EMOTION AND REASON IN POLITICAL LANGUAGE*

Gloria Gennaro and Elliott Ash

This paper studies the use of emotion and reason in political discourse. Adopting computational-linguistics
techniques to construct a validated text-based scale, we measure emotionality in six million speeches given
in U.S. Congress over the years 1858–2014. Intuitively, emotionality spikes during times of war and is
highest in speeches about patriotism. In the time series, emotionality was relatively low and stable in earlier
years but increased significantly starting in the late 1970s. Across Congress members, emotionality is higher
for Democrats, for women, for ethnic/religious minorities, for the opposition party and for members with
ideologically extreme roll-call voting records.

An emotional speaker always makes his audience feel with him, even when there is
nothing in his arguments; which is why many speakers try to overwhelm their audience
by mere noise.
  (Aristotle, Rhetoric (350 B.C.E), chapter 7. Translation comes from Rhetoric, W.D.
Ross (2010), page 129)

In politics, when reason and emotion collide, emotion invariably wins.
  (Drew Westen, The Political Brain: The Role of Emotion in Deciding the Fate of the
Nation (2007), pg. 35)

In his treatise on *Rhetoric*, Aristotle suggested that persuasion can be achieved through either
logical argumentation or emotional arousal in the audience; success depends on selecting the most
appropriate strategy for the given context. Cultivated by these early ideas, the classic dichotomy
between emotions and affect (*pathos*) on the one side and rationality and cognition (*logos*) on
the other has informed all realms of social sciences, from social psychology (LeDoux, 1998),
to political philosophy (Elster, 1999), to economics (Frank, 1988). In the day-to-day of political

debate, politicians resort to a mix of emotion and reason and search for the right balance between these two elements.

The extent to which politicians engage with this trade-off, and what institutional, political and psychological factors underlie their choices, is largely unknown. Providing empirical evidence on these questions has been difficult due to the lack of a reproducible, validated and scalable measure of emotionality in political language. In this paper, we propose a measure that satisfies these requirements, and we extensively validate it against human judgement. We then use it for a variegated description of how politicians in U.S. Congress have used emotion in their rhetoric over the last 150 years.

Our approach builds on recently developed computational linguistics tools, which represent semantic dimensions in language as geometric dimensions in a vector space. The algorithm for this purpose, word embedding, transforms words and phrases to vectors, where similar words tend to co-locate and directions in the space (dimensions) correspond to semantically meaningful concepts (Collobert and Weston, 2008; Mikolov *et al.*, 2013; Pennington *et al.*, 2014). Our goal is to construct a dimension in this space corresponding to reason at one pole and emotion on the other. To this end, we take validated word lists for emotion and reason and construct the poles as the average vectors for these semantically coherent word groups. The relative emotionality of a word is the proximity to the emotion pole, relative to the reason pole. In turn, the emotionality of a document is the relative proximity of the document vector to the emotion pole. We compute scores for six million floor speeches reported in the *U.S. Congressional Record* for the years 1858 through 2014.

Our measure of emotionality in political language convincingly survives a rigorous sequence of validation steps, consistent with rising standards in empirical work using text data (Quinn *et al.*, 2010; Grimmer and Stewart, 2013; Goet, 2019; Rodman, 2020; Rodriguez and Spirling, 2022; Osnabrügge *et al.*, 2021a). First, we qualitatively inspect the words and sentences that are most associated with the ends of the emotion-rationality spectrum. The inspected examples are intuitive and satisfying. Second, we undertake a substantial human validation effort and ask human annotators to assess the relative emotionality of thousands of ranked sentence pairs. The ranking provided by our preferred measure agrees with human judgement over 90% of the time, a superior accuracy to that obtained applying commonly used dictionary-based methods that count relevant terms.

Given the lengthy historic time period covered by our corpus, we also seek to validate comparisons over time. A targeted human validation shows that the emotionality score is historically valid for the whole time period of the *Congressional Record* back to the 1850s. In addition, we show that our measure of emotionality in politics is distinct from the political topics chosen (Quinn *et al.*, 2010), from positive and negative sentiment (e.g., Rheault *et al.*, 2016), from changes in the sophistication of political language (e.g., Benoit *et al.*, 2019) and from emotionality trends in the broader society (e.g., Morin and Acerbi, 2017). These checks allow us to confidently make comparisons over time and to attribute the observed results to dynamics that are specific to emotion in political language.

In the empirical analysis section, we provide a rich description of how emotion and reason have varied over time, by topic, and across speakers in Congress. First, we look at long-run rhetorical history using our 150-year time series. Intuitively, emotional expression spikes in times of war. Furthermore, we find a significant increase in emotionality since the late 1970s, coinciding with the introduction of televised congressional floor debates via C-SPAN. This descriptive evidence is consistent with C-SPAN motivating the use of more emotional rhetoric, in line with previous

work on how television has reshaped politics (Gentzkow, 2006; DellaVigna and Kaplan, 2007; Martin and Yurukoglu, 2017; Durante *et al.*, 2019).

Second, we compare emotionality across topics. We find, intuitively, that patriotism, foreign policy and social issues are discussed with the most emotion, while procedure and federal organisation are discussed with the least emotion. Within the realm of economic policy, issues related to taxation and redistribution have increased the most in emotionality in recent years (especially for Republicans), coincident with the post-Reagan increase in economic inequality. In light of recent debates on the ideological foundations of inequality (McCarty *et al.*, 2016; Piketty, 2020), it is illuminating that Republicans use emotional rather than rational appeals when arguing about redistributive policies.

Third, we assess how emotionality varies across politicians' personal characteristics and institutional factors. Democrats, women and racial/religious minorities tend to use more emotive language than Republican, male, white protestants serving in the same chamber and year. Furthermore, members of Congress use more emotional language when in the opposition (minority) party. Overall, emotion appears in situations of disempowerment, not just in terms of political party control, but also for disadvantaged identity groups. This result is consistent with explanations of emotional appeals from both behavioural economics and political economy: emotion may help politicians deal with loss of control or frustration of expectations (e.g., MacLeod, 1996; Lin *et al.*, 2006), or it may serve to push policy positions (Jerit *et al.*, 2009) and complement minority representation strategies (see Swers, 2002; Tate, 2018).

Finally, we look at how emotionality is related to partisan polarisation. We find that politicians with highly partisan roll-call-voting records (on either the left or the right) use more emotion than their more moderate colleagues. This higher emotional expressiveness is driven both by the speech topics chosen but also by how the same topics are framed. Hence, trends in emotive rhetoric are linked to increasing polarisation in U.S. politics (McCarty *et al.*, 2016; Gentzkow *et al.*, 2019b).

Overall, the paper provides both a methodological and a substantive contribution. Methodologically, we push forward the use of text analysis in economics and political economy (Gentzkow *et al.*, 2019a). The focus of most previous work has been partisan differences in language, taking a supervised learning approach (Gentzkow and Shapiro, 2010; Jensen *et al.*, 2012; Ash *et al.*, 2017; Gentzkow *et al.*, 2019b). Baker *et al.* (2016) and Enke (2020) each used a dictionary approach to respectively analyse policy uncertainty and moral norm priorities. Hansen *et al.* (2018) used a topic model to analyse central bank communications. Our new method, using word embeddings to scale emotionality, addresses the technical limitations of dictionary methods while still targeting a specific dimension of discourse.

Substantively, we add to the literature on rhetorical choices in political communication, and in particular the role of emotions in politics (e.g., Marcus, 2000; Lau and Rovner, 2009). The previous literature has shown that political speech sentiment and emotional intensity respond to economic conditions (Rheault *et al.*, 2016), ideological divisions (Kosmidis *et al.*, 2019), institutional context (Osnabrügge *et al.*, 2021b) and the characteristics of the speaker (Dietrich *et al.*, 2019; Hargrave and Blumenau, 2021; Boussalis *et al.*, 2021). Consistent with these findings, we find coherent descriptive evidence that emotional rhetoric corresponds to prevailing political opportunities and conditions, including party control, personal identity and ideological polarisation.

From a more historical perspective, a number of studies have attended to the long-run evolution of rhetoric in parliaments. The emerging theme is that of increasingly polarised language,

accompanied by a general simplification. Upward trends in divisive language around party lines in U.S. Congress have been repeatedly replicated (Jensen *et al.*, 2012; Gentzkow *et al.*, 2019b; Rheault and Cochrane, 2020), with comparable trends also seen in UK Parliament (Peterson and Spirling, 2018; Goet, 2019). Meanwhile, the linguistic sophistication of political speeches has decreased over time (Lim, 2002; Benoit *et al.*, 2019), and confidence among politicians has increased (Jordan *et al.*, 2019). We add an important piece to this picture by observing that the secular trends in polarisation, simplification and confidence have been accompanied by a more intense expression of emotion. All of these trends can be understood as a coherent shift toward a rhetoric that addresses voters rather than fellow politicians and elites.

　　More generally, this research adds to a long tradition on the dichotomy of emotion and reason in social theory and social science (Damasio, 1995; LeDoux, 1998; Elster, 1999). A classic view from economics is Frank (1988), who explored how various emotions support both self-interested and socially conscious decision-making. A subsequent line of work in behavioural economics has shown the role of emotions in supporting pro-social behaviour, for example through motivating costly punishment (MacLeod, 1996; Bosman and Van Winden, 2002; Xiao and Houser, 2005; Van Buskirk *et al.*, 2012). Overall, emotions are complementary with rationality in supporting human decisions and communication (e.g., Elster, 1999; Loewenstein, 2000; Kahneman, 2011; Lerner *et al.*, 2015; Wälde and Moors, 2017). Thus, it is not surprising to observe a pivotal role for emotions, along with reason, in political discourse.

1. Measuring Emotion and Reason in Text

This section outlines the approach to measuring dimensions of emotion and reason in unstructured text. After giving some details on the political speeches corpus, we describe the word lists for identifying emotion and cognition dimensions. Then, we introduce word embeddings and how they allow us to scale documents in this emotion-reason dimension.

