Introducing the Congressional Support for the Use of Military Force (CSUMF) Dataset

Patrick Hulme

April 8, 2023

Abstract

The effect of congressional support for, or opposition to, the use of military force on the ultimate intervention decision of the executive branch is a matter of debate in political science and related disciplines such as history and law. The dominant Imperial Presidency thesis suggests Congress has little influence over such decisions, but historians often highlight congressional sentiment as being decisive in individual case studies. There are, moreover, serious normative implications of how closely U.S. policy corresponds to the will of the American people's representatives.

Measuring congressional sentiment toward possible military interventions is difficult, however, as lawmakers strategically avoid taking votes "on the record" on use of force decisions. Because of this, there has thus far yet to exist any comprehensive dataset covering sentiment in the legislature toward the use of military force across crises over time. Nevertheless, legislators often strongly express their views of deployment decisions in speech and print. A new dataset capturing such sentiment—the Congressional Support for the Use of Military Force (CSUMF) dataset—is introduced in this paper. The dataset utilizes floor speeches from the congressional record to measure sentiment toward the possible use of force in crises from the end of the Second World War until today. It not only includes overall congressional sentiment in nearly two hundred crises after 1945, but also provides a break-down by political party, as well as dynamic scores allowing for variance within crisis over time. Validation tests comparing the scores to roll call votes and public opinion polls suggest they are superior proxies to any other alternative—including widely used measures such as the portion of Congress consisting of Republicans (McManus 2017) or presidential copartisans (Howell & Pevehouse 2007, Kriner 2010).

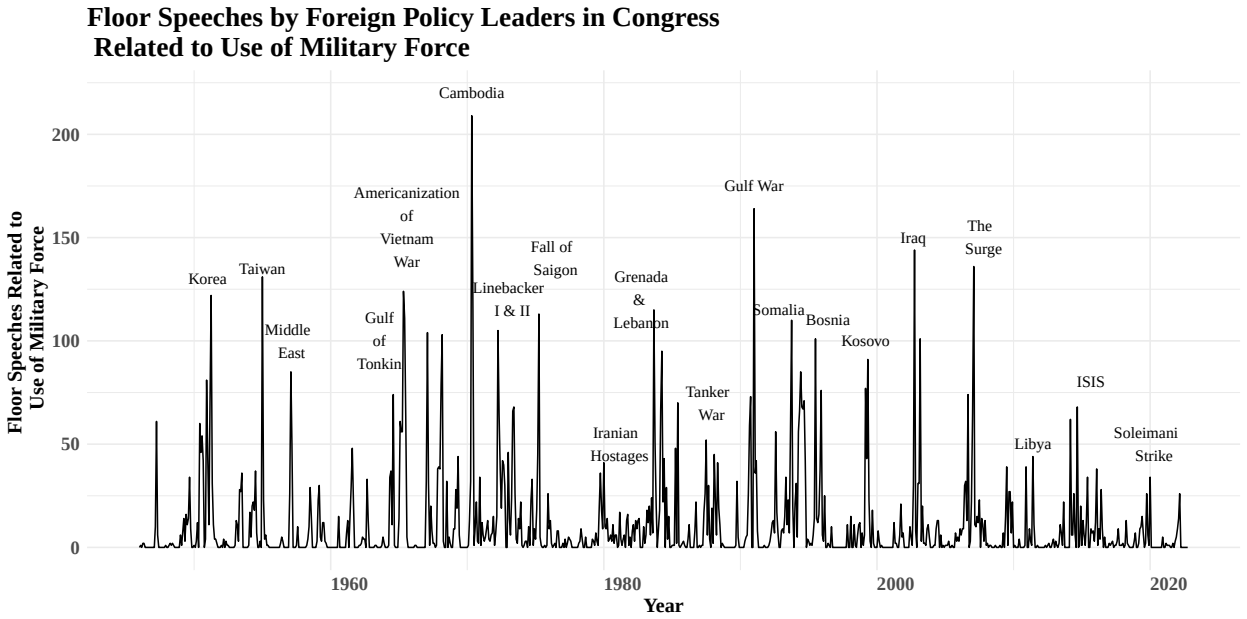
The Problem

Democratic control over the use of military force is of both positive and normative interest to academics, policy-makers, and even the public more broadly. From a positive perspective, political scientists, historians, and legal scholars have debated how much influence Congress has over use of military force decisions (Wildavsky 1966, Howell & Pevehouse 2007, Kriner 2010, Sofaer 1976, Cox 1984), with many concluding that lawmakers have little power when it comes to decisions in military crises (Lindsay 1994, Burns 2019, Schlesinger 1973). Normatively, many scholars, pundits, politicians, and everyday Americans desire that their representatives in government have a strong say over when the U.S. goes to war. Regrettably, the conventional wisdom of the Imperial Presidency suggests that since the end of the Second World War, much of this desire left unfulfilled (Schlesinger 1973, Fisher 2013).

Determining the extent of Congress's influence over use of military force decisions requires, however, a sufficient measure of the legislator attitudes in specific crises. Most attention is given to whether uses of military force are undertaken pursuant to formal authorization from Congress, with uses of force undertaken unilaterally often presumed to be against congressional will (Fisher 2013, Burns 2019, e.g.). Data merely focusing on use of force votes by legislatures (Ostermann & Wagner 2023, e.g.), however, is insufficient in the American context for determining the extent of congressional influence precisely because use of force votes are relatively rare in the United States, and—if they take place at all—tend to take place only after the crisis has elevated to a relatively high level. A myopic focus on formal votes misleadingly suggests members of Congress give little attention to questions regarding the use of American military forces.

Yet, Members of Congress are actually quite active in making their opinions over possible uses of force known both to the White House and the general public. Plotted below is a frequency plot of floor speeches given by foreign policy leaders in Congress[1] related to the possible use of military force in crises since the Second World War. Instead of abdicating responsibility for foreign policy, the data shows congressional leaders have given substantial attention to the use of American military force in virtually every crisis encountered by the United States.

[1] More on this below.



It is important to measure the sentiment conveyed by congressional leaders, moreover, because determining how Imperial of a Presidency we have rests on how close the policy ultimately implemented by the White House diverges from the desires of lawmakers. To give a concrete example, President Obama began an extended air campaign against ISIS in the fall of 2014, and did so without getting specific formal authorization from Congress. Legislators, moreover, failed to pass a law authorizing the campaign even after requested by the White House. Proponents of the Imperial presidency, therefore, chalk the episode up as yet another example of congressional impotence in foreign policy. Yet this is an odd interpretation of the episode, because members of Congress on both sides of the aisle widely supported the intervention. Indeed, Republicans vocally berated Obama for referring to ISIS as the "J.V. team" and for leaving a vacuum in the region by withdrawing American forces three years prior. Foreign policy leaders such as Senators John McCain, Lindsey Graham, and Marco Rubio of the Foreign Relations and Armed Services Committee, for example, had relentlessly berated the administration after the Fall of Mosul to ISIS forces in June 2014—loudly calling for a massive American intervention two months before the White House chose to do so.[2] It would be strange to classify President Obama's actions as undemocratic when he was reluctantly doing precisely what lawmakers were demanding of him. If anything, the crisis was a

[2]https://www.politico.com/story/2014/06/john-mccain-iraq-criticize-barack-obama-107780

perfect example of executive responsiveness to congressional sentiment, even if the intervention was undertaken unilaterally from a formal legal perspective. Hence, in order to systematically study the relative power of the President and Congress over the use of military force, we need an adequate measure of congressional support for, or opposition to, the use of military force in specific crises over time.

Measuring Congressional Support for the Use of Military Force

While legislative support for, or opposition to, the employment of American military force in a crisis is theoretically quite important (Schultz 2001, e.g.), actually measuring it is difficult. One technique commonly utilized in the political science literature is to simply use the proportion of Congress made up of copartisans of the President (Howell & Pevehouse 2005, Howell & Pevehouse 2007, Kriner 2010) or Republicans (McManus 2017) as a proxy for congressional support. Both of these proxies, however, have significant drawbacks. First, as shown below, both of these measures are very poor predictors of, e.g, the few use of force votes that do exist or public opinion polls regarding the use of force in specific crises. It is not difficult to identify major cases in which non-copartisans gave greater support to the President in using force (e.g., Vietnam under Johnson, or the Afghanistan "Surge" under Obama), or in which otherwise hawkish Republicans opposed the use of force (e.g., Haiti or Bosnia under Clinton). Moreover, the fixed nature of these proxies does not readily allow for a change in congressional support during a crisis[3] or across crises occurring within the same congressional session.[4] Furthermore, the composition of Congress is easily observable *ex ante* and thus is less likely to affect the outcome of crises we observe.[5] Moreover, many prominent theories focusing on the effect of opposition party signalling on crisis outcomes (Schultz 1998, Schultz 2001, Ramsay 2004) cannot even be tested with proxies such as these because they give us no information about the stance of the opposition party.

[3]For example, congressional support for U.S. military operations in Somalia in the early 1990's declined rapidly in the fall of 1993.

[4]For example, in 1973 the same Congress that was highly adverse to re-engagement in Vietnam was quite willing to take a strong stand supporting Israel in the Yom Kippur War (Kissinger 2011*b*).

[5]I.e., whatever effect the composition of Congress might have should be "priced in" to the decision to start a crisis, but is less likely to affect the outcome of the crisis (Fearon 1994).

Additionally, there is no consensus over whether ideology or partisan politics drives the foreign policy preferences and behavior of members of Congress. In other words, if one has to choose between utilizing the "% copartisans" and "% Republicans" proxies, it is not immediately clear which is more appropriate. While Howell and Pevehouse argue that partisan politics is key (2005, 2007), more recent work by McManus (2017) and Bendix and Jeong (2022) suggests that ideological beliefs are predominant. Others show that *both* are important (Böller 2021).

