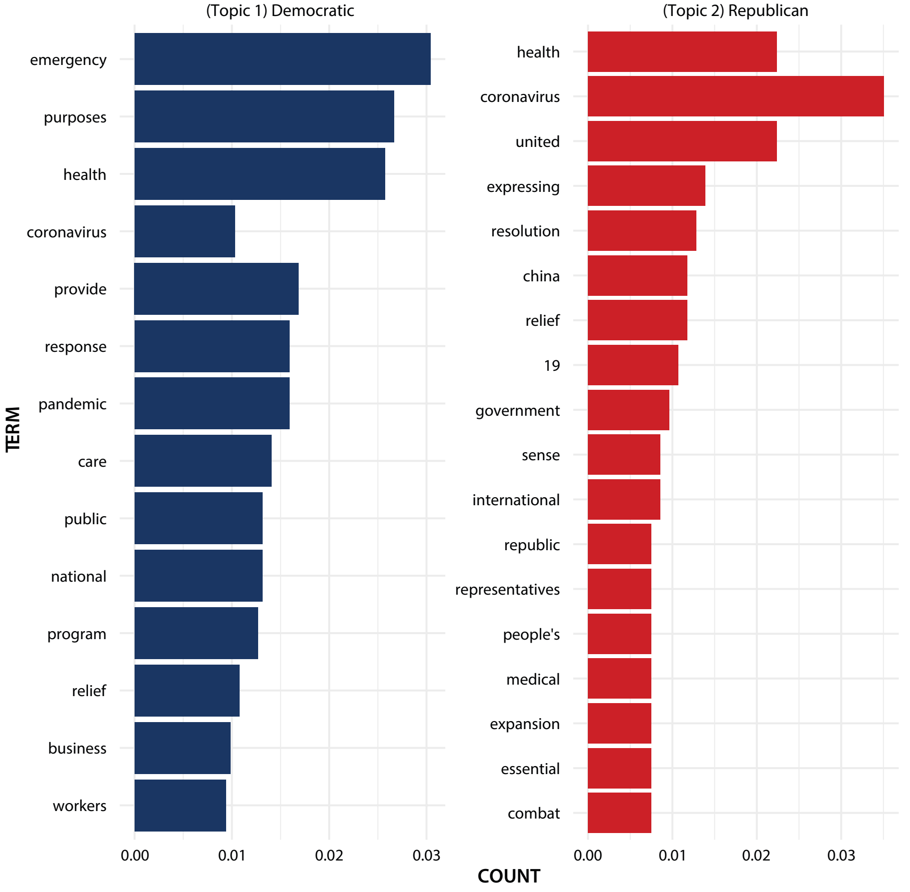


FIGURE 12.6 Topics in the COVID era, by party.

"workers," whereas Republicans' bills include terms like "China," "combat," and "international." Though the motivation driving the use of these terms is unable to be obtained from the current analysis, of greatest value for present purposes is the partisan distinction in types of bills sponsored in responding to the same apolitical pandemic of COVID-19.

This stage gives a closer look at the partisan question, suggesting that policymaking as an expression of elite responsiveness in the current congressional climate is one distinguished by party division. Whether this is a normatively "good" or "bad" trend is beyond the scope and goal of this project. Rather, this project is interested in exploring these data in search of natural patterns. The unsupervised nature of the modeling strategy allows the structure to emerge. And the emergent structure points to partisan forces at work in policymaking in response to COVID-19.

Exploring Networks of Partisan Term Co-Occurrence in the COVID Era

In the final stage of analysis, I continue with focus on the COVID era. I build on the previous findings that the parties approach policymaking in response to COVID-19 differently. Now, I am interested in understanding the structure of term usage within and across both major parties. To do so, I leverage bigram networks. I weight the edges of the network connecting use of bigrams to capture the frequencies of co-occurrence of terms. I break down term usage by party and present networks in Figures 12.7 and 12.8 for Democrats and Republicans, respectively.

