

Do You Walk the Walk, Talk the Talk, or Tweet the Tweet? Legislators' Ideal Points Across Venues

SoRelle W. Gaynor, Kristina Miler, Pranav Goel, Alexander M. Hoyle, and Philip Resnik*

Prepared for the Women in Legislative Studies Research Seminar, Fall 2022
October 28, 2022

Abstract: Members of Congress face consistent pressure—and ample opportunity—to express their ideological positions. The most commonly measured outlet is congressional floor votes, but it is not the only one. This research develops a novel ideal point measure across three different ideological expressions—votes, floor speeches, and Tweets—to capture an understudied interaction between ideology, communication style, and audience. We examine the differences in congressional ideal points in the aggregate, as well as provide novel evidence of how individual legislators' position themselves differently depending on the venue. We find that legislators use speech and tweets to convey nuance in their ideological positions and to differentiate themselves in ways that voting does not allow. In both theory and methods, this work engages the considerable literatures on the measurement of ideology and congressional representation. This research contributes to our understanding of legislator behavior, ideological positioning, and introduces three ideal points to measure the ideology of members of the U. S. Congress: a vote-based ideal point and two text-based ideal points based on congressional speech and official Tweets.

**Gaynor is an Assistant Professor of Political Science at College of the Holy Cross (sgaynor@holycross.edu), Miler is an Associate Professor of Political Science at the University of Maryland, College Park (kmiler@umd.edu). They are the corresponding authors.*

Goel and Hoyle are PhD students in Natural Language Processing and Computer Science at the University of Maryland, College Park. Resnik holds a joint appointment as Professor in the University of Maryland Institute for Advanced Computer Studies and the Department of Linguistics, and an Affiliate Professor appointment in Computer Science.

Funding statement: This material is based upon work supported by the National Science Foundation under Grant No. (2008761)

Introduction

Members of Congress face consistent pressure—and ample opportunity—to express their ideological positions. The most commonly measured outlet is congressional floor votes, but it is not the only one. Members also have speaking opportunities on the chamber floor to discuss policy on the congressional record, and Twitter provides legislators with an opportunity to communicate their beliefs briefly and directly. Each of these three actions—voting, floor speeches, and Tweets—allow members to express their ideological positions to different audiences. Yet, in part due to measurement challenges, scholars lack an understanding of how members’ ideological expressions differ by medium.

This research introduces three new ideological measures that allow us to examine the differences across congressional ideal points in the aggregate and provide novel evidence of how individual legislators’ position themselves differently depending on the venue. This comparison allows us to evaluate an understudied interaction between ideology, communication style, and audience. Ultimately, we find that members express different ideological positions across these different venues, and we argue this is due to variation in audience. Text-based measures in particular allow us to capture ideological nuance that allows legislators to differentiate themselves from their peers in ways that voting does not allow. Voting is constrained by both the dichotomous nature of the “yea or nay” structure inherent to voting, as well as the strong partisan divide that exists in the modern Congress. As a result, ideal points derived from roll-call votes provide legislators with a more limited expression of their ideological positioning. Conversely, floor speeches and social media provide members an opportunity to explain their legislative decisions, and allows them to take more moderate or more polarized positions. Furthermore, across these three venues, members of Congress knowingly contend with different audience

considerations when making ideological decisions or communicating their position. We evaluate constituent effects, including the partisanship, demographics, and policy interests of their district, as well as institutional goals such as party status or personal responsibilities of their position in their chamber, whether in leadership, committee, or caucuses, on an individual legislator's ideological positioning. Lastly, we consider individual demographic attributes of the legislator, such as gender and race.

In both theory and methods, this work engages the considerable literatures on the measurement of ideology and congressional representation, paying particular attention to the demands of district and chamber responsibilities. Theoretically, this paper contributes to improving our understanding of the motivations behind legislative behavior, specifically the ways in which ideological positions differ by forum. Methodologically, we utilize a novel approach that allow researchers to scale the ideology of texts without prior model training, allowing for easy adoption to future research questions. Our models can infer ideological frames on discovered themes, addressing an ongoing challenge in computational text analysis and political science. We apply these models to develop three reliable estimates of ideology for members of the U. S. Congress: a vote-based ideal point and two text-based ideal points based on congressional speech and official Tweets. By ranking members across these three measures, we observe if, and how, members shift their ideological position across these three mediums and audiences. We apply these measures to an original dataset of members of the 115th and 116th Congress that includes demographic, institutional, and district information.

We find that different means of legislative expression do produce different ideal points, both at the aggregate and individual level. The distribution of House members' ideal points differs considerably when measured by votes, speech, and tweets. Across our dataset, floor

speeches and tweets allow legislators to differentiate themselves from their co-partisans and to take more nuanced positions. In particular, we find that ideal points expressed via social media offer legislators the opportunity to express more ideologically extreme positions as compared to vote-based ideal points. Interestingly, these effects differ by gender and race. We also find notable differences among the two major political parties that are reflective of the demographic and ideological challenges facing the two caucuses. The Republican caucus is more united in both vote and speech, compared to the Democratic caucus. Overall, we argue that the ability to adjust their relative ideological position is valuable because it allows legislators to better represent their districts. Thus, the development of multiple measures of ideology has implications for assessing ideological congruence and representation in Congress.

Literature

There is a large congressional literature on the meaning, role, and measurement of ideology (e.g., Jackson and Kingdon 1992; Krehbiel 1993; Hill 2001; Bishin 2003, Clinton et al. 2012). Legislators' ideology is generally described as their position along the liberal-conservative spectrum that reflects the policies they would implement if they could (e.g., Poole and Rosenthal 1985,1991; Clinton et al. 2004). Put differently, legislators' ideology is seen as their expressed policy preferences and measured by observable actions.

The primary audiences for this ideological positioning include legislators' constituents and congressional colleagues. Members of Congress are famously reelection-minded, and use numerous opportunities to communicate their ideological positions to their constituents (Mayhew 1974). For constituents, legislator-led communication serves as the connecting link between home and Washington, D. C. Given the wide range of policymaking activity, constituents are

reliant on members of Congress to pass along relevant information to their districts and states. Even the most attentive constituent (Arnold 1990) will find it challenging to keep up with every committee hearing, floor proceeding, or bill introduction. Members utilize constituent communication to share information about legislative activities (Hall 1987, Curry 2015), constituent services (Grimmer 2013, Parker and Goodman 2009), partisan messages (Jacobson and Carson 2019), and how and why they voted (Ansolabehere and Jones 2010). Increasingly, a large part of constituent communication focuses on members' ideological beliefs. (Cormack 2016, Russell 2017, Smith and Russell 2022). Members are eager to promote not only their legislative and partisan accomplishments, but provide ideological reasoning for these decisions, too.

Over time, the changes in communications and technology have created new opportunities for messaging. Constituents no longer rely on newspapers and nightly news for updates on their member of Congress. Anyone can go online and find the roll-call outcome of any vote, on any bill or amendment. The advent of C-SPAN and the near-constant television presence on the floor and in committees has made floor speeches accessible to constituents even if not widely watched. Constituents can, and do, directly and immediately contact their members via social media and email. Over the past decade, social media has also become a prominent form of congressional communication: In 2020, every member of Congress had a dedicated Facebook page for communication, 99 percent ran an official Twitter account, and 86 percent of Senators (79 percent of House members) had an official Instagram account (Quorum Analytics).

Members also communicate with one another. Classic studies underscore that legislators look to their colleagues for cues when issues come before Congress, and are a factor in understanding legislative decisions (e.g., Kingdon 1989; Matthews and Stimson 1975). Actions

taken inside Congress, therefore, also can be used to signal one's position to congressional peers. Votes serve as a signal to party leaders that have control over the agenda. Floor speeches and social media allow members to add more depth to their ideological expression. Research has found that some members use verbal communication, like floor speeches, as a way to signal their institutional status. Lerner and Shaffer (2020) found that the most prominent speakers in committees were also more effective lawmakers. Research on social media in Congress suggests that it too may play a role in intra-congressional relations (e.g., Barbera 2015; Barbera et al. 2019; Shapiro et al. 2012). Unlike votes or floor speeches, Twitter is an equal-opportunity communication source. All members have the ability to Tweet any time of day, regardless of position, or whether Congress is in session or not, and about (nearly) any topic.¹ Thus, communication is used by lawmakers to send a signal back home or ideological disposition to congressional peers. Ideological positioning (and party alliance) can be demonstrated on social media, just as it can be through votes and floor speeches.

Existing measurements of ideology

Political science has tackled the measurement of ideology in several ways. The most common and widely accepted application are spatial models that rely on roll call votes (ex: Poole and Rosenthal 1991). Roll call vote models have been applied to understand the ideology of U.S. Congress members, as well as state legislatures (Shor and McCarty 2011). These models establish a continuum, from liberal to conservative, that allows researchers to orient individual legislators or groups (such as a state legislature) on an ideological spectrum. While there are limitations to this approach—most notably that the data is both binary (yes or no vote) and is

¹ Official Twitter accounts are monitored by the House Committee on Ethics and House Communications Standards Commission and are subject to rules on campaign fundraising and personal information.

limited to the legislation that is brought forward for a vote, they are easily applicable and well-established. Other iterations of spatial models have provided insight beyond chamber dynamics, using external relationships such as campaign donations (Bonica 2013) or interest group ratings (Berry et al. 2010).

As the accessibility of computational text analysis has increased, spoken and written words have been used to establish ideology, sometimes in tandem with legislative votes. Text allows researchers to consider policy positions beyond floor votes, and contributes to a better understanding of ideology, particularly for policies that are not subject to roll call votes. Spatial models using texts have been applied to the U. S. Congress to provide more nuanced understanding of legislators' positioning, including expressed preferences by policy area (e.g., Diermeier et al. 2012; Gerrish and Blei 2011; Grimmer and Stewart 2013; Kim et al. 2018; Vafa et al. 2020, Ebanks et al. 2022). Language has also been used in conjunction with vote behavior to give a more holistic view of legislative dynamics. For example, Nguyen et al. (2015) develop a model of legislative votes, speech, and bill text to examine substantive policy differences between the Tea Party caucus and establishment Republicans. Relatedly, Davoodi et al. (2020) use relationships between bill text, donor contributions, and legislator attributes to understand the factors that characterize (dis)agreement within state legislatures.

