Moderates on Capitol Hill:

Volume, Tenor, and Topics of House Floor Speeches, 1985–2016

Shu Fu*

February 2, 2024

Abstract

How do moderates communicate on the floor of Capitol Hill? Congressional floor speeches are overlooked in the study of American institutions. This paper examines the Congressional Record from 1985 to 2016, quantifying the volume, partisan tenor, and topic prevalence of House members' floor speeches, to unveil the differences in speech patterns between moderates and extremists that have evolved over decades. I use a dictionary method—"Fightin' Words" Model—to create a comprehensive measure of the partisan tenor, and a supervised learning method—Keyword Assisted Topic Model—to classify the topics in all floor speeches. Descriptively, moderates consistently talk less in total volume, speak moderately in partisan tones and focus on selective issues. And, the voices of moderation grow quieter over time. Importantly, I demonstrate that moderates are quieter on the floor not because they have electoral concerns, but because they are the pivotal players who want to bargain for more. This paper deepens our understanding of the political communication of legislators and has wide implications on the importance of moderates in legislative deliberation.

*__Shu Fu__ is an Assistant Professor in School of International and Public Affairs at Shanghai Jiao Tong University, fushu@sjtu.edu.cn.

I thank Ruth Bloch Rubin, Josh Byun, Anthony Fowler, Bobby Gulotty, Will Howell, Minju Kim, Sihang Luo, Molly Offer-Westort, Eric Oliver, Jon Rogowski, Ilana Ventura and Adam Zelizer for valuable suggestions. I also thank seminar participants at the University of Chicago, University of California, Los Angeles, Georgetown University in Qatar, University of Kentucky, and Shanghai Jiao Tong University, and participants at Midwest Political Science Association Conference for helpful feedback.

Introduction

Recent literature on American political institutions has focused on the voting decisions of ideological moderates in Congress compared with those of ideological extremists (Canes-Wrone, Brady, and Cogan 2002; Poole and Rosenthal 1997; Theriault 2008). But there is more to congressional behavior than how moderate legislators vote on bills. Moderate lawmakers also invest time and energy to communicate and debate on the floor before they cast their votes.

Congressional scholars have broadly investigated legislative communication in different venues, such as press releases (Grimmer 2013; Grimmer, Messing, and Westwood 2014), news media (Wagner and Gruszczynski 2018), social media (Barberá et al. 2021; Fu and Howell 2020; Jones et al. 2018; Russell 2017), congressional committee hearings (Ban, Park, and You 2022), and back-channel communication (Ritchie 2018). However, how moderates deliver floor speeches compared to extremists has received relatively less attention. Existing studies on floor speeches have focused on one-minute speeches and special-order addresses (Harris 2005; Maltzman and Sigelman 1996; Morris 2001; Wilkerson and Casas 2017), expressions of bipartisanship (Westwood 2022), partisanship trends (Gentzkow, Shapiro, and Taddy 2019), and interbranch messaging (Nobel 2023.; Fu n.d.). Yet, scholarship lacks an understanding of to what extent moderates and extremists are different in terms of their speech patterns in legislative deliberation.

Thus, I ask the following question in this work: To what extent do moderate legislators exhibit behaviors that differ from those of extremists in terms of their involvement in legislative deliberation on the congressional floor? Theory of legislative speechmaking is ambiguous. Researchers note competing expectations about the speech patterns of moderate lawmakers. On the one hand, moderate legislators, who are more likely to represent marginal districts and are often cross-pressured, might talk *less* often and be selective in their communications. In other words, moderates would rather keep silent to dodge the consequences of supporting legislation that may upset their constituencies. On the other hand, studies on

elections show that, all else being equal, moderates perform well electorally (Canes-Wrone, Brady, and Cogan 2002; Fowler et al. 2023; Hall 2015; Hall and Thompson 2018), therefore, moderates should advertise their moderation and we should hear from moderates *more* often. It is not obvious, *ex ante*, what speech patterns of moderate legislators we would theoretically expect to see.

In this paper, I systematically unveil the speech patterns of House members on the floor over three decades and discuss mechanisms that undergird these patterns. I use text data from the US Congressional Record from 1985 to 2016 (Gentzkow, Shapiro, and Taddy 2018) to quantify the volume, partisan tenor, and topic prevalence of each legislator's floor speeches, and characterize the differences between moderates and extremists that have evolved over time. I calculate the total speech volume by the logged word count, measure partisan tenor by a dictionary method with Bayesian shrinkage and regularization, and use a Keyword Assisted Topic Model to classify the speeches into policy issue domains.

Descriptively, moderates consistently deliver fewer floor speeches compared to extremists, and the voices of moderation grow quieter over time. Furthermore, moderates speak with muted partisan tones and concentrate more on bipartisan or day-to-day issues, compared with their extremist counterparts. I also explore why moderates stay characteristically silent. I find empirical evidence that moderates are quieter in legislative discussions not because they have electorally concerns but because they are pivotal and arguably want to bargain for more.

The paper proceeds as follows. The first section characterizes competing expectations about the House speeches of moderate legislators. The subsequent section introduces the data and methods. In the empirical sections that follow, I first explore whether legislative moderates talk less in total volume on the floor; then, I offer a measure of partisan tenor and provide descriptive results on the partisanship of legislative moderates' speeches; next, I measure topic prevalence across domain and discover what issues moderates discuss more often in their speeches. After a rich description of the speech patterns of moderates on

the floor, I provide empirical evidence testing the mechanisms for the so-called "silent moderates," which suggests that moderates stay quiet not because they are weak in electoral competition but because they are strong in bargaining. The final section discusses possible interpretations of my findings and concludes.

Background

How often and how much do moderates talk on the congressional floor? Scholarly literature provides different—and often contrasting—answers. On the one hand, moderates are pivotal players in legislation (Krehbiel 1998) and their views on legislation should be valuable to their colleagues. Therefore, we should hear from them *more* often. Democracy thrives through debate. Legislative debate is often considered as a vehicle for sharing and eliciting information (Austen-Smith 1990); it "enables members to gain a better understanding of complex issues, and it may influence the collective decisions of the House" (Oleszek 1984, 130). Given their placement on the ideological spectrum, legislative moderates are potentially pivotal in legislation (Downs 1957). Consequently, moderates—who are strategically valuable or legislatively important members of the chamber—might recognize opportunities to make their policy views known to their colleagues so their opinions can be reflected in the substantive content of legislation.

Floor speeches can also be a particularly important way for moderates to demonstrate to their constituencies the effectiveness and importance of achieving policy goals. Legislators want their ideologies to be reflected in legislation (Diermeier et al. 2011) but they also want to show their constituencies that they are policy-makers. Demonstrating their critical role in policy-making could help them gain favor among the electorate, similar to the way preferred committee assignments are rewarded by constituencies (Grimmer and Powell 2013). Also, scholars have shown that everything being equal, moderates have an edge over extremists in general elections (Hall 2015; Hall and Thompson 2018), so moderates should advertise their

moderation to their voters.

Compared to other platforms for congressional communications (e.g., press releases, news media, and social media), the congressional floor provides a particular venue for legislators to publicly express their views on legislation with their colleagues. Although since the introduction of the C-SPAN network that records and broadcasts every speech on the floor, floor speeches are now exercised more for legislative deliberation among lawmakers than direct communication with constituencies. Therefore, moderates on the floor would contribute more during a debate in which they are more likely to affect the content of the legislation.

On the other hand, moderates are more likely to represent marginal districts, so they could risk upsetting their constituents or revealing that they are out of step with them on some issues. Thus, moderates might rather talk *less* often to dodge the consequences of supporting legislation that may disturb their uneasy constituencies. We know legislators use floor speeches as a means to express policy positions (Zaller and Chiu 1996). However, rather than pushing their own agendas, moderates are often running scared of the volatility of their electoral coalitions (Chaturvedi 2021). Therefore, we may also expect that moderates may participate less often in legislative debates on the floor.

Several empirical studies on other platforms of congressional communication undergird this expectation. Grimmer (2013) argues that moderates or marginals tend to discuss pork more than policy; therefore, moderates would talk less on the congressional floor, which is center stage for policy-making. Ritchie (2017) claims that cross-pressured senators choose less visible, back-channel means for pursuing policy goals to avoid conflicts between party and constituency interests. Similarly, Chaturvedi (2021) finds evidence in amendment activities that moderates remain quiet to avoid traceability, regardless of how they feel about a bill. Fu and Howell (2020) investigate how congressional candidates targeted President Trump in their social media posts and find that Republican moderates who publicly opposed Trump were punished in short-term fundraising.

In sum, these internal and external incentives, which bear on how legislators might approach the strategy of speech-making, derive competing expectations of how their speech patterns could look like. In this paper, I attempt to comprehensively investigate the speech patterns of moderates on the congressional floor over time. With the full text of the Congressional Record, I am able to assess how actively moderates participated on the floor compared to extremists. Moreover, I will compare the difference in partisan tenor between moderates and extremists, and the topic preferences of moderates and extremists. Lastly, I will empirically assess mechanisms undergirding these speech patterns.

Data on House Floor Speech

Congressional floor speech occupies the center stage in legislative communication. Representatives use floor speeches to express their perspectives on policy (Zaller and Chiu 1996), take positions on key issues (Mayhew 1974), and advance their career profiles (Bernhard and Sulkin 2018). Given the high visibility of floor speeches (Proksch and Slapin 2012), especially after the introduction of the C-SPAN network (Garay 1984), legislators also use their floor speeches as a means to connect with voters (Hill and Hurley 2002), either through news media coverage (Maltzman and Sigelman 1996) or self-promoted clips on their websites and social media platforms (Grimmer 2013). Representatives strategically spend effort preparing and executing floor speeches, expressing rhetoric designed to make a legislative impact, actively gain media coverage, and ultimately improve their chances of reelection. Moreover, unlike other communication platforms (i.e., press releases and social media) that might favor either the electoral or bargaining theses as mechanisms, the Congressional Record provides a somewhat more neutral territory for both (or one) mechanism(s) to be operative.[1] For these reasons, congressional floor speeches provide a useful opportunity to investigate

[1]For instance, press releases and social media posts would seem more about communicating with voters as opposed to other politicians, so it wouldn't give the bargaining story a fair shake. Likewise, committee hearings or behind-the-door communication might be perceived as too far into the bargaining camp to really give the electioneering mechanism a shot.

the general communication patterns of moderates and extremists over time.

My primary data source is the Congressional Record in the House of Representatives from the 99th Congress to the 114th Congress. I obtain the raw text from the *U.S. Congressional Record*, compiled by Gentzkow, Shapiro, and Taddy (2018). The data include the text of every speech given by every legislator while Congress is in session. I exclude Extensions of Remarks, which are printed, unspoken additions that are not germane to the day's proceedings. I dropped all non-substantive speeches—those with fewer than 60 words.[2] In total, there are 533,582 speeches in this analysis. It is worth noting that the time range analyzed in this paper is from 1985 to 2016. These 32 years include three instances of party shifts in the White House and cover four eight-year presidential administrations. So from the presidential agenda-setting perspective, it is quite balanced in terms of which party occupies the White House.

In the scope of this paper, I primarily investigate floor speeches in the House of Representatives. I focus on House members for the following reasons. First, compared to Senators, House members have a shorter term of two years, so they basically start running for reelection the minute they are declared winners. This means that, at some level, concern for reelection interests is constant (Mayhew 1974). And, since the elimination of "the disappearing quorum" in the House in the late 19th century[3], the House has a simple majority rule where the median legislator is the pivotal player (Krehbiel 1998). Thus, making a distinction between pivotal moderates and non-pivotal moderates is relatively straightforward. For the scope of this paper, I do not analyze speeches on the Senate floor, however, I believe the speech patterns of moderates and the underneath mechanism that I discover in the House can carry over to the Senate. Further discussion on this point can be found in the conclusion of this paper.