1.1. *Congressional Speeches Corpus*

Our empirical corpus comprises digitised transcripts of the universe of speeches in the U.S. House and Senate between 1858 and 2014 ($N = 7, 336, 112$ speeches). The corpus includes all speeches from the *U.S. Congressional Record*, after removing those speeches that contain readings of pieces of legislation.

　　The corpus pre-processing can be summarised as follows (see Online Appendix F.1 for additional details). Each speech in the corpus is first segmented into sentences. To extract the most informative tokens, we tag parts of speech and take only nouns, adjectives and verbs. Punctuation, capitalisation, digits and stopwords (including names for states, cities, months, politicians and procedural words) are removed. Tokens are stemmed using the Snowball stemmer. After filtering out rare stems (those occurring in less than ten speeches), we have 113,055 token types left in the vocabulary.

1.2. *Dictionaries for Emotion and Cognition*

The first ingredient in our method is to use thematically categorised lists of words for emotion and reason. To build lists of emotive and cognitive words, we start with Linguistic Inquiry and Word

Count (LIWC), a leading set of categorised dictionaries validated by linguistic psychologists (Pennebaker *et al.*, 2015). LIWC researchers have collected coherent sets of words, word stems and idiomatic expressions that map onto various structural, cognitive and emotional components of language.

From LIWC, we take two word lists. First, to get at reasoning we use the 'Cognitive Processing' category, consisting of 799 words, phrases and wildcard expressions. This category embraces concepts of insight, causation, discrepancy, certainty, inhibition, inclusion and exclusion. Second, to get at emotion, we use the 'Affective Processing' category, comprising 1,445 tokens, phrases and wildcard expressions. This category refers to emotions, moods and other affective states—both positive (joy, gratitude) and negative (anxiety, anger, sadness).

We reviewed the raw LIWC dictionaries and adapted them to analysis of congressional speeches (see Online Appendix F.2 for details). We excluded a number of inappropriate patterns (e.g., emojis, punctuation, digits, multi-word expressions), and a number of words that do not translate well to the Congressional Record (e.g., 'admir*' matching to 'admiral'). At the end of the process, we have a list of stemmed nouns, verbs and adjectives representing affective processing (629 tokens) and cognitive processing (169 tokens). Let A and C represent these word lists. Online Appendix F.5 provides the two final dictionaries and the frequency of each dictionary word in the corpus.

The word lists for emotion and cognition can already be used to produce a dictionary-based measure of emotionality. This type of measure, where one counts the words from the dictionary to detect semantic domains in documents, is the previous standard in social science (Kosmidis *et al.*, 2019; Osnabrügge *et al.*, 2021b). For our analysis below, we produce such a measure based on the relative frequency of these words in each speech (Online Appendix B.1 describes how this measure is calculated). In the human validation below, we show that a dictionary method compares poorly with human judgements about the emotionality of congressional speech snippets.

The problem with the dictionary approach, in our setting as in others, is that the method relies too heavily on the presence or absence of the particular listed words. The dictionary approach requires that the dictionary is reliably specified. This puts a lot of pressure on the researcher to identify all emotive words and their variants. This task could be especially difficult in historical contexts where the set of probative words might not be clear from a contemporary perspective.

A second problem is that the dictionary approach assumes that each word in the emotionality list is treated as equally indicative of emotionality, while each word not in the list is treated as equally indicative of emotionlessness. But this model of language is clearly wrong. For example, the word 'like' could refer to preference or to similarity, while the word 'dislike' only refers to preference (with 'unlike' reserved for dissimilarity). A properly constructed emotion scale would give more weight to 'dislike' than 'like', but a dictionary approach assumes binary categories and cannot scale words continuously.

1.3. *Embedding Approach to Scaling Emotionality*

Word embeddings are well suited to addressing the main problems with the dictionary approach. Rather than require that all words with emotive content be identified, word embeddings only require that a representative sample of emotive words are identified. In addition, word embeddings can flexibly learn from the corpus the intensity with which words are emotively associated, with

no assumptions of discrete categories. In particular, if some emotive words in an historical period are missing from the specified list, their emotive association can still be learned by the model and accounted for by the resulting scale.

Beyond lexicon construction, a continuous scale can capture more subtle linguistic cues implied by full sentences. In contrast, a word-counting approach relies on the sparse, explicit and intentional placement of emotion-laden words. Thus, as shown in Caliskan *et al.* (2017) and Ash *et al.* (2021), word embedding dimensions tend to reveal more about social attitudes than do word counts (see also Garg *et al.*, 2018; Kozlowski *et al.*, 2019).

1.3.1. *Word embeddings*

More formally, word embeddings are a tool from natural language processing for learning numerical representations of words based on co-occurrence statistics in a given corpus (Mikolov *et al.*, 2013; Pennington *et al.*, 2014). A word, normally a string object drawn from a high-dimensional list of categories, is 'embedded' in a lower-dimensional space, where the geometric location encodes semantic meaning. Semantically related words (e.g., 'happy' and 'joyful') will tend to have geometrically proximate vectors. Semantically unrelated words (e.g., 'happy' and 'econometrics') will tend to have geometrically distant vectors.

In the context of word embedding algorithms, semantic relatedness means that the words appear in similar contexts. The key intuition is: 'You shall know a word by the company it keeps' (Firth, 1957). Take the sentence, 'I was _____ to learn that I had won re-election.' While *happy* and *joyful* would fit in nicely, *econometrics* would not. A word embedding algorithm learns word locations that predict which words would best complete any given sentence. The useful result is that directions in the embedding space correspond to semantic dimensions of language (e.g., emotion and rationality dimensions).

Thus we learn a vector w corresponding to each word w in the vocabulary based on how words co-occur in congressional speeches. More technically, we learn word embeddings using the Word2Vec algorithm from Mikolov *et al.* (2013) applied to the full corpus. We use the python gensim implementation with three-hundred-dimensional vectors, an eight-word context window and training for ten epochs. These are all standard hyperparameter choices from the applied natural language processing literature. Rodriguez and Spirling (2022) and Ash *et al.* (2021) showed that results produced from word embeddings are generally robust to variation in those choices.

1.3.2. *Scaling congressional speeches by emotion and cognition*

We now can use the embeddings and our dictionaries to scale speeches in the Congressional Record with an emotionality score. First, the word embeddings are combined with the thematic word lists to isolate directions in embedding space corresponding to emotion and reason. The vector A representing *emotion* is the average of the vectors w for the words in the emotion word list, $w \in A$. The vector C for *cognition* is defined analogously.[1]

Second, we produce vector representations for each congressional speech, using the same specification as done for the emotion and rationality poles. Let the vector d_i for speech i be the

---

[1] Some recent papers have used this approach to extrapolate word lists more effectively to the political domain. Word embedding models can expand the dictionaries to larger lists of sentiment or emotion words (e.g., Rheault *et al.*, 2016; Rice and Zorn, 2021; Osnabrügge *et al.*, 2021b). This approach can address the problem of missing words in the dictionary, but it does not address the problem that dictionaries assume discrete categories of words, rather than a continuous scale of emotion.

average of the vectors w of the words w in the speech.[2] Thus we construct a three-hundred-dimensional vector for each speech in the Congressional Record.

Taking these ingredients together, we can scale texts along the emotion and cognition dimensions. Our measure for the emotionality of speech i is

$$Y_i = \frac{\text{sim}(\boldsymbol{d_i}, \boldsymbol{A}) + b}{\text{sim}(\boldsymbol{d_i}, \boldsymbol{C}) + b}, \tag{1}$$

where $\text{sim}(\boldsymbol{v}, \boldsymbol{w}) = (\boldsymbol{v} \cdot \boldsymbol{w})/(\|\boldsymbol{v}\|\|\boldsymbol{w}\|)$ is the cosine similarity between vectors $\boldsymbol{v}$ and $\boldsymbol{w}$. The addition of a constant b in the numerator and denominator is for smoothing outliers; we set $b = 1$ but it can be set to any small positive number. An increase in Y_i indicates a shift towards the emotion pole relative to the cognition pole. In addition, we can produce separate measures for emotion and cognition with $\text{sim}(\boldsymbol{d_i}, \boldsymbol{A})$ and $\text{sim}(\boldsymbol{d_i}, \boldsymbol{C})$, respectively.

A speech that is equally emotive and cognitive would take value $Y_i = 1$. Online Appendix Figure A4 shows the distributions of the measure and its emotive and cognitive components. Online Appendix Table A4 and Online Appendix Figure A3 show the distribution of Y_i and how that evolved over time in U.S. Congress.

1.3.3. *Alternative emotionality measures*

For robustness and to assess better the performance of our measure, we calculate two alternative measures of emotionality in speeches based on the previous literature. First, as already mentioned, we compute a count-based measure using the frequencies that the words in our dictionaries appear in each speech. The count-based measure turns out to produce a quite different ranking of speeches than the embedding-based measure from (1). As shown in Online Appendix Figure A4, the distribution of the count measure is highly sparse and skewed because it relies on the presence or absence of words in the dictionaries. The correlation coefficient is 0.15 with our baseline measure in the full dataset. In addition, the count-based measure performs much worse in the human validation. Still, many of our central results hold when using the count-based measure (Online Appendix B.1).

Second, we compute an alternative distance metric from the embeddings, based more closely on Kozlowski *et al.* (2019) and Ash *et al.* (2021) and using the feature that analogous dimensions in vector space can be constructed with vector differences. In these previous papers, a gender dimension is constructed as the 'male' vector minus the 'female' vector. Correspondingly, we isolate an emotion-to-cognition dimension as the vector difference $\boldsymbol{AC} = \boldsymbol{A} - \boldsymbol{C}$ (the vector for 'affect' minus the vector for 'cognition'). Then the emotion score for document i is the cosine similarity to the differenced vector, $\text{sim}(\boldsymbol{d_i}, \boldsymbol{AC})$.