Text-based analysis also has proven especially valuable in establishing ideology in institutions such as the judicial branch, in which quantifiable measures such as votes or campaign donations are limited (e.g., Bonica and Sen 2021; Hausladen et al. 2020; Lauderdale and Clark 2014), as well as electoral contexts such as debate reactions (Argyle et al. 2021). Text-analysis has also been used to assess ideological positions of executive branch officials (Bertelli and Gross 2011) and presidents (Treier 2010). The application of text-analysis ideal points is not

limited to the U.S. context as scholars have examined party platforms and legislative speeches in numerous countries (e.g., Laver et al. 2003; Benoit and Laver 2012; Lowe et al. 2011).

Ideology and representation

Scholars' enduring interest in legislators' ideological positioning in part reflects the importance of questions about congressional representation, and the promise of ideal points to measure the extent to which members of Congress accurately reflect the preferences of their constituents (Miller and Stokes 1963). Consequently, there is a sizeable literature focused on ideological congruence, or whether ideologically more liberal (conservative) districts are represented by more liberal (conservative) legislators (e.g., Ansolabehere and Jones 2010, Carson et al. 2010; Hall 2015; Miler 2018). The increase in polarization in American politics raises questions of whether members of Congress are more ideologically extreme than their voters, or whether legislators are responding to the preferences of their constituents (see Tausanovitch and Warshaw 2013; Battista et al. 2021). There is some evidence that voters punish legislators whose preferences are out of line with the preferences of the district (Ansolabehere et al. 2001; Canes-Wrone et al. 2002), but the electoral penalty for being out of step may be declining (Bonica and Cox 2018; Highton 2019), and there is evidence that legislators are more extreme than the public (Bafumi and Herron 2010).

The lack of a clear relationship may reflect the fact that studies of ideological representation generally look at vote-based ideal points, which may not be the behavior we expect to be most reflective of constituents. In part, the scarcity of constituent ideal points that are comparable to a vote based ideal point hampers the study of ideological congruence. For instance, whereas a legislator votes on a specific piece of environmental legislation that will

change the accepted levels of a specific pollutant, public opinion surveys generally ask broad questions like whether the government should prioritize doing more on the environment. Additionally, efforts to measure ideological representation will falter when considering issues where the public and legislators have preferences, but those issues do not receive a roll-call vote. Thus, the expanded breadth of measurement of legislative ideal points put forth here should also facilitate a re-examination of ideological representation by engaging the question of which ideological expression would be expected to reflect constituents. We argue that text-based ideological expressions are taken with an eye to an audience of constituents and therefore offer a valuable tool for better understanding congressional representation.

The Case for Multiple Measures of Ideology

Legislators make choices about how to present themselves and communicate their actions to wider audiences (e.g., Fenno 1978; Mayhew 1974; Grimmer 2013; Blum et al. 2022), and we expect that this behavior extends to how they express their ideological positions. We build on existing work by Vafa et al. (2020) to develop new, comparable ideal point models across three forms of behavior: votes, legislative speech, and social media expression. Our novel research design then examines the motivations behind each type of ideal point as well as the dynamics of shifts in an individual legislators' ideal points across venues. While votes provide the most limited choice and ideological expression, they present a clear signal of support for the party. Voting is primarily an inside-Congress audience, including party leaders and colleagues. And while larger or controversial legislation does attract the attention of attentive voters, these are often highly partisan constituents. Floor speeches allow for greater ideological variation, but still

face time and scheduling restrictions subject to party leaders and member travel.² The audience is limited to those in the chamber (or watching C-SPAN), unless the member proactively promotes their floor speech through another medium, or a third-party publicizes the speech more broadly. Social media has the fewest limitations in terms of content and timing, but a short word count limits the actual text a member can share, particularly on Twitter. However, Twitter provides individual members with equal footing with party leaders, allowing members to speak directly to a national audience. Some rank-and-file members have easily surpassed the attention and audience of the formal leaders of their party.

Thus, we expect that members will take audience variation into consideration, and present different ideological stances across these three mediums. Given that floor votes are determined (and inherently limited) by party leaders, we expect that vote-based ideal points will be the most polarized ideal point, shaped mainly by party and chamber dynamics. We expect floor speech-based ideal points to be shaped by a combination of both institutional and district motivations, thus introducing more variation across legislators' ideal points. For example, we consider how party and chamber responsibilities, like seniority and leadership positions, may lead some members to reflect a more unified, party-centered message. However, given the public-facing nature of floor speeches (and the tendency to use them in press releases) floor speeches also allow members to appeal to their district's nuanced ideology—whether that be more moderate or more extreme. Conversely, because tweets carry little institutional weight, we expect tweet-based ideal points to reflect a legislator's positioning for a more national constituency. Given previous scholarship on extremism on social media (Banks et al. 2020, Conover et al. 2011, Barberá et al. 2015, Ballard et al. 2022), we anticipate greater variation in

² Party leaders control the floor schedule, including when members are able to speak on the floor. Furthermore, in the House, the majority of speeches are limited to one minute.

social media posts: some legislators' tweet-based ideal points are likely to be more ideologically extreme, while other legislators may deliberately take more moderate positions.

These expectations not only shape how we expect the three ideal point measures to compare with one another, but also how we expect individual legislators to use these distinct opportunities to communicate their ideological position. We predict that a member will shift their ideological position as expressed through speech on the House floor or on Twitter (as compared to roll-call votes) to appeal to constituent and partisan audiences. However, we also expect institutional factors—such as party leadership and seniority—to impact ideological positioning as well. These party-loyal members will be less likely to vary from their vote-based position.

Data and Methods

Data

We develop three measures to compare ideology across votes, speech, and Tweets. We build off of Vafa et al.'s (2020) Text-Based Ideal Points (TBIP) approach, to create stable, easily applicable models that automatically infer ideal point estimates from text alone. By extending existing work on more traditional vote-based ideal points as well as Text-Based Ideal Points (TBIP), our approach reliably infers ideological frames, as confirmed by subject-matter experts. Crucially, our model requires only the collection of text documents and who the author is for each document—it does not need party or any other labels for either the documents or the authors. As a result, the unsupervised setup allows for easy adoption on different datasets and contexts (only a collection of texts and their authors is needed). This also helps the model to be applied to large text collections, as well as large collections of smaller texts (like Tweets).

To build these ideal points, we utilize votes, congressional records, and Tweets from the official accounts of House members from the 115th and 116th Congress (2017-2020). We create

an original dataset with these three TBIP measures for each member alongside biographical and district information for all members of the 115th and 116th Congress. Thus, unlike prior research, we include information from two Congresses, across all policy issues.

The result is a dataset with 507 members.³ For each member, we collected biographical information such as gender, race, party, the number of terms they have served in the House, whether they are in a position of party or committee leadership or on a high-ranking committee (Appropriations, Energy and Commerce, and Ways and Means), and if they are a member of a major intra-party caucus.⁴ We also collect district information, including the percentage of minorities in the district, the district unemployment rate,⁵ a categorical variable for district density that indicates the urban, suburban, or rural nature of the district,⁶ and electoral information on both the House member (percent of the vote received in the most recent general election) and the 2016 Presidential election (percent of the vote received by the Republican presidential candidate), which we consider to be a reflection of citizens' national political preferences (Tausanovitch and Warshaw 2013). This legislator-specific information serves as our independent variables in our pursuit of understanding the motivations behind ideological communication strategies.

To acquire the text data we use to apply TBIP, we use the Congressional Record parser made available by Judd (2017) to download floor speeches, extracting the speech text and

³ The majority of representatives are members of both the 115th and 116th Congress. This dataset also includes those members who departed from the 115th Congress or were new members in the 116th Congress.

⁴ This includes the Problem Solvers Caucus, Republican Study Committee (RSC), Freedom Caucus, Blue Dog Democratic Caucus, New Democratic Coalition, and Progressive Caucus. Caucus information was collected via archival research using the Congressional Yellow Books (Gaynor 2021). While most representatives' membership spans across both congressional sessions, legislators are coded as being a member of the caucus if they were a member at any time in the dataset—i.e., the 115th *or* the 116th Congress.

⁵ District demographic information is collected using the Census Bureau's American Community Survey data, 2016.

⁶ We use a simplified measure of the CityLab Congressional Density Index. Rural = pure rural, rural-suburban mix; Suburban = sparse suburban, dense suburban; Urban = urban-suburban mix, pure urban.

speaker Bioguide ID (to correlate with biographical information of lawmakers) from the json files.⁷ For Tweet data, we acquire the Twitter user IDs for House representatives⁸ and download the accounts' text using Python package "tweepy" (Roesslein 2020). We scrape roll-call vote data using the public domain data collectors for US Congress, primarily maintained by the Sunlight foundation and GovTrack.⁹

Preparing Text Data

Using the text processing settings described in Vafa et al. (2020) as our guide, we removed legislators with a low number of speeches or tweets. From the text itself, we also remove the names of cities, states, and representatives, as well as common English stopwords and corpus-specific stopwords such as procedural terms that are unlikely to be relevant or useful in distinguishing documents and speaker. To reduce noise, we omit terms spoken by only a handful of representatives. Lastly, we remove procedural speeches, which offer no ideological or policymaking value. Other major processing decisions along with the final vocabulary size, number of documents, and number of legislators are found in the Appendix.

For Tweets, we engaged in the standard pre-processing for text, retaining hashtags. The original number of downloaded non-empty tweets was about 2.67 million, which was not computationally tractable to run with the TBIP model, and therefore, we randomly sampled about 300,000 Tweets, maintaining the original distribution of number of tweets per author.