Table 1: Competing Measures of Congressional Support for the Use of Military Force

	% Congress Copartisan[6]	**% Congress Republican**[7]	**Roll Call Votes**	**Congressional Support Scores**[8]
Coverage:	Universal	Universal	Low	High
Difficulty to Measure:	Low	Low	Medium	High
Easily Observable *Ex Ante*:	Yes	Yes	Sometimes	Sometimes
Captures Intraparty Heterogeneity:	No	No	Yes	Yes
Can Measure Opposition Party Stance:	No	No	Yes	Yes
Allows for Change over Course of Crisis:	No	No	Only if multiple votes	Yes
Allows for Different Sentiment in Different Concurrent Crises:	No	No	Only if multiple votes	Yes
Correlates Well with Actual Use of Military Force Votes:	Weakly	No	-	Strongly
Correlates Well with Public Opinion Polls in Crises:	No	No	-	Strongly

Another commonly utilized method of capturing congressional sentiment over a given issue—the analysis of roll call votes—is also of relatively limited utility in this issue area because members of Congress are quite reluctant to vote on use of military force decisions. This means that the vast majority of crises lack any relevant vote from which we could deduce congressional sentiment.

[6]Howell & Pevehouse (2007); Kriner (2010).
[7]McManus (2017).
[8]Introduced below.

In order to measure congressional sentiment over potential uses of force in a much broader set of crises than the select few in which formal votes are taken, we instead focus on what members of Congress *say*. Even when Congress declines to formally vote on the use of military force, members of Congress frequently make their positions publicly known through their rhetoric—be it on cable news, through op-eds, via press releases, or by way of speeches on the floor of the Senate.

This paper describes how "Congressional Support Scores"—i.e. a metric representing congressional support or opposition to the use of military force in specific crises since World War II—were measured. First, it describes how a "universe" of nearly two hundred postwar U.S. relevant crises were identified. Next, it discusses how congressional floor speeches from foreign policy leaders in Congress were collected and scored in each of these crises. It then describes how an aggregate "Congressional Support Score" for each crisis was calculated before introducing descriptive statistics of the dataset. Validation tests comparing the dataset to public opinion polls and actual use of force votes by the legislature are then presented, showing that the Congressional Support Scores far outperform any other proxy. Lastly, the dataset is utilized to show that the amount of force a President is willing to utilize is seemingly highly constrained by sentiment in the legislature—suggesting the Presidency is far less Imperial than commonly believed (Schlesinger 1973).

Coverage: Identifying U.S.-relevant crises in the Postwar Period

To begin with, a "universe" of cases in which the use of American military force might have reasonably been considered had to be identified. The aspiration for this dataset was to be as comprehensive as possible—not only to include events included in datasets such as the Militarized Interstate Dispute (MIDs) or the International Crisis Behavior (ICB) datasets, but even beyond to include as many use of military force decisions as possible. The Interstate Crisis Behavior dataset contains around seventy crises in which the U.S. is deemed to have been a crisis actor, but we need to also consider "dogs not barking"—i.e., cases in which military action was considered but decided against. To give one example, the U.S. is not considered by the ICB dataset to have been a crisis actor during the 1975 Fall of Saigon,[9] but this was not due to any lack of serious consideration

[9] http://www.icb.umd.edu/dataviewer/?crisno=258

(Ford 1979).[10] In order to capture cases in which the U.S. might reasonably be considered to have had a substantial security interest, but in which it ultimately decided against intervention, we additionally include all ICB crises in which at least one crisis actor experienced economic—or greater—involvement by the United States in the crisis. Economic involvement in a conflict suggests the U.S. had some interest in the outcome of the crisis, even if its interest was not great enough to justify direct kinetic military intervention. Requiring this minimal level of U.S. involvement in the crisis helps eliminate cases in which the U.S. had negligible interests involved. In the ICB dataset, this includes cases in which UNSINV is rated at "4" or higher at the actor-level, and thus also includes cases of higher U.S. involvement, such as propaganda involvement, covert involvement, U.S. semi-military involvement (military aid or advisors, without participation in actual fighting) and cases in which American military forces were directly involved.[11] This yields 141 ICB-defined crisis between 1945 and 2016.

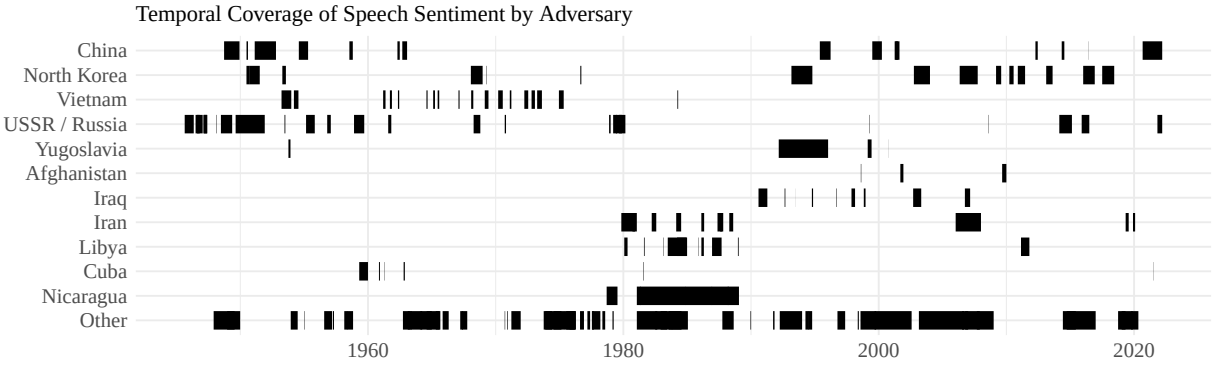
An even more inclusive list of U.S. crises was developed further in order to also include non-interstate crises (e.g., Somalia in the early 1990's or ISIS in the mid-2010's), crises occurring after 2017 (such as the 2020 Soleimani strike), and major decisions made during a war (for example the "Surge" decisions in Iraq and Afghanistan decided in early 2007 and late 2009, respectively). Cases from the ICB dataset in which the U.S. was not coded as being even "economically" involved, but for which we nonetheless have evidence that American intervention was considered, are also added. For example, we know from memoirs and news reporting that the White House closely followed the Russian invasion of Georgia in 2008, but consciously ruled out direct intervention. Lastly, the MIDs dataset was examined for disputes in which the U.S. was an actor, and each of these (except for fishing disputes or inadvertent airspace boundary violations) was also included. Thus, the dataset here includes no only virtually every potentially U.S.-relevant crisis from the ICB dataset, by also each of all but the most *de minimis* U.S. MIDs from the Correlates of War dataset. Altogether, this yields a grand total of 191 crises between 1945 and 2022. The list is meant to be as comprehensive

[10]Indeed, there is good reason to believe the administration's omission to use force was substantially due to congressional opposition to re-engagement in Indochina after the 1973 Paris Peace Accords.

[11]For similar approaches in identifying U.S. opportunities to use military force, compare to *Assessing Trade-Offs in U.S. Military Intervention Decisions: Whether, When, and with What Size Force to Intervene* (Frederick, Kavanagh, Pezard, Stark, Chandler, Hoobler & Kim 2021) and *Informing the Leader: Bureaucracies and International Crises* (Schub 2022).

as possible, and in future iterations will be expanded further. For the time being, the list should include virtually every crisis encountered by the White House in the time period in which force was even somewhat seriously considered.

The plot below depicts the distribution of crises over time, organized by the adversary faced by the United States. Many of the adversaries of the U.S. today—especially China, Russia, Iran, and North Korea—-are covered by a long history of crises extending all the way to the beginning of the Cold War. Crises with China extend all the way back to the Chinese Civil War in the late 1940's, through the Taiwan Straits crises of 1955, 1958, and 1996 to more recent crises over Taiwan up to 2022. Crises with North Korea similarly cover the beginning of the Korean War in 1950 until recent nuclear crises on the peninsula. Disputes with Russia extend from confrontations over Iran and Turkey shortly after the conclusion of World War II to the Ukrainian War starting in 2022.



Congressional Floor Speeches

In order to measure congressional sentiment over the possible use of force in these crises, we focus on floor speeches in the congressional record. While other sources—such as press releases, op-eds, interviews, and speeches outside Congress—could clearly also be used to convey opinions over possible uses of force, focusing on speeches available in the Congressional Record allows us to consult a single source whose relevance has seemingly remained high over the past eighty years.

Consider the following examples of floor speeches expressing support or opposition to the potential use of military force in Table 2. The first four examples convey sentiment in favor of the use of force.

Table 2: Speech Examples: Supporting the Use of Military Force

	Supporting Force
First Taiwan Strait Crisis (1955)	"[T]here can be no question that we should say to the world that we now propose to make our position clear. We must say that we will not be parties to the placing of Formosa and the Pescadores in unfriendly hands." —Senator Walter George (D-GA)
Libya (2011)	"[I] urge the President of the United States to take long overdue action to prevent the massacres that are taking place in Libya as we speak" —Senator John McCain (R-AZ)
ISIS (2014)	"Like it or not the American military is second to none. The special forces capability we have can really be decisive in this fight. To every American, this is not only about them over there: this is about us here. The better and the sooner that ISIL is defeated, the more decisive ISIL is defeated." —Senator Lindsey Graham (R-SC)
Venezuela (2019)	"We cannot let evil triumph in Venezuela. It would be a failure of leadership with disastrous consequences... It is becoming clear that we will have to consider the use of American military assets ... Our safety, national security, and the peace of our hemisphere demand that we take action." —Senator Rick Scott (R-FL)

Sometimes sentiment over the use of military force is conveyed during debate over legislation authorizing (or prohibiting) the use of force—as is the case in the example excerpt from the First Taiwan Strait Crisis. More often, however, such sentiment is conveyed outside the context of the consideration of specific legislation. For each of the other speech excerpts—related to the 2011 Libyan intervention, the counter-ISIS campaign beginning in 2014, and the 2019 Venezuela crisis[12]—a policy preference is being conveyed even if a specific congressional resolution is not being considered. Sometimes the support for the use of force is quite explicit, as illustrated by the McCain excerpt from 2011. Other times, intervention may not be specifically called for, but the implicit sentiment conveyed is clearly quite hawkish. In the Graham excerpt from the ISIS crisis, for example, the Senator never actually directly calls for intervention, but the sentiment in favor of the use of force is clear.