[2]I found that speeches with fewer than 60 words were generally procedural, such as members yielding time, appreciating the speaker, recording a vote, etc. This approach is also used by Fu and Howell (n.d.) and Nobel (n.d) in their work that uses the Congressional Record dataset.

[3]Before 1889, the House retained a procedural loophole that functioned much like the contemporary filibuster. For a more detailed discussion, please see, Krehbiel 2017 and Fu and Howell 2023

Research Design

This paper offers a rich description of House floor speeches through three attributes of interest—total volume, partisan tenor, and topic prevalence—and these are the dependent variables in my empirical investigation. By volume, I mean whether moderates talk less or more on the floor.[4] Measuring the total volume—or quantity—of congressional speech, I use logged word counts per year to gauge the total amount of speeches delivered by each legislator each year. Due to the high skewness, I take the word counts in log transformation. Next, I investigate whether moderates talk in a more or less partisan tenor. Here I borrow the term "tenor," which means high or low pitch in music, to define the high or low partisan tone in floor speeches. I formulate a comprehensive measure of partisan tenor by a dictionary method—"Fightin' Words" model (Monroe, Colaresi, and Quinn 2009). After generating a universal partisan tenor dictionary, I calculate the partisan tenor of each floor speech by a weighted average, then aggregate to the legislator-year level. Details will be introduced in the later section. Finally, I study whether moderates concentrate on specific issues over others in floor discussions. I classify topics in each speech by a supervised learning method—Keyword Assisted Topic Model (Eshima, Imai, Sasaki 2020). The outcomes are aggregated to the legislator-year level that shows the prevalence of each topic discussed by each legislator per year.

This study focuses on the communication patterns of moderates and extremists. Moderates and extremists refer to the legislators in terms of their ideology. This paper does not aim to provide alternative approaches for measuring legislators' ideology. Instead, I use a pre-existing ideological measure to shed light on communication patterns between moderates and extremists. Specifically, I rely on Nokken-Poole scores, 1st dimension (Boche et al. 2018; Lewis et al. 2022). Compared to the NOMINATE scores, 1st dimension, which

[4]The word choice of "volume" here refers to the total volume of speeches or the amount of speeches delivered on the floor. It does not refer to the power of the sound or the degree of loudness. Scholars have started to investigate the possibility of audio as data. In terms of the vocal pitch and the emotional intensity of congressional speech, please see Dietrich, Hayes, and O'Brien (2019).

is the most common measure of legislative behavior, Nokken-Poole estimates allow legislators' ideal points to move freely over time, and thus make less restrictive assumptions about legislator ideological fixedness (Nokken and Poole 2014). Because Nokken-Poole estimates vary Congress by Congress, this measure offers an opportunity to compare speech patterns within a legislator. It's worth noting that for most legislators, their ideology doesn't change drastically, so NOMINATE estimates and Nokken-Poole estimates are similar. In Online Appendix Figure A.1, I provide robustness checks that measure ideology by the NOMINATE scores.

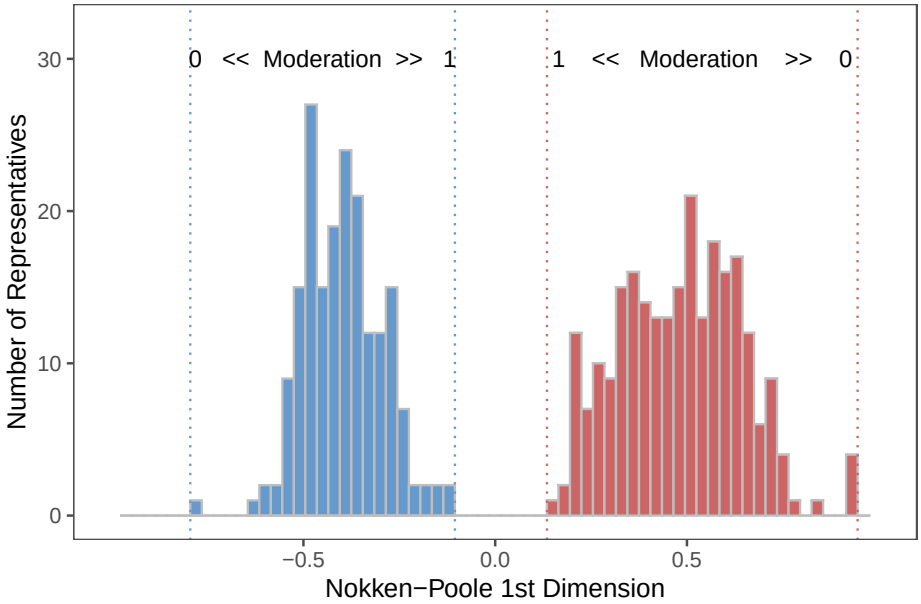
There is no conventional definition of moderates. Being "moderate" or "extreme" is a relative comparison in terms of a legislator's ideology. Thus, I create such a measure of how moderate a legislator is, "Moderation," by rescaling the Nokken-Poole scores, 1st dimension, by party and by Congress. The rescaled Nokken-Poole scores range from 0 (extreme) to 1 (moderate), with a higher value indicating the moderateness of legislators. This is the independent variable for the main analyses. This measure gauges how moderate each legislator is compared to her co-partisan members in the same session of Congress.[5] For instance, Figure 1 demonstrates the distribution of Nokken-Poole scores in the 114th Congress. Among Democrats, for example, the most liberal member, Raúl Grijalva (D-AZ-3) has a value of zero in moderation, while the most conservative, Collin Peterson (D-MN-7) has a value of one in moderation.

To assess the relationship between legislators' moderateness and their speech patterns, I have a set of incrementally demanding specifications. Each specification adds new elements. The first model is a cross-sectional comparison. Formally,

$$\text{Volume/Tenor/Topic}_{i,t,p,d} = \beta \text{Moderation}_{i,t,p,d} + \delta_{t,p} + \epsilon_{i,t,p,d},$$

[5]Admittedly, moderates today are different from moderates in the late 1980s. This measure of moderation is based on the ideology of a legislator relative to the other co-partisan MCs serving in the same Congress. In robustness checks, I have alternative measures of who is moderate in absolute terms. Please see Online Appendix Table A.1 and Table A.2. First, I choose $[-0.2, 0.2]$ as a threshold that delimits moderates from non-moderates and applies to each Congress. Second, I look at the full distribution of legislator ideologies across the 1985–2016 period and see who is within the interquartile range for that period.

Figure 1: Moderation of House Representative in the 114th Congress



where i refers to the legislator, t is year, p is party, d is congressional district. This model simply regresses the volume, tenor, or prevalence of a specific topic that each legislator has on how moderate a member is. The inclusion of year-by-party fixed effects, $\delta_{t,p}$, controls the secular time and restricts the comparison within a party. Hence, it is a cross-sectional comparison that compares legislators' speech patterns across the chamber in a given year.

The second specification adds covariates to control for the potential confounders at the individual level. These covariates, $X_{i,t}$, consist of pre-treatment socio-economic factors, such as age, gender, and race. In these first two specifications, I calculate the robust standard errors clustered by party-year. Formally,

$$\text{Volume/Tenor/Topic}_{i,t,p,d} = \beta \text{Moderation}_{i,t,p,d} + \delta_{t,p} + \Phi X_{i,t} + \epsilon_{i,t,p,d}.$$

The third model adds congressional district fixed effects, γ_d. This within-district estimate shows, conditional on the same constituencies, whether different speech patterns emerge from members with different ideologies, even representing the same voters. The

robust standard errors are clustered by congressional district. Formally,

$$\text{Volume/Tenor/Topic}_{i,t,p,d} = \beta \text{Moderation}_{i,t,p,d} + \delta_{t,p} + \Phi X_{i,t} + \gamma_d + \epsilon_{i,t,p,d}.$$

The last specification is the most demanding one, which includes legislator fixed effects, α_i. The individual-level covariates are subsumed in the legislator dummies, α_i. Since the Nokken-Poole estimates of the same legislator have variations across time, this can reveal how moderation within a legislator is associated with her speech patterns. In other words, this estimate is about when a legislator becomes relatively moderate in her roll-call voting, which is used to measure her ideology, do we see her speech patterns change? The robust standard errors are clustered by legislator. Formally,

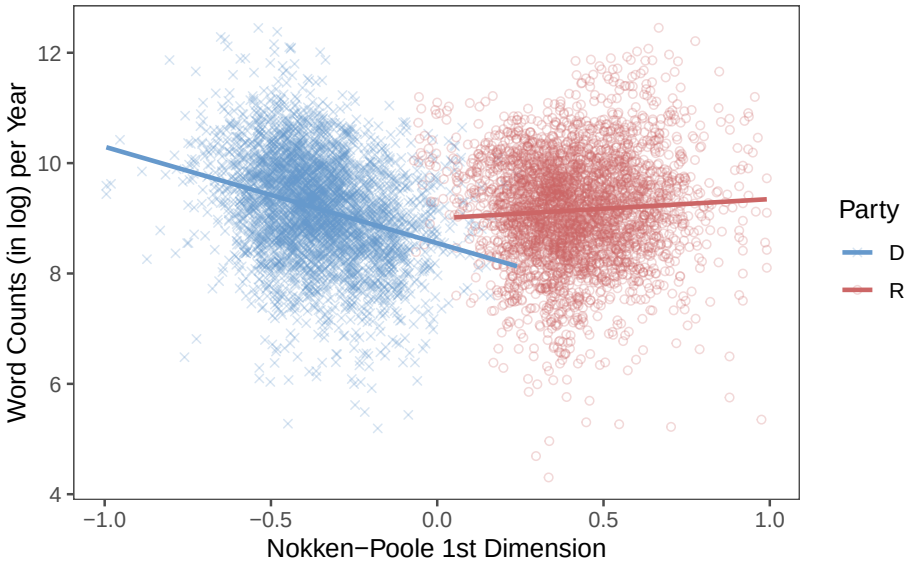
$$\text{Volume/Tenor/Topic}_{i,t,p,d} = \beta \text{Moderation}_{i,t,p,d} + \delta_{t,p} + \alpha_i + \epsilon_{i,t,p,d}.$$

Volume on the Floor

Do moderates talk less or more on the floor? I measure floor speech volume by the logged word counts per year. Word counts of legislators are aggregated to the year level. Since the word counts are highly skewed, I take the log transformation of the word counts. Figure 2 shows the relationship between yearly floor speech volume and the ideology of House members. The x-axis is the 1st dimension of Nokken-Poole scores, with liberal on the left and conservative on the right, and y-axis is logged word count per year. Each point refers to a House member in a given year, colored by the party. I use OLS regression fit lines to delineate the relationship between speech volume and the ideology of two-party lawmakers.

At first glance, moderates—those in the middle of the ideological spectrum—speak fewer words compared to extremists. The steeper slope for Democrats indicates that the more moderate they are on ideology, the quieter they are on the floor. Among Democrats, we observe a clear negative relationship between moderation and the volume of floor speech.

Figure 2: Yearly Speech Volume and Ideology of House Members (1985–2016)



Notes: Each point represents a House member. Blue crosses and red circles indicate the partisanship of legislators. Smooth fit lines are drawn by OLS regressions by party.