Formal Development of the Text Based Ideal Points

⁷ <https://github.com/unitedstates/congressional-record>

⁸ 115th Congress: <https://dataverse.harvard.edu/dataset.xhtml?persistentId=doi:10.7910/DVN/UIVHQR>; 116th Congress: <https://dataverse.harvard.edu/dataset.xhtml?persistentId=doi:10.7910/DVN/MBOJNS>

⁹ <https://github.com/unitedstates/congress>

In general, Bayesian ideal point models consider legislators voting yea or nay on a shared set of bills, and posit a legislator's vote can be modeled in terms of a per-legislator scalar latent variable and per-bill scalar latent variable, resulting in an ideal point of the legislator indicating their polarity on a political spectrum. The resulting spatial output should be familiar to congressional scholars: when lawmakers share the same sign, we can assume they are more likely to vote in a similar way. When lawmakers differ in their spatial direction, they are less likely to vote together (Vafa et al. 2020).

Text-based ideal points are a generative model of text, where individual word probabilities are adjusted by author- and topic- specific latent variables according to the word's polarity. However, using a corpus of textual documents is very different from using voting records (Vafa et al. 2020). Votes have a cleaner structure: they are generally two-dimensional and associated with specific bills. Language data is high dimensional, unstructured, and ambiguous. The meaning and significance of a word is dependent on its context in multiple ways and may not even be relevant to political ideology.

TBIP counteracts this challenge by running a generalized topic model step using Poisson Factorization. This gives us the initial topic intensities in documents as well as neutral topics which are then used in the ideal point estimation step. In order to obtain stable results, TBIP must be initialized with topic-word distribution β and document-topic distributions θ estimated from a separate topic model. In the standard setup, these are obtained through a Poisson factorization model. However, the outputs of models estimated with variational inference, like Poisson factorization, are less interpretable and stable than those provided by latent Dirichlet allocation (LDA) estimated with Gibbs sampling (Hoyle et al. 2021; Hoyle et al., forthcoming).

To rectify this issue, we initialize with the outputs of such an LDA model.¹⁰ This procedure makes the original TBIP procedure more flexible and adaptable by allowing the use of a different topic model of the practitioner’s choice to get initial topic estimates. An advantage of using MALLET’s LDA implementation to initialize topics is that the topics are generally relatively stable (an important consideration for content analysis and empirical research with text data).

In addition, we utilized subject-matter experts to validate the coherence and polarization of the computational models’ outputs. Two political science graduate students independently reviewed the topics produced by the text-based models, and first, analyzed whether the computational model topics, top documents, and words representative of the topics represented an easily identifiable category or meaningful concept. Students rated the coherence of a topic on a scale from 1 (not identifiable) to 3 (easily identifiable topic). If the topic was rated 2 (somewhat identifiable) or 3, students then rated the topic’s expected polarization, given the top words used as well as their knowledge of American politics.¹¹ This process allowed us to remove any non-coherent topics if necessary and validate the computational outputs of our TBIP modes. Ultimately, we found that our models performed very well: the overwhelming majority of topics were not only coherent, but performed along ideological lines as expected—the topics were not only polarized (or not) as expected, but the words identified as liberal or conservative by the models were validated by our independent coders.¹²

¹⁰ However, β and θ are sampled from a Gamma distribution in Poisson factorization and Dirichlet distributions in LDA. To transform the posterior estimates of β and θ from the LDA model, we use the identity that for samples $y_1, \dots, y_n \sim \text{Gamma}(\alpha, 1)$ and $[x_1, \dots, x_n] \sim \text{Dirichlet}(\alpha)$, then $x_i = \frac{y_i}{\sum_j y_j}$. To estimate the normalization factor ($\sum_j y_j$) for the document-topic and topic-word estimates, we run the Poisson factorization model N times and take the mean. We then scale the LDA estimates by the corresponding factor.

¹¹ Full details of the validation process can be found in the Appendix.

¹² For floor speeches, 81.82 percent of initial topics were coherent and ideologically polarized as expected; for Tweets, 72.73 percent of initial topics were coherent and ideologically polarized.

We evaluate the selection of words by legislators on given “topics”, derived from texts authored by these legislators rather than votes on proposed congressional bills. During inference, latent variables capture the topic-specific polarity of individual words.¹³ For example, for a topic concerning abortion, the words “women,” “unborn,” and “child” will likely have larger absolute values than for a topic about the environment. A conservative author may use words like “life” or “unborn” with higher frequency than those who refer to abortion as a medical procedure (see Figure 1). Conversely, a liberal author will use these words less frequently when discussing the same topic, while perhaps using terms like “women” or “reproductive health” more (see Figure 2). If we assume that estimation assigns conservatives a positive ideal point (>0), and liberals a negative one (<0), then “unborn” will have a positive sign and “women” a negative one. This neatly ties with the concept of framing which selects some particular perspectives within an issue to make more salient in discussions on that issue (Entman and Rojecki 1993, Chong and Druckman 2007; Nelson et al. 1997).

¹³ Formally expressed as: $y_{dv} \sim f(d; \exp(\log k_v + x_{ad} k_v))$. This mirrors Vafa et al. (2020), who use the Poisson distribution for f , and use the above formulation to draw a parallel with the Bayesian vote-based ideal point estimation, with the $\log k_v$ term being the intercept for the “popularity” of the word v in topic k , while k_v represents the “polarity” of the word v in topic k . d_k ensures that the topics discussed in the document (d) are contributing to the sum that estimates word counts for that document. k_v being zero would mean the word is not politicized in that particular topic. In case on a non-zero k_v , sharing the same sign as with the author's ideal point x_{ad} would mean increasing the (Poisson rate of) word count for v in the document d , and decreasing the (Poisson rate of) count if the signs are opposite.

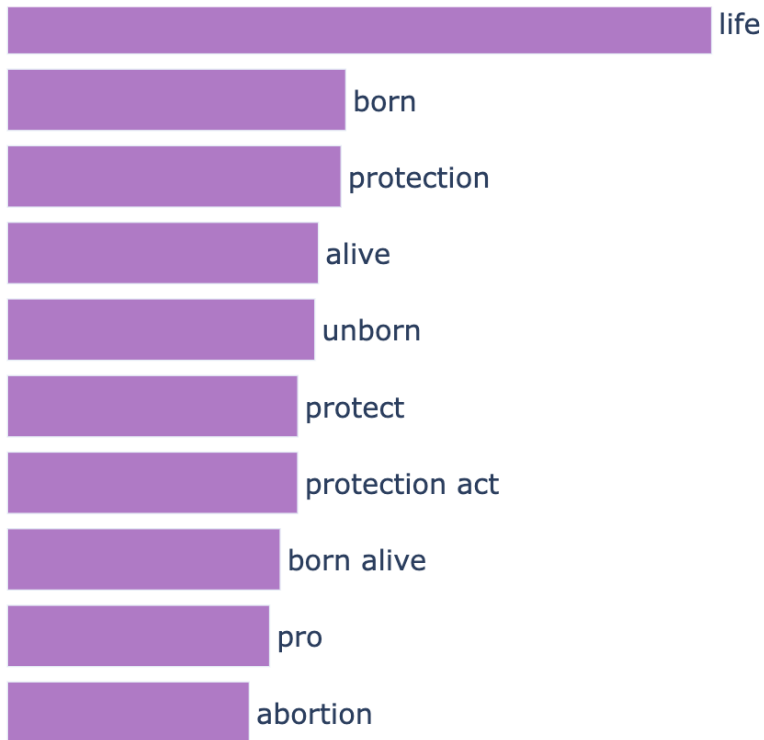


Figure 1: Illustration of “Conservative” Word Usage on Abortion

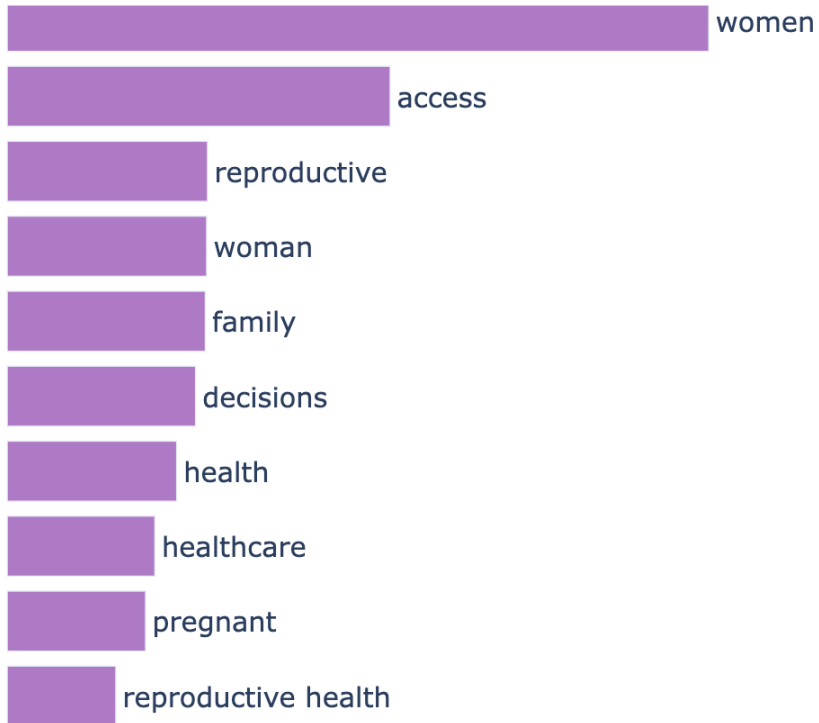


Figure 2: Illustration of “Liberal” Word Usage on Abortion

Finally, we normalize the ideal points to compare the members across congressional sessions. Although this obscures differences in the relative dispersion of the three measures, the distributions are similar, as is, importantly, the ordering of a member’s ideological positioning within the chamber.¹⁴ Given our multi-faceted approach, including the additional step of subject-matter expert evaluation discussed above, we ultimately find that the model is very stable in its ideal point estimation. Even with different, random initializations, the texts (and legislators) converge to similar values.