Conceptually, this 'geometric' measure directly leverages the 'analogy-solving' capacity of word embeddings. In subtracting cognition from emotion, the scale implicitly assumes a semantic dimension along which an increase in the relatedness to one concept pole (i.e., emotion) corresponds to a decrease in relatedness to the other concept pole (i.e., reason). In contrast, the

---

[2] In the preferred specification, the document vector averages (as well as the emotion and cognition vector averages) are weighted by the smoothed inverse frequency of each word, as done in Arora *et al.* (2016). That is, words that appear relatively often are down weighted, while words that are relatively rare—and therefore distinctive—are up weighted. This weighting improved performance of the metric in human validation, but does not make a difference in the downstream results. See Online Appendix F.3 and Online Appendix Table A6. We note, further, that this step of dimension reduction using word embeddings is an alternative to the regularisation approach taken by Gentzkow *et al.* (2019b) to address sparsity in their high-dimensional n-gram representation of the congressional speeches.
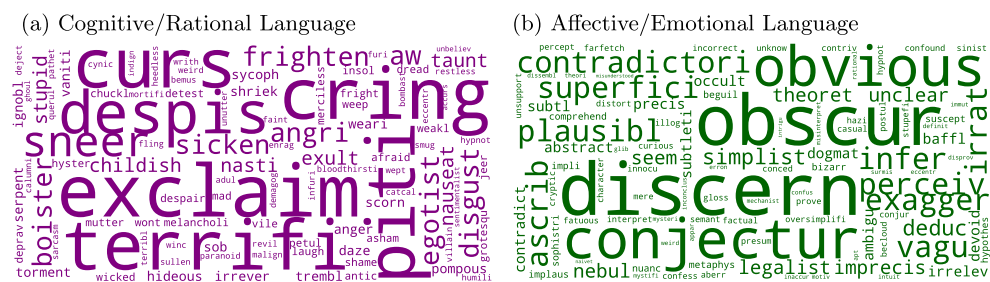
(a) Cognitive/Rational Language   (b) Affective/Emotional Language

Fig. 1. *Semantic Poles for Rationality and Emotion.*
Notes: The wordclouds show the dictionary words that are closest to the respective 'poles' of the dimension in the embedding space corresponding to rationality/cognition (a) and affect/emotion (b). Size denotes closeness to the respective word-vector centroid.

baseline ratio measure from (1) relies only on the assumption that vector distances proxy for semantic distances, allowing for cases where an increase in emotion does not imply a decrease in reason. Thus we prefer the ratio measure as the baseline in our setting. Moreover, it performs slightly better than the geometric measure in the human validations. Practically, however, the data reveal that the two measures are highly correlated in the congressional speeches. The geometric measure's correlation coefficient with our main score is 0.95 in the full sample of speeches, and thus the measures can be used as substitutes. Unsurprisingly, our empirical results are robust to using the geometric measure instead (Online Appendix B.2).

2. Validation

This section reports our multiple validation exercises (as in Quinn *et al.*, 2010; Goet, 2019; Osnabrügge *et al.*, 2021a). First, we show qualitative evidence that our approach captures distinctive semantic dimensions that correspond to emotion and cognition. Second, we compare our measure to human judgements about the emotionality of short speech segments. The pairwise rankings provided by our embedding-based measure agree with human rankings over 90% of the time, much higher than that for a more standard count-based measure.

2.1. *Qualitative Evaluation of the Semantic Dimensions*

We first ask: do the vector dimensions underlying our measure capture qualitatively coherent and distinctive semantic dimensions in language? A simple test for semantic validity is to inspect the language associated with the geometric poles for cognitive and emotional language. For each word in the vocabulary outside the lexicon, we compute the relative similarity to the cognitive and emotive poles. This gives a ranking of the words along a single cognitive-to-emotive dimension.

Figure 1 shows clouds for the words that are closest to the cognitive (panel (a)) and emotive (panel (b)) centroids, where larger word size indicates closer proximity to the centroid. The word clouds illustrate the clear, intuitive and distinct flavours of language captured by each linguistic pole. Cognitive language includes logical concepts such as conjecture, discernment and contradiction. The emotional dimension includes emotive actions such as cringe, terrify and exclaim.

To evaluate these dimensions of language in context, we next inspect prototypical speech snippets that correspond to the emotional and rational poles. After sampling speeches from the top and the bottom of the distribution, we then sample the most emotive and cognitive sentences within those speeches, for a qualitative analysis.[3]

Online Appendix Table A1 provides lists of example sentences for the most emotional and most cognitive speeches. Consistent with the word clouds, there is a clear differential in the tone, following intuitive language for logic and emotion. For example, the emotional sentences feature tributes to colleagues and to veterans, while the cognitive sentence include dry enumerations of policy details.[4] Differences in emotionality also emerge within specific topics: as illustrative examples, we report the most emotional and most cognitive sentences about taxation (Online Appendix Table A2) and abortion (Online Appendix Table A3).

A potential question with our measure is that it might capture positive and negative sentiment, as opposed to cognition and emotion. Hence, we would like to demonstrate that emotionality is a separable dimension of language from sentiment. For this purpose, we construct positive and negative sentiment dimensions in our embedding space using our centroid method, with positive and negative seed lexicons taken from Demszky *et al.* (2019) (see Online Appendix A.4 for details). We can then assess whether emotionality and sentiment dimensions work independently.

First, we inspect the 2×2 semantic context around four centroids in our embedding space: cognitive-positive, cognitive-negative, emotive-positive and emotive-negative. Online Appendix Figure A1 shows word clouds for the closest vectors to these four poles, revealing intuitive and distinctive words in each of these groups. The cognitive dimension has both positive tone (discern, knowledge, insight) and negative tone (contradict, vague, irrelevant). For emotion, the positive (serene, smile, thrill) and negative (frighten, disgust, sicken) are even more divergent.

Similarly, in our dataset of speeches, emotionality and sentiment are separable. Online Appendix Figure A2 provides a scatter plot of speeches across the two dimensions and shows that they are only weakly positively correlated. The R^2 from regressing emotionality on sentiment is just 0.011.

2.2. *Validation with Human Judgement*

This subsection reports the results of a human annotation task to assess the validity of our score in capturing emotion and cognition in language (e.g., Lowe and Benoit, 2013). The task is as follows. Coders are provided with pairs of sentences extracted from the corpus. For each pair, they are asked which sentence is more emotional and which sentence is more cognitive. In particular, the coder is provided with three options: (*i*) sentence A is more emotional[logical] than sentence B, (*ii*) sentence B is more emotional[logical] than sentence A and (*iii*) the sentences are equivalent or I don't understand one or both of the sentences. No additional information is provided about where the snippets come from.

In the baseline validation check, sentence pairs are constructed as follows. We start by selecting the five thousand most and least emotional speeches for each decade. From those speeches, we extract and score all the sentences. Finally, we randomly pair sentences that come from the top

---

[3] Specifically, we select speeches that fall within the 1st and 99th percentiles for the score distribution. We then exact ten random sentences among the highest and lowest scoring sentences within the sample.

[4] Online Appendix Table A5 shows additional examples where we have excluded any sentences containing a word from the lexicons. These sentences are still clearly and intuitively related to emotion and logic, respectively, yet they would be missed by a lexicon-based approach.

and bottom 5% of the score distribution. Pairs are always formed from sentences that come from the same decade, and all decades are roughly equally represented in the set of annotated snippets.

The annotators are Amazon Mechanical Turk workers born in the USA and whose primary language is English. Each coder is asked to code ten sentence pairs (twenty sentences). To assess inter-coder reliability, each pair of sentences is annotated by two different coders. In addition, each coder took a simple English comprehension test, which asked them to correctly separate a set of unambiguously emotion and cognition words into two groups.[5]

We obtain 1,714 annotations in total. The coders chose option (iii) (could not understand the snippets or judge the relative emotionality) for only 3.5% of the sentence pairs. These pairs are not considered for the computed accuracy statistics.

Table 1, panel A reports the results for the main validation exercise. The top row ('Overall') shows the statistics for the full sample of annotated pairs. In the full sample of annotations (columns (1)–(3)), our score agrees with human judgement 87% of the time. When we restrict to the sample of coders who passed the English comprehension test (columns (4)–(6)), our score agrees with human judgement 92% of the time. If, alternatively, we restrict to annotated pairs where both assigned coders agree on the ranking (columns (7)–(9)), accuracy reaches 93%.

Even if the measure is accurate overall, we must still confront the possibility that it is not valid for the older time periods in our sample. In particular, inconsistency may come from the use of modern seed dictionaries to evaluate dimensions of language in the past, since word meanings shift over time (Hamilton *et al.*, 2016; Garg *et al.*, 2018). We address this issue directly. In panel B, we report analogous statistics when subsetting the pairs by the decade when they were spoken in Congress, starting from the first decade, i.e., 1858–1868, up until the sixteenth (incomplete) decade, i.e., 2008–2014. Importantly, there are no significant drops in the accuracy of our score in earlier decades. This temporal validation addresses a major concern with our method: although it relies on recently developed seed dictionaries that use modern understandings of emotional and cognitive language, the final score produces a time-consistent measure of emotionality. Therefore we can produce meaningful long-run historical comparisons.

In panel C, we compare the performance in human validation for the two alternative measures of emotionality, described above in the methods section. First, the geometric measure $\mathrm{sim}\overrightarrow{AC}$ refers to the cosine similarity between each document vector $\boldsymbol{d}$ and the cognition-to-emotion dimension $\boldsymbol{AC}$, as done in Kozlowski *et al.* (2019). This vector-distance alternative obtains very similar accuracy to our baseline measure in the human validation task. Second, *word count* refers to the count-based measure giving the ratio of emotion words to cognition words. The performance for the count-based measure is much worse than the embedding-based measures and comparable to random guessing.[6]

3. Empirical Analysis

This section reports the results of our descriptive analysis of emotionality in U.S. Congress. We first explore whether emotional expression varies over time and across topics. Then, we show that members of the minority party resort systematically to more emotional rhetoric. The use of

---

[5] These words are, for emotion, love, afraid, glad, disgust, joy; and for cognition, consequence, therefore, discern, obvious, contradiction.

[6] Online Appendix A.7 provides some additional results on the human annotation validation. Online Appendix Table A6 reports a set of complementary assessments using alternative sentencing pairing procedures based on variants of the emotionality measure.