However, quieter moderate Republicans are less obvious. A number of moderate Republicans deliver relatively more speeches, which flattens the Republican slope. Of these active moderate Republicans—who locate around 0 on the ideology score—almost all appear in the early periods of the data. To name a few, Stewart McKinney (R-CT4) passed away in 1987, Claudine Schneider (R-RI2), Silvio Ottavio Conte (R-MA1) left office in 1991, and Frank Horton (R-NY29) retired from Congress in 1993.

To examine in depth how actively moderates participate in legislative discussion on the floor, the regressions deliver more accurate estimates. Table 1 shows the results of these regressions. I find consistent results that moderate House members each year talk in statistically fewer speeches. Column (1) estimates, among co-partisans in a given year, to what extent being an ideologically moderate legislator is associated with her total volume delivered on the floor. Column (2) includes additional socio-economic covariates. These two cross-sectional comparisons reveal that the most moderate congressperson should deliver

about 68% less content on the floor, compared to the most extreme congressperson.[6] For example, the most extreme legislators, who have values of zero in my moderation measure, speak around 24,000 words per year, which is equivalent to 3 hours of speeches. Therefore, the most moderate members would be predicted to say only about 1 hour on the floor per year. Column (3) reports the within-district estimates. Representing the same congressional district, the moderate representative delivers a statistically lower volume of speeches on the floor than her extreme predecessors or successors. The magnitude of the association is similar to the cross-sectional models.

Table 1: Estimating Floor Speech Volume of Moderates

	Dependent Variable: Logged Word Count			
	Cross-Sectional		Within-District	Within-Legislator
	(1)	(2)	(3)	(4)
Moderation	-1.144^{***}	-1.166^{***}	-1.023^{***}	-0.347^{**}
	(0.095)	(0.097)	(0.132)	(0.174)
Year×Party FE	✓	✓	✓	✓
Covariates		✓	✓	
District FE			✓	
Legislator FE				✓
$e^{\beta} - 1$	−68.1%	−68.8%	−64.1%	−29.3%
Observations	13,982	13,982	13,982	13,982
R^2	0.115	0.131	0.350	0.619

Notes:
1. *p<0.1; **p<0.05; ***p<0.01.
2. Most extreme legislators, on average, deliver around 24,000 words per year (≈ 3 hours)

In Column (4), I compare within-legislator whether being relatively moderate to the self across the tenure served in the House is correlated to the annual amount of floor speeches

[6] I use the logarithm of word counts because word counts are highly skewed and because this allows the coefficients to be approximately interpreted as proportionate changes. Since the coefficients here are too large, the approximate interpretation of beta coefficients as proportionate change deviates too far from the actual value. From the definition of the natural log, the predicted proportionate change associated with moderation is $e^{\beta} - 1$, which is printed at the bottom of the table.

delivered. Specifically, I utilize the variation in the Nokken-Poole scores within each legislator to identify the relationship between moderation in roll-call voting and floor speech volume. Evidence shows that when a member becomes more moderate in her voting records, her total volume of floor speech becomes statistically smaller at the same time. Note that given the measure of moderation that is rescaled within party and within Congress, the variation of moderation within-legislator model might be potentially driven by some combination of individual member's ideological journey across time and the composition of party caucus changing across time. However, using original Nokken-Poole scores without rescaling and simply flipping the signs of Nokken-Poole scores (see Online Appendix Table A.3), I recover similar results in the within-legislator model. I also rescale NOMINATE scores by party and by Congress, since NOMINATE scores are static for each legislator, the rescaled NOMINATE scores capture only the variation caused by the composition of the party caucus changing over time. Taking the rescaled NOMINATE scores in the same within-legislator model (see Online Appendix Table A.3), I find null results. This indicates that the variation in the member's ideology explains more of the variation of her speech volume than the party composition change does.

Voices of Moderation Grow Quieter Over Time

The result of quieter moderates is also robust to other considerations. Particularly, I investigate variation over the years. As shown in Online Figure A.2, when moderates occupy a shared ideological space with members from the opposite party, it is not obvious that moderates talk less; however, when moderate Democrats and moderate Republicans are no longer occupying a shared ideological space, the moderates speak less.

As mentioned previously, moderates did not always stay quiet, especially during the late 1980s and the early 1990s. Online Figure A.3 demonstrates the floor speech volume of moderates over time. Using the cross-sectional model with controls to run separate regressions for each congressional session over time, Online Figure A.3 (top) visualizes the

coefficients for each session of Congress, I find the magnitude of effects of moderates on floor speech volume became increasingly larger as time goes by, except the last three sessions of Congress during the Obama administration. Further breaking this down by party, shown in Online Figure A.3 (bottom), Democratic moderates consistently deliver less amount of floor speeches compared to their extreme co-partisans; while Republican moderates did not mute their volume during the Reagan, Bush 41, and Clinton administrations, but they grew quieter in the 21st century.

A potential implication of this finding is that it shifts the dynamics of lawmaking to a world in which a Democratic leader could try to buy off the most conservative member of her party or the most liberal members of the opposition party. Therefore, moderates from both parties speak more often to voice their opinions. In the 1980s and early 1990s, moderates in the minority party participated in the legislative discussion as actively as their extreme colleagues.

Partisan Tenor on the Floor

Measure of Partisan Tenor

In this section, moving beyond the total volume, do moderates substantively talk in a less partisan way? I further investigate the partisanship of legislators' speeches. This is not the first study on the partisanship of floor speeches. Westwood (2022) measures bipartisanship by culling floor speeches that contain the word, "bipartisan," and Gentzkow, Shapiro, and Taddy (2019) quantify the magnitude of partisan differences in speech and characterize partisan trends over time. Here, I create a comprehensive measure of the partisanship of floor speeches, which I name "Partisan Tenor." I use a dictionary method with Bayesian shrinkage and regularization—"Fightin' Words" model (Monroe, Colaresi, and Quinn 2009)—to calculate the partisan tenor of each individual floor speech.

The analytic elements are bigrams or two-word phrases (hereafter, simply "phrases"),

instead of unigrams or one-word (hereafter, "word"), because phrases can capture the nuances of opposing parties talking about the same issue. For example, when Democrats and Republicans talk about the issue of macroeconomics, the word "tax" is used by both parties but lawmakers from opposing parties would use different phrases. When talking about tax reduction, Democrats would usually criticize it as a "tax break" for the wealthy, while Republicans would emphasize it as "tax relief" that will stimulate economic growth.

Intuitively, phrases used more often by Democratic (Republican) congresspeople yield higher Democratic (Republican) tenor. Legislators have, in theory, complete control of what phrases they use in their speeches. Therefore, I model the choice of phrases as a function of the political party. The strategy is to model phrase choice in the full collection of speeches and then assess how the usage of a particular phrase differs by a speaker's political party.

To calculate the count of phrases used in each speech, I follow pre-processing steps performed by Gentzkow, Shapiro, and Taddy 2019.[7] It is worth noting here that floor speech is full of Congress-specific procedural phrases, such as "consent request,", "order recognition," etc, therefore, I have dropped all these phrases in calculating the partisan tenor dictionary. Therefore, these procedural phrases will have no partisan weights and will not influence the calculation of the partisanship of floor speech. Then, I use the total counts of phrases by both parties from the 99th to 114th Congress to calculate a universal weight for each phrase that allows me to compare across time.

I use j to denote a phrase, specify $\boldsymbol{\pi}_{\text{Dem}}$ as the prior probability that a specific phrase is used by a Democratic legislator, and $\boldsymbol{x}_{\text{Dem}}$ as the number of times Democrats use this phrase during the years in this study (1985–2016). N_{Dem} corresponds to the total number of phrases used by Democrats. Given the prior probability of each phrase being used by a

[7]These pre-processing steps include: (i) delete hyphens and apostrophes; (ii) replace all other punctuation with spaces; (iii) remove non-spoken parenthetical insertions; (iv) drop a list of stop words; (v) reduce words to their stems according to the Porter2 stemming algorithm; (vi) drop Congress-specific procedural phrases (e.g., "consent requir," "order recognit,", "extend debat," etc.); and (vii) restrict attention to phrases spoken at least 10 times in at least one session, spoken in at least 10 unique speaker-sessions, and spoken at least 100 times across all sessions.

Democratic member of Congress has a multinomial distribution. Formally:

$$\begin{aligned} \boldsymbol{\pi}_{\text{Dem}} &\sim \text{Dirichlet}(\boldsymbol{\alpha}) \\ \boldsymbol{x}_{\text{Dem}}|\boldsymbol{\pi}_{\text{Dem}} &\sim \text{Multinomial}(N_{\text{Dem}}, \boldsymbol{\pi}_{\text{Dem}}), \end{aligned}$$

where $\boldsymbol{\alpha}$ is a vector describing the variation in phrase frequency for each phrase. The full Bayesian estimate using the Dirichlet prior[8] yields the function:

$$\boldsymbol{\pi}^{*}_{\text{Dem},j} = \frac{x_{\text{Dem},j} + \alpha_j}{N_{\text{Dem}} + \sum_{j=1}^{J} \alpha_j}.$$

The above notations apply to the Republican party by simply replacing Dem with Rep. Next, I use the log odds ratio of phrase j weighted by its variance to measure how the rate at which phrase j is used varies by political party. The log odds ratio and its variance[9] are:

$$\begin{aligned} \text{Log Odds Ratio}_j &= \log\left(\frac{\pi_{\text{Dem},j}}{1-\pi_{\text{Dem},j}}\right) - \log\left(\frac{\pi_{\text{Rep},j}}{1-\pi_{\text{Rep},j}}\right) \\ \text{Var}(\text{Log Odds Ratio}_j) &\approx \frac{1}{x_{jD}+\alpha_j} + \frac{1}{x_{jR}+\alpha_j}\ , \end{aligned}$$

where α_j is regularized penalty, which here is set $\alpha_j = 1$. Now I can evaluate not only the point estimates of the phrase partisan weight by also the certainty about those estimates. Specifically, I use the standardized log odds ratio to measure the partisan weight of a phrase.

$$\text{Std. Log Odds Ratio}_j = \frac{\text{Log Odds Ratio}_j}{\sqrt{\text{Var}(\text{Log Odds Ratio}_j)}}$$

Based on how often Democratic and Republican lawmakers use a particular phrase, the "Fightin' Words" algorithm assigns a particular level of partisan weight to each phrase.

[8]See Monroe, Colaresi, and Quinn 2009 for a detailed derivation.
[9]See Monroe, Colaresi, and Quinn 2009 for a detailed derivation.