Ideal Point Measures

The result of this data collection and modeling are three ideal points—one for votes, floor speeches, and Tweets from members of the 115th and 116th Congresses. Consistent with convention, negative scores are more “liberal” ideal points, and positive, higher scores are more “conservative.” The three ideal point measures capture different dynamics across the parties. Figure 3 shows the distribution of member ideal points using roll-call votes. As expected, the two parties are highly polarized, and within the two parties, vote-based ideal points are the most condensed of the measures. The median for the Democratic Party is -1.109 with a range of 1.483, and the median ideal point among Republican House members is 0.962, with a range of 1.856.

Given the centralized control over what proposals come to a floor vote, as well as the binary

¹⁴ The full process is documented in Appendix A2, along with the procedure to find an optimal number of topics for Twitter tweets. We also establish the stability of the text based ideal points with multiple runs for different random seeds. Given the resulting topics of the first step, text based ideal point estimates produced across multiple runs are very stable as similar values are learned for legislators (see Appendix A4). Finally, we normalize the ideal points to compare the members across congressional sessions. Although this obscures differences in the relative dispersion of the measures—for example, it may be the case that text-based ideal points show greater variation than votes, perhaps indicating increased polarization of speech—in practice the distributions are similar.

choice they present, there are few opportunities for individual members to express nuanced ideological positions, and hence Figure 3 illustrates the concentration of ideal points by party.

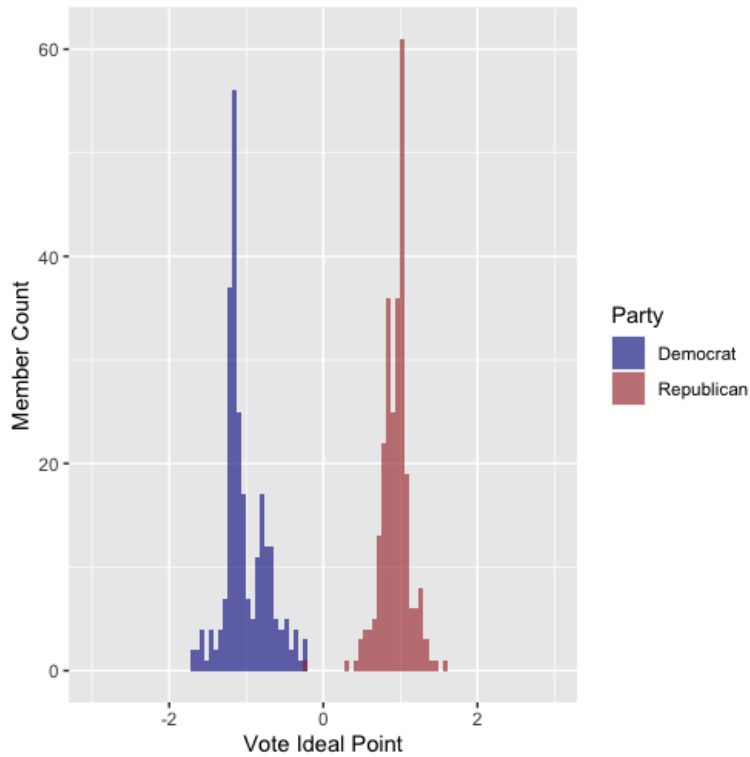


Figure 3: Member Distribution of Vote Ideal Points, 115th- 116th Congress

The distribution of speech-based ideal points is showcased in Figure 4, and reveals that although the two parties are polarized, there is notable overlap of members with more “moderate” ideal points. The median speech-based ideal point for both parties is also less extreme than when measured by votes: the median score for Democratic members is -0.867, and for GOP members is 0.794. Not only are the two parties closer together ideologically, but there is also greater variation in speech ideal points within the two parties: Democratic members are spread across a 4.143 range, and GOP members across a 5.572 range, indicating greater variation in how co-partisans position themselves ideologically in their speech as compared to votes. In

other words, measuring ideal points through speech uncovers more conservative Democrats and more liberal Republicans than are identified by vote-based ideal points.

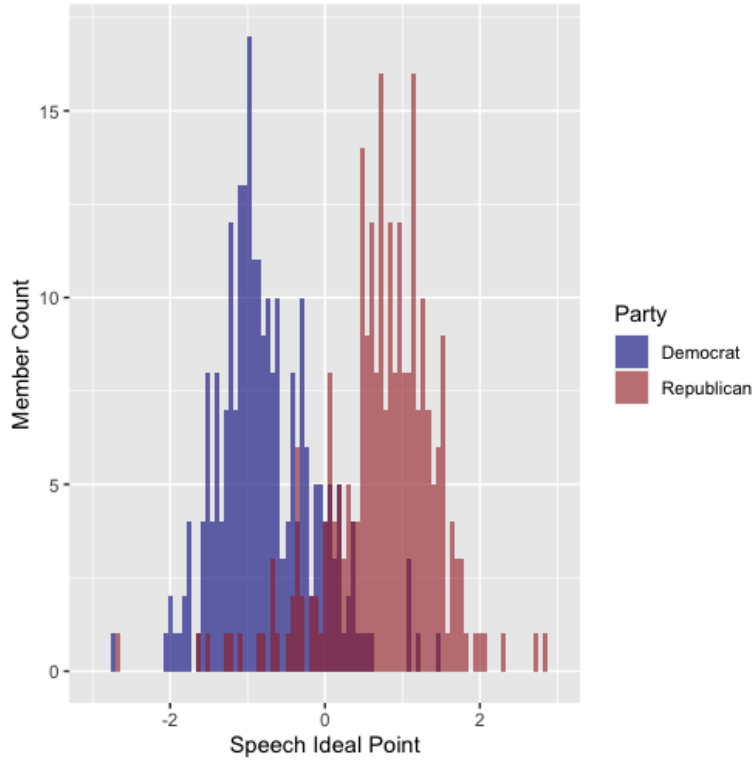


Figure 4: Member Distribution of Speech Ideal Points, 115th- 116th Congress

Lastly, Figure 5 presents the distribution of Tweet-based ideal points. Here we see two polarized parties, but with more within-party variation than was captured with vote-based ideal points, and less variation than was revealed with the speech-based ideal points. Reflective of this dynamic, the partisan medians are -0.994 for Democrats, and 0.984 for Republicans, and the ranges are 2.912 and 3.371 respectively. This is consistent with Twitter as both having the potential to be less polarized because it is outside the partisan structures of Congress, but also prone to more extreme appeals to a national polarized audience.

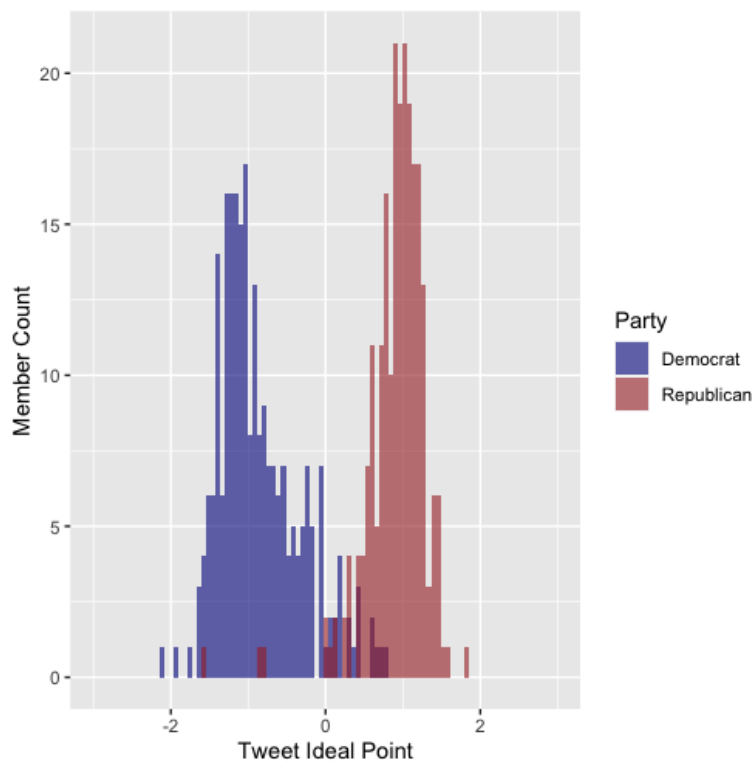


Figure 5: Member Distribution of Twitter Ideal Points, 115th- 116th Congress

Altogether, our vote-based ideal points look like what we expect of Congress – polarized parties with tight intraparty cohesion and little overlap between the parties. Floor speech-based ideal points offer more space for ideological expression and show a Congress that is polarized, but less extreme with wide distributions of preferences within parties and space for moderates in

both parties. Twitter-based ideal points present a Congress that also allows for more ideological expression, but that shows more partisan polarization than floor speech.

Given the differences revealed across the three ideal point measures, we examine how they reflect congressional and district pressures across the chamber (or party) as a whole. More specifically, we predict the impact of district, chamber, and individual considerations on each ideal point measure using ordinary least squares (OLS) regression models. To avoid concerns about comparison across ideal points (discussed above), each ideal point is run as its own model and the analyses are separated by party for ease of interpretation.¹⁵ At its most simple interpretation, a negative value means that the resulting ideal point is more liberal, holding all else constant. A positive value reflects a more conservative ideal point. Thus, for the Democratic party, while a negative value indicates that members are reflecting a more liberal ideal than other members of the party, a positive value below indicates a moderating effect on a member's ideal point—i.e., more towards the “center” of the chamber. Conversely, a negative value for Republicans indicates a moderating influence towards the “center” of the chamber, while a positive value indicates a more conservative impact.