The next four speech excerpts come from the same four crises, but instead convey sentiment in opposition to the use of military force. As is the case for sentiment in favor of the use of force, sentiment opposed to military intervention can occur in the context of debate over congressional

[12]Esper (2022).

Table 3: Speech Examples: Opposing the Use of Military Force

	Opposing Force
First Taiwan Strait Crisis (1955)	"every Senator who votes for this resolution is authorizing the President ... to send troops anywhere —possibly thousands of miles into the mainland of China...[I am] not going to vote at this time to give a blank check to the President"—Senator William Langer (R-ND)
Libya (2011)	"Madam Speaker. we are in the midst of a foreign policy and constitutional crisis. The administration has committed our Nation to a war against Libya in violation of the Constitution of the United States." —Rep. Dennis Kucinich (D-OH)
ISIS (2014)	"I completely agree with the cautionary notes that have been cited about just sending in the U.S. military to do it. I think the risks there are enormous, and it would not be successful because it would unite... a fair number of Sunnis and radicals against us"—Rep. Adam Smith (D-WA), Chairman of the House Armed Services Committee
Venezuela (2019)	"The organizing principle of American policy seems to be the need to drive Maduro from power. What if Maduro is not really in power right now? What if the people who are really calling the shots in Venezuela are a group of transnational criminal organizations...? And what if their entire purpose is to draw the United States into an ill-advised war?...certainly, as a Congress, we need to be very critical in our thinking to not get our Nation in another ill-advised war." —Rep. Matt Gaetz (R-FL)

legislation—as depicted in the example from the First Taiwan Strait crisis. Another common focus of opponents to utilizing force is that such action violates the Constitution. Thus—as shown by Dennis Kucinich's opposition to the 2011 Libya Intervention—lawmakers can attack the authority of the executive (Christenson & Kriner 2020). Alternatively, as shown by the excerpts from the 2014 ISIS and 2019 Venezuela crises, legislators can focus on policy criticism instead of constitutional attacks (Christenson & Kriner 2017).

Note, further, that in each of the examples, above, copartisans of the President *opposed* the use of military force. This is especially notable in the first three examples because there were simultaneously opposition party legislators that supported the President in the use of force. Even seemingly close ideologues and allies of the President can have vastly different positions on the possible the use of military force. Matt Gaetz—whom the Washington Post dubbed the "Trumpiest Congressman"—frequently opposed the use of military force by the Trump Administration while

otherwise being one of the closest allies to the Administration.[13]

The Stanford Social Science Data Collection provides parsed speeches from the congressional record from 1873 through 2016. In order to, first, identify potentially relevant congressional speeches, the congressional record was searched for a specific keyword(s) during the time period of the crisis and the three months prior to the beginning of the crisis.[14] The total number of "matches" for all crises by all members of Congress was over one hundred thousand.

Automated text analysis tools—such as supervised machine learning—are not well suited for measuring policy prescriptions (Schub 2022), so qualitative hand-coding was employed.[15] This, however, made the manual review of all speeches beyond the resources available. In order to decrease the number of speeches to a more manageable level, speeches from key foreign policy leaders in Congress were identified for review. Such foreign policy leaders consisted of the following:

Senate:

- Senate Majority Leader
- Senate Minority Leader
- Chairman and Ranking Member of Foreign Relations Committee
- Chairman and Ranking Member of Armed Services Committee
- Chairman and Ranking Member of Intelligence Committee
- Chairman and Ranking Member of Appropriations Committee
- All Members of Senate Foreign Relations Committee
- All Members of Senate Armed Services Committee

House of Representatives:

- Speaker of the House
- House Minority Leader
- Chairman and Ranking Member of Foreign Affairs Committee
- Chairman and Ranking Member of Armed Services Committee
- Chairman and Ranking Member of Intelligence Committee
- Chairman and Ranking Member of Appropriations Committee

In total, this amounted to approximately 40 lawmakers in any given Congress, and for a total of roughly thirty thousand speeches. Notably, despite making up only 7-8% of all members of

[13] https://www.washingtonpost.com/nation/2020/01/10/gaetz-war-powers/

[14] McManus (2017) similarly includes presidential speeches not only from the time period of the dispute, but also prior to it.

[15] As described below, machine learning was utilized to predict labels for the non-hand-labeled speeches as a robustness check.

Congress, these foreign policy leaders were responsible for roughly one quarter of all speeches containing the keyword(s) during the relevant crisis time periods.

Coding Procedures

All such "matches"—thus containing both the keyword(s) and being spoken by a foreign policy leader in Congress—were then read by an individual from a team of undergraduate research assistants and hand-scored.[16] Each speech was first determined to be relevant or not to the crisis under consideration. If any of the questions below—focused on the speaker's policy preferences as well as the possible legal authority to use force—were answered in the affirmative, the speech was considered relevant. In contrast, if none of the following questions were answered positively, then the speech was considered irrelevant. Removing irrelevant speeches is important in order to avoid drawing conclusions from random noise. Schub, for example, removes from the corpus all text-portions he deems to be non-pertinent to his theory (Schub 2022, pg. 7). 73% of speeches were determined to the irrelevant, showing that speech coders were thus quite discerning in deciding whether a speech actually contained sentiment regarding the use of military force.

If a speech was determined to be relevant, coders then answered a series of binary questions related to both 1) policy preferences and 2) authorization status. The policy-preference questions were presented as follows:

Support:

- Does the speaker seemingly advocate for, or would support the use of, American military force *in general* in this crisis? (Gen^{Supp}: Yes = 1, No = 0)
- Does the speaker seemingly advocate for, or would support the use of, American *ground forces* in this crisis? (Grd^{Supp}: Yes = 1, No = 0)
- Does the speaker seemingly advocate for, or would support the use of, American *air assets* in this crisis? (Air^{Supp}: Yes = 1, No = 0)

[16] While the vast majority of speeches were labelled by a single coder, 10% of the speeches were labelled by a second coder in order to determine the intercoder reliability of the task. With a Cohen's Kappa of 0.60, there was moderate to substantial agreement. Moreover, because of the way the aggregate "congressional support score" for each crisis is calculated (it excludes speeches marked "irrelevant"), the primary worry is not if one coder judged a speech irrelevant and another did not. Rather, the largest threat was if one coder marked a speech as in favor of the use of force and another marked it as opposed to the use of force. This was found to be the case in less than 2% of speeches. Lastly, the validation tests illustrated below show that the Congressional Support Scores far better predict use of force votes than any available alternative

- Does the speaker seemingly advocate for, or would support the use of, American *naval assets* in this crisis? (Nav^{Supp}: es = 1, No = 0)
- Does the speaker seemingly advocate for, or would support the use of, American *aid or sanctions* in this crisis? (Aid^{Supp}: Yes = 1, No = 0)

Opposition:

- Does the speaker seemingly advocate against, or oppose, the use of American military force *in general* in this crisis? (Gen^{Opp}: Yes = 1, No = 0)
- Does the speaker seemingly advocate against, or oppose, the use of American *ground forces* in this crisis? (Grd^{Opp}: Yes = 1, No = 0)
- Does the speaker seemingly advocate against, or oppose, the use of American *air assets* in this crisis? (Air^{Opp}: Yes = 1, No = 0)
- Does the speaker seemingly advocate against, or oppose, the use of American *naval assets* in this crisis? (Nav^{Opp}: Yes = 1, No = 0)
- Does the speaker seemingly advocate against, or oppose, the use of American *aid or sanctions* in this crisis? (Aid^{Opp}: Yes = 1, No = 0)

Next, questions regarding the legal or constitutional authority of the President to potentially utilize force were also considered.

Legal and Constituional Authority:

- Does the speaker seemingly express the belief that the President *already has sufficient authority* to use military force in this crisis? ($Possess^{Does}$: Yes = 1, No = 0)
- Does the speaker seemingly express the belief that the President does not *already has sufficient authority* to use military force in this crisis? ($Possess^{Not}$: Yes = 1, No = 0)
- Does the speaker seemingly advocate *for granting* the President more authority to use military force in the crisis? ($Grant^{Supp}$: Yes = 1, No = 0)
- Does the speaker seemingly advocate *against granting* the President more authority to use military force in the crisis? ($Grant^{Opp}$: Yes = 1, No = 0)
- Does the speaker seemingly advocate *for restricting* the President's authority to use military force in the crisis? ($Restrict^{Supp}$: Yes = 1, No = 0)
- Does the speaker seemingly advocate *against restricting* the President's more authority to use military force in the crisis? ($Restrict^{Opp}$: Yes = 1, No = 0)

Calculating Congressional Support Scores

Once the speech codings were made, aggregate "Congressional Support Scores" could be calculated for each crisis. There are theoretically many different ways one might desire to create an aggregate score from the individual speech codings, and individual researchers are encouraged to

do so from the underlying speech codings. Described here, however, is precisely how the aggregate "Congressional Support Score" data presented here were calculated. In order to not overweight the sentiment of one speaker making repeated speeches, an Individual Support Score was calculated for each individual lawmaker who gave at least one relevant speech during the crisis. A support score for an individual lawmaker (i) in the crisis was calculated simply by creating a fraction with expressed support for the use of force across their speeches in the numerator and all expressed sentiment in favor of, or opposed to, the use of force in the denominator. Thus, the Individual Support Score of lawmaker i, ISS_i, is calculated by summing their sentiment across the relevant speeches they made during the crisis as follows:

$$ISS_i = \frac{\sum Gen_i^{Supp} + \sum Grd_i^{Supp} + \sum Air_i^{Supp} + \sum Nav_i^{Supp}}{(\sum Gen_i^{Supp} + \sum Grd_i^{Supp} + \sum Air_i^{Supp} + \sum Nav_i^{Supp}) \ + \ (\sum Gen_i^{Opp} + \sum Grd_i^{Opp} + \sum Air_i^{Opp} + \sum Nav_i^{Opp})}$$

For example, if Lawmaker A made two speeches seemingly favoring the use of force in general, they would have an Individual Support Score of 100%. In contrast, if Lawmaker B made two speeches seemingly opposed the use of force in general, they would have an Individual Support Score of 0%. And if Lawmaker C made two speeches, one favoring the use of force and one opposing the use of force, they would have an Individual Support Score of 50%. The Individual Support Score can also account for domain-dependent sentiment. For example, if Lawmaker D expressed support for bombing ISIS ($Air^{Supp} = 1$) but opposition to the reintroduction of ground forces in Iraq ($Ground^{Opp} = 1$) in a single speech, this would yield an overall Individual Support Score of 50%. Similarly, if Lawmaker E expressed support for bombing ISIS ($Air^{Supp} = 1$) in one speech, and then opposition to the reintroduction of ground forces in Iraq ($Ground^{Opp} = 1$) in another speech, this would also yield an overall Individual Support Score of 50%.