Figure 3: Partisan Weights of Phrases (99th–114th Congresses)

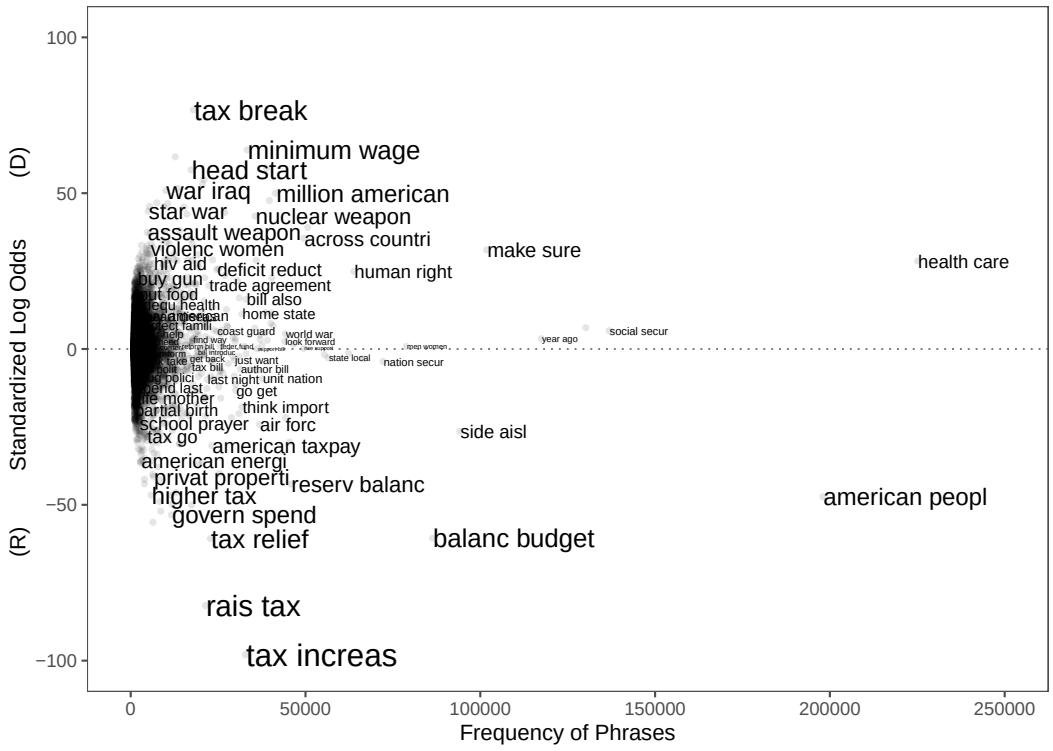


Figure 3 demonstrates the universal partisan weights of phrases from the 99th Congress to the 114th Congress, along with how frequently each phrase was used. Phrases that appear above the horizontal line of zero carry Democratic tones and phrases below represent Republican tones. For example, "minimum wage" and "head start" are measured as more Democratic in partisan weights than "balanc budget" and "govern spend." And this measure also picks up nuanced differences when legislators from the two major parties talk about the same issues. For instance, tax-related phrases (i.e., "tax increas," "rais tax," "tax relief," etc.) are used more often by Republicans, but Democrats use the phrase, "tax break" (for educational use), to make their appeals on tax issues.

Using the partisan weights assigned to each phrase, I can assign a partisanship score to each floor speech. For each speech i, I calculate how often each phrase appeared along with the total number of phrases. The partisanship of each phrase is denoted $\boldsymbol{\theta} = (\theta_1, \theta_2, \ldots, \theta_J)$, where $\theta_j \in \Re$. The count of each phrase by each speech is denoted $\boldsymbol{X}_{ij} = (X_{i1}, X_{i2}, \ldots, X_{iJ})$.

The partisan tenor of each floor speech i given by a House member is calculated using the following formula:

$$\text{Partisan Tenor}_i = \frac{\sum_{j=1}^{J} \theta_j X_{ij}}{\sum_{j=1}^{J} X_{ij}}.$$

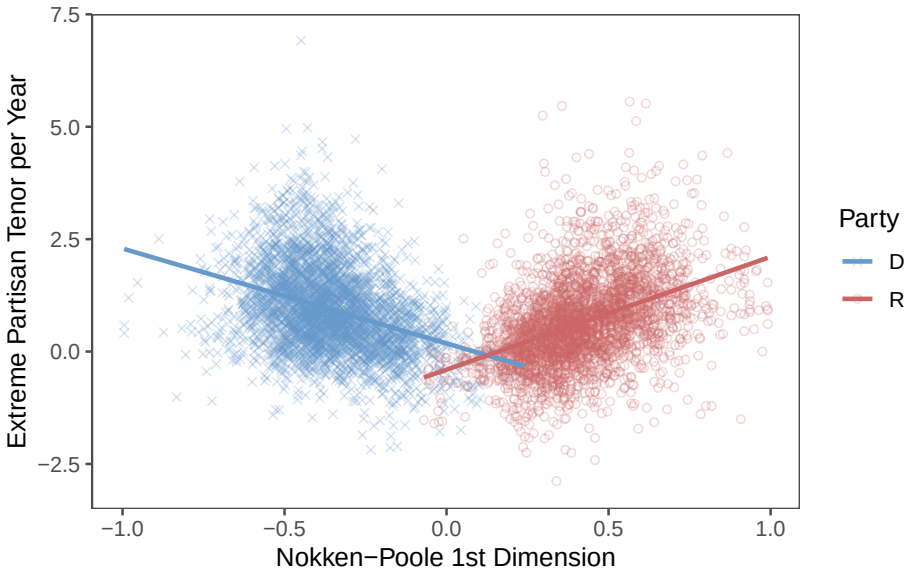
Thus, the partisan tenor is a continuous variable. A speech having Partisan Tenor $_i > 0$ reflects a Democratic tone; conversely, a speech having Partisan Tenor$_i < 0$ reflects a Republican tone. In the analysis below, I aggregate the partisan tenor to the legislator-by-year level.

Moderates in Lower Partisan Tenor

Figure 4 illustrates the relationship between partisan tenor and the ideology of House members. Since the measure of the partisan tenor is a continuum from an extreme Republican to an extreme Democratic tone, I recreate a new variable, "Extreme Partisan Tenor," by flipping the sign of partisan tenor for all Republican legislators. Thereby, the extreme partisan tenor measures the partisan tones within each party, where a higher value indicates a higher partisan tone. Evidence in Figure 4 shows a clear pattern that moderates talk in a lower partisan tenor compared to their co-partisan extremists. And this pattern is almost the same for both Democrats and Republicans. For the moderates whose Nokken-Poole scores are close to zero, their values of partisan tenor are around zero as well, meaning that moderates talk in a non-partisan way on the floor.

Table 2 presents the results from the regression analysis that more accurately estimates the extent to which moderates speak in a less partisan way. The model specifications are the same as in the previous section. I find consistent results that moderates, on average, talk in a lower partisan tenor compared to their co-partisan extremists. Cross-sectionally, holding the socio-economic covariates constant, moderate legislators speak in statistically lower partisan tenor than extremists. In specific, the partisan tenor of a most moderate member is around two standard-deviation lower than a most extreme person. I find similar

Figure 4: Yearly Partisan Tenor and Ideology of House Members (1985–2016)



Notes: Each point represents a House member. Blue crosses and red circles indicate the partisanship of legislators. Smooth fit lines are drawn by OLS regressions.

results in within-district estimates. Representing the same congressional district, a moderate talks in a statistically lower partisan tenor than her extreme predecessors or successors. Even within-legislator, evidence shows that when the ideology of a member becomes relatively moderate, the partisan toner in her floor speeches decreases simultaneously.

This results is robust to an assessment over the years. Online Figure A.4 delineates the partisan tenor trend over time. The figure highlights the average partisan tenor of extreme Democrats, extreme Republicans, moderate Democrats, and moderate Republicans within each Congress. It is obvious that moderates consistently speak in a lower partisan tenor than their co-partisan extremists. It is interesting that the sharply increased partisanship of floor speeches in recent years (Gentzkow, Shapiro, and Taddy 2019) is mainly driven by Democrats. Democratic House members express higher and higher average partisan tones over the years. However, Republicans' partisan tenor generally stays constant, except for Clinton's and Obama's initial two years when they become more partisan. In general, moderates of both parties talk in a less partisan way than their co-partisan extremists and

Table 2: Estimating Partisan Tenor of Moderates

	Dependent Variable: Extreme Partisan Tenor			
	Cross-Sectional		Within-District	Within-Legislator
	(1)	(2)	(3)	(4)
Moderation	-2.099^{***}	-2.088^{***}	-2.121^{***}	-0.961^{***}
	(0.077)	(0.080)	(0.115)	(0.175)
Year×Party FE	✓	✓	✓	✓
Covariates		✓	✓	
District FE			✓	
Legislator FE				✓
Observations	13,982	13,982	13,982	13,982
R^2	0.252	0.272	0.392	0.548

Notes:
1. *p<0.1; **p<0.05; ***p<0.01.
2. DV has a mean of 0.797 and a standard deviation of 1.12.

this is consistent over time.

Topics on the Floor

Besides the participation and partisanship, do legislative moderates and extremists cover different policy issues in floor speeches? To classify all the floor speeches delivered by House members into policy domains, I use Keyword Assisted Topic Models (keyATM) (Eshima, Imai, Sasaki 2020). Compared to unsupervised topic models, such as Latent Dirichlet Allocation (Blei, Ng, and Jordan 2003) and Structural Topic Models (Roberts, Stewart, and Tingley 2017), which have unavoidably high researcher degrees of freedom in labeling the topics and choosing the topic number, keyATM is a supervised topic model that pre-sets the topic numbers and clear topic issues. Furthermore, the keywords associated with each topic supervise machine learning to achieve more accurate classification.

Congressional floor speech mainly discusses legislation that covers a wide range of policy issues. Here, I use State of the Union Addresses to derive keywords for topics. The

Policy Agendas Project labeled each sentence in the State of the Union addresses as one of 20 major topics.[10] Performing the same pre-processing steps on the text of the State of the Union Addresses (1985–2016), I pick the most frequent 15 words of each topic as keywords. Online Table A.4 presents all keywords that are associated with the 20 major topics.[11]

To prepare the text of the floor speech, I use preprocessing steps that are similar to the dictionary method described in the previous section. As a major distinction, however, here I use words (unigrams) rather than phrases (bigrams) because words are sufficient to attain the main goal of classifying topics. To prepare the keywords, I also remove capitalization, reduce keywords to their stems, and store them as a list object. The input for the model is a document-term matrix (DTM), where each row refers to each floor speech, and columns include the most frequent 3,000 words (or unigrams).

The input for the keyATM is all 533,582 floor speeches. I run keyATM by each congressional session, with the same set of keywords. I use a total of $K = 20$ topics. I do not include any additional topics without keywords because these pre-labeled topics—drawn from presidents from both Democratic and Republican parties—should encompass all possible topics on both sides of the aisle.[12]

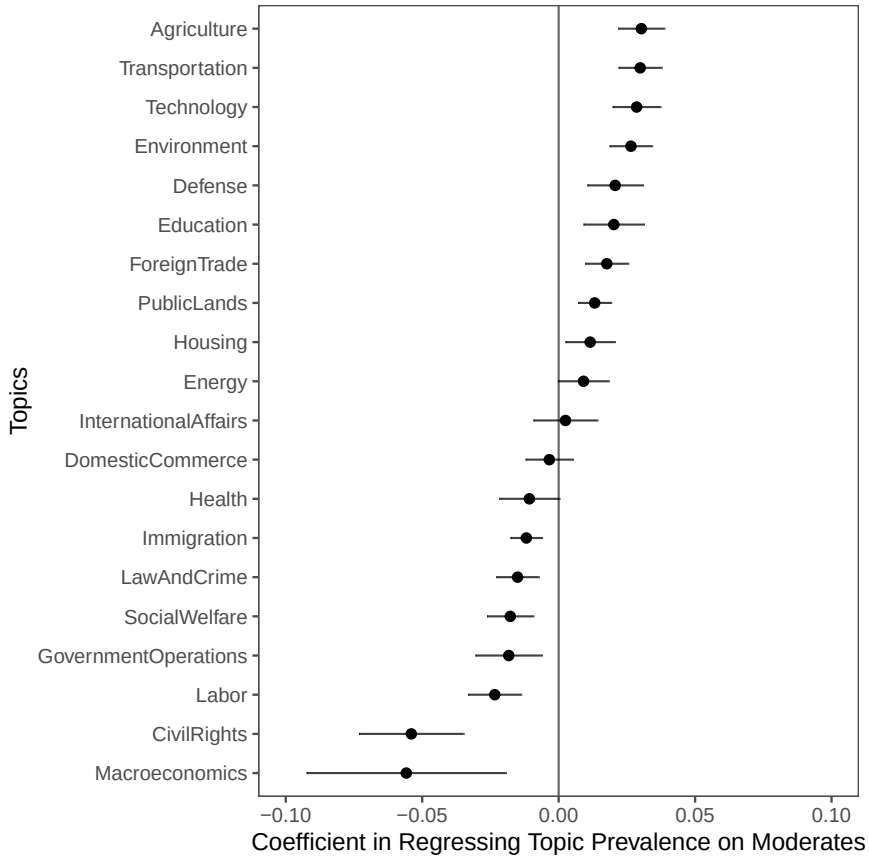
The outcome of the keyATM includes prevalence across the 20 topics for each floor speech. After estimating the topic prevalence of each floor speech, I similarly aggregate the data at the legislator-year level. Online Figure A.5 provides a general picture of the relations between prevalence across the topic domains and legislators' ideology. It shows the smooth fits between legislators' Nokken-Poole scores and the prevalence of these 20 topics. It is apparent that moderates and extremists do not distribute their attention across the topics evenly.