Tables 1 and 2 present the results of the Democratic party and Republican party models, respectively. Across both parties, legislators' ideal points are largely determined by district and institutional variables and this is true whether measured by vote, speech, or tweets. Perhaps the most consistent finding is that increased district support for Donald Trump in the 2016 presidential election predicts more conservative ideological positions for both Democrats and Republicans. While there is some variation in exactly which district factors and institutional

¹⁵ Full models that consider party as an interaction variable can be found in the Appendix. The results are consistent across specification and we present the split sample here for ease of interpretation.

position predict ideal point measures, a striking difference is that legislators' gender and race are not predictors of vote-based ideal points, but do play a significant role in understanding ideological positions based on speech and tweets.¹⁶ For Democrats, gender is an important determinant of ideological positioning as women are more liberal than their male colleagues all else equal. This suggests that female Democratic legislators use floor speech and tweets to establish positions that voting alone does not accommodate. This finding is consistent with previous work on female legislators' use of speech in the US and cross-nationally (e.g., Vallejo and Gomez 2022; Pearson and Dancey 2011, Russell 2021). Table 1 also reveals that non-white Democratic legislators take more liberal positions than their white colleagues when looking at tweet-based ideal points, which indicates that minority legislators employ Twitter to carve out a more personalized ideological position whereas white legislators use the flexibility of Twitter to take a more conservative position. These relationships also hold true among Republican legislators as female and non-white Republicans have more liberal tweet-based ideal points than their co-partisans, indicating the use of social media to differentiate oneself within the party. It is worth noting, however, that the small number of legislators in these categories warrants caution in drawing conclusions about the dynamics within the Republican party.¹⁷

¹⁶ We use a binary definition of gender and an indicator variable to denote female legislators. For legislators' racial identity, we code all legislators as white or non-white, where non-white includes African American, Latino, Asian American and Pacific Islander and Native American legislators.

¹⁷ There are 24 female and 10 non-white Republican legislators in the dataset.

Table 1: Democratic Party Ideal Points Across Venues, 115th – 116th Congress

<i>Independent Variable</i>	<i>Dependent Variable</i>		
	Vote Ideal Point	Speech Ideal Point	Tweet Ideal Point
Female MC	0.0184 (0.025)	-0.168*** (0.065)	-0.084* (0.045)
Non-White MC	0.009 (0.028)	0.005 (0.069)	-0.097* (0.054)
Number of terms served	-0.014*** (0.003)	-0.002 (0.008)	0.005 (0.008)
Percent vote in last election	-0.001 (0.001)	-0.005 (0.005)	0.001 (.003)
District vote for Trump	0.007*** (0.002)	0.010** (0.005)	0.014*** (0.004)
Suburban district	0.037 (0.039)	-0.335*** (0.091)	-0.126 (0.078)
Urban district	0.052 (0.045)	-0.148 (0.113)	-0.196** (0.090)
District percent white	-0.001 (0.001)	-0.001 (0.002)	-0.0001 (0.002)
District unemployment rate	-0.053*** (0.016)	0.033 (0.031)	-0.001 (0.030)
Problem Solvers Caucus	-0.048 (0.032)	0.099 (0.092)	0.069 (0.068)
Progressive Caucus	-0.017 (0.028)	-0.226*** (0.074)	-0.184** (0.053)
New Democratic Coalition	0.039 (0.030)	0.054 (0.081)	-0.017 (0.063)
Blue Dog Coalition	0.333*** (0.046)	0.579*** (0.108)	0.504*** (0.090)
Party leader	-0.067* (0.036)	-0.233** (0.110)	-0.186** (0.085)
Committee chair	0.034 (0.050)	0.247* (0.134)	-0.030 (0.121)
Top committee member	-0.048** (0.022)	-0.200*** (0.064)	-0.101** (0.048)
Constant	-0.915*** (0.159)	-0.543 (0.442)	-1.140*** (0.324)
Observations	249	241	244
R ²	0.599	0.477	0.537

This includes members from the 115th and 116th Congress

Estimated with OLS regression

p<0.1; **p<0.05; *p<0.01*

Table 2: Republican Party Ideal Points Across Venues, 115th – 116th Congress

<i>Independent Variable</i>	<i>Dependent Variable</i>		
	Vote Ideal Point	Speech Ideal Point	Tweet Ideal Point
Female MC	-0.023 (0.027)	-0.034 (0.075)	-0.115* (0.062)
Non-White MC	-0.020 (0.037)	0.067 (0.121)	-0.337** (0.149)
Number of terms served	-0.008*** (0.003)	-0.040*** (0.008)	0.008 (0.006)
Percent vote in last election	0.004** (0.002)	-0.002 (0.004)	-0.004 (0.003)
District vote for Trump	0.005*** (0.002)	0.010** (0.004)	0.011*** (0.003)
Suburban district	0.024 (0.028)	-0.225*** (0.081)	-0.012 (0.040)
Urban district	-0.051 (0.051)	-0.493*** (0.137)	-0.848*** (0.323)
District percent white	-0.003*** (0.001)	-0.005 (0.003)	-0.0001 (0.002)
District unemployment rate	-0.042*** (0.010)	0.035 (0.031)	0.054*** (0.021)
Problem Solvers Caucus	-0.093* (0.050)	-0.053 (0.078)	0.059 (0.059)
Republican Study Committee	0.072*** (0.029)	0.027 (0.071)	0.067 (0.043)
Freedom Caucus	0.063** (0.027)	-0.244** (0.105)	-0.030 (0.054)
Party leader	0.013 (0.030)	-0.061 (0.105)	0.166*** (0.040)
Committee chair	-0.033 (0.032)	-0.326** (0.144)	0.005 (0.050)
Top committee member	-0.028* (0.017)	-0.152* (0.087)	0.120*** (0.041)
Constant	0.801*** (0.155)	1.101*** (0.386)	0.253 (0.211)
Observations	257	243	227
R ²	0.440	0.318	0.481

This includes members from the 115th and 116th Congress

Estimated with OLS regression

p<0.1; **p<0.05; *p<0.01*

More generally, the examination of tweet-based ideal points reveals the factors that push legislators towards more ideologically extreme positions on Twitter. When looking at vote and floor speech based ideal points, we find that some factors push legislators in a more extreme direction while other factors encourage moderation. But when examining tweet-based ideal

points, the factors that affect legislators' positions overwhelmingly push towards greater ideological extremism (e.g., Republicans being more conservative and Democrats being more liberal). There are just two notable exceptions to this pattern: Republicans from more urban districts take more moderate positions than their colleagues and Democrats from districts highly supportive of Donald Trump's 2016 presidential campaign also take more moderate positions. The pattern of more ideologically extreme positions is consistent with Twitter's more national audience, as well as the general tendency of social media to encourage extremism and partisanship.

Legislator-level factors on ideal points

Given chamber-wide trends across the three measures of legislative ideal points indicate ideological differentiation across our three mediums, we now turn to the question of which individual legislators shift their ideological positioning and under what conditions. We expect legislators to deliberately deviate from their vote-based positions when necessary to more accurately reflect their ideal point. Floor speeches and tweets both allow legislators to adjust their positions in order to differentiate themselves from their party and to take a more personalized position. The motivations for such fine-tuning of ideological position may be rooted in constituency representation and the electoral incentives for legislators to more accurately reflect their district. Legislators may also use speeches or tweets to more fully reflect their own preferences, or their position (or aspirations) within the institution. Our arguments and analyses allow for all of these motives to play a role.

In order to evaluate individual-level shifts in ideological positions across venues, we first address comparability across the three ideal point measures. Since the three ideal point estimates are generated using independent models, and are not directly comparable, we remedy this by

comparing legislators' relative positioning on the scale for each measure. For each ideal point measure, we rank legislators by ideal point. This produces an ordinal ranking for members in order of liberal to conservative position based on vote-based ideal points, another ranking based on speech-based ideal points, and a final ranking based on tweet-based ideal point. We use these rankings to evaluate which members are more or less conservative (or liberal) in how they talk or Tweet, compared to how they vote. For example, Rep. Kevin Brady (R-TX), Republican chair of the Ways and Means Committee is ranked as one of the more conservative voters in the chamber (413), but in his floor speeches to his peers, he is 70 points more moderate (343). Yet when he addresses his national audience on Twitter, he presents a more conservative ideological lean (ranked at 470).

In order to facilitate depicting the degree of relative movement in ideal point positions, we generate a categorical variable that captures how far an individual member shifted in their ideological presentation. Members who shifted more than 0.5 standard deviations (44 spots), are coded as 1 (or a one-unit shift in a liberal/ conservative direction). Members who shift one standard deviation (88 spots) are coded as a 2, or a two-unit shift in a liberal/ conservative direction, and so on. We use this categorical variable in Figures 4 and 5 below to illustrate how prevalent changes in relative ideological positions are at the individual legislator level.

When interpreting the below figures, a positive shift is indicative of a liberal direction (a more conservative vote-based ideal point minus a more liberal text-based ideal point), while a negative shift is indicative of a more conservative shift. Each categorical increase (decrease) is a half-point standard deviation increase (decrease), towards a liberal (conservative) direction. As Figures 6 and 7 indicate, members are most likely to stay put ideologically—their text-based

ideal points are similar to their vote-based ideal points.¹⁸ However, even with this conservative metric, there is still a great deal of variation in ideological presentation among members. The vote-to-speech ideological shift is more widespread and common than the vote-to-Tweet ideological shift. The two parties use text-based ideal points to moderate their vote-based ideal points, with Republican members more likely to move in a liberal direction and Democratic members in a conservative one. However, Republican members show greater variation in their ideological presentation than their Democratic counterparts. Republicans are more likely to use speeches and Tweets to present a different ideological position—both in a liberal and conservative direction.¹⁹

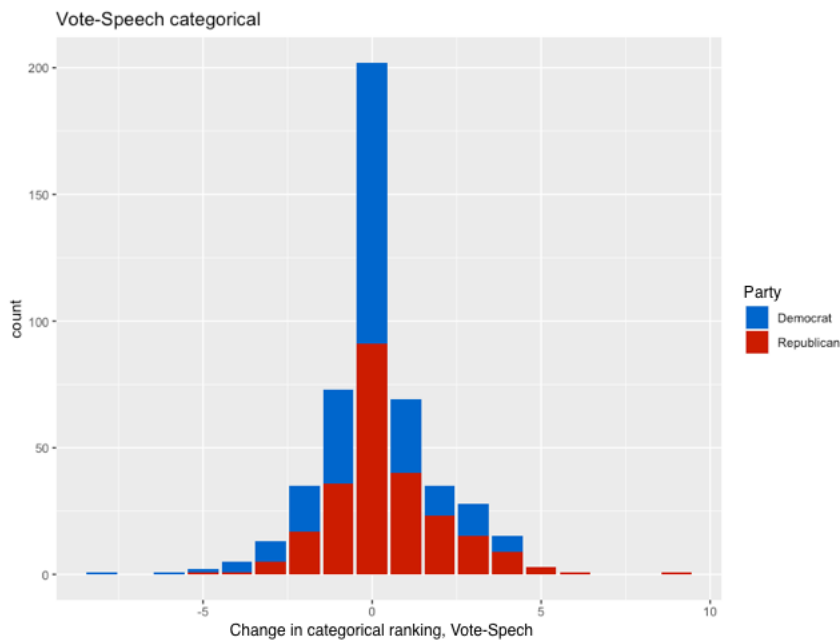


Figure 6: Categorical shift from vote-based ideal points to floor-speech-based ideal points, 115th and 116th Congress

¹⁸ We recognize that even a “zero” ranking likely captures a slight shift (<44), but we consider this relatively stable, particularly given the normal distribution of our categories.