An overall Congressional Support Score was then created by averaging the Individual Support Scores of the lawmakers who gave relevant speeches during the crisis. For example, together Lawmakers A, B, C, D, and E's Individual Support Scores would yield an overall Congressional Support Score of 50%.

$$CSS = Average(ISS)$$

This simple calculation yielded an aggregate Congressional Support Score ranging between 0 (uniform opposition) and 1 (uniform support) for each crisis. This is then re-centered at 0 (and thus ranges from -0.5 to +0.5), so that scores with a positive value signify overall support in Congress, while scores below 0 suggest there is more opposition in the legislature.[17] Crises with few speakers also are "penalized" so that they are not given extreme values based on the sentiment of just a couple of legislators. Throughout the results presented below, we will require a minimum of five lawmakers to speak before assigning the raw score calculated as shown above. Crises with less than five speakers have their polarity prorated toward 0.[18] Moreover, in order to examine the sentiment of lawmakers *before* a use of force decision was concluded by the White House, speeches were limited to those made before combat commenced for crises in which the United States ended up actually utilizing military force.[19]

Description of the Data

The plot below shows the concentration of congressional attention across the world in the time period since World War II. The map seemingly suggests legislators have been quite attentive to the major trends of American foreign policy in the time period. A clear ring, for example, is seen around the Sino-Soviet bloc—reflecting congressional attention to crises over Berlin, in the Middle-East, and in East Asia during the Cold War (with large amounts of attention seen in spots Korea, Vietnam, and in the Taiwan Strait.). Cold War attention to Cuba and Central America (primarily Nicaragua and El Salvador) is also seen. Focus on former Yugolsavia—including crises in Bosnia

[17] I.e. -0.5 corresponds to uniform opposition, +0.5 corresponds to uniform support, and 0 corresponds to overall neutral (or, perfectly split) sentiment in Congress.

[18] I.e., neurtral sentiment. E.g., if 5 lawmakers in this elite subset gave speeches uniformly opposing a potential intervention, this will meet the threshold to yield the lowest possible sentiment: -0.5. If only one lawmaker spoke, however, it would be more difficult to deduce overall congressional sentiment. In this case, since there is only one out of the five required, we divide the valence by 5—yielding a congressional support score of -0.1.

There are, of course, many different cutoffs one could use. 10, 20, 40, and even 0 were all tried and yielded nearly identical results. Data utilizing this 5-member cutoff is presented here for two reasons: first, it yielded the best model fit statistically. Second, while it seems important to have *some* cutoff—it would be peculiar to have a single member speak out against an operation and then assume all of Congress was in opposition to it—it also seems intuitive that there would be diminishing returns to speaking out when several of your colleagues already have and there is a uniform consensus. Thus, if five members of this group of lawmakers speak out unanimously in favor of an operation (or in opposition to it), and there is no dissent whatsoever from any other member, it seems more reasonable to conclude that this is a well represented view.

[19] This is specifically done in order to see where members stood before the use of force commenced. This restriction can be relaxed, and similar statistical results are still yielded.

and Kosovo—is also evident, as is clear attention to the greater Middle East region, in line with major interventions in Somalia in the early 1990's and in Iraq and Afghanistan in the early 2000's.

The histograms in Figure 1, below, depict key metrics of the speech data. The first shows the total number of relevant speeches given in Congress for specific crises. Note that the histogram follows a power law distribution, with many crises having less than ten relevant speeches from the sample of lawmakers. The tail of the distribution is quite long, with some crises having over three hundred relevant floor speeches from foreign policy leaders in Congress.

Because the number of speeches, however, is likely highly correlated with the temporal length of the crisis, it is important to also consider the distribution of crisis days. Like the number of speeches, the distribution of crisis days follows a power law distribution. Roughly a third of the crises are less than fifty days long, while the majority are longer than this benchmark. In order to determine the relative magnitude of congressional sentiment targeted toward the potential use of military force in each crisis, a speeches-per-crisis-day metric is calculated for each crisis.[20] Again, a

[20]Note, this only includes speeches found to be *relevant* to whether force should or should not be used.

power law distribution is evident, with many crises having relatively less attention in congressional speeches.

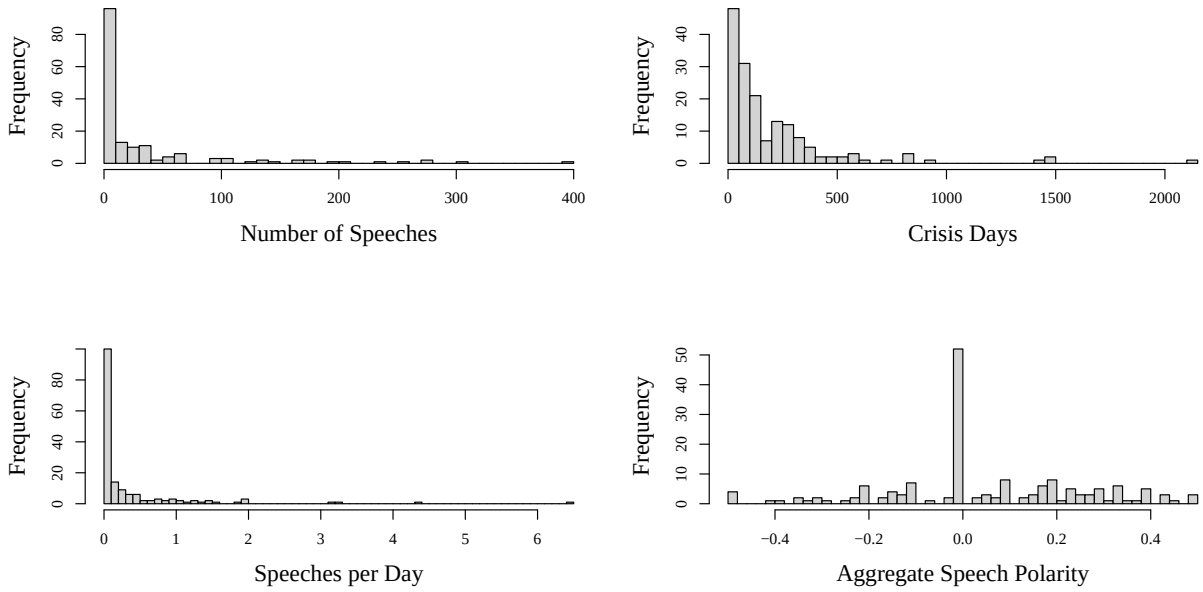


Figure 1: Histograms of Speech Data

Lastly, a histogram of the calculated congressional support scores is presented. Note again that scores are re-centered at 0, so that while a score of 0 equates to balanced or neutral overall sentiment, a negative score—with a minimum of -0.5—corresponds to aggregate congressional sentiment against the use of military force. Similarly, positive scores—with a maximum of 0.5—equate to average congressional sentiment in favor of the use of military force. The modal—and median—crisis has neutral polarity (0), primarily due to roughly one quarter of the crises having no relevant speeches from foreign policy leaders in Congress.[21] Aside from this collection of crises at 0, we see crises of varying congressional support across the range of sentiment from -0.5 all the way to +0.5. A handful of crises is shown below, and in general evinces some face validity given conventional understandings of congressional and popular will in each of these events.

As might be expected given the enormous rally-around-the-flag effect witnessed after the 9/11

[21] Additionally, some crises with relevant speeches happened to have equal amounts of positive and negative sentiment conveyed, resulting in a neutral polarity score.

Afghanistan (2001)	0.455
Korea (1950)	0.426
Iraq (2003)	0.287
Panama (1989)	0.262
Vietnam (1964)	0.257
Cuba Missile Crisis (1962)	0.245
Libya (2011)	0.200
Gulf War (1991)	0.132
Syria Red Line (2013)	0.088
Yom Kippur War (1973)	-0.125
Haiti (1994)	-0.211
Iran (2007)	-0.220
Bangladesh War (1971)	-0.500
Angola (1975)	-0.500

terrorism attack, the invasion of Afghanistan receives a very high Congressional Support Score. Similarly, the initial intervention in the Korean War in 1950 garnered significant bipartisan support—a fact recognized by politicians at the time and historians in later decades, despite the fact the Truman administration chose to not seek formal authorization from Congress (Acheson 1969, Schlesinger 1973, Blomstedt 2016). The invasion of Iraq in 2003 had strong, but far from uniform support in Congress as nearly all Republicans but only around half of democrats supported the operation. Nevertheless, George W. Bush actually had substantially greater support for the 2003 action then his father has in 1991 against the same enemy (with CSS's of 0.287 and 0.132, respectively). Notably, the elder Bush face much stronger resistance from Senate Democrats, and the formal authorization vote from Congress only passer the upper chamber by a handful of votes (Bush & Scowcroft 1999). More recently, President Obama balked at enforcing a "red line" in Syria in 2013 when congressional support (CSS of 0.088) proved to be lower than expected.