[10] State of the Union Speeches. The Policy Agendas Project at the University of Texas at Austin, 2022. www.comparativeagendas.net

[11] The 20 major topics are "Macroeconomics," "Civil Rights," "Health," "Agriculture," "Labor," "Education," "Environment," "Energy," "Immigration," "Transportation," "Law and Crime," "Social Welfare," "Housing," "Domestic Commerce," "Defense," "Technology," "Foreign Trade," "International Affairs," "Government Operations," and "Public Lands."

[12] I also attempted to add two no-keyword topics, but I found that these two topics do not capture more informative or meaningful issue topics and the prevalence of these two additional topics is close to zero.

Figure 5: Estimating Topic Prevalence of Moderates



Notes: The points are estimates of moderates on the prevalence of each topic and the lines refer to the 95% confidence interval.

To assess whether moderates talk about specific issues more or less compared to extremists, I separately regress the topic prevalence of each of the 20 topics on my measure of moderation, adding year-by-party fixed effects. The cross-sectional model shows that within the same chamber and within the same party, do moderates and extremists concentrate on different topic issues on the floor? Figure 5 displays the results. The y-axis lists all 20 topics and the x-axis refers to the coefficient of moderation in each cross-sectional regression, with the lines indicating the 95% confidence interval. Since the independent variable of interest has a range of zero to one, the coefficients mean how much more proportion in floor speech, on average, the most moderate legislators talk on a specific topic than the most extremists. The topics with positive coefficients indicate that the more moderate a member is, the more

proportion she talks about those topics; the topics with negative coefficients mean that the more extreme a legislator is, the more she discusses those issues. As Figure 5 shows, it is evident that moderates and extremists do not talk about all issues equally. Moderates prefer to talk more on topics, such as Transportation, Technology, Environment, Foreign Trade, Defense, Agriculture, and Education. While extremists focus more on Civil Rights, Macroeconomics, Law and Crime, Social Welfare, and Labor.

Mechanism for the Silent Moderates

The previous sections have provided a rich description of the speech patterns of moderates on the congressional floor. In terms of the partisan tenor and topic prevalence, it is fair to say that moderates talk in moderate ways. It would be strange if they did not. However, it is a mystery why moderates speak less on the floor. Why do moderates mute their voice in legislative deliberation on the congressional floor? This section discusses the mechanism for the silent moderates.

There are lots of speculations about the silent moderates and two main stories might explain why moderates stay quieter on the floor. The first is an election story. Moderates, who are more likely to represent marginal districts, happen to be electorally vulnerable; and extremists, who represent aligned and homogeneous districts, do not have to face the pressure of opposing partisans (Abramowitz, Alexander, and Gunning 2006; Grimmer 2013; McCarty, Poole, and Rosenthal 2009). For this reason, moderates would attempt to be more cautious in delivering floor speeches, especially in the C-SPAN era when every speech on the floor is recorded and broadcast. So, they would talk less in order to dodge contentious issues and avoid upsetting their split voters.

Another explanation for the silent moderates is a bargaining story. Being the pivotal player in legislation (Krehbiel 1998), moderates have a strong incentive to keep silent. They do not want to express what they really want, at least on the floor, so they intentionally

keep their positions ambiguous. Because ambiguity gives them more bargaining power. For example, if a moderate is pivotal with respect to the final passage of a bill, she would be less willing to discuss who she really wants on the floor and on the record. Rather, she would be more likely to resort to behind-the-door conversations in order to either add pet projects to the bill or just enjoy the media attention for a moment. The following subsections will empirically test these two stories.

Testing Election Story

On the election story, the empirical hypothesis is that moderates should stay quieter on the floor when they are electorally vulnerable. I come up with two measures on the electoral vulnerability of each legislator in a given year. The first is a binary measure, "cross-pressured," which means the partisanship of a legislator is at odds with the majority of home-district voters in the most recent presidential election. The second is a continuous measure, "two-party vote share of co-partisan presidential candidate" in the most recent presidential election. The two-party vote share has a range from zero to one, where a lower value indicates being more vulnerable in elections.

The research design to test the election story is to simply take the measure of electoral vulnerability into the same regressions shown previously. Table 3 shows the results. Columns (1) and (4) duplicate the previous results, showing that moderation has a statistically significant negative association with the total volume delivered on the floor, both cross-sectionally and within-legislator. Columns (2) and (5) demonstrate that even conditional on whether legislators are cross-pressured or not, the relationship between moderation and floor speech volume remains statistically significant and stays in a similar magnitude. Also, the cross-pressured indicator has null effect on the word counts. Similarly, in Columns (3) and (6), even controlling for a continuous measure, the association of legislators' moderation and their floor speech volume still holds. On top of the moderate ideology, the co-partisan presidential vote share explains very little on the floor speech volume.

Table 3: Testing Election Story on Volume of Floor Speech

	Dependent Variable: Logged Word Count					
	Cross-Sectional			Within-Legislator		
	(1)	(2)	(3)	(4)	(5)	(6)
Moderation	-1.166^{***}	-1.132^{***}	-1.129^{***}	-0.347^{**}	-0.347^{**}	-0.350^{**}
	(0.097)	(0.098)	(0.101)	(0.174)	(0.173)	(0.173)
Cross-Pressured		−0.044			−0.002	
		(0.033)			(0.036)	
Co-Partisan Pres. Vote Share			0.001			−0.0005
			(0.001)			(0.001)
Year×Party FE	✓	✓	✓	✓	✓	✓
Covariates	✓	✓	✓			
Legislator FE				✓	✓	✓
Observations	13,982	13,982	13,982	13,982	13,982	13,982
R^2	0.131	0.132	0.132	0.619	0.619	0.619

Notes:
1. *p<0.1; **p<0.05; ***p<0.01.
2. Robust standard errors are clustered by party-year in Columns (1)–(3) and clustered by legislator in Columns (4)–(6).

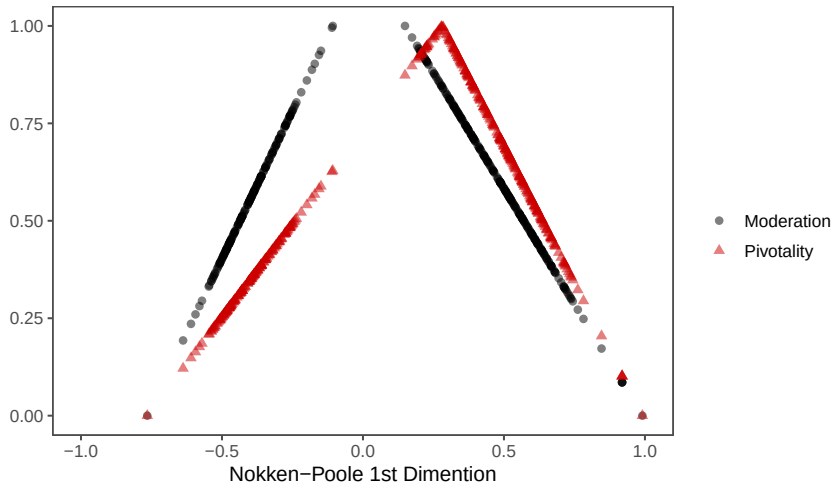
An auxiliary analysis (see Online Appendix Table A.5) investigates whether the phenomenon of silent moderates varies during election seasons. I disaggregate to legislator-day panel data and originally collect primary dates of congressional elections by state. Even interacting with the primary election seasons (dates between the filing dates and primary dates) and the general election seasons (dates between the primary dates and general dates), the fact that moderates stay quieter on the floor, both cross-sectionally and within-legislator, barely changes across the electoral calendar. On net, I do not find evidence to support the election story. Moderates being silent on the floor is not related to their electoral connection, and there is something else in play.

Testing Bargaining Story

An empirically testable hypothesis related to the bargaining story is that moderates are quieter when they are closer to the chamber median (being pivotal). The logic is that when a moderate is playing a pivotal role in the passage of a bill, she has more bargaining power entitled to her pivotality. The pivotal member would be less likely to express what she really wants on the floor; instead, she might resort to communications behind the closed door. Here, I measure how pivotal legislators are by absolute Nokken-Poole distance to the chamber median. Then, I create a measure, "pivotality," which rescales the absolute Nokken-Poole distance to the median to the range of zero and one by Congress. This pivotality measure is distinct from the moderation measure. For example, Figure 6 lays out the moderation and pivotality of House members in the 114th Congress. Each point is a House member. X-axis is the Nokken-Poole score and y-axis is the value of these two measures. The black dots demonstrate the moderation measure, and the red triangles refer to the pivotality score. The peak of pivotality depends on which party controls the House and how much the margin is. Since the GOP dominated the 114th House, among those moderate Republicans is where I could get identification. Because some moderate GOP members are higher in moderation but lower in pivotality, and some are higher in pivotality but lower in moderation.

I run a horse-race regression to assess which measure actually explains the variation in floor speech volume more effectively. Table 4 shows the results. Still, Column (1) and Column (3) display the results from the previous analysis. Column (2) presents the horse-race regression result in the cross-sectional model, and Column (4) presents the result in the within-legislator model. Comparing the two variables with the same scale, the variable of moderation does not reach the significance level anymore; instead, it is how pivotal a legislator is that explains the most variation in the total volume of floor speech.

Figure 6: Moderation and Pivotality of House Members in the 114th Congress



Notes: Each point indicates a House member. Black circles refer to the moderation score of each legislator and red triangles represent the pivotality score of each legislator.

Table 4: Testing Bargaining Story on Volume of Floor Speech

	Dependent Variable: Logged Word Count			
	Cross-Sectional		Within-Legislator	
	(1)	(2)	(3)	(4)
Moderation	−1.166***	−0.507	−0.347**	0.126
	(0.097)	(0.372)	(0.174)	(0.300)
Pivotality		−0.675*		−0.515*
		(0.341)		(0.275)
Year×Party FE	✓	✓	✓	✓
Covariates	✓	✓		
Legislator FE			✓	✓
Observations	13,982	13,982	13,982	13,982
R^2	0.131	0.132	0.619	0.619

Notes:
1. *p<0.1; **p<0.05; ***p<0.01.
2. Robust standard errors are clustered by party-year in Columns (1)–(2) and clustered by legislator in Columns (3)–(4).