¹⁹ See the Appendix for individual-level shifts, sans categorical deviations (Figs. A1, A2).

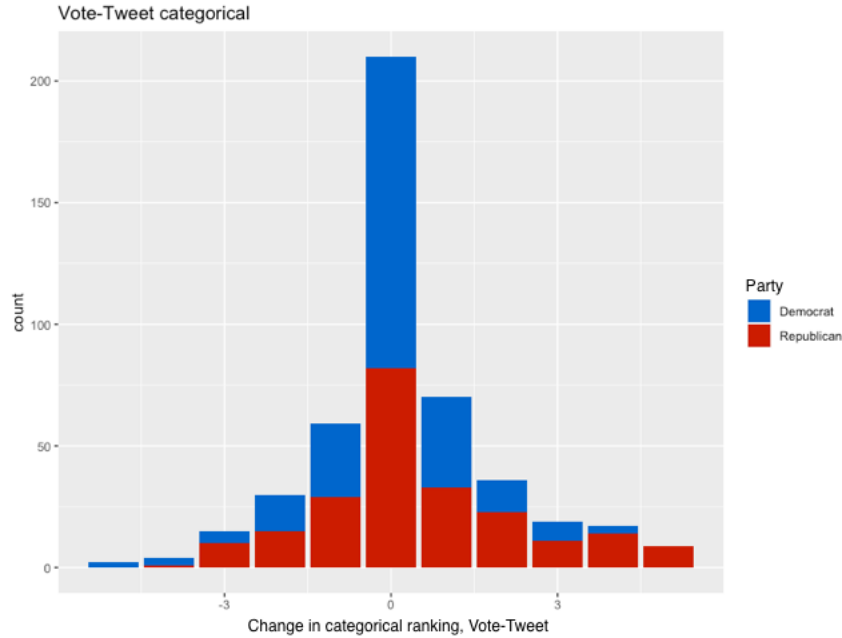


Figure 7: Categorical shift from vote-based ideal points to Tweet-based ideal points, 115th and 116th Congress

To examine why legislators shift their ideological position across votes, floor speeches, and tweets in greater detail, we estimate the change in legislators’ relative ideological rank within their party and compare vote-based ideal point rankings to each alternate measure – speech-based ideal points and tweet-based ideal points. For each model, we evaluate the impact of individual-level factors such as legislators’ identity and seniority in the chamber, district considerations such as electoral security, partisanship, and composition of the district, and institutional considerations such as leadership positions, membership on top committees, and party caucus affiliations to determine the conditions under which legislators are more liberal or more conservative when measured by speech or tweets as compared to vote behavior. The results of these analyses are presented in Table 3 and reveal how legislators use behaviors other than voting to adjust their ideological position within their party to reflect their district and institutional roles.

Table 3: Changes in Ranked Ideal Point Position, 115th – 116th Congress

	Democrats		Republicans	
	Shift Vote to Speech	Shift Vote to Tweet	Shift Vote to Speech	Shift Vote to Tweet
Female MC	18.709 (11.683)	15.573* (8.846)	-12.246 (15.854)	10.324 (20.268)
Non-White MC	6.379 (14.557)	17.029* (10.425)	-12.593 (22.322)	59.155* (33.253)
Number of terms served	-3.062** (1.530)	-3.739*** (1.365)	3.547*** (1.271)	-6.424*** (1.485)
Vote in last election	0.337 (0.535)	-0.433 (0.502)	1.273** (0.658)	1.408 (0.927)
District vote for Trump	0.452 (0.971)	-0.396 (0.761)	0.311 (0.930)	0.638 (1.099)
Suburban district	50.364*** (16.591)	25.325* (14.531)	31.920** (13.281)	6.871 (13.140)
Urban district	32.234 (21.347)	40.091** (17.169)	49.505** (23.793)	81.856 (59.217)
District percent white	0.336 (0.455)	0.095 (0.395)	-1.031* (0.553)	-2.131*** (0.581)
District unemployment rate	-12.127 (7.516)	-6.720 (6.771)	-27.265*** (4.980)	-36.407*** (5.916)
Problem Solvers Caucus	-28.163** (12.594)	-25.213*** (10.183)	-15.670 (16.037)	-32.341** (16.001)
Progressive Caucus	21.389* (12.594)	14.242 (11.231)		
New Democratic Coalition	7.381 (14.176)	13.615 (11.509)		
Blue Dog Caucus	4.293 (18.269)	7.683 (9.532)		
Republic Study Committee			24.492** (11.875)	9.782 (13.963)
Freedom Caucus			73.036*** (15.825)	29.549* (16.724)
Party leader	11.264 (18.269)	10.630 (17.525)	25.750 (19.315)	-27.521* (14.775)
Committee chair	-40.698* (24.282)	0.741 (23.602)	14.792 (21.978)	-25.665 (16.521)
Top committee member	10.535 (10.590)	5.448 (8.625)	2.787 (12.265)	-44.748*** (11.841)
<i>Constant</i>	<i>-58.703</i> <i>(73.826)</i>	<i>30.130</i> <i>(62.401)</i>	<i>33.481</i> <i>(68.359)</i>	<i>221.295</i> <i>68.070</i>
Observations	246	244	243	227
R ²	0.188	0.207	0.322	.408

*This includes members from the 115th and 116th Congress.
OLS regression. * $p < 0.1$; ** $p < 0.05$; *** $p < 0.01$.*

The analyses in Table 3 reveal several important dynamics. First, the type of district a legislator represents is important in determining how a legislator will position themselves through floor speech and tweets. Notably, we find consistent evidence that urban and suburban districts drive legislators to express relatively more liberal positions than captured by vote-based ideal points. This is consistent with cities generally leaning more liberal and Democratic, but the dynamics of the suburbs is very much a contested political landscape and both recent academic and journalistic accounts suggest movement in the liberal direction (Samuels 2022, Tam Cho et al. 2013). This liberal shift means that Democrats in urban and suburban districts shift towards the ideological extreme of their party, but when Republicans represent these districts, the shift means a more moderate position or movement towards the ideological center. It is also worth noting that the moderating effect on Republicans only exists when looking at speech-based ideal points, and the urban or suburban nature of the district does not compel a more liberal (moderate) position in their tweets.

For Republicans, the racial composition and economic conditions of their district also shapes how they position themselves when comparing votes to speech and tweets. Republican members who represent districts with a higher percentage of white residents, as well as those from districts with higher unemployment rates lead legislators to talk in ways that communicate a more conservative ideological position relative to their colleagues than their voting behavior would indicate. This shift is consistent with the rise of populist messaging on the political right and the use of racist undertones in some conservative politics (Kelly 2010, Hartman 2021). While normatively problematic, this shift is expected given the different—and increasingly partisan—audiences across the three venues. The effect of unemployment rates on the likelihood of conservative messaging also echoes a well-established relationship between Republicans and

economic anxiety, genuine or not (Pierson 2017, Hacker and Pierson 2005). It is also consistent with the institutional constraints on roll-call voting (i.e., leaders' agenda setting powers, the Hastert rule, etc.) that promote less extreme ideological expression through voting and therefore contribute to the need for legislators to stake out an ideological position that votes do not capture by shifting their position through their language.

Second, legislators' own lived experience and their experience in the chamber affects how they use language to differentiate their position from their votes. More senior legislators of both parties use floor speeches to position themselves more moderately than their voting behavior, such that Republicans move in a more liberal direction and Democrats in a more conservative direction. This moderating effect of seniority is repeated among Democrats when comparing their vote-based ideal points with their tweet-based ideal points. However, more senior Republicans actually take more conservative positions through their tweets than their votes. The finding that more senior legislators moderate their positions in their floor speech likely reflects the long-standing (but arguably declining) norms of the chamber, including civility, institutional loyalty, and coalition-building (Matthews 1959, Hanges et al. 2020).

Additionally, as suggested by the aggregate patterns, legislator's gender and race influence their positioning on Twitter as compared to their voting behavior. Among Democrats, women and non-white legislators shift their ideological position in the more liberal direction when comparing tweet-based ideal points to vote-based ideal points, consistent with work on legislative ideology and gender (Reingold 2008; Osborn et al. 2019, Thomsen and Sanders 2019). Similarly, non-white Republican legislators also shift their ideological position in a more liberal direction when comparing Twitter to roll-call votes.²⁰ These findings suggest that, at least

²⁰ Again, a caveat here is warranted given the small number (10) of non-white Republicans who served in the House during this period.

for Democratic women and minority members, party leadership does not bring to a vote that sufficiently allow expression of their ideological positions, and so they use other forums outside the constraints of Congress (i.e., Twitter) to convey these positions.

Lastly, our data show that institutional position drives some shifts in ideological positions. Notably, intraparty caucuses are important predictors of shifts in legislators' relative ideological positions, and this is more pronounced for Republicans than Democrats (Gaynor 2021). Republicans who belong to the Republican Study Committee (RSC) or the Freedom Caucus, two conservative caucuses in the chamber, moderate their positions in their speech and take relatively more liberal positions than their votes convey. This finding is somewhat unexpected given the reputation of conservative House members for engaging in ideologically extreme rhetoric. While further examination is called for, one consideration is that conservative Republican legislators stake their identity on their largely oppositional votes to the Democrats' legislative agenda, even on bills that receive bipartisan support. In this way, their vote-based ideal points may sufficiently capture their conservative ideological position allowing them to moderate slightly when speaking on the floor.