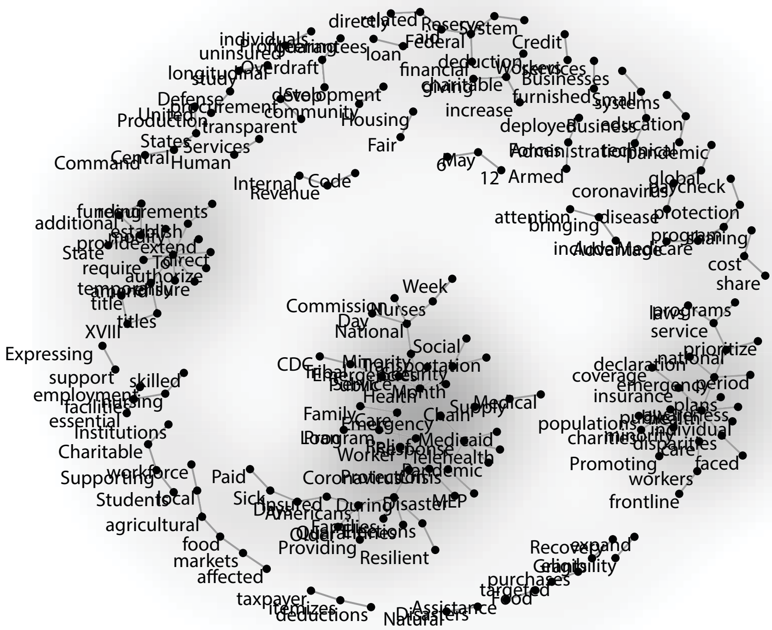
Substantively, this approach allows for visualizing usages of the terms by both major parties in the COVID era to understand the topology of how terms are used together. The goal of this final stage is to place the broader topic trends found in the previous sections into context, which is exploration of patterns *within* party ranks. Cross-party comparisons are also possible. But the focus of this section is to offer a window into how parties use and recycle certain words in their proposed policies, giving another angle of policymaking dynamics in the era of COVID-19.

For both figures, the network is an undirected, weighted graph with shading varying by weighted edges, such that darker shades mean greater frequencies of bigram usage.

In Figure 12.7, the volume of bigrams Democrats used as well as their interconnection is much greater than that of the Republicans, shown in Figure 12.8. Some

DEMOCRATIC BILLS

Bigram Network Representations

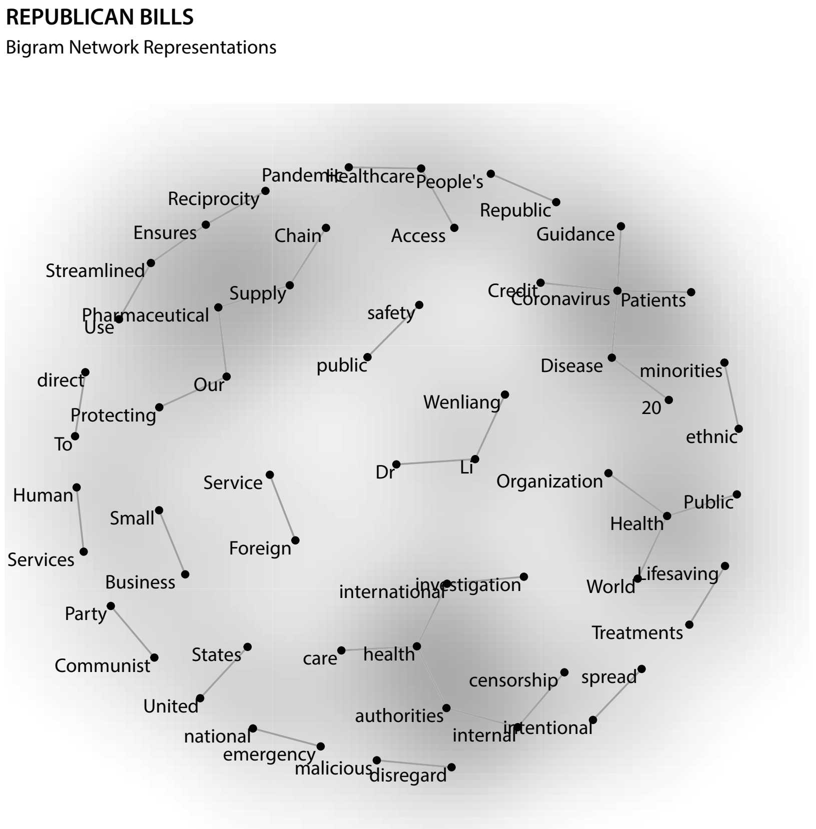


Note: Shading indicates greater density of co-occurrence.

FIGURE 12.7 Bigram networks of Democratic bills.

of the dense regions in Figure 12.7 for Democrats involve discussion on extending and authorizing governmental funding (seen in the upper left of Figure 12.7). The other dense region in Figure 12.7 is toward the middle, which focuses on worker relief, Medicaid, and families. These results are consistent with findings to this point.