Table 1. *Human Validation of the Text Emotionality Measure.*

	Full sample			Restricted sample English comprehension			Restricted sample consistent coding		
	(1) Accuracy	(2) Blank	(3) Sample	(4) Accuracy	(5) Blank	(6) Sample	(7) Accuracy	(8) Blank	(9) Sample
Panel A: main analysis									
Overall	0.874	0.035	1,714	0.923	0.029	1,158	0.927	0.013	1,388
Panel B: main analysis by decade									
Decade 1	0.842	0.056	72	0.893	0.037	54	0.929	0	56
Decade 2	0.853	0.062	96	0.94	0.024	82	0.905	0.028	72
Decade 3	0.912	0.026	78	0.944	0.038	52	0.926	0	68
Decade 4	0.812	0.067	90	0.894	0.031	64	0.909	0.031	64
Decade 5	0.836	0.081	62	0.843	0.062	48	0.927	0.025	40
Decade 6	0.859	0.083	72	0.871	0.107	56	0.94	0.042	48
Decade 7	0.856	0.069	130	0.863	0.08	88	0.902	0.037	108
Decade 8	0.915	0.008	128	0.944	0	72	0.97	0.01	100
Decade 9	0.876	0.025	118	0.925	0.015	66	0.881	0.009	108
Decade 10	0.957	0.009	114	1	0	72	0.971	0.01	104
Decade 11	0.873	0	126	0.976	0	82	0.907	0	108
Decade 12	0.889	0.029	140	0.969	0.021	94	0.949	0.017	116
Decade 13	0.827	0.024	124	0.831	0.035	86	0.89	0	100
Decade 14	0.869	0.022	134	0.936	0	78	0.915	0.017	116
Decade 15	0.843	0.061	114	0.902	0.051	78	0.963	0	80
Decade 16	0.931	0	116	1	0	86	0.96	0	100
Panel C: alternative measures									
sim $\overrightarrow{AC}$	0.856	0.031	1,272	0.931	0.029	872	0.922	0.012	956
Word count	0.508	0.142	1,306	0.52	0.136	866	0.465	0.092	928

Notes: This table reports the results of the human validation. Panel A reports the main analysis with pairs formed by sentences with high and low emotionality scores. Panel B reports the breakdown of the main analysis by decade. Panel C reports results from alternative measures. *Full sample* indicates the full set of annotated sentences. *Restricted sample English comprehension* indicates a sample including only responses from coders who passed the English comprehension test. *Restricted sample consistent coding* indicates a sample including only responses consistently coded by two independent coders. *Accuracy* indicates the share of correct guesses over all guesses. *Blank* indicates the share of questions left blank over the total number of questions. *Sample* is the number of sentences in the sample.
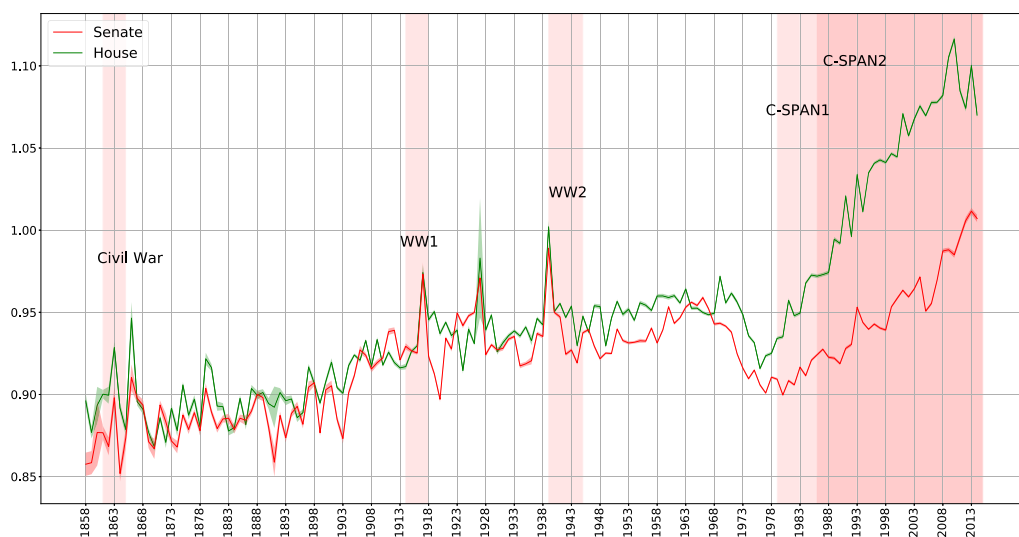
© The Author(s) 2022.

Fig. 2. *Emotionality in U.S. Congress by Chamber, 1858–2014.*
Notes: Time series of emotionality in the Senate (red) and the House of Representatives (green).

emotional language also differs across individual politician's characteristics, such as their gender, race and religion, and is positively correlated with ideological extremism.

3.1. *Emotionality over Time*

An initial descriptive question is how the relative use of emotion and reason has shifted over time. We use the long temporal range of our data to show the evolution of emotive language since the start of our data in 1858. These results add to other recent work looking at evolution of party polarisation in congressional speeches over this period; see Gentzkow *et al.* (2019b). Outside politics, Garg *et al.* (2018) used word embeddings to analyse the evolution of gender and ethnic stereotypes since 1910.

Our main descriptive results for emotionality over time are reported in Figure 2. The two time series show the average emotion score of speeches by year in the House of Representatives and in the Senate. Overall, we observe a generally increasing trend towards higher emotionality in political language, punctuated by some sudden spikes.

First let us consider the spikes in emotion in light of the intuition that political leaders express more emotions at pivotal moments in history (e.g., De Castella *et al.*, 2009). In our data, the first observed spike in the use of emotional language appears around the Civil War and its immediate aftermath (1861 to 1866). Two more major spikes occur in 1917 and 1939. These two years correspond to the entry of the United States into World War I (with President Wilson's declaration of war against Germany being approved by Congress) and the beginning of World War II (with Germany's invasion of Poland). The presence of higher emotionality during these events is intuitive and adds credibility about the behaviour of the measure.

Next, consider the broader trends. Emotionality makes a slow but steady increase up until the 1950s, then drops a bit in the early 1970s, and then starts a more rapid increase starting in the late

1970s that continues until the present. This striking pattern is seen for both chambers. The trend break is especially salient for the House of Representatives and is followed with some delay by the Senate.

We highlight that the emotionality trend is quite different from trends in text polarisation— that is, differences in the language used by Democrats and Republicans. Jensen *et al.* (2012) and Gentzkow *et al.* (2019b) found historically low levels that then increase only starting in the mid 1990s. Thus, the rise of emotive rhetoric pre-dates the current wave of polarised language. In addition, as shown in Online Appendix B.4, we can rule out that the shift is due to changes in the readability or simplicity of language (Benoit *et al.*, 2019). Emotional language is a distinct dimension of political rhetoric that evolves independently from other salient dimensions.

A potential concern is that these trends reflect changes in language generally, rather than changes in the political sphere. To check for this possibility, Online Appendix B.5 provides a comparison trend in emotionality for a more general historical corpus: Google Books. Emotional language in Google Books actually declines up until the 1980s, after which it shows a small rebound (Online Appendix Figure A12).[7] Thus, the trends we see in congressional speeches appear to be specific to politics. Online Appendix Figure A13 shows that we can normalise the congressional measures by the general-corpus emotionality and the qualitative trends are unchanged.[8]

Given that these trends are indeed politics specific, that makes the trend break in the late 1970s especially noteworthy. An intriguing possible explanation is the introduction of C-SPAN, a public television network for Congress that started broadcasting from the House in 1979 and from the Senate in 1986. Zooming in on this time period, we note that the first Congress elected after the founding of C-SPAN takes office in 1977. This is the precise timing of the trend break in emotional language. It could be that, when television comes online in Congress, that increases the marginal benefit to emotional language in floor speeches as more voters are now viewing them. While such a mechanism deserves additional investigation, it would be consistent with previous empirical work on the effectiveness of emotional appeals in influencing voters (Brader *et al.*, 2008; Gross, 2008; Renshon *et al.*, 2015).

3.2. *Emotionality and Topics*

Our second descriptive analysis is to look at how emotive-cognitive content for congressional debates varies by topic. In particular, the observed emotional variation over time may be driven by the selection of different topics over time, and that politicians talk more about emotionally charged issues in recent years. Alternatively, politicians may have changed their rhetoric style in how the same topics are framed.

To understand the relationship between emotionality and topics, we apply an unsupervised topic model (latent Dirichlet allocation or LDA; see, e.g., Blei, 2012). This is the same approach used by Hansen *et al.* (2018) to analyse the content of Federal Reserve committee transcripts.

---

[7] This trend in emotional expression is similar to that estimated by Morin and Acerbi (2017), who also used Google Books but focused on fiction. They wrote: 'Our data confirm that the decrease in emotionality in English-speaking literature is no artefact of the Google Books corpus, and that it pre-dates the twentieth century, plausibly beginning in the early nineteenth century'. Acerbi *et al.* (2013) found similar results.

[8] Note that the decreasing trend in Google Books also addresses another potential issue with our measure: that it is built with LIWC, a dictionary based on contemporary language as of 2015. On top of the consistent rates of human validation across decades (Table 1), this confirms again that our measure is not just picking up increasing use of the language used in LIWC; if that were the case, we would also see a similar increase in Google Books.

To summarise, LDA works by assuming a structural model for language, where documents are distributions over topics and topics are distributions over words. The parameters of these distributions are learned from the corpus and then produce interpretable topics to assist unpacking of text results.

We apply LDA to the full pre-processed corpus, with speeches treated as documents and assuming 128 topics. To get at non-emotive dimensions in language, we drop from the vocabulary all words in our emotive-cognitive lexicon. Online Appendix Table A11 lists the topics learned by the topic model and the most representative words for each topic. Overall, the quality is good and 119 of the 128 topics are recognisable as a coherent topic. For ease of interpretation, we inspected the individual topics and aggregated them into eleven larger categories (also indicated in Online Appendix Table A11).

Using the trained model, we assign to each speech the topic with the highest probability based on the speech content. Online Appendix Figure A17 shows the historical proportions of the eleven broad topic categories in congressional speeches over time. Speeches concerning procedural aspects of decision-making comprise the largest single category.[9] The share of procedural speeches shrinks slightly over time, mostly in favour of speeches on social issues and speeches that hinge upon a national narrative, historical heritage or patriotism. Given the proportional importance of procedure, Online Appendix Figure A19 shows robustness of our main time series results to dropping procedural speeches.