Another auxiliary analysis (see Online Appendix Table A.6) considers the deliberation for major legislation. The logic here is that if moderates talk less on the floor because they are pivotal, this should be particularly acute when they are discussing major laws. Again, I disaggregate the data to the legislator-by-day level. I originally collect data on the dates when congressional priority agendas (Curry and Lee 2020) are under debate on the floor. I find that during those days when major legislation is fully discussed on the House floor, moderate legislators would even stay quieter than on normal days. Even focusing only on the days when major legislation is under debate, see Online Table A.7, the more pivotal a legislator is, the quieter she stays on the floor. In sum, evidence suggests that moderates are silent because they are pivotal and they want to bargain for more.

Conclusion

In the current hyper-polarized era, it is the impression that political communication is dominated by extremists, while moderate legislators mostly stay quiet. Previous literature has competing expectations about the communication patterns of members of Congress in the middle of the ideological spectrum. This paper, drawing historical evidence from all speeches on the House floor, demonstrates a holistic description of the communication patterns of moderates over the years. Descriptively, on the congressional floor, moderates have systematically different speech patterns, compared to extremists. Moderates on average deliver fewer floor speeches in total volume. Comparing legislators across the chamber, within the same district, and within themselves over time, I find consistent and robust evidence that moderation in ideology is associated with a less total volume of floor speeches. Moreover, moderates consistently talk in a lower partisan tenor than extremists and have particular sets of policy topics they prefer to talk about more often. In terms of why moderates keep silent in the venue of legislative deliberation, I find evidence that election considerations are limited in explaining the silent moderates; instead, moderates stay quiet on the floor because

they are pivotal and want to bargain for more.

To what extent this bargaining story can be extended to the Senate? More research needs to be done, but I expect this mechanism should be more acute on the other side of Capitol Hill. It might be a bit complicated in terms of who the pivotal player is in the Senate because of the filibuster rule. But since a bulk of major legislation is passed in the Senate as budget reconciliations, which are not subject to the filibuster, the pivotal Senator is quite obvious in a 50-50 scenario, such as Joe Manchin (D-WV) and Kyrsten Sinema (D-AZ) in the current Senate. Given the higher stakes, these moderate Senators are reasonably more likely to hold their cards close to chests in order to bargain for more. For example, in 2022, with the Climate and Tax Deal in sight, the moderate Democrat, Kyrsten Sinema, had been her party's final holdout on the package. She intentionally stayed silent for several days before Democratic leaders made changes. In the final passage of the bill (Inflation Reduction Act of 2022), she used her silence to bargain for more, such as including 1% excise tax on stock buybacks, adding $5 billion in drought resiliency, and whipping her colleagues to swiftly confirm Roopali Desai, a Sinema-backed circuit court judge.

References

Abramowitz, Alan I., Brad Alexander, and Matthew Gunning. 2006. "Incumbency, Redistricting, and the Decline of Competition in US House Elections." *Journal of Politics* 68 (1): 75–88.

Arnold, Douglas R. 2004. *Congress, the Press, and Political Accountability.* Princeton, NJ: Russell Sage.

Austen-Smith, David. 1990. "Information Transmission in Debate." *American Journal of Political Science* 34 (1): 124–152.

Ban, Pamela, Ju Yeon Park, and Hye Young You. 2022. "How Are Politicians Informed? Witnesses and Information Provision in Congress." *American Political Science Review.*

Barberá, Pablo, Andreu Casas, Jonathan Nagler, Patrick J. Egan, Richard Bonneau, John T. Jost, and Joshua A. Tucker. 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.

Bernhard, William and Tracy Sulkin. 2018. *Legislative Style.* University of Chicago Press.

Berry, Christopher R. and Anthony Fowler. 2018. "Congressional Committees, Legislative Influence, and the Hegemony of Chairs." *Journal of Public Economics* 158:1–11.

Blei, David, Andrew Ng, and Michael Jordan. 2003. "Latent Dirichlet allocation." *Journal of Machine Learning and Research* 3: 993–1022.

Canes-Wrone, Brandice, David W. Brady, and John F. Cogan. 2002. "Out of Step, Out of Office: Electoral Accountability and House Members' Voting." *American Political Science Review* 96 (March): 127–40.

Chaturvedi, Neilan S. 2021. *Life in the Middle: Marginalized Moderate Senators in the Era*

of Polarization. Oxford University Press.

Cox, Gary W. and Mathew D. McCubbins. 2005. *Setting the Agenda: Responsible Party Government in the U.S. House of Representatives*. Cambridge: Cambridge University Press.

Curry, James M. and Frances E. Lee. 2020. *The Limits of Party: Congress and Lawmaking in a Polarized Era*. The University of Chicago Press.

Diermeier, Daniel, Jean-Francois Godbout, Bei Yu, and Stefan Kaufmann. 2011. "Language and Ideology in Congress." *British Journal of Political Science* 42 (1): 31–55.

Dietrich, Bryce J., Matthew Hayes, and Diana Z. O'Brien. 2019. "Pitch Perfect: Vocal Pitch and the Emotional Intensity of Congressional Speech." *American Political Science Review* 113 (4): 941–62.

Downs, Anthony. 1957. *An Economic Theory of Democracy*. New York: Harper.

Eshima, Shusei, Kosuke Imai, and Tomoya Sasaki. 2020. "Keyword Assisted Topic Models." Working Paper, arXiv:2004.05964.

Fenno, Richard F. 1978. *Home style: House Members in their districts*. Boston: Little, Brown.

Fu, Shu. "Against and Alongside the Bully Pulpit: Interbranch Messaging in a Polarized Era." Working Paper.

Fu, Shu and William G. Howell. 2020. "The Behavioral Consequences of Public Appeals: Evidence on Campaign Fundraising from the 2018 Congressional Elections." *Presidential Studies Quarterly* 50 (2): 325–347.

Fu, Shu and William G. Howell. 2023. "The Filibuster and Legislative Discussion." *The Journal of Politics* 85 (4): 1575–1580.

Garay, Ronald. 1984. *Congressional Television: A Legislative History*. Westport: Green-

wood Press.

Gentzkow, Matthew, Jesse M. Shapiro and Matt Taddy. 2018. *Congressional Record for the 43rd-114th Congresses: Parsed Speeches and Phrase Counts*. Palo Alto, CA: Stanford Libraries.

Gentzkow, Matthew, Jesse M. Shapiro and Matt Taddy. 2019. "Measuring Group Differences in High-Dimensional Choices: Method and Application to Congressional Speech." *Econometrica* 87 (4): 1307–1340.

Grimmer, Justin. 2013. *Representational Style in Congress: What Legislators Say and Why It Matters*. Cambridge University Press.

Grimmer, Justin and Eleanor Neff Powell. 2016. "Money in Exile: Campaign Contributions and Committee Access." *The Journal of Politics* 78 (4): 974–988.

Grimmer, Justin, Sean Westwood and Solomon Messing. 2014. *The Impression of Influence: Legislator Communication, Representation, and Democratic Accountability*. Princeton University Press.

Hall, Andrew B. 2015. "What Happens When Extremists Win Primaries?" *American Political Science Review* 109 (1): 18–42.

Hall, Andrew B. and Daniel M. Thompson. 2018. "Who Punishes Extremist Nominees? Candidate Ideology and Turning Out the Base in U.S. Elections." *American Political Science Review* 112 (3): 509–524.

Harris, Douglas B. 2005. "Orchestrating Party Talk: A Party-Based View of One-Minute Speeches in the House of Representatives." *Legislative Studies Quarterly* 30 (1): 127–141.

Hill, Kim Quaile and Patricia A. Hurley. 2002. "Symbolic speeches in the US Senate and their representational implications." *The Journal of Politics* 64 (1): 219–231.

Jones, Kevin L., Sharareh Noorbaloochi, John T. Jost, Richard Bonneau, Jonathan Nagler,

and Joshua A. Tucker. 2018. "Liberal and Conservative Values: What we can Learn from Congressional Tweets." *Political Psychology* 39 (2): 423–443.

Krehbiel, Keith. 1998. *Pivotal Politics: A Theory of U.S. Lawmaking*. Chicago: University of Chicago Press.

Krehbiel, Keith. 2017. "Majoritarianism, Majoritarian Tension, and the Reed Revolution". In *Governing in a Polarized Age: Elections, Parties, and Political Representation in America*, ed. Alan Gerber and Eric Schickler. New York, NY: Cambridge University Press pp. 328–370.

Lee, Frances E. 2009. *Beyond Ideology: Politics, Principles, and Partisanship in the U.S. Senate*. Chicago, IL: University of Chicago Press.

Lewis, Jeffrey B., Keith Poole, Howard Rosenthal, Adam Boche, Aaron Rudkin, and Luke Sonnet. 2022. *Voteview: Congressional Roll-Call Votes Database*. https://voteview.com/

Lipinski, Daniel. 2004. *Congressional Communication: Content and Consequences*. Ann Arbor: University of Michigan Press.

Maltzman, Forrest, and Lee Sigelman. 1996. "The Politics of Talk: Unconstrained Floor Time in the U.S. House of Representatives." *The Journal of Politics* 58: 819–830.

Mayhew, David R. 1974. *Congress: The Electoral Connection*. New Haven, CT: Yale University Press.

McCarty, Nolan, Keith T. Poole, and Howard Rosenthal. 2006. *Polarized America: The Dance of Ideology and Unequal Riches*. Cambridge, Massachusetts: MIT Press.

McCarty, Nolan, Keith T. Poole, and Howard Rosenthal. 2009. "Does Gerrymandering Cause Polarization?" *American Journal of Political Science* 53 (3): 666–80.

Morris, Jonathan D. 2001. "Reexamining the Politics of Talk: Partisan Rhetoric in the 104th House." *Legislative Studies Quarterly* 26: 101–121.

Noble, Benjamin S. "Presidential Cues and the Nationalization of Congressional Rhetoric, 1973-2016." *American Journal of Political Science*.

Nokken, Timothy P. and Keith T. Poole. 2004. "Congressional Party Defection in American History." *Legislative Studies Quarterly* 29 (4): 545–568.

Oleszek, Walter. 1984. *Congressional Procedures and the Policy Process*. Washington, D.C.: Congressional Quarterly Press.

Poole, Keith T. and Howard Rosenthal. 1997. *Congress: A Political-Economic History of Roll Call Voting*. Oxford University Press.

Proksch, Sven-Oliver, and Jonathan B. Slapin. 2012. "Institutional Foundations of Legislative Speech." *American Journal of Political Science* 56 (3): 520–537.

Quinn, Kevin, Burt L. Monroe, Michael Colaresi, Michael H. Crespin, and Dragomir R. Radev. 2010. "How to Analyze Political Attention with Minimal Assumptions and Costs." *American Journal of Political Science* 54 (1): 209–28.

Ritchie, Melinda N. 2018. "Back-Channel Representation: A Study of the Strategic Communication of Senators with the US Department of Labor." *The Journal of Politics* 80 (1): 240–253.

Roberts, Margaret, Brandon Stewart, and Dustin Tingley. 2017. stm: R Package for Structural Topic Models. http://www.structuraltopicmodel.com.

Russell, Annelise. 2017. "U.S. Senators on Twitter: Asymmetric Party Rhetoric in 140 Characters." *American Politics Research* 46 (4): 695–723.

Theriault, Sean M. 2008. *Party Polarization in Congress*. New York: Cambridge University Press.