Regarding the patterns for Republican bills in Figure 12.8, not only are fewer terms used, implying greater homogeneity and consistency within their ranks, but there is also a relatively high, consistent density across the full space (i.e., darker shading in most of the network). This pattern means that these bigrams are used together and frequently, reflecting a possible strategy within the party (e.g., sticking to party-derived talking points, organized policy priorities, and so forth). Some of the terms in the Republican policies include "foreign service,"



Note: Shading indicates greater density of co-occurrence.

FIGURE 12.8 Bigram networks of Republican bills.

"lifesaving treatments," and "small business." While these terms might be expected in the context of such a massive pandemic, some other frequently used terms are unique to Republicans, such as "Communist Party," "protecting our pharmaceutical supply chain," and "People's Republic" (possibly "of China").

Building on topic model results in Figure 12.6, Republicans seem more focused on responding to COVID in the contexts of securing the domestic economy and the international aspects of COVID (e.g., "People's Republic," "World Health Organization," "Dr. Li Wenliang," and so forth). Democrats, inversely, focus virtually zero attention on such international aspects and instead focus efforts very broadly on domestic politics and policies, in sum resulting in clear partisan differences in responding to COVID-19.

DISCUSSION AND CONCLUSION

To recap, a few key patterns emerged across the five stages of analysis. Regarding the time dimension and the question of evolutionary dynamics, there were two dominant trends. First, the "what" question pertaining to the topics of focus tend to remain relatively stable over time. Policymakers tend to address the given epidemic with epidemic-specific terms in their policies, implying virtually no evolutionary, time-dependent process. However, the more prominent differences that point to evolutionary dynamics were the shifts in overall tone of the proposed policies. In the earlier days of the study period (1970s–1980s), the tone was relatively muted, with few positive- or negative-toned policies being offered by either party or chamber. In the 1990s and early 2000s, this tone, both positive and negative, significantly ticks upward, where more extreme terms are used in policy descriptions. This pattern culminates in the most recent era of COVID-19 (2019–2020), where the negative tone defines policymaking and is starker than in any other period and across both chambers *and* parties.

Regarding the partisan dimension, in the COVID era specifically, the parties cast their solutions to COVID in starkly different lights, highlighting different realms and focus within their party ranks. Democrats highlighted domestic responses on average, while Republicans highlighted international actors and responses to a greater degree. This suggests that there is indeed a partisan flavor to policymaking regarding COVID-19. Yet, whether this qualifies as "bitter" or "polarized" politics and policymaking is a trickier question and is addressed more below.

Though they are exploratory, from these results it is clear that, perhaps as expected, the two major American political parties are different in their approaches to governing in the time of COVID-19. Yet, despite these partisan differences, the intensity and negativity of tone both at the chamber and party levels has been steadily growing since the 1970s. This suggests that there is an evolutionary dynamic to epidemic policymaking, which is at a climax in the current era of COVID-19.

Limitations

Though patterns from the sentiment analysis appear to have been evolutionary and growing in intensity, this may not be a reflection of division or bitter policymaking but rather a reflection of the grave nature of COVID-19. Such a negative epidemic could certainly be accompanied by an increase in negative-toned legislation.

Yet while this may be the case, it would make sense that negative-toned legislation should characterize virtually all epidemics across all periods given the scope and nature of these types of social problems. Indeed, epidemics are cast as emergencies and issues of prime importance for the government to address, seen at the first stage in Figure 12.1, where frequently used terms in both pre-COVID and COVID eras implied that epidemics are emergent issues.

The tone of related legislation, then, should also be more negative than positive if tone is a function of subject and rather than an era of harsh or bitter policymaking. Yet there are numerous dramatic spikes in positive sentiment that grow over time. This could be a reflection of the approach to branding the policy response (e.g., a triumph over the epidemic in question). Given the plausibility of numerous explanations underlying these patterns, future research should take up the question drivers behind tone and linguistic patterns in policymaking through a targeted causal study to shed light on the "why" behind these trends.

Concluding Remarks

In sum, this project is an exploratory effort focused on uncovering and understanding the contours of government policymaking as a formal response to epidemics over a long period of time. The duration of time, as well as these data being the clearest signal of government priorities, make this an ideal place from which to launch an exploration of many other related topics. For example, future work might consider the role of media and reporting on government responses in times of epidemics, or the presence of partisan division in policymaking surrounding epidemics like the opioid crisis or COVID-19.