For Democrats, there is evidence that members of the Progressive Caucus shift to take a more liberal position in the floor speech than their voting records captures, but this dynamic does not carry over to their ideological expressions on Twitter. Nor do we see evidence of membership in other intra-Democratic party caucuses (e.g., New Democrats or Blue Dog Caucus) affecting whether legislators shift their ideological positions when measured through speech or tweets. In contrast, Democratic members of the explicitly bipartisan Problem Solvers Caucus take steps to moderate their vote positions in both floor speech and Twitter. This suggests that although they may vote with their party, they use other venues to differentiate

themselves and establish a more moderate position consistent with the Problem Solvers Caucus. However, membership in the Problem Solvers Caucus does not have this effect for Republicans who position themselves as more conservative on Twitter than their vote-based ideal points. While inconsistent with the goals of the Problem Solvers Caucus, the national audience of Twitter may create incentives for Republican members to establish more conservative positions for an audience outside of Congress even while espousing principles of bipartisanship within the chamber.

Taken together, these analyses illustrate how individual legislators take different ideological positions across votes, floor speeches, and tweets. This deliberate adjustment to their relative ideological position throughout the chamber reflects legislators' incentives to adapt to these distinct venues that have their own rules, audiences, and legislative implications. The factors that explain why some legislators are more or less liberal in floor speeches or tweets as compared to their votes highlight the importance of constituency considerations and the importance of institutional position. As expected, the more restricted nature of voting – both in the control over what comes to a vote as well as the limited choice of voting for or against – means that sometimes it cannot capture all the nuance of legislators' ideological expression. Legislators, therefore, use other opportunities such as floor speeches or communicating on Twitter to express their ideological positions, often deviating from their vote-based positions.

Discussion and Conclusion

To many observers of Congress, the conclusion that legislators behave differently across different settings is not surprising nor controversial. Indeed, what these data reveal is conventional wisdom on Capitol Hill. However, the academic study of legislative ideal points has sometimes assumed a more fixed view of ideology where vote-based ideal points capture the

singular underlying position of a legislator. We agree that vote-based ideal points are an important and accurate reflection of legislators' ideological position, but we contend that they are one of multiple possible measures of ideological position. Looking at a legislator's ideal points across venues and understanding the shifts in positioning provides a more complete picture of how legislators express their preferences to colleagues and constituents. The addition of text-based ideal points also expands scholars' ability to evaluate ideological representation, as well as examine ideological positions of those who do not have voting records such as first term members of Congress.

Of even broader interest, we believe, is the applicability of these new ideal points. The ideal points developed here rely on more data points (over 435 House members versus 100 Senators) and over a longer time span (two congressional sessions versus one) than prior work by Vafa et al. (2020). Although this results in less-structured models than prior iterations, we find cohesive and replicable topics across speakers. Notably, even for short, social-media posts, we derive reliable estimates of ideology. In addition to measuring modeling stability computationally, we also assess the reliability of the text-based ideal point method via original, extensive human annotation and validation efforts. Ultimately, our models automatically infer high-quality ideal point estimates by authors. The automated discovery of topics and frames across ideology is a notable and open challenge in computational text analysis and computer science research more generally. Our research shows how improvements in the computational domain can shed new light on theoretical questions. Of course, like all work that uses text as data, questions about the authenticity of an ideological expression come to mind. But regardless of this caveat, we do believe that these new ideal point models present significant potential for the study of ideology in Congress (and beyond). This unsupervised setup allows for easy

adoption on different datasets and contexts (only a collection of texts and their authors is needed).

Ultimately, these novel measures indicate that members of Congress shift their ideological position, depending on the venue used to express their ideology, and these shifts are reflective of both district pressures and institutional responsibilities. While legislators do not all move in the same direction, nor do they shift to the same degree, there are underlying dynamics to these individual-level adjustments in ideological positions. First, the different rules and constraints of each venue shape the expression of ideological position. Second, we find evidence that legislators use these three venues to reach different audiences and express different ideological positions better tailored for their constituencies.

When looking at the individual-level data, captured by how much a legislator's ideological position across venues changes when compared to their peers, we find evidence that lawmakers are catering to their district, and in some cases, the institution. At the district-level, we find that lawmakers representing urban and suburban areas express more liberal positions in speech and Tweet. Thus, for Republicans representing these areas, speeches and Tweets have a moderating effect on their voting records. Also, for Republicans, the demographic and economic makeup of their district impacts their ideological expressions: Republican members that represent majority white and areas of high relative unemployment are more conservative in how they speak to their constituents—a likely reflection of the populist direction of the Republican Party.

We also find evidence that institutional expectations impact ideological positioning. Senior lawmakers moderate their (likely party-line) votes when speaking to their peers, but still use Twitter to express more conservative positions. This is similar to Republican members of

congressional caucuses, including the bipartisan Problem Solvers Caucus—while floor speeches were more moderate than votes, Twitter still provides Republican lawmakers a chance to be more conservative than their voting record allowed. To us, this indicates the staying power of norms within the chamber, juxtaposed against the increasing importance for Republicans to cater to the national base.

Lastly, we found that Democratic women and non-white members, and non-white Republican members expressed more liberal positions on Twitter than their vote. Like other findings reiterated here, this is indicative of the limitations of roll-call votes to express an ideological position. Personal attributes, district pressures, and institutional positions lead lawmakers to use floor speeches and social media to differentiate themselves ideologically. When only given the choice of ‘yes’ or ‘no’, text-based expressions allow for greater nuance that lawmakers gladly take.

Notably, Republican members, across our models, were more likely to use these outlets to express conservative positions. Future work should build on this, considering how these findings interact with existing work on asymmetrical polarization between the two parties. As social media in particular becomes more impactful and important in the political dialogue of elites and constituents alike, understanding how lawmakers are using these outlets to encourage ideological extremism could potentially become more important than votes.

Scholars have long accepted that Congress is becoming more polarized, but the roll-call mechanism by which these assumptions are made is inherently limited. We find that by using textual data, we can more accurately capture the ideological positioning of a member of Congress. When comparing the ideal points of votes, speeches, and Tweets, this produces notable differentiation that speaks to the influence of district and institutional pressures, as well

as demographic attributes. These models, and the ideal-point outputs they produce can easily be applied to other questions of ideology, even by policy area. The potential for application throughout the field is significant, and our hope is these models offer researchers an easily applicable alternative to measuring ideology.

Citations

- Ansolabehere, S., & Jones, P. E. (2010). Constituents' responses to congressional roll-call voting. *American Journal of Political Science*, 54(3), 583-597.
- Ansolabehere, S., Snyder Jr, J. M., & Stewart III, C. (2001). The effects of party and preferences on congressional roll-call voting. *Legislative Studies Quarterly*, 533-572.
- Argyle, D., Argyle, L. P., Eidelman, V., & Resnik, P. (2021). Debate Reaction Ideal Points: Political Ideology Measurement Using Real-Time Reaction Data. *Statistics, Politics and Policy*, 12(1), 5-28.
- Arnold, D. R. (1990). *The Logic of Congressional Action*. Yale University Press.
- Bafumi, J., & Herron, M. C. (2010). Leapfrog representation and extremism: A study of American voters and their members in Congress. *American Political Science Review*, 104(3), 519-542.
- Ballard, A. O., DeTamble, R., Dorsey, S., Heseltine, M., & Johnson, M. (2022). Dynamics of polarizing rhetoric in congressional tweets. *Legislative Studies Quarterly*.
- Banks, A., Calvo, E., Karol, D., & Telhami, S. (2021). # polarizedfeeds: Three experiments on polarization, framing, and social media. *The International Journal of Press/Politics*, 26(3), 609-634.
- Barberá, P. (2015). "Birds of the same feather tweet together: Bayesian ideal point estimation using Twitter data." *Political Analysis* 23(1): 76-91.
- Barberá, P., Casas, A., Nagler, J., Egan, P.J., Bonneau, R., Jost, J.T., & Tucker, J.A. (2019). "Who leads? Who follows? Measuring issue attention and agenda setting by legislators and the mass public using social media data." *American Political Science Review* 113(4): 883-901.
- Battista, J. C., Peress, M., & Richman, J. (2021). Estimating the locations of voters, politicians, policy outcomes, and status quos on a common scale. *Political Science Research and Methods*, 1-17.
- Benoit, K., & Laver, M. (2012). The dimensionality of political space: Epistemological and methodological considerations. *European Union Politics*, 13(2), 194-218.
- Berry, W. D., Fording, R. C., Ringquist, E. J., Hanson, R. L., & Klarner, C. E. (2010). Measuring citizen and government ideology in the US states: A re-appraisal. *State Politics & Policy Quarterly*, 10(2), 117-135.
- Bertelli, A. M., & Gross, C. R. (2011). The lengthened shadow of another institution? Ideal point estimates for the Executive Branch and Congress. *American Journal of Political Science*, 55(4), 767-781.
- Bishin, B. G. (2003). Independently validating ideology measures: A look at NOMINATE and adjusted ADA scores. *American Politics Research*, 31(4), 404-425.