To show topic-level variation in emotional expression, we residualise out time fixed effects (to adjust for secular trends) and then compute the average topic-specific emotionality. Figure 3(a) plots this variation for the eleven topic categories since 1970, producing an intuitive ranking.[10] The most emotional category corresponds to *National Narrative*, a 'patriotism' topic including references to American history, heritage, values, as well as to the sacrifice of American soldiers. Second, the *Foreign Policy* category includes highly emotive speeches on human right violations and the Cold War threat. The ranking of *Social Issues* (e.g., crime, abortion), *Party Politics* and *Immigration* as emotional is similarly sensible. On the other side of the spectrum, it is not surprising that speeches referring to internal *Procedure* and *Governance* (government organisation) tend to rank low and to use more cognitive language.

How does this topic-level variation in emotionality vary by political party affiliation? Figure 3(b) reports the ratio of Republican emotionality to Democrat emotionality by topic. First, and perhaps most strikingly, fiscal policy is the most Republican-slanted topic in its emotional content, with the Republican score being 2.5 times larger than the Democrat score. In comparison, most other topics are quite similar across parties in emotive content. The exceptions are two Democrat-slanted topics: social issues, which makes sense in light of Democrats' defence of civil rights and women's rights, and economic policy, a topic that is focused on regulation of corporate misbehaviour. Thus, emotionality helps capture partisan differences in policy priorities.

Next, we explore the time series in emotionality by topic. As illustrated in Figure 4(a), speeches about *Procedure* are the least emotionally charged and their low level remains constant over time. *National Narrative* scores the highest throughout the period, and follows the generally increasing trend. Recalling the discussion above on the emotive trend break in the late 1970s in concert with the arrival of C-SPAN, it is notable that *Economy* and *Society* have the steepest relative increases starting at that point in history. It makes sense that, when the public becomes a more salient audience, congressmen start speaking more emotionally about topics of general interest.

[9] See Online Appendix Figure A18 for the time series of topic shares after excluding procedural speeches.
[10] For the rankings of all 128 individual topics, see Online Appendix Figure A16.
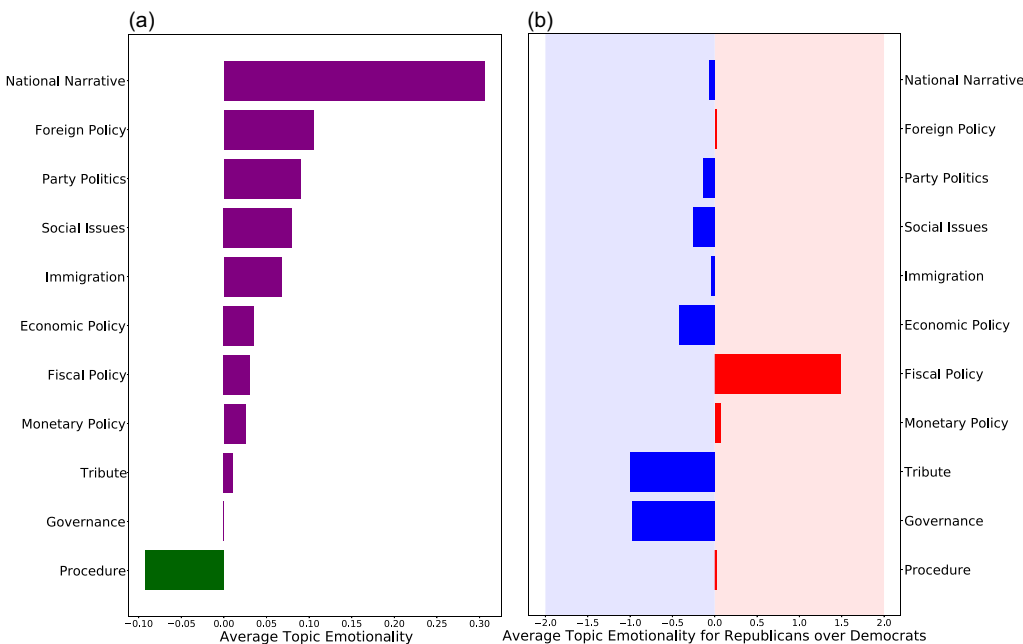
Fig. 3. *Emotionality by Topic and Party, and Average Emotionality by Topic, 1970–2014.*
Notes: Panel (a) reports the emotionality score by topic. Panel (a) reports the ratio of average emotionality
by topic for Republicans over Democrats, centred at 0. Values larger than 0 indicate the Republicans use
more emotionality than Democrats when talking about the given topic. Values smaller than 0 indicate the
Democrats use more emotionality than Republicans when talking about the given topic. The emotionality
score is demeaned by the average level of emotionality in each Congress number. We include only policy
topics. The topic *Other* is excluded.

Figure 4(b) focuses on speeches from the larger category of *Economy* and shows the breakdown
by the three main components: fiscal, monetary and economic (regulatory) policies. In the earlier
decades, there was a persistent emotive ranking from regulatory policy to monetary policy to fiscal
policy. Around the 96th Congress (late 1970s), however, the trend break for fiscal policy is most
intense. By the 102nd Congress (1991–1992), fiscal policy had become the most emotionally
charged topic among economic issues. An intriguing feature of this time period is that it coincides
with the Reaganite transformation of fiscal policy and the associated shifts in income and wealth
inequality. This result resonates with the partisan slant in emotionality about fiscal policy, and
suggests that Republicans use emotional rhetoric to defend inequality-increasing fiscal policies.
In light of the evidence that economic inequality increases political polarisation (Garand, 2010;
McCarty *et al.*, 2016; Piketty, 2020), it makes sense that divisive issues related to redistribution
have become more emotionally charged.

3.3. *Emotionality and Politician Characteristics*

So far, we have looked at the broad temporal, topical and partisan factors explaining emotion
and reason in U.S. legislative politics. In this section we assess how emotionality varies across
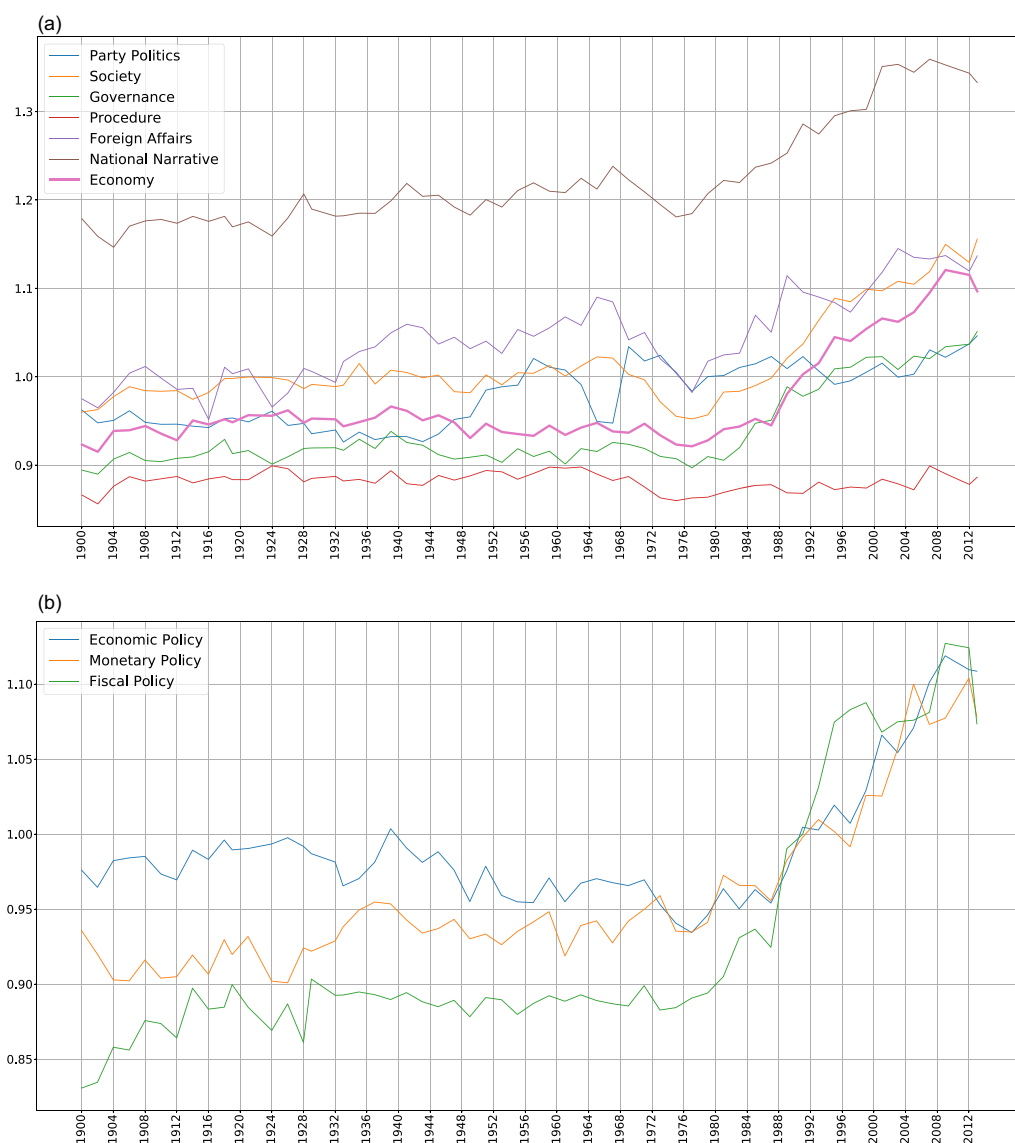
© The Author(s) 2022.

Fig. 4. *Emotionality by Topic Over Time and Time Series of Emotionality by Topic, 1900–2014.*
Notes: Panel (a) reports all topics, excluding *Other*. Panel (b) reports the breakdown of the *Economy* topic
into its three components, i.e., *Fiscal Policy*, *Monetary Policy* and *Economic Policy*.

politicians. First, we look at party opposition status. Second, we attend to politician identity
characteristics. Third, we compare to partisanship in voting records.