Wagner, Michael W. and Mike Gruszczynski. 2018. "Who Gets Covered? Ideological Extremity and News Coverage of Members of the U.S. Congress, 1993 to 2013." *Journalism & Mass Communication Quarterly* 95 (3): 670–690.

Westwood, Sean. 2022. "The Partisanship of Bipartisanship: How Representatives Use Bipartisan Assertions to Cultivate Support." *Political Behavior* 44: 1411–1435.

Yiannakis, Diana Evans. 1982. "House Members' Communication Styles: Newsletter and Press Releases." *The Journal of Politics* 44 (4): 1049–71.

Zaller, John and Dennis Chiu. 1996. "Government's Little Helper: US Press Coverage of Foreign Policy Crises, 1945–1991." *Political Communication* 13 (4): 385–405.

Moderates on Capitol Hill:

Volume, Tenor, and Topics of House Floor Speeches, 1985–2016

Shu Fu

Online Appendix

A. Robustness Checks

- Alternative Measure 1: NOMINATE scores
- Alternative Measure 2: In Absolute Terms
- Alternative Measure 3: Non-rescaled Nokken-Poole and Rescaled NOMINATE
- Estimates Over the Years

B. Text-as-Data Methods

- Topics and Keywords
- KeyATM Topic Prevalence

C. Auxiliary Analysis on Mechanism

- Primary and General Election Seasons
- Debating Major Legislation

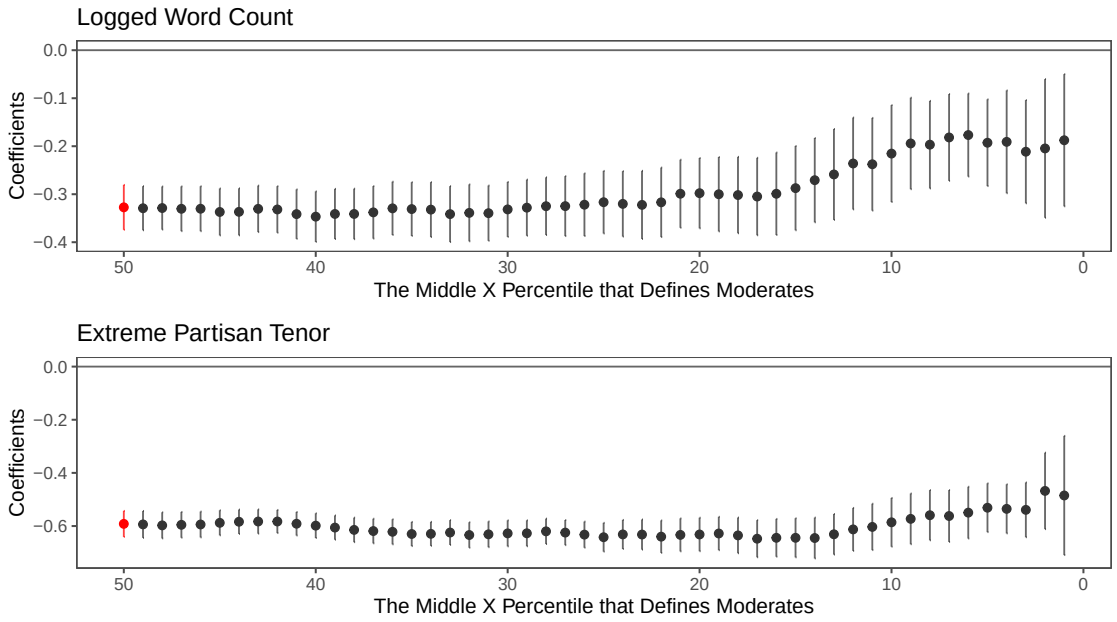
A. Robustness Checks

Alternative Measure 1: NOMINATE scores

In this section, I discuss alternative measures to demonstrate the robustness of the previous findings. At the heart of this paper is the inquiry of what it means to be a "moderate." Intuitively, a moderate legislator is someone situated in the middle of the ideological spectrum between the liberal and conservative poles. However, "moderate" is a relative term compared to extremists; there is no convention defining the exact range containing them in the ideological center. Therefore, first and foremost, I rely on NOMINATE scores, 1st dimension, and code an indicator variable, NOMINATE Moderates, as "1" if the NOMINATE score of a legislator is between the interquartile range in a given Congress. This measure divides all House members within the same Congress into one-half moderates and one-half extremists. Then, I vary the middle percentile that defines moderates from the 50th percentile to the 1 percentile. And I run the cross-sectional model specification with covariates and incrementally decrease the percentile in the middle of the NOMINATE score in each Congress that defines moderates.

Figure A.1 shows the results. The findings are consistent and robust in that moderates deliver fewer floor speeches and smaller total volume if I define moderates from the middle 50th to the middle 1st percentile. But given the sheer number of those legislators (i.e., 44 for the middle 10th percentile or 22 for the middle 5th percentile), the uncertainty of the estimate increases. In terms of partisan tenor, the evidence shows that no matter how narrow the range for moderates, they consistently speak in a lower partisan tenor than their extremist counterparts.

Figure A.1: Measuring Moderates with Different Percentiles



Notes: The points are estimates of moderates and the lines refer to the 95% confidence interval. The highlighted red items refer to the middle 50th percentile that defines moderates.

Alternative Measure 2: In Absolute Terms

I use a couple of ways to measure moderates in absolute terms. First, I define legislators with NOMINATE scores between -0.2 to 0.2 as moderates, and apply to each Congress. Second, I define legislators with NOMINATE scores within the interquartile range across 1985-2016 as moderates. The tables below show the cross-sectional estimates on the volume and partisan tenor.

Table A.1: Cross-Sectional Estimates of Moderates in Absolute Terms 1

	Dependent Variable	
	Logged Word Count	Extreme Partisan Tenor
	(1)	(2)
NOMINATE Moderates	-0.346^{***}	-0.675^{***}
(NOMINATE scores, 1st dimension $\in [-0.2, 0.2]$)	(0.043)	(0.039)
Year×Party FE	✓	✓
Covariates	✓	✓
Observations	13,984	13,984
R^2	0.149	0.213

Notes: 1. *p<0.1; **p<0.05; ***p<0.01. 2. Robust standard errors clustered by year-party.

Table A.2: Cross-Sectional Estimates of Moderates in Absolute Terms 2

	Dependent Variable	
	Logged Word Count	Extreme Partisan Tenor
	(2)	(3)
NOMINATE Moderates	-0.351^{***}	-0.600^{***}
($\in$ all-time interquartile range, $[-0.376, 0.392]$)	(0.024)	(0.026)
Year×Party FE	✓	✓
Covariates	✓	✓
Observations	13,984	13,984
R^2	0.163	0.243

Notes: 1. *p<0.1; **p<0.05; ***p<0.01. 2. Robust standard errors clustered by year-party.

Alternative Measure 3: Non-rescaled Nokken-Poole and Rescaled NOMINATE

In the within-legislator estimates, I replace the variable of interest with two alternative measures. The first one, Nokken-Poole Moderation (non-rescaling) uses the raw Nokken-Poole scores without rescaling and simply flips the signs of Nokken-Poole scores. The second one, NOMINATE Moderation, is the rescale moderation score using NOMINATE scores. Since NOMINATE scores are static for each legislator, the rescaled scores capture only the variation caused by the composition of the party caucus changing over time.

Table A.3: Alternative Measures of Moderates in Within-Legislator Model

	Dependent Variable: Logged Word Count	
	Within-Legislator	
	(1)	(2)
Nokken-Poole Moderation (non-rescaled)	-0.637^{***}	
	(0.200)	
NOMINATE Moderation		0.328
		(0.231)
Year-Party FE	✓	✓
Legislator FE	✓	✓
Observations	13,982	13,982
R^2	0.620	0.619

Notes:
1. *p<0.1; **p<0.05; ***p<0.01.
2. Robust standard errors clustered by legislator.

Estimates Over the Years

Figure A.2: Floor Speech Volume, by Presidential Terms

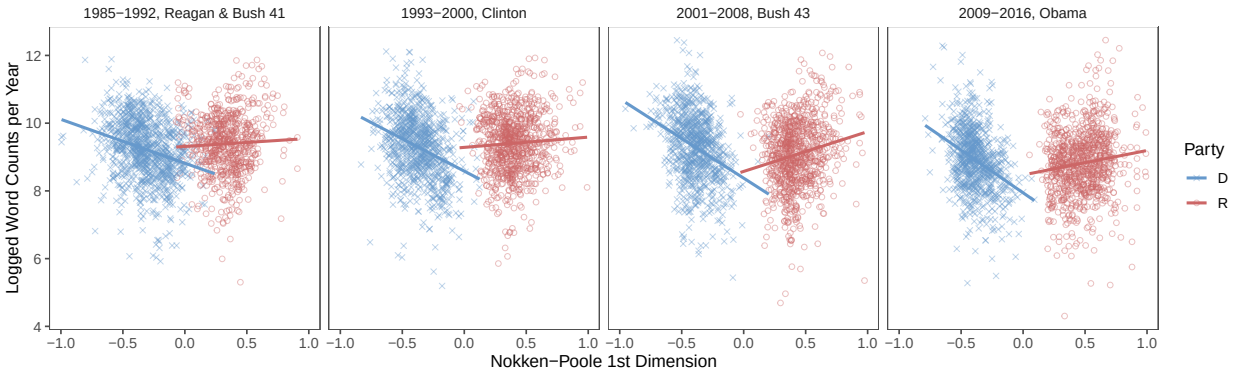
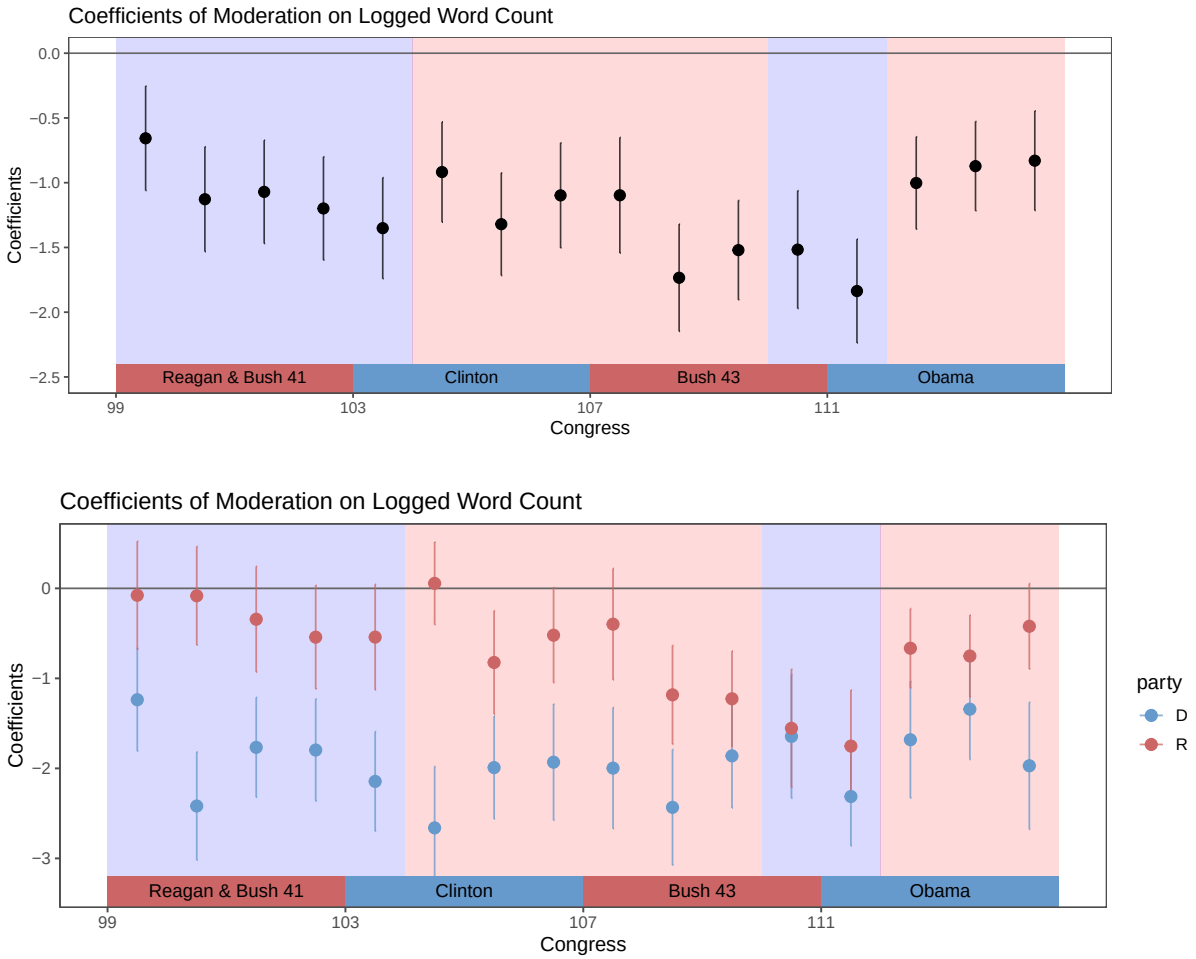
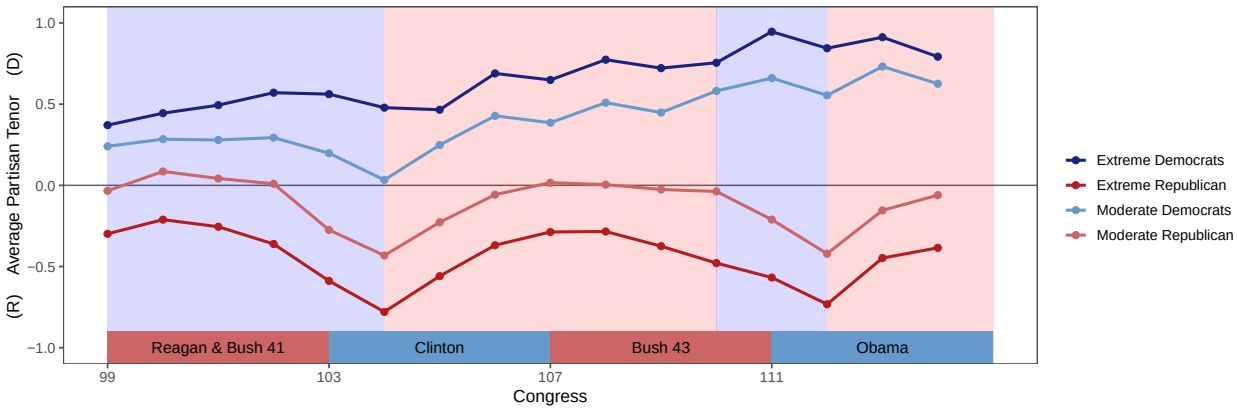