- Blum, R., Cormack, L., & Shoub, K. (2022). Conditional Congressional communication: how elite speech varies across medium. *Political Science Research and Methods*, 1-8.
- Bonica, A. (2013). Ideology and interests in the political marketplace. *American Journal of Political Science*, 57(2), 294-311.
- Bonica, A. & Cox, G.W. (2018). "Ideological Extremists in the U.S. Congress: Out of Step but Still in Office", *Quarterly Journal of Political Science*: Vol. 13: No. 2, pp 207-236.
<http://dx.doi.org/10.1561/100.00016073>
- Bonica, A., & Sen, M. (2021). Estimating Judicial Ideology. *Journal of Economic Perspectives*, 35(1), 97-118.
- Canes-Wrone, B., Brady, D. W., & Cogan, J. F. (2002). Out of step, out of office: Electoral accountability and House members' voting. *American Political Science Review*, 96(1), 127-140.
- Carson, J. L., Koger, G., Lebo, M. J., & Young, E. (2010). The electoral costs of party loyalty in Congress. *American Journal of Political Science*, 54(3), 598-616.
- Chong, D., & Druckman, J. N. (2007). A theory of framing and opinion formation in competitive elite environments. *Journal of communication*, 57(1), 99-118.
- Clinton, J. D., Bertelli, A., Grose, C. R., Lewis, D. E., & Nixon, D. C. (2012). Separated powers in the United States: The ideology of agencies, presidents, and congress. *American Journal of Political Science*, 56(2), 341-354.
- Clinton, J., Jackman, S., & Rivers, D. (2004). The statistical analysis of roll call data. *American Political Science Review*, 98(2), 355-370.
- Congressional Management Foundation, 2015. "How The Internet Has Changed Citizen Engagement".
<https://www.congressfoundation.org/projects/communicating-with-congress/how-the-internet-has-changed-citizen-engagement>
- Congressional Management Foundation, 2019. "The Future of Citizen Engagement: Coronavirus, Congress, and Constituent Communications."
https://www.congressfoundation.org/storage/documents/CMF_Pubs/cmf_citizenengagement_covid-19.pdf
- Conover, M., Ratkiewicz, J., Francisco, M., Gonçalves, B., Menczer, F., & Flammini, A. (2011). Political polarization on twitter. In *Proceedings of the international aaai conference on web and social media* (Vol. 5, No. 1, pp. 89-96).
- Cormack, L. (2016). Extremity in congress: communications versus votes. *Legislative Studies Quarterly*, 41(3), 575-603.

- Curry, J. M. (2015). *Legislating in the Dark*. University of Chicago Press.
- Davoodi, M., Waltenburg, E., & Goldwasser, D. (2020, July). Understanding the language of political agreement and disagreement in legislative texts. In *Proceedings of the 58th Annual Meeting of the Association for Computational Linguistics* (pp. 5358-5368).
- Diermeier, D., Godbout, J. F., Yu, B., & Kaufmann, S. (2012). Language and ideology in Congress. *British Journal of Political Science*, 42(1), 31-55.
- Ebanks, D., Yan, H., Alvarez, R. M., Das, S., & Sinclair, B. (2021). Leadership Communication and Power: Measuring Leadership in the US House of Representatives from Social Media Data.
- Entman, R. M., & Rojecki, A. (1993). Freezing out the public: Elite and media framing of the US anti-nuclear movement.
- Fenno, R. F. (1978). *Home style: House Members in their districts*. Little, Brown.
- Gaynor, S. W. (2021). The (Financial) Ties That Bind: Social Networks of Intraparty Caucuses. *Legislative Studies Quarterly*.
- Gerrish, S. M., & Blei, D. M. (2011, October). Predicting legislative roll calls from text. In *Proceedings of the 28th International Conference on Machine Learning, ICML 2011*.
- Grimmer, J. (2013). *Representational style in Congress: What legislators say and why it matters*. Cambridge University Press.
- Grimmer, J., & Stewart, B. M. (2013). Text as data: The promise and pitfalls of automatic content analysis methods for political texts. *Political analysis*, 21(3), 267-297.
- Hacker, J. S., & Pierson, P. (2005). Abandoning the middle: The Bush tax cuts and the limits of democratic control. *Perspectives on Politics*, 3(1), 33-53.
- Hall, R. L. (1987). Participation and purpose in committee decision making. *American Political Science Review*, 81(1), 105-127.
- Hanges, P., Lee, F., Miler, K., & Wessel, J. (2020). "Report on the Organizational Climates of Congress." University of Maryland.
- Hausladen, C. I., Schubert, M. H., & Ash, E. (2020). Text classification of ideological direction in judicial opinions. *International Review of Law and Economics*, 62, 105903.
- Highton, B. (2019). Issue accountability in US House elections. *Political Behavior*, 41(2), 349-367.
- Hill, K. Q. (2001). Multiple-method measurement of legislators' ideologies. *Legislative Studies Quarterly*, 263-274.

- Hoyle, A., Goel, P., Hian-Cheong, A., Peskov, D., Boyd-Graber, J., & Resnik, P. (2021). Is automated topic model evaluation broken? the incoherence of coherence. *Advances in Neural Information Processing Systems*, 34, 2018-2033.
- Jackson, J. E., & Kingdon, J. W. (1992). Ideology, interest group scores, and legislative votes. *American Journal of Political Science*, 805-823.
- Jacobson, G. C., & Carson, J. L. (2019). *The politics of congressional elections*. Rowman & Littlefield.
- Kingdon, J. W. (1989). *Congressmen's voting decisions*. University of Michigan Press.
- Krehbiel, K. (1993). Where's the Party?. *British Journal of Political Science*, 23(2), 235-266.
- Lauderdale, B. E., & Clark, T. S. (2014). Scaling politically meaningful dimensions using texts and votes. *American Journal of Political Science*, 58(3), 754-771.
- Laver, M., Benoit, K., & Garry, J. (2003). Extracting policy positions from political texts using words as data. *American political science review*, 97(2), 311-331.
- Lee, F. E. (2009). *Beyond ideology: Politics, principles, and partisanship in the US Senate*. University of Chicago Press.
- Lerner, J. Y., & Shaffer, R. (2020). Setting the Committee Agenda: Measuring Speaker Influence in Congressional Hearings.
- Lowe, W., Benoit, K., Mikhaylov, S., & Laver, M. (2011). Scaling policy preferences from coded political texts. *Legislative studies quarterly*, 36(1), 123-155.
- Matthews, D. R. (1959). The folkways of the United States Senate: Conformity to group norms and legislative effectiveness. *American Political Science Review*, 53(4), 1064-1089.
- Matthews, D. R., & Stimson, J. A. (1975). *Yeas and nays: Normal decision-making in the US House of Representatives*. Wiley-Interscience.
- Mayhew, D. R. (1974). *Congress: The electoral connection*. Yale university press.
- Miler, K. C. (2018). *Poor representation: Congress and the politics of poverty in the United States*. Cambridge University Press.
- Miller, W. E., & Stokes, D. E. (1963). Constituency influence in Congress. *American political science review*, 57(1), 45-56.
- Nelson, T. E., Clawson, R. A., & Oxley, Z. M. (1997). Media framing of a civil liberties conflict and its effect on tolerance. *American Political Science Review*, 91(3), 567-583.

- Osborn, T., Kreitzer, R.J., Schilling, E.U., & Clark, J.H. (2019). Ideology and Polarization Among Women State Legislators. *Legislative Studies Quarterly*, 44(4): 647-680.
- Parker, D. C. W., & Goodman, C. (2009). Making a Good Impression: Resource Allocation, Home Styles, and Washington Work. *Legislative Studies Quarterly*, 34 (4), 493–524.
- Pearson, K., & Dancey, L. (2011). Speaking for the underrepresented in the House of Representatives: Voicing women's interests in a partisan era. *Politics & Gender*, 7(4), 493-519.
- Pierson, P. (2017). American hybrid: Donald Trump and the strange merger of populism and plutocracy. *The British journal of sociology*, 68, S105-S119.
- Poole, K. T., & Rosenthal, H. (1991). Patterns of congressional voting. *American journal of political science*, 228-278.
- Reingold, B. (2008). Women as Office Holders: Linking Descriptive and Substantive Representation.” In *Political Women and American Democracy: Critical Perspectives on Women and Politics Research*, ed. Wolbrecht, C., Beckwith, K., & Baldez, L. New York: Cambridge University Press. 128-147.
- Roesslein, J. (2020). Tweepy: Twitter for python. URL: <https://github.com/tweepy/tweepy>, 957, 84.
- Russell, A. (2017). US senators on Twitter: Asymmetric party rhetoric in 140 characters. *American Politics Research*, 46(4), 695-723.
- Russell, A. (2021, December). Gendered Priorities? Policy Communication in the US Senate. In *Congress & the Presidency* (Vol. 48, No. 3, pp. 319-342). Routledge.
- Samuels, A. (2022). Why Republicans in Blue Cities are Increasingly Outliers. *FiveThirtyEight*. April 11, 2022. <https://fivethirtyeight.com/features/why-republicans-in-blue-cities-are-increasingly-outliers/>
- Shapiro, M. A., Hemphill, L., & Otterbacher, J. (2012). "Doing what I say: connecting congressional social media behavior and congressional voting." Paper prepared for presentation at the Annual Meeting of the Midwest Political Science Association. Chicago, IL, April 2012.
- Shor, B., & McCarty, N. (2011). The ideological mapping of American legislatures. *American Political Science Review*, 105(3), 530-551.
- Smith, S. A., & Russell, A. (2022). Different Chambers, Divergent Rhetoric: Institutional Differences and Policy Representation on Social Media. *American Politics Research*, 1532673X221113017.
- Tam Cho, W. K., Gimpel, J. G., & Hui, I. S. (2013). Voter migration and the geographic sorting of the American electorate. *Annals of the Association of American Geographers*, 103(4), 856-870.

- Tausanovitch, C., & Warshaw, C. (2013). Measuring constituent policy preferences in congress, state legislatures, and cities. *The Journal of Politics*, 75(2), 330-342.
- Thomsen, D. M., & Sanders, B. K. (2020). Gender differences in legislator responsiveness. *Perspectives on Politics*, 18(4), 1017-1030.
- Treier, S. (2010). Where does the president stand? Measuring presidential ideology. *Political Analysis*, 18(1), 124-136.
- Vafa, K., Naidu, S., & Blei, D. M. (2020). Text-based ideal points. *arXiv preprint arXiv:2005.04232*.
- Vallejo Vera, S., & Gómez Vidal, A. (2022). The politics of interruptions: Gendered disruptions of legislative speeches. *The Journal of Politics*, 84(3), 000-000.