Nixon and Kissinger threatened the Soviets with intervention during the October War in 1973, but Congress was opposed to an American intervention just at its extricated itself from Vietnam—a fact well-recognized by the White House at the time (Nixon 1986). Congressional resistance to using military force in Haiti in 1994 was acknowledged by President Clinton (Clinton 2005), and memoirs from the Second George W. Bush Administration make clear the pressure they felt from Congress (Bush 2011, Gates 2015)—especially a then Chairman of the Senate Foreign Relations

Committee, Joe Biden—to not use military force against the Islamic Republic. On the far-negative end of the spectrum, possible American interventions in South Asia and Southern Africa in the 1970's were virtually uniformly opposed by lawmakers. In the case of the Bangladesh War, Nixon and Kissinger attempted to feign a willingness to use the U.S. Seventh Fleet in the Indian Ocean, but the strong resistance in Congress led Indians observers to call the bluff (Kissinger 2011*a*). In Angola 1975, Congress was so opposed to U.S. intervention that legislators even banned military aid to groups in the country (Stevenson 2007).

Partisan differences can also be examined. Plotted below is a subset of roughly fifty of the most prominent crises in postwar U.S. foreign policy, organized by level of support for the use of force ranging from uniform opposition (CSS = -0.5) to uniform support (CSS = +0.5) . The plot on the left shows both the overall Congressional Support Score in the crisis, but also scores specific to political parties in Congress. Dark grey points indicate the overall score in Congress, while blue points represent sentiment among Democrats and Red among Republicans.

Several crises have seen overwhelming opposition in Congress to the use of force, and hence little difference between parties. For example, while there exists debate in the legislature over U.S. support for the Saudi war effort in Yemen, there has been a consensus that *American* forces should not be utilized in the conflict. Similarly, while George W. Bush considered striking into Syria in 2007 (both because of its nascent nuclear weapons program and the route is served for foreign fighters pouring into Iraq) (Bush 2011), there was little support in Congress from this as Democrats were firmly opposed and little support existed even amongst congressional Republicans. In other crises, the parties were similarly strong in their support for the use of force. For example, in the successive Berlin crises under Eisenhower and Kennedy, there was virtually no difference in party support for taking a strong stand in the face of Khrushchev's ultimatums. Similarly, the initial invasion of Afghanistan received enormous support from both Republicans and Democrats. Other actions evinces a clear split in sentiment between Democrats and Republicans. For example, Bush's decision to "Surge" U.S. forces after the 2006 midterm elections received solid support amongst Republicans but strong opposition from Democrats. More recently, Trump's strike against Soleimani in early 2020 evinced hyper-partisan reactions with Democrats and Republicans being

nearly uniform in their opposition and support, respectively.

The plot on the right, in contrast, again separates lawmakers by party but focuses on whether the President was of the same party of not. Dark grey points again indicate the overall score in Congress, while light grey points correspond to lawmakers of the same party as the President and purple points represent lawmakers in the opposition party. While copartisans seem to give more support for the use of force than non-copartisans, there are clear counter-examples to this. For example, support for Obama's Surge in Afghanistan announced in late 2009 saw much stronger support among Republicans than Democrats (Woodward 2010). Non-copartisans can allow attack a President for omitting to act. Republicans famously attacked the Truman administration for the Nationalist failure in the Chinese Civil War and waged a campaign asking "Who lost China?" And while Democrats for the most part supported the Obama administration's anti-ISIS campaign launched in the second half of 2014, congressional Republicans pushed the intervention much more forcefully (taking the President to task for calling the group the "J.V. Team", and resisting Democratic efforts to pass a use of force resolution that would have seemingly limited the Presidents powers in the conflict).

More common, however, is copartisan support for the use of military force and attacks from non-copartisans. President Clinton, for example, encountered stiff opposition from Congressional Republicans toward interventions after the infamous Battle of Mogadishu—a sentiment manifesting in GOP opposition to intervening to stop the Rwandan Genocide in 1994, the occupation of Haiti later that same year, and the deployment over several thousand U.S. peacekeepers to Bosnia at the end of 1995 (Clinton 2005). Only a few years before in the "Tanker War" in the Persian Gulf during the late 1980's, copartisan Republicans had, for the most part, supported the Reagan Administrations re-flagging and escort of Kuwaiti oil tankers while Democrats had shown significant opposition to the move. More recently, Republicans strongly supported President Trump's "maximum pressure" campaign against Iran—including his risky targeted killing of Iranian General Qasem Soleimani in early 2020—while Democrats opposed starting a war with Iran.

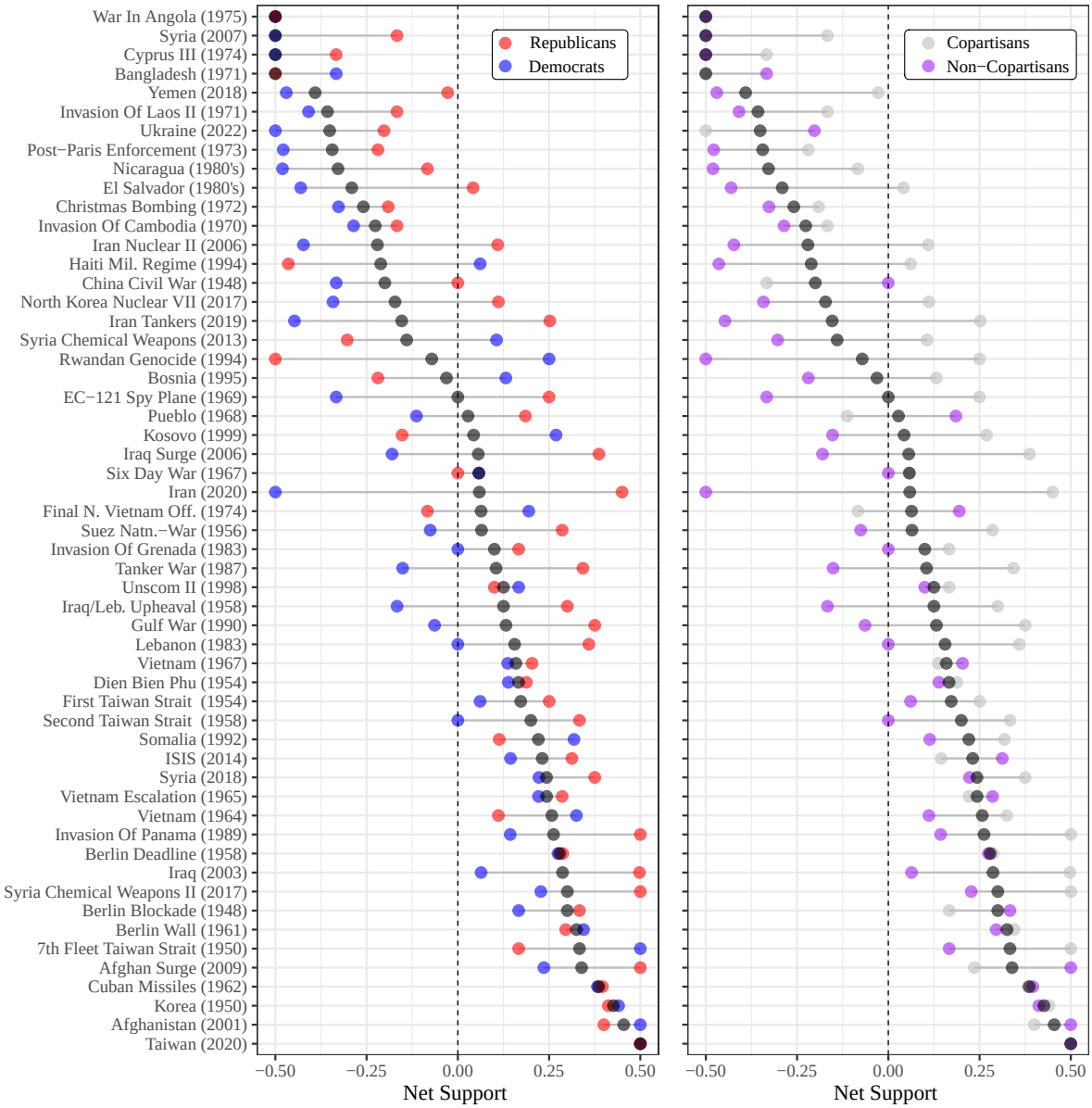


Figure 2: Examples of Crises—Overall Congressional Support Score, and Party Breakdown

Validation

While the plots and accompanying narrative above provide some face validity to the scores, there are obvious concerns about how accurate a measure of congressional sentiment this method might actually produce. Several drawbacks clearly make this method imperfect: speeches from less than 10% of congressional members are even considered; speeches from Senators are disproportionately represented in the sample; many members of Congress give no relevant speech at all. Validation is thus key to increasing our confidence in the ability of this method to sufficiently measure congressional sentiment over the use of military force. Two strategies of validation are presented here: first, a test of face validity is presented by comparing Congress Support Scores taken during successive "crises" in the Vietnam War with floor votes and public opinion polls during the conflict. Second, a more systematic test of the scores is presented by comparing them to over sixty roll call votes and public opinion polls related to the use of military force in crises since World War II. The evidence presented suggests that Congressional Support Scores adequately capture congressional sentiment in these crises, and far outperform any available alternative measurement, including proxies commonly utilized in the literature.

The Vietnam War

Examining the Congressional Support Score data in the context of the Vietnam War is useful because the well-known experience in Vietnam provides a helpful background context against which to test the validity of the measurement technique. Depicted on the plot below is the calculated Congressional Support Scores over a series of crises ranging from the 1964 Gulf of Tonkin Incident to the signing of the Paris Peace Accords in January 1973 and the passage of the Cooper-Church Amendment in the summer of 1973 prohibiting the use of the American military in Southeast Asia. The solid line represents the Congressional Support Scores yielded in these successive crises. The overall trend is familiar: support for the war was high at the beginning of the conflict, and gradually dropped until the end of the war in 1973. As is well recognized, the drop from aggregate support to opposition occurs around the Tet Offensive in 1968 (Rusk, Rusk & Papp 1991).