ACKNOWLEDGMENTS

Funding was made possible in part by the C-SPAN Education Foundation. I am exceedingly grateful for the feedback and questions raised by Center for C-SPAN Scholarship & Engagement conference participants, and to Purdue University for hosting the 2020 CCSE conference. All remaining errors are my own.

NOTE

1. Technical details of the algorithm are omitted due to space and goals of the project but are available on request.

REFERENCES

Berinsky, A. J. (2007). Assuming the costs of war: Events, elites, and American public support for military conflict. *The Journal of Politics*, *69*(4), 975–997. https://doi.org/10.1111/j.1468-2508.2007.00602.x

Blei, D. M., Ng, A. Y., & Jordan, M. I. (2003). Latent dirichlet allocation. *Journal of Machine Learning Research*, *3*(March), 993–1022. https://dl.acm.org/doi/10.5555/944919.944937

C-SPAN (Producer). (2020a, April 23). U.S. House of Representatives: Representatives McCarthy, Neal, Jayapal on coronavirus bill [Video]. https://www.c-span.org/video/?471431-12/representatives-mccarthy-neal-jayapal-coronavirus-bill

C-SPAN (Producer). (2020b, May 18). U.S. Senate: Senator John Barrasso on House coronavirus bill [Video]. https://www.c-span.org/video/?472227-5/senator-john-barrasso-house-coronavirus-bill

Druckman, J. N., Peterson, E., & Slothuus, R. (2013). How elite partisan polarization affects public opinion formation. *American Political Science Review*, *107*(1), 57–79. https://doi.org/10.1017/S0003055412000500

Grossmann, M., & Hopkins, D. A. (2016). *Asymmetric politics: Ideological Republicans and group interest Democrats*. Oxford University Press.

Jones, B. D., Larsen-Price, H., & Wilkerson, J. (2009). Representation and American governing institutions. *The Journal of Politics*, *71*(1), 277–290. https://doi.org/10.1017/S002238160809018X

Larcinese, V., Puglisi, R., & Snyder Jr., J. M. (2011). Partisan bias in economic news: Evidence on the agenda-setting behavior of US newspapers. *Journal of Public Economics*, *95*(9–10), 1178–1189. https://doi.org/10.1016/j.jpubeco.2011.04.006

Layman, G. C., Carsey, T. M., Green, J. C., Herrera, R., & Cooperman, R. (2010). Activists and conflict extension in American party politics. *American Political Science Review*, *104*(2), 324–346. https://doi.org/10.1017/S000305541000016X

Liu, B. (2012). Sentiment analysis and opinion mining. *Synthesis lectures on human language technologies*, *5*(1), 1–167. https://doi.org/10.2200/S00416ED1V01Y201204HLT016

Perryman, M. R. (2019). Biased gatekeepers? Partisan perceptions of media attention in the 2016 U.S. presidential election. *Journalism Studies*, *20*(16), 2404–2421. https://doi.org/10.1080/1461670X.2019.1598888

Robison, J., & Mullinix, K. J. (2016). Elite polarization and public opinion: How polarization is communicated and its effects. *Political Communication*, *33*(2), 261–282. https://doi.org/10.1080/10584609.2015.1055526

Souva, M., & Rohde, D. (2007). Elite opinion differences and partisanship in congressional foreign policy, 1975–1996. *Political Research Quarterly*, *60*(1), 113–123. https://doi.org/10.1177/1065912906298630

Waggoner, P. D. (2019). Do constituents influence issue-specific bill sponsorship? *American Politics Research*, *47*(4), 709–738. https://doi.org/10.1177/1532673X18759644