3.3.1. *Opposition status*
First, we explore whether U.S. politicians resort to emotionality more when they are in the
opposition. As discussed in Green (2015) and Lee (2016), minority-party politicians are engaged
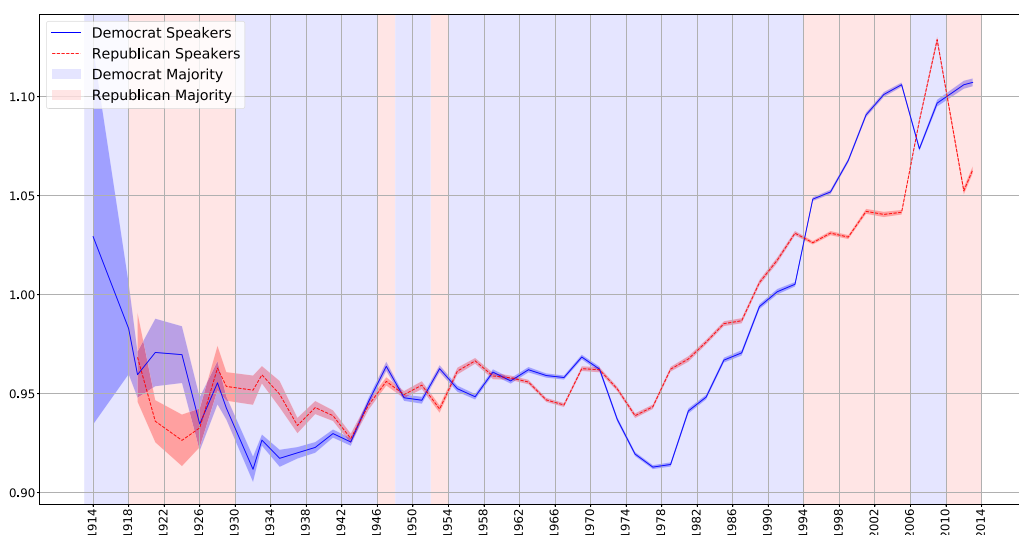
Fig. 5. *House Member Emotionality by Party and by Party Majority.*
Notes: Time series of emotionality in the House of Representatives for Democrats (blue) and Republicans (red), 1900–2014. Blue and red areas indicate Democratic majorities in the House of Representatives.

in crafting a national message to accrue electoral gains in upcoming campaigns. Emotional language can be used to communicate large and consensual values (Jerit, 2004), and it is more likely to be reported by traditional and social media (Bennett, 2016; Brady *et al.*, 2017). Thus, politicians may use more emotional language when they are in the minority party.

As initial visual evidence on this point, Figure 5 plots the average level of emotionality by politician party in the House of Representatives. The background colour indicates the party with majority control of the chamber. We see that, overall, Democrats and Republicans do not differ much in their use of emotional language. However, members of the minority party are systematically more emotional than members of the majority party, a striking trend that we see consistently flip as the party in control flips. In the long term of Democrat control in the second half of the twentieth century, Republicans consistently used more emotional language. In turn, after Republicans retook the House in 1994, Democrats were more emotive. Throughout the time series, changes in House majorities correspond to changes in relative emotionality in the two parties.[11]

3.3.2. *Partisanship in voting*
To delve further into the role of emotional rhetoric in political division, we next explore its relation to ideological policy choices. Previous work has shown that ideological extremists use dissent with their own party to appeal to extreme voters (Kirkland and Slapin, 2018). Extremism

---

[11] Online Appendix Table A9 reports the estimates from a series of ordinary least squares regressions for the effect of opposition status on the emotion score, for both the House and the Senate. The regressions include chamber-year fixed effects and standard errors are clustered by politician. These results confirm that the dynamic relation noted in Figure 5 is statistically significant when looking at both chambers. Including politician fixed effects reveals that the same politician uses more emotional appeals when her party is in a minority position, relative to her personal average level. Results are not driven by the choice of different topics, for example due to mechanical differences in responsibility for procedural functions.

© The Author(s) 2022.

Table 2. *How Emotionality Varies by Politician Characteristics.*

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
$(DWnom1)^2$	0.016***					0.014***	0.017***
	[0.003]					[0.003]	[0.002]
Democrat		0.027***				0.012**	0.018***
		[0.006]				[0.006]	[0.004]
Female			0.268***			0.253***	0.158***
			[0.016]			[0.016]	[0.010]
Black				0.168***		0.136***	0.067***
				[0.026]		[0.027]	[0.017]
Hispanic				0.109***		0.081**	0.040*
				[0.034]		[0.035]	[0.022]
Asian				−0.039		−0.074*	−0.095***
				[0.038]		[0.042]	[0.025]
Catholic					0.043***	0.038***	0.016***
					[0.010]	[0.010]	[0.006]
Jewish					0.053***	0.056***	0.001
					[0.017]	[0.016]	[0.010]
Chamber-year FEs	Y	Y	Y	Y	Y	Y	Y
Topic FEs							Y
Observations	5,593,863	5,593,863	5,593,863	5,593,863	5,593,863	5,593,863	5,593,863
R^2	0.05	0.05	0.05	0.05	0.05	0.05	0.37

Notes: Each column shows the OLS regression of the standardised emotionality score in a given speech on individual politician characteristics. The sample is composed of all speeches pronounced by Democrat and Republican members of Congress between 1858 and 2014. All specifications include chamber-year fixed effects. Column (6) also includes topic fixed effects. SEs are clustered at the politician level. *, **, *** denote significance at the 10%, 5% and 1% levels, respectively.

can also be associated with simpler sentences and longer speeches (Slapin and Kirkland, 2020). We test whether members of Congress that are more ideologically polarised are also more likely to use emotional rhetoric.

We measure ideological extremism using DW-NOMINATE, a standard measure constructed from roll call votes. DW-NOMINATE summarises the tendency of a congressman to vote with Republicans versus with Democrats. As initial visual evidence for a relationship, we see in Online Appendix Figure A21 that there is a U-shaped pattern between emotionality and vote partisanship. Congressmen with more extreme ideological positions (either left or right) tend to use more emotionally charged language in their floor speeches.

We test the statistical significance of this relationship by regressing emotionality on the squared DW-NOMINATE score, such that it takes larger values for more extreme roll call voting on either the left or the right. Both emotionality and partisanship are standardised to standard deviation one to facilitate interpretation of the coefficients. Chamber-year fixed effects are included to adjust for any chamber-level time-varying factors influencing rhetorical choices, and standard errors are clustered by politician to allow for serial correlation in the error term by politician across speeches and over time.

Table 2 reports the results in the first row of estimates. There is a significant and positive relationship between ideological voting and emotionality (column (1)), which holds when adjusting for politician demographics (column (6)). The ideology effect is of the same magnitude even when conditioning on topic fixed effects (column (7)), showing that polarised rhetoric comes through the framing of topics rather than selection of topics.[12]

---

[12] Online Appendix Table A10 shows robustness of these results when including controls for sentiment, speech length and minority status.

3.3.3. *Identity characteristics*

The next question is: are there observable *individual* characteristics of Congress members who tend to use more emotional language? For example, members from demographic groups that are underrepresented in Congress are typically associated with distinctive policy position and representation choices (e.g., Swers, 2002; Tate, 2018). We explore whether members of under-represented groups are also more likely to use emotional language.

Demographic information on members of Congress comes from the CQ Press Congress Collection. The explanatory variables of interest are indicator variables for party, gender, race and religion. The regression specification is the same as that used above for DW-NOMINATE, with chamber-year fixed effects used to compare Congressmen to their contemporaneous colleagues.

As reported in Table 2, we find that Democrat members of Congress tend to use more emotional language than Republicans (column (2)). Column (3) shows that females tend to use more emotional language compared to males, while column (4) shows higher emotion for historically disadvantaged minorities in Congress (blacks and Hispanics). Finally, column (5) shows that religious minorities (Catholics and Jews) use more emotional language than Protestants. These factors remain statistically significant when taken together in a single regression (column (6)) or when conditioning on topic fixed effects (column (7)). In terms of magnitudes, the effects of political party are relatively small (0.01 SD in column (6)), while the effects for gender are relatively substantial (0.25 SD in column (6)).[13]

Considering also the estimate for DW-NOMINATE, this evidence suggests that identity characteristics are more pivotal than political factors for variation in emotion across politicians. An important question, then, is whether the long-run changes in emotionality from Figure 2 are due to a changing composition of politician types. We explore this issue in Online Appendix Figure A8, where we compare the unconditional time trends with the emotionality score residualised on demographic variables (gender, race, religion). We can see that after this adjustment for demographics, the trends in emotionality remain virtually unchanged.[14]

4. Conclusion

This paper has provided an analysis of emotion and reason in the language of U.S. members of Congress. We produced a new measure of emotive speech, which combines dictionary methods with word embeddings to look at the relative use of affective and cognitive language. We then analysed how that measure evolves over time, varies across individuals and changes in response to electoral and media pressures.

This paper's substantive contribution is, first, in showing that the secular trends in increasing polarisation and increasing simplification of speaking styles have been accompanied by increasing intensity in the expression of emotion. This new rhetoric is more populist in nature, addressing polarised voters rather than fellow politicians, bureaucrats or elites. In line with this key idea, we

---

[13] Online Appendix Table A10 reproduces the main results controlling for the length of the speech and positive or negative sentiment. The results are unaffected by the inclusion of these controls on rhetoric styles: this suggests that the group differences in emotionality are not driven by a different use of language. Online Appendix Table A10 also shows that these relationships are roughly constant over time. To provide additional visual support for these estimates, Online Appendix Figure A20 shows the time series of emotionality by gender and race. The differences in the use of emotional language across demographic groups is constant over time. Online Appendix Table A8 reports the same results on the two separate components of the emotionality score—emotion by itself, and reason by itself. See Online Appendix C for a discussion of the differences in those results.