As a test of convergent validity, the estimated Congressional Support Scores are plotted in

comparison to measures of public support for the war. Specifically, this data comes from a series of Gallup polls taken from 1965 to 1973 (twenty-three in total) asking respondents "In view of the developments since we entered the fighting in Vietnam, do you think the U.S. made a mistake sending troops to fight in Vietnam?".[22] This particular times series of data has been utilized widely in the literature examining public support for the Vietnam War (Gelb & Betts 2016, Lunch & Sperlich 1979, e.g.).[23]

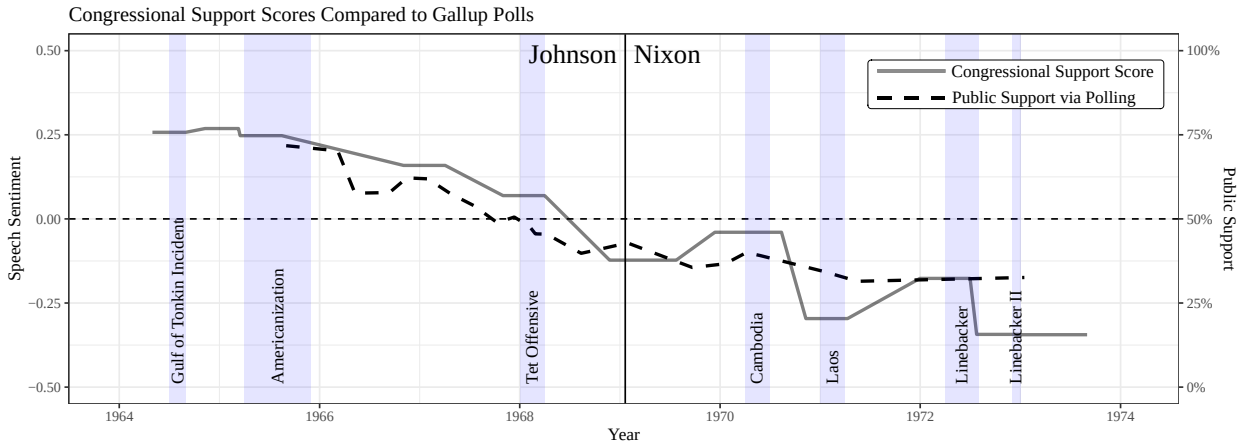


Figure 3: Congressional Support Scores Compared to Gallup Polls of Public Support

As can be seen from the plot, the two trend lines closely follow each other. While support in the public at the beginning of the Americanization of the conflict in 1965 was relatively higher (around 75%), this gradually faded over time. By the end of American involvement in the conflict in 1973, support had dropped to around 25%. This closely follows the trend illustrated by the Congressional Support Scores, and thus suggests the calculated proxy for sentiment amongst legislators is performing as expected.

Another test of face validity of is to compare the congressional support scores to vote shares in Congress for key votes during the Vietnam War. Plotted below is, again, the Congressional Support Score estimates—with the general trend in dark gray and estimates separated by party in red (Republicans) and blue (Democrats). Key votes are shown as black dots. These votes included

[22] https://news.gallup.com/poll/2299/americans-look-back-vietnam-war.aspx

[23] Just as the Congressional Support Score metric is calculated by dividing positive sentiment by all sentiment expressed (positive or negative), the public support metric plotted here is calculated by dividing the number of respondents who disagreed that the Vietnam War was a mistake (i.e., those who supported the war) by all respondents who either agreed or disagreed with the statement.

the most prominent resolutions passed in the conflict, including the 1964 Gulf of Tonkin Resolution, a 1965 defense supplemental specifically used by the administration as a proxy for congressional support, votes on the Cooper-Church Amendment after the Cambodian Incursion in the Spring of 1970, the Mansfield Amendment passed by the Senate in 1971, and votes in both Houses on the 1973 Case-Church Amendment that finally cut off all funding for operations in Southeast Asia by U.S. forces.[24] Like public opinion, we would expect a fair measure of congressional sentiment to vary alongside use-of-force vote shares.

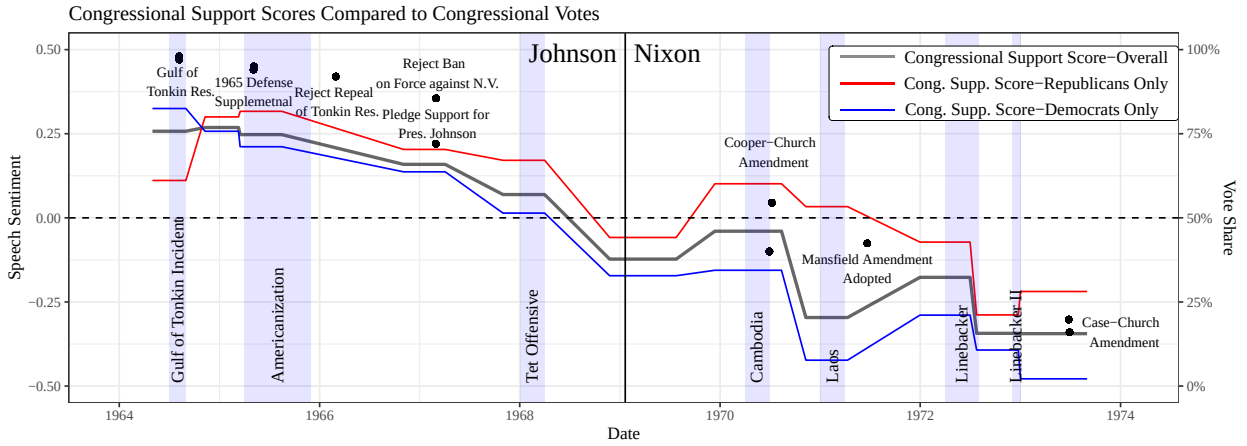


Figure 4: Level of Force Employed by Support in Congress for Use of Military Force in Crisis

Similar to the public opinion data, the vote shares supporting the war in Vietnam start quite high but decline sharply over time. The trend shown by the Congressional Support Score data closely follows this. Note that while the point estimates do not perfectly align—for example, vote shares seem consistently higher than the Congressional Support Score yielded in the first three years of the war—it is not necessarily the case that this reflects "error" in the Congressional Support Score estimates. Instead, there is a strong argument that the Congressional Support Score estimates potentially better reflect congressional sentiment than the vote share data. Specifically, it is well recognized in the historiography of the Vietnam War than many of those lawmakers who voted in favor of the Gulf of Tonkin Resolution in August 1964 had serious reservations about getting involved in a war in Southeast Asia (Gibbons 2014). This lack of uniform support for the

[24]Two other key votes identified by Congressional Quarterly Almanac—rejections of proposals in 1966 and 1967 to repeal the Gulf of Tonkin Resolution and to ban the use of force in North Vietnam, respectively.

war is much better captured by the Congressional Support Score data than by looking at simple vote shares, because the speech data utilized to estimates these scores takes account of reservations expressed by the speaker. Thus, while merely looking at vote shares—especially at the beginning of the war—paints a misleadingly high picture of congressional sentiment of the war, the Congressional Support Scores arguably more actually capture the feelings in Congress at the time.

The Congressional Support Score estimates, moreover, closely follow the vote shares exhibited in key votes later in the war, including votes on the 1970 Cooper-Church amendment to ban the use of American ground forces in Cambodia and the 1971 Mansfield Amendment passed by the Senate—a non-binding but high-profile vote declaring it to be the policy of the United States to "terminate at the earliest practicable date all U.S. military operations" in Southeast Asia. The 1973 Case-Church Amendment formally barred all funding of U.S. military operations Vietnam, and its overwhelming passage corresponds closely to the highly negative sentiment expressed in Congress toward the use of force at the time.

The party trend lines, as well, conform to the conventional wisdom of the conflict: the parties were relatively similar in their support for the conflict (with hawkish Republicans supporting the war under Lyndon Johnson) until Richard Nixon came into office. At this point, support from Democrats quickly evaporated as the lack of a copartisan in office allowed them to return to their more dovish predispositions and call for an end to the conflict.[25] Republicans, in contrast, stayed more supportive of the war as a fellow Republican now occupied the White House. Thus, we see a large divergence between Republicans and Democrats between 1970 and 1972. Republicans were, on balance, supportive of the 1970 Cambodian Incursion while Democrats opposed it, for example. By 1973 and the signing of the Paris Peace Accords, however, members of both parties were highly opposed to American re-intervention in the conflict (Kissinger 2011*b*).

[25]Senator Robert Kennedy—seen as the likely Democratic nominee for President in 1968 before his assassination the same year—was highly critical of the Johnson Administration's war effort, and Senator George McGovern—the 1972 presidential nominee for the party—was a long critic of the Vietnam War.

Systematic Tests: Votes and Polls

The underlying benchmarks utilized in the Vietnam example—roll call votes and public opinion polls—can also be utilized more systemically across the crises in the dataset. As noted, roll call votes are relatively rare. While there are approximately 65 roll call votes related to the use of force in the postwar period, this includes both houses of Congress and thus covers only around 20% of all crises. Separately, public opinion polls were identified for 67 of the crises (35% of the dataset). The Congressional Support Score data and alternative proxies one might consider for congressional sentiment toward the use of force are then tested against these benchmarks.

Use-of-Force Votes

A way to more systematically "test" how well this approach works is to compare the congressional support score yielded by the speech data to the handful of roll call votes we do have regarding use of military force decisions. How well these scores predict vote outcomes can then be examined both in an absolute sense and in relationship to other commonly used measures.