APPENDIX

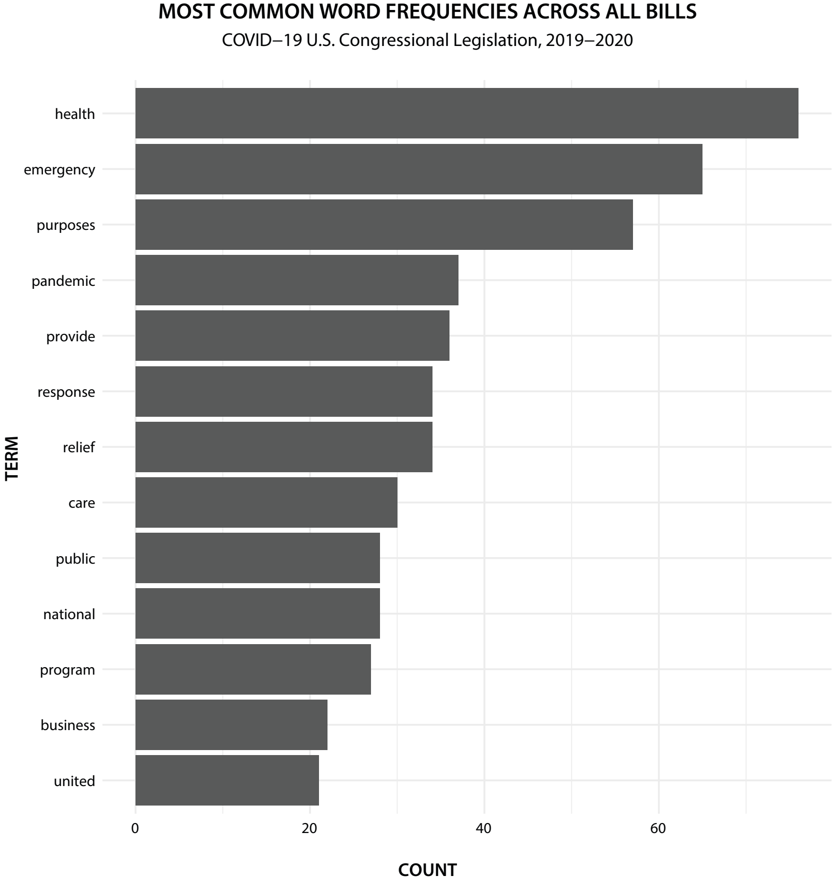


FIGURE A.1 COVID top terms used.

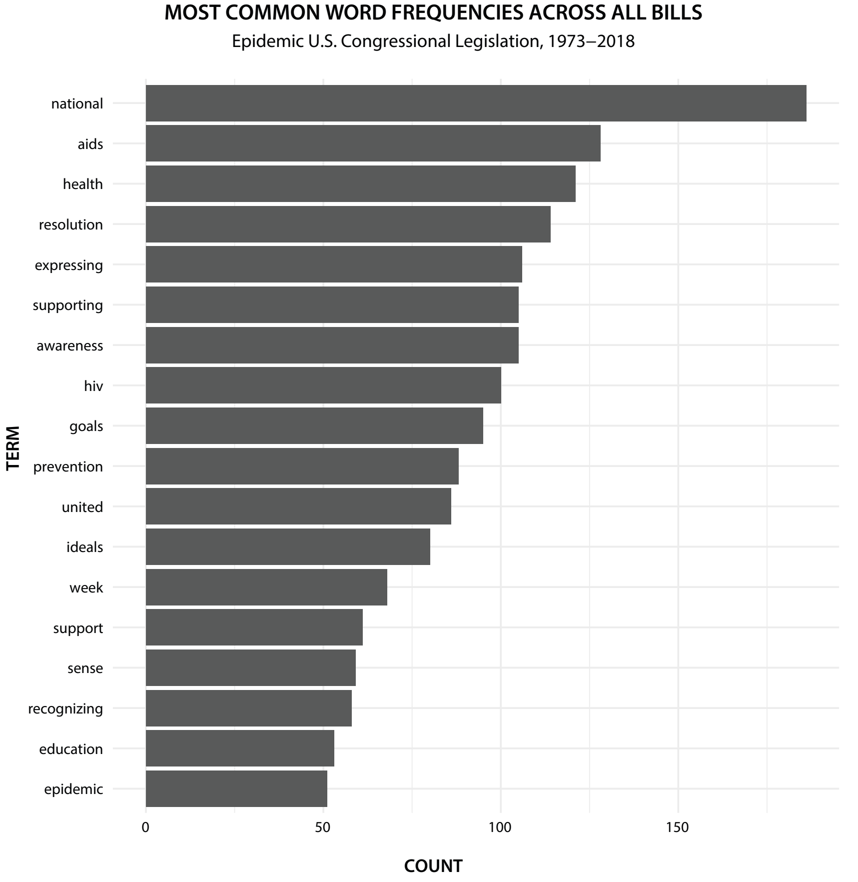


FIGURE A.2 Pre-COVID top terms used.