[14] In Online Appendix Figure A9 we also include topic fixed effects and can explain some, but not all, of the long-run changes in emotionality.

find that emotionality has been increasing over time in Congress while it has been decreasing in the broader culture. The steep increase since the 1970s appears to be related to the introduction of televised congressional debates via C-SPAN. These trends speak to the importance of media technology in the strategic value of emotional rhetoric, just as previous work has shown television's connection with partisanship and populism (DellaVigna and Kaplan, 2007; Martin and Yurukoglu, 2017; Durante *et al.*, 2019).

Second, we produce a series of results on how emotional rhetoric is related to power imbalance and conflict. Emotionality is higher for less empowered political minorities: women, Hispanics, blacks, Jews and Catholics. The status of being in the minority party and therefore having less power over policy increases emotional language. Relatedly, we find evidence for emotions as a response for conflict. They increase during wars. Income inequality is an ingredient for class conflict over redistribution, which we can observe in high emotional intensity on fiscal policy. And finally, we find evidence that the more divisive and ideologically polarised members of Congress tend to use more emotional rhetoric.

The new measurement approach and initial descriptive results set the stage for a number of further empirical studies. Notably, further research is needed to understand the role of television in increasing emotional rhetoric. This work could go beyond the introduction of C-SPAN, for example to focus on partisan cable news (e.g., Clinton and Enamorado, 2014; Arceneaux *et al.*, 2016; Martin and Yurukoglu, 2017). Such an investigation should consider how electoral incentives interact with new visibility obtained through television to influence rhetorical choices of politicians.

Another important open question concerns the relation between emotionality and polarisation. As affective polarisation in the electorate is on the rise in the United States and Europe alike (Iyengar *et al.*, 2019), more attention should be devoted to understand the possible feedback loops between polarisation and emotive speech in parliaments. Do politicians use more emotion when discussing people, policies or principles?

Beyond these substantive avenues, the new emotionality metric could itself be a useful tool to be applied in other empirical contexts. In Congress, analysing committee debates would be a natural next step to delve deeper into congressional dynamics. Measuring emotional expression in newsletters that congressmen send to their constituents would provide for interesting insights into linkages between a politician and her constituency. Outside of politics, news articles or television transcripts would be perfect candidates to provide evidence on how expressed emotion is used for different persuasive and professional purposes.

Finally, our methodology may inform experimental studies of how emotionality in political language influences voters. Using the emotion metric combined with generative language models (e.g., Radford *et al.*, 2019; Brown *et al.*, 2020), it is possible to identify or generate comparable political arguments that differ in their use of emotive language. More causal analysis of how emotions influence voters is needed to validate the mechanism of emotional rhetoric as a strategic response to voter preferences.

ETH Zurich, Switzerland
ETH Zurich, Switzerland

Additional Supporting Information may be found in the online version of this article:

Online Appendix
Replication Package

References

Acerbi, A., Lampos, V., Garnett, P. and Bentley, R.A. (2013). 'The expression of emotions in 20th century books', *PloS One*, vol. 8(3), e59030.

Arceneaux, K., Johnson, M., Lindstädt, R. and Vander Wielen, R.J. (2016). 'The influence of news media on political elites: Investigating strategic responsiveness in congress', *American Journal of Political Science*, vol. 60(1), pp. 5–29.

Arora, S., Liang, Y. and Ma, T. (2017). 'A simple but tough-to-beat baseline for sentence embeddings', *Proceedings of the 2017 International Conference on Learning Representations*, Toulon: OpenReview.net.

Ash, E., Chen, D.L. and Ornaghi, A. (2021). 'Gender attitudes in the judiciary: Evidence from US circuit courts', Center for Law & Economics Working Paper Series 2019(02).

Ash, E., Morelli, M. and Van Weelden, R. (2017). 'Elections and divisiveness: Theory and evidence', *The Journal of Politics*, vol. 79(4), pp. 1268–85.

Baker, S.R., Bloom, N. and Davis, S.J. (2016). 'Measuring economic policy uncertainty', *The Quarterly Journal of Economics*, vol. 131(4), pp. 1593–636.

Bennett, W.L. (2016). *News: The Politics of Illusion*, Chicago: University of Chicago Press.

Benoit, K., Munger, K. and Spirling, A. (2019). 'Measuring and explaining political sophistication through textual complexity', *American Journal of Political Science*, vol. 63(2), pp. 491–508.

Blei, D.M. (2012). 'Probabilistic topic models', *Communications of the ACM*, vol. 55(4), pp. 77–84.

Bosman, R. and Van Winden, F. (2002). 'Emotional hazard in a power-to-take experiment', ECONOMIC JOURNAL, vol. 112(476), pp. 147–69.

Boussalis, C., Coan, T., Holman, M. and Müller, S. (2021). 'Gender, candidate emotional expression, and voter reactions during televised debates', *American Political Science Review*, vol. 115(4), pp. 1242–57.

Brader, T., Valentino, N.A. and Suhay, E. (2008). 'What triggers public opposition to immigration? Anxiety, group cues, and immigration threat', *American Journal of Political Science*, vol. 52(4), pp. 959–78.

Brady, W.J., Wills, J.A., Jost, J.T., Tucker, J.A. and Van Bavel, J.J. (2017). 'Emotion shapes the diffusion of moralized content in social networks', *Proceedings of the National Academy of Sciences*, vol. 114(28), pp. 7313–8.

Brown, T.B., Mann, B., Ryder, N., Subbiah, M., Kaplan, J., Dhariwal, P., Neelakantan, A., Shyam, P., Sastry, G., Askell, A., Agarwal, S., Herbert-Voss, A., Krueger, G., Henighan, T., Child, R., Ramesh, A., Ziegler, D., Wu, J., Winter, C., Hesse, C., Chen, M., Sigler, E., Litwin, M., Gray, S., Chess, B., Clark, J., Berner, C., McCandlish, S., Radford, A., Sutskever, I. and Amodei, D. (2020). 'Language models are few-shot learners', Preprint, https://arxiv.org/abs/2005.14165v4.

Caliskan, A., Bryson, J.J. and Narayanan, A. (2017). 'Semantics derived automatically from language corpora contain human-like biases', *Science*, vol. 356(6334), pp. 183–6.

Clinton, J.D. and Enamorado, T. (2014). 'The national news media's effect on Congress: How Fox News affected elites in Congress', *The Journal of Politics*, vol. 76(4), pp. 928–43.

Collobert, R. and Weston, J. (2008). 'A unified architecture for natural language processing: Deep neural networks with multitask learning', in *Proceedings of the 25th International Conference on Machine Learning*, pp. 160–7, New York: Association for Computing Machinery.

CQPress. (2021). *CQ Congress Collection Database*, CQ Press Library Database Collection.

Damasio, A.R. (1995). 'Toward a neurobiology of emotion and feeling: Operational concepts and hypotheses', *The Neuroscientist*, vol. 1(1), pp. 19–25.

De Castella, K., McGarty, C. and Musgrove, L. (2009). 'Fear appeals in political rhetoric about terrorism: An analysis of speeches by Australian Prime Minister Howard', *Political Psychology*, vol. 30(1), pp. 1–26.

DellaVigna, S. and Kaplan, E. (2007). 'The Fox News effect: Media bias and voting', *The Quarterly Journal of Economics*, vol. 122(3), pp. 1187–234.

Demszky, D., Garg, N., Voigt, R., Zou, J., Gentzkow, M., Shapiro, J. and Jurafsky, D. (2019). 'Analyzing polarization in social media: Method and application to tweets on 21 mass shootings', Preprint, https://arxiv.org/abs/1904.01596.

Dietrich, B.J., Hayes, M. and O'Brien, D.Z. (2019). 'Pitch perfect: Vocal pitch and the emotional intensity of congressional speech', *American Political Science Review*, vol. 113(4), pp. 941–62.

Durante, R., Pinotti, P. and Tesei, A. (2019). 'The political legacy of entertainment TV', *American Economic Review*, vol. 109(7), pp. 2497–530.

Elster, J. (1999). *Alchemies of the Mind: Rationality and the Emotions*, Cambridge: Cambridge University Press.

Enke, B. (2020). 'Moral values and voting', *Journal of Political Economy*, vol. 128(10), pp. 3679–729.

Firth, J.R. (1957). 'A synopsis of linguistic theory, 1930–1955', *Studies in Linguistic Analysis*. Special Volume of the Philological Society, pp. 1–31.

Frank, R.H. (1988). *Passions Within Reason: The Strategic Role of the Emotions*, New York: Norton & Company.

Garand, J.C. (2010). 'Income inequality, party polarization, and roll-call voting in the US senate', *The Journal of Politics*, vol. 72(4), pp. 1109–28.

Garg, N., Schiebinger, L., Jurafsky, D. and Zou, J. (2018). 'Word embeddings quantify 100 years of gender and ethnic stereotypes', *Proceedings of the National Academy of Sciences*, vol. 115(16), pp. E3635–44.

Gentzkow, M. (2006). 'Television and voter turnout', *The Quarterly Journal of Economics*, vol. 121(3), pp. 931–72.

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

Gentzkow, M. and Shapiro, J.M. (2010). 'What drives media slant? Evidence from US daily newspapers', *Econometrica*, vol. 78(1), pp. 35–71.

Gentzkow, M., Shapiro, J.M. and Taddy, M. (2019b). 'Measuring group differences in high-dimensional choices: Method and application to congressional speech', *Econometrica*, vol. 87(4), pp. 1307–40.

Goet, N.D. (2019). 'Measuring polarization with text analysis: Evidence from the UK House of Commons, 1811–2015', *Political Analysis*, vol. 27(4), pp. 518–39.

Green, M.N. (2015). *Underdog Politics: The Minority Party in the US House of Representatives*, New Haven, CT: Yale University Press.

Grimmer, J. and Stewart, B.M. (2013). 'Text as data: The promise and pitfalls of automatic content analysis methods for political texts', *Political Analysis*, vol. 21(3), pp. 267–97.

Gross, K. (2008). 'Framing persuasive appeals: Episodic and thematic framing, emotional response, and policy opinion', *Political Psychology*, vol. 29(2), pp. 169–92.