Figure A.3: Floor Speech Volume of Moderates Over Time



Notes: The points are estimates of moderates on logged word counts and the lines refer to the 95% confidence interval. The shaded color indicates the majority party in the House of Representatives, and the color bars at the bottom of each figure present the president and party occupying the White House.

Figure A.4: Partisan Tenor Trend Over Time

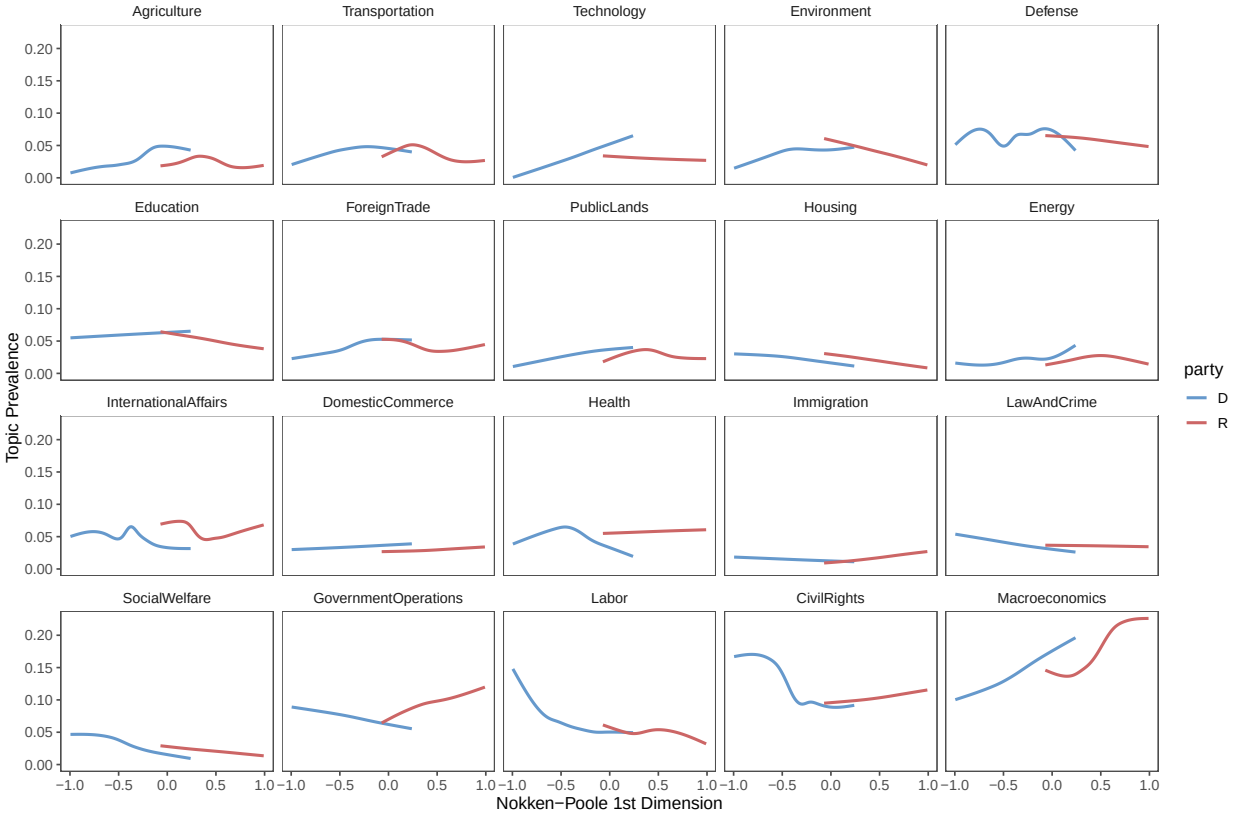


Text-as-Data Methods

Table A.4: Topics and Keywords

Topics	**Keyword Stems**
Macroeconomics	economi, deficit, job, spend, econom, growth, pai, invest, grow, billion, monei, recoveri, credit
Civil Rights	discrimin, marriag, pai, race, children, children, gai, human, share, black, love, violenc, white
Health	medicar, medic, children, drug, doctor, patient, pai, prescript, research, human, cancer, choic
Agriculture	farmer, market, water, food, clean, drink, safeti
Labor	job, worker, wage, train, economi, skill, commun, unemploy, pai, earn, expand, growth, invest
Education	student, children, teacher, child, commun, learn, math, parent, teach, classroom, graduat, kid
Environment	clean, climat, pollut, commun, challeng, water, environment, emiss, toxic, global, greenhous
Energy	clean, ga, technologi, fuel, car, invest, effici, research, solar, cleaner, wind
Immigration	border, antiimmigr, illegal, comprehens, alien, amnesti, deport, dream
Transportation	infrastructur, project, speed, worker, invest, rebuild, attract, auto, auto, built, highwai, car, rail
Law and Crime	crime, drug, children, gun, parent, commun, child, violent, street, violenc, bradi, kid
Social Welfare	child, children, parent, worker, commun, poverti, poor, monei, pai, earn, medicar
Housing	commun, zone, empower, poor, homeless, ownership, homeown, neighborhood
Domestic Commerce	credit, job, invest, mortgag, street, wall, entrepreneur, crisi, lend, monei
Defense	nuclear, iraq, weapon, troop, alli, fight, iraqi, terrorist, threat, iran, nato
Technology	research, space, technologi, internet, invest, comput, gener, centuri, explor, digit, tech
Foreign Trade	trade, job, market, worker, economi, global, china, econom, expand, partner, growth, invest
International Affairs	terrorist, democraci, weapon, terror, iraq, east, alli, fight, afghanistan, threat, regim, soviet
Government Operations	spend, bipartisan, lobbyist, bureaucraci, monei, democraci
Public Lands	park, everglad, project, acr, ag, banner, beauti, rock, yellowston

Figure A.5: Topics and Ideology: 99th–114th Congresses



Notes: These LOESS fit lines, colored by party, delineate the relationship between legislators' ideology and topic prevalence in their speeches.

C. Auxiliary Analysis on Mechanism

Primary and General Election Seasons

Table A.5: Silent Moderates Across the Electoral Calendar

	Dependent Variable: Logged Word Count			
	Cross-Sectional		Within-Legislator	
	(1)	(2)	(3)	(4)
Moderation	-0.849^{***}	-0.843^{***}	-0.448^{***}	-0.470^{***}
	(0.019)	(0.021)	(0.038)	(0.040)
Primary Season		0.006		-0.052^{**}
		(0.026)		(0.025)
General Season		0.126^{***}		-0.071^{***}
		(0.027)		(0.026)
Moderation × Primary Season		0.031		0.070^{*}
		(0.044)		(0.042)
Moderation × General Season		-0.069		0.065
		(0.045)		(0.042)
Party×Day FE	✓	✓	✓	✓
Covariates	✓	✓		
Legislator FE			✓	✓
Observations	1,160,826	1,157,874	1,160,826	1,157,874
R^2	0.069	0.069	0.138	0.138
Note:	*p<0.1; **p<0.05; ***p<0.01.			

Debating Major Legislation

Table A.6: Silent Moderates When Major Legislation Under Debate

	Dependent Variable: Logged Word Count			
	Cross-Sectional		Within-Legislator	
	(1)	(2)	(3)	(4)
Moderation	-0.783^{***}	-0.769^{***}	-0.562^{***}	-0.551^{***}
	(0.015)	(0.015)	(0.142)	(0.142)
Moderation ×		-0.186^{**}		-0.142^{*}
Debating Major Legislation		(0.072)		(0.073)
Party×Day FE	✓	✓	✓	✓
Covariates	✓	✓		
Legislator FE			✓	✓
Observations	1,772,438	1,772,438	1,772,438	1,772,438
R^2	0.074	0.074	0.137	0.137
Note:				*p<0.1; **p<0.05; ***p<0.01

Table A.7: Subsetting Days for Debating Major Legislation

	Dependent Variable: Logged Word Count	
	Cross-Sectional	Within-Legislator
	(1)	(2)
Moderation	−0.038	−0.137
	(0.113)	(0.189)
Pivotality	−0.241**	−0.135
	(0.114)	(0.180)
Party×Day FE	✓	✓
Covariates	✓	
Legislator FE		✓
Observations	32,962	32,962
R^2	0.057	0.207
Note:		*p<0.1; **p<0.05; ***p<0.01