Each of the figures below plots on the X-axis the percent of lawmakers in favor of the use of force in sixty-five roll call votes in postwar crises. The votes include, for example, authorizations for the use of military force in the early Cold War (Formosa, the Middle East, Cuba, and Southeast Asia), and more recent AUMF's such as those passed prior to the Gulf War and the 2001 and 2002 AUMF's. Also included are prominent votes against the use of force, such as the 1973 Case-Church Amendment banning the use of military force in Southeast Asia.[26]

The Y-axis for each of the nine figures, in contrast, varies and consists of different possible metrics used as a proxy for congressional support for the use of military force. An optimally performing measure of congressional support would (theoretically) provide a trend line from the bottom left of the square to the top right (represented by the blue dashed line in each plot), with observations close to the line. Given the continuous nature of the variables on each plot, ordinary least squares regression is used to determine how well each measure predicts use of force vote share. A black trend line for the observations is shown in each figure, and a simple R-squared metric is

[26]For resolutions opposing the use of force, the votes shares are, of course, "flipped".

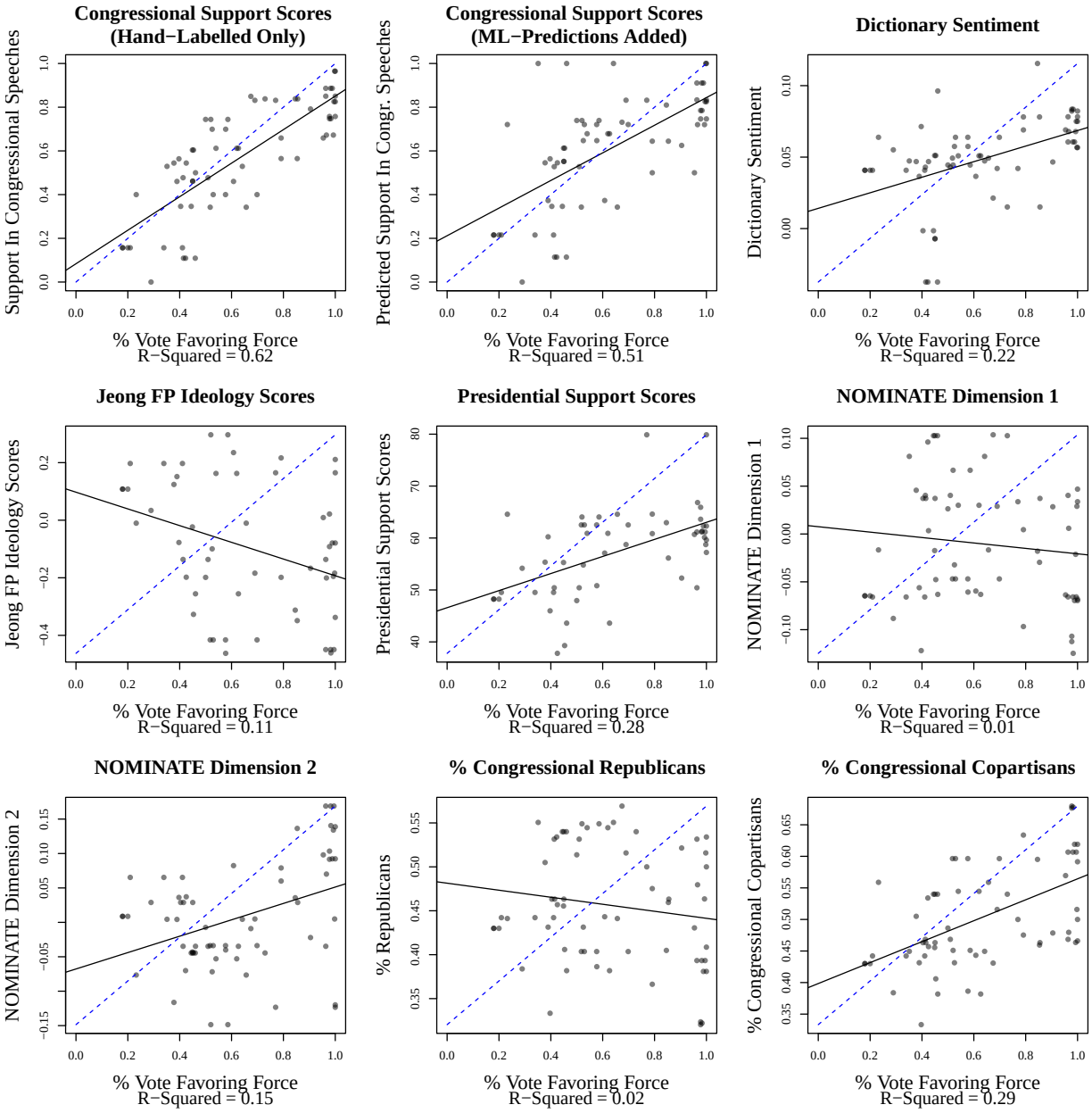


Figure 5: Accuracy of Competing Measures in Predicting Use of Military Force Votes

presented below each plot. The R-squared is a useful measure of how well each variable performs as a proxy for congressional support.

The first plot in the upper left corner shows how well the "congressional support scores" from the speech data predict vote share in these key war votes. While predictions are clearly imperfect,[27] the R-squared of 0.62 is relatively high. The black trend line is quite close to the blue dashed line representing a "perfect" trend line, and the observations are relatively close to the line. More importantly, looking either at the plot or the R-squared, the measure clearly outperforms each of its potential competitors.

The second plot similarly utilizes congressional support scores calculated as described above, but here also includes predicted classifications from a supervised machine learning model that was utilized to predict labels for the approximately 75% of congressional speeches that were made by legislators outside the group that were hand-labelled. These members outside the elite group of foreign policy leaders have much less influence over foreign policy, but given their sheer numbers, it is not inconceivable that their opinions could alter the findings presented. To this end, a series of Transformer models from the HuggingFace library were trained and tested on the hand-labelled speeches. After hyperparameter tuning, the models were compared for out-of-sample predictive accuracy. Ultimately, a DistilRoBERTa base model was found to maximize predictive accuracy while minimizing computation time. The speech classifications (i.e., whether a speech was in favor of or opposed to the use of force, or whether it was irrelevant) from the model were then used as predicted labels for the non-hand-labelled speeches. Aggregate congressional support scores were then calculated using these labels. Note that others have recently shown (Schub 2022, for example) supervised and dictionary methods are not optimal for determining policy prescriptions—the very task explored here—and that hand-labelling is more appropriate (Schub 2022, pg. 8-10).[28] Nonetheless, while this measure does not perform as well as the scores from the hand-labelled data

[27]Note, however, that even when the sentiment score does not closely match the vote share, it is not necessarily the case that the sentiment score is the metric missing the true congressional opinion at the time. For example, it is well understood that several of the members of Congress who voted in favor of the 1964 Gulf of Tonkin Resolution had serious reservations. The congressional support score shows this, while the actual vote does not.

[28]Nonetheless, as discussed below, even when including machine-predicted labels in calculating a "congressional support score" for each crisis as a robustness check, the statistical results in the models presented below are nearly identical.

alone, they are a clear second-best with an R-squared of 0.46.

The third plot (top right) utilizes the predicted sentiment of speeches from the popular "sentimentr" package in R. This utilizes a standard dictionary method to create an overall polarity score for a speech based off of positive-sentiment and negative-sentiment words and phrases. Standard sentiment packages and dictionaries such as these have a difficult time deciphering support and opposition to the use of military force because politicians advocating for military action often do not utilize "happy" words. This proxy performs far worse than the first two, with a much lower R-squared (0.22). Others have utilized custom dictionaries to good effect in the use of force context (McManus 2017), but for a different aim. McManus uses the dictionary method to measure statements of resolve from Presidents by utilizing both a customized dictionary from prior work (Wood 2012) and additional terms added for her specific purpose. The major difference between presidential rhetoric and congressional speech, however, is that Presidents in a crisis tend to avoid conveying a lack of resolve publicly—regardless of their private thoughts or intentions (Fearon 1995). Thus, because the rhetoric tends to only go in one direction, McManus can reasonably argue that the word "fail" will much more likely be used by the President to say "we will not fail" than "we will fail" (McManus 2017, pg. 211).[29] Members of Congress, on the other hand, are much more willing to speak out against the use of force. The word "fail" in a congressional speech very well could be used in a context conveying opposition to the use of force. It is partially for this reason that others have recently argued dictionary methods are not well suited to determining policy prescriptions in speech data (Schub 2022).

Foreign policy ideal point estimates from Jeong are tested in the next figure (middle row, left) (Jeong 2018, Jeong & Quirk 2019). These scores are available for each member of Congress from 1945-2010, and are intended to represent the hawkishness or dovishness of each member. The median score can be used to represent the hawkishness of the Congress as a whole, and thus plausibly represent congressional support for the use of military force in a crisis occurring during that congressional session. This measure turns out to be a poor proxy, however. Not only is the R-squared quite modest—at 9% compared with the 62% yielded by the speech data—but the

[29] "For example, Presidents often say, 'we will not fail,' but almost never say, 'We will fail.' Therefore, 'fail' is considered to be a word associated with resolve." (McManus 2017, pg. 211)

predicted relationship is actually in the wrong direction. Here, the more "hawkish" Congresses are predicted to be more opposed to the use of force.

A series of other off-the-shelf metrics are similarly tested, including presidential support scores (Lewis, Poole, Rosenthal, Boche, Rudkin & Sonnet 2022), and NOMINATE dimensions 1 and 2 (Lewis et al. 2022).[30] Of these, presidential support scores perform the best, yielding an R^2 of 0.26, and in the correct direction. Nominate Dimension 1, however, exhibits little relationship with supporting the use of force in these key war votes, while Dimension 2 yields an R^2 of only 0.16. Again, the "congressional support score" estimates from the speech data thus yield a better proxy of vote share in actual use of force votes than any of these possible alternatives.

Lastly, we consider two key metrics that have been used in the political science literature as deliberate proxies for congressional support for the use of force. The first is the partisan composition of Congress, in terms of Democrats vs. Republicans. The theory behind this measure is that Republican members of Congress have a reputation for being more hawkish than their counterparts across the aisle.[31] Arguably, using the percent of Congress made up of Republicans can be viewed as a proxy for how supportive Congress is for the use of military force (McManus 2017). Interestingly, however, with an R^2 of 0.02 (and in the wrong direction) the partisan composition of Congress seems to exhibit little relationship with congressional support for the use of force revealed in roll call votes. Others have recently noted that while Democrats have a reputation of being more dovish than their Republican counterparts, the actual empirical evidence of this claim is limited (Kertzer, Brooks & Brooks 2021).