Hamilton, W.L., Leskovec, J. and Jurafsky, D. (2016). 'Diachronic word embeddings reveal statistical laws of semantic change', Preprint, https://arxiv.org/abs/1605.09096.

Hansen, S., McMahon, M. and Prat, A. (2018). 'Transparency and deliberation within the FOMC: A computational linguistics approach', *The Quarterly Journal of Economics*, vol. 133(2), pp. 801–70.

Hargrave, L. and Blumenau, J. (2021). 'No Longer Conforming to Stereotypes? Gender, Political Style, and Parliamentary Debate in the UK', *British Journal of Political Science*, forthcoming.

Iyengar, S., Lelkes, Y., Levendusky, M., Malhotra, N. and Westwood, S.J. (2019). 'The origins and consequences of affective polarization in the United States', *Annual Review of Political Science*, vol. 22, pp. 129–46.

Jensen, J., Naidu, S., Kaplan, E., Wilse-Samson, L., Gergen, D., Zuckerman, M. and Spirling, A. (2012). 'Political polarization and the dynamics of political language: Evidence from 130 years of partisan speech [with comments and discussion]', *Brookings Papers on Economic Activity*, pp. 1–81.

Jerit, J. (2004). 'Survival of the fittest: Rhetoric during the course of an election campaign', *Political Psychology*, vol. 25(4), pp. 563–75.

Jerit, J., Kuklinski, J.H. and Quirk, P.J. (2009). 'Strategic politicians, emotional citizens, and the rhetoric of prediction', in (E. Borgida, C.M Federico and J.L. Sullivan, eds.), *The Political Psychology of Democratic Citizenship*, pp. 100–24, New York: Oxford University Press.

Jordan, K.N., Sterling, J., Pennebaker, J.W. and Boyd, R.L. (2019). 'Examining long-term trends in politics and culture through language of political leaders and cultural institutions', *Proceedings of the National Academy of Sciences*, vol. 116(9), pp. 3476–81.

Kahneman, D. (2011). *Thinking, Fast and Slow*, London: Penguin.

Kirkland, J.H. and Slapin, J.B. (2018). *Roll Call Rebels: Strategic Dissent in the United States and United Kingdom*, Cambridge: Cambridge University Press.

Kosmidis, S., Hobolt, S.B., Molloy, E. and Whitefield, S. (2019). 'Party competition and emotive rhetoric', *Comparative Political Studies*, vol. 52(6), pp. 811–37.

Kozlowski, A.C., Taddy, M. and Evans, J.A. (2019). 'The geometry of culture: Analyzing the meanings of class through word embeddings', *American Sociological Review*, vol. 84(5), pp. 905–49.

Lau, R.R. and Rovner, I.B. (2009). 'Negative campaigning', *Annual Review of Political Science*, vol. 12, pp. 285–306.

LeDoux, J. (1998). *The Emotional Brain: The Mysterious Underpinnings of Emotional Life*, New York: Simon & Schuster.

Lee, F.E. (2016). *Insecure Majorities: Congress and the Perpetual Campaign*, Chicago: University of Chicago Press.

Lerner, J.S., Li, Y., Valdesolo, P. and Kassam, K.S. (2015). 'Emotion and decision making', *Annual Review of Psychology*, vol. 66, pp. 799–823.

Lewis, J.B., Poole, K., Rosenthal, H., Boche, A., Rudkin, A. and Sonnet, L. (2019). 'Voteview: Congressional roll-call votes database', https://voteview.com/ (last accessed: 27 July 2018).

Lim, E.T. (2002). 'Five trends in presidential rhetoric: An analysis of rhetoric from George Washington to Bill Clinton', *Presidential Studies Quarterly*, vol. 32(2), pp. 328–48.

Lin, C.H., Chuang, S.C., Kao, D.T. and Kung, C.Y. (2006). 'The role of emotions in the endowment effect', *Journal of Economic Psychology*, vol. 27(4), pp. 589–97.

Loewenstein, G. (2000). 'Emotions in economic theory and economic behavior', *American Economic Review*, vol. 90(2), pp. 426–32.

Lowe, W. and Benoit, K. (2013). 'Validating estimates of latent traits from textual data using human judgement as a benchmark', *Political Analysis*, vol. 21(3), pp. 298–313.

MacLeod, W.B. (1996). 'Decision, contract, and emotion: Some economics for a complex and confusing world', *Canadian Journal of Economics*, vol. 29(4), pp. 788–810.

Marcus, G.E. (2000). 'Emotions in politics', *Annual Review of Political Science*, vol. 3(1), pp. 221–50.

Marcus, G.E., Neuman, W.R. and MacKuen, M. (2000). *Affective Intelligence and Political Judgement*, Chicago: University of Chicago Press.

Martin, G.J. and Yurukoglu, A. (2017). 'Bias in cable news: Persuasion and polarization', *American Economic Review*, vol. 107(9), pp. 2565–99.

McCarty, N., Poole, K.T. and Rosenthal, H. (2016). *Polarized America: The Dance of Ideology and Unequal Riches*, Cambridge, MA: MIT Press.

Michel, J.B., Shen, Y.K., Aiden, A.P., Veres, A., Gray, M.K., Team, G.B., Pickett, J.P., Hoiberg, D., Clancy, D., Norvig, P., Orwant, J., Pinker, S., Nowak, M. A. and Aiden, E. L. (2011). 'Quantitative analysis of culture using millions of digitized books', *Science*, vol. 331(6014), pp. 176–82.

Mikolov, T., Sutskever, I., Chen, K., Corrado, G. and Dean, J. (2013). 'Distributed representations of words and phrases and their compositionality', in (M.I. Jordan, Y. LeCun and S.A. Solla, eds.), *Advances in Neural Information Processing Systems*, pp. 3111–19.

Morin, O. and Acerbi, A. (2017). 'Birth of the cool: A two-centuries decline in emotional expression in anglophone fiction', *Cognition and Emotion*, vol. 31(8), pp. 1663–75.

Osnabrügge, M., Ash, E. and Morelli, M. (2021). Cross-Domain Topic Classification for Political Texts, *Political Analysis*, pp. 1–22, https://www.cambridge.org/core/journals/political-analysis/article/crossdomain-topic-classification-for-political-texts/F074564984969CE168BCBCF5E7D931C8.

Osnabrügge, M., Hobolt, S.B. and Rodon, T. (2021). 'Playing to the gallery: Emotive rhetoric in parliaments', *American Political Science Review*, vol. 115(3), pp. 1–15.

Pennebaker, J.W., Boyd, R.L., Jordan, K. and Blackburn, K. (2015). 'The development and psychometric properties of LIWC2015', University of Texas at Austin, https://repositories.lib.utexas.edu/bitstream/handle/2152/31333/LIWC2015_LanguageManual.pdf?sequence=3&isAllowed=y.

Pennington, J., Socher, R. and Manning, C.D. (2014). 'GloVe: global vectors for word representation', in *Proceedings of the 2014 Conference on Empirical Methods in Natural Language Processing*, pp. 1532–43, Association for Computational Linguistics.

Peterson, A. and Spirling, A. (2018). 'Classification accuracy as a substantive quantity of interest: Measuring polarization in westminster systems', *Political Analysis*, vol. 26(1), pp. 120–28.

Piketty, T. (2020). *Capital and Ideology*, Cambridge, MA: Harvard University Press.

Quinn, K.M., Monroe, B.L., Colaresi, M., Crespin, M.H. and Radev, D.R. (2010). 'How to analyze political attention with minimal assumptions and costs', *American Journal of Political Science*, vol. 54(1), pp. 209–28.

Radford, A., Wu, J., Child, R., Luan, D., Amodei, D. and Sutskever, I. (2019). 'Language models are unsupervised multitask learners', *OpenAI Blog*, vol. 1(8), p. 9.

Renshon, J., Lee, J.J. and Tingley, D. (2015). 'Physiological arousal and political beliefs', *Political Psychology*, vol. 36(5), pp. 569–85.

Rheault, L., Beelen, K., Cochrane, C. and Hirst, G. (2016). 'Measuring emotion in parliamentary debates with automated textual analysis', *PloS One*, vol. 11(12), e0168843.

Rheault, L. and Cochrane, C. (2020). 'Word embeddings for the analysis of ideological placement in parliamentary corpora', *Political Analysis*, vol. 28(1), pp. 112–33.

Rice, D.R. and Zorn, C. (2021). 'Corpus-based dictionaries for sentiment analysis of specialized vocabularies', *Political Science Research and Methods*, vol. 9(1), pp. 20–35.

Rodman, E. (2020). 'A timely intervention: Tracking the changing meanings of political concepts with word vectors', *Political Analysis*, vol. 28(1), pp. 87–111.

Rodriguez, P.L. and Spirling, A. (2022). 'Word embeddings: What works, what doesn't, and how to tell the difference for applied research', *Journal of Politics*, forthcoming.

Slapin, J.B. and Kirkland, J.H. (2020). 'The sound of rebellion: Voting dissent and legislative speech in the UK House of Commons', *Legislative Studies Quarterly*, vol. 45(2), pp. 153–76.

Swers, M.L. (2002). *The Difference Women Make: The Policy Impact of Women in Congress*, Chicago: University of Chicago Press.

Tate, K. (2018). *Black Faces in the Mirror: African Americans and their Representatives in the US Congress*, Princeton, NJ: Princeton University Press.

US Congress. (2021). *United States Congressional Record: 35th to 113th Congress*, United States Government Publishing Office.

Van Buskirk, J., Candolin, U. and Wong, B. (2012). 'Behavioural plasticity and environmental change', in (U. Candolin and B.B.M. Wong, eds.), *Behavioural Responses to a Changing World. Mechanisms and Consequences*, pp. 145–58, Oxford: Oxford University Press.

van Winden, F.A. and Ash, E. (2012). 'On the behavioral economics of crime', *Review of Law & Economics*, vol. 8(1), pp. 181–213.

Wälde, K. and Moors, A. (2017). 'Current emotion research in economics', *Emotion Review*, vol. 9(3), pp. 271–8.

Xiao, E. and Houser, D. (2005). 'Emotion expression in human punishment behavior', *Proceedings of the National Academy of Sciences*, vol. 102(20), pp. 7398–401.