Second, others have focused not on the absolute partisan composition of Congress, but instead on the amount of copartisans a President has on the Hill (Howell & Pevehouse 2007, Kriner 2010).[32] Howell and Pevehouse find that a larger number of copartisans in Congress predicts an increased likelihood of initiating major uses of force in the postwar period (2007).[33] Kriner similarly finds

[30] Again, the median scores for Congress as a whole at that time are utilized.

[31] As but one example, consider the Afghanistan "surge" early in the Obama Administration. Obama anticipated "There won't be universal applause on Capitol Hill", because "Everyone knew that the Democrats were going to be the biggest naysayers, and the Republicans the biggest supporters," (Woodward 2010, pg. 326).

[32] Recent research suggests both copartisanship with the President and Republican Party identification separately predict support for the use of force by members of Congress (Böller 2021).

[33] Gowa, however, finds no relationship between divided government and proclivity to use military force (Gowa 1998).

the duration of American uses of force are predicted by the percent of the legislature composed of lawmakers in the President's party (2010). The proportion of Congress made up of presidential copartisans does yield a positive relationship with the support shown for the use of force in roll call votes, and this measure seemingly outperforms each of its off-the-shelf competitors. Nonetheless, with an R^2 of 0.29 this still leaves much of the variance unexplained and suffers from many of the drawbacks highlighted above.

Public Opinion Polls

A similar test of performance can be undertaken by comparing the potential proxies to public support polls taken during the crises. While congressional sentiment need not necessarily align with that in the public at large, institutional incentives for lawmakers mean that the two measures should correlate.

The plots below test the same nine proxies for their performance in predicting public support for military intervention in each of sixty-seven crises, ranging from the 1948 Berlin Airlift to the 2022 crisis in Ukraine. Note that while some of these crises are also covered by the roll call votes explored above, the majority are not. Thus, this is a substantially different test of performance than that analyzed above.

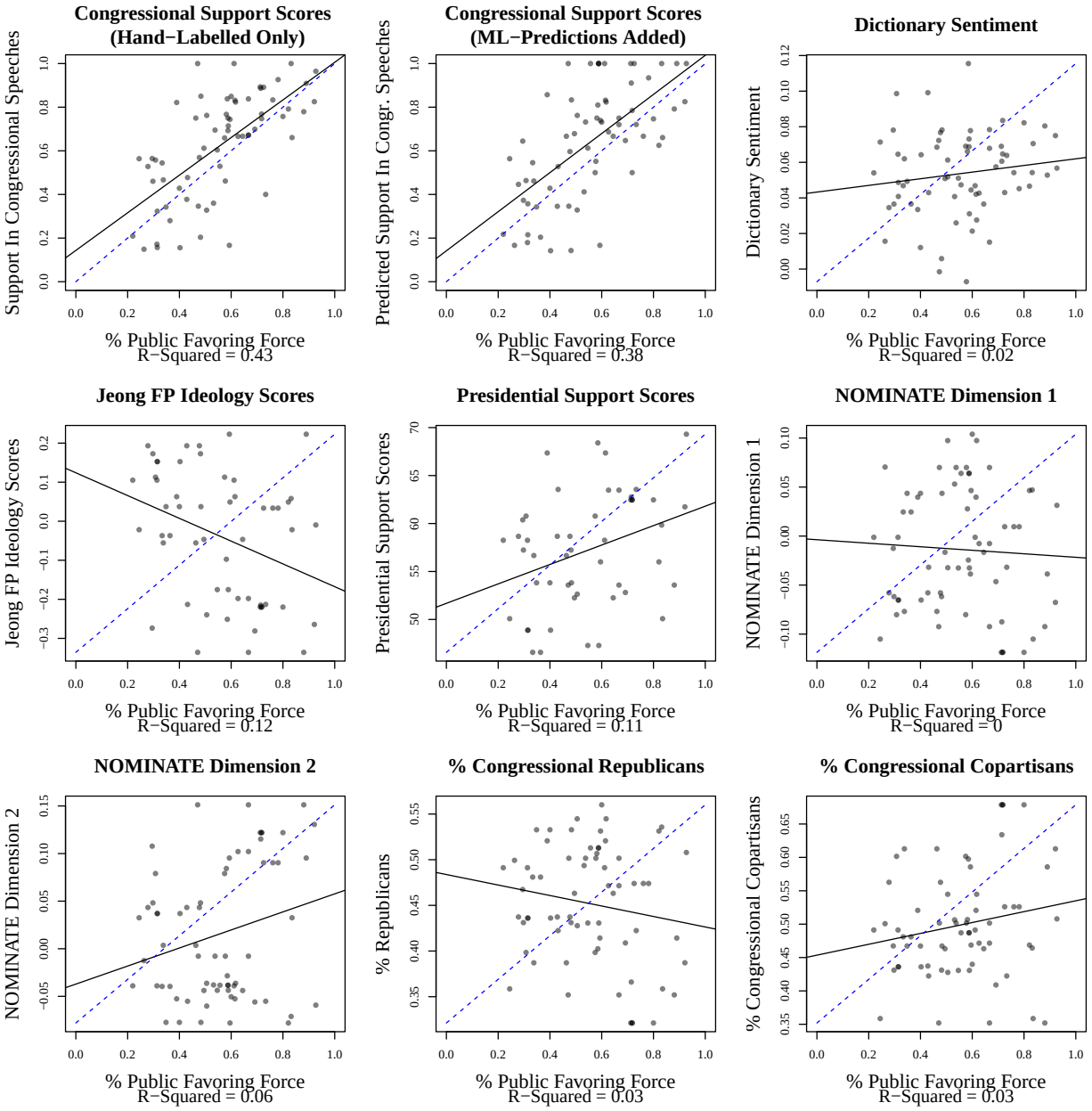


Figure 6: Accuracy of Competing Measures in Predicting Public Support for the Use of Military Force

In general, each of the proxies has a more difficult time predicting public opinion than congressional roll call votes—although this is not surprising given that congressional sentiment need not perfectly align with popular sentiment, and roll call votes would thus likely be a better proxy than public opinion polls. Nevertheless, the relative performance of each of the nine proxies reflects a

similar pattern as the prior figure: the Congressional Support Scores far outperform any of the potential other measures.

The scores derived from the hand-labelled data of Foreign Policy Leaders in Congress yields an R-squared of 0.43, while those yielded by additionally taking account of the machine-predicted scores on the non-hand-labelled speeches exhibits similar performance with an R-squared of 0.38. Of the other seven measures, the next closest performance is given by Presidential Support Scores with a R-squared of 0.11. The rest of the available proxies either barely correlate with public opinion—dictionary based sentiment analysis (R-squared of 0.02), NOMINATE Dimension 1 (R-squared of 0.00), NOMINATE Dimension 2 (R-squared of 0.06), % Copartisans (R-squared of 0.03)—or correlate negatively with public opinion votes. Foreign policy ideology scores and the percent of Republicans in Congress both show a negative correlation with public support for the use of military force in these historical crises.

In sum, each of the potential off-the-shelf proxies for congressional support for the use of force is substantially outperformed by the scores derived from the speech data—regardless of whether we use roll call votes or public opinion polls as our baseline to measure proxy performance. While these "congressional support scores" are imperfect, they are far better than any other measure currently available.

Congressional Support and Presidential Willingness to Use Force

The figure below depicts U.S.-relevant crises plotted by the sentiment expressed in Congress for the use of force versus the amount of force actually employed.[34] The Y-axis ranges from crises in which the U.S. took no action whatsoever (or, when it limited its reaction to mere diplomatic protest) to full scale war involving more than one thousand U.S. combat fatalities. Crises above the horizontal dashed line represent conflicts in which American forces engaged in actual combat,

[34]Note that if the ICB dataset split a war into multiple crises—the Korean and Vietnam Wars, for example—only the first crisis is included on the plot. In the case of Korea, congressional support remained quite strong through the end of the conflict, while for Vietnam congressional support slowly eroded from strong support to strong opposition. Consistent with the theory, as opposition in Congress grew, crisis responses by American Presidents became more and more tailored to avoid American casualties (e.g., the 1971 Laos intervention or the Linebacker and Linebacker II operations in 1972 that lacked American "boots-on-the-ground").

while those below the line consist of crises in which American action was limited to that short of armed conflict. Actions short of armed conflict are arranged in order of escalating risk of American casualties.

First, note that lack of observations in the upper left quadrant of the figure. This quadrant is where we should see evidence of the Imperial Presidency—i.e., Congress expressing opposition to the use of force, but the Commander-in-Chief choosing to use it anyway. Instead, we see no evidence of this when it comes to initial use of force decisions.[35] Instead, we see that—almost without exception—American engagement in combat was only undertaken when it was supported by a majority in Congress *even if members of Congress did not formally vote on the use of force.* The 1999 Kosovo and 2011 Libya interventions are good illustrations of this. While neither received formal, legally binding, approval from Congress, both had clear bipartisan support. The Korean War—notably the *only* major use of force undertaken by a President without formal approval—had enormous congressional support across political parties.

When facing a lack of congressional support, in contrast, Presidents appear highly reluctant to engage in operations anticipated to create American combat fatalities. While Clinton deployed military force in two crises with little congressional support—the 1994 Haiti intervention clearly contradicted the will of Congress, while the post-Dayton Accords deployment to Bosnia had divided support (Schultz 2003)—neither yielded a single American fatality, nor even saw American

[35]The place where we do find more evidence of it is in a major war that Congress initially supported but then turned against—e.g., Vietnam and Iraq. But because we are only considering the decision to enter conflict here, decision-points made after war has already begun are excluded. The Vietnam sentiment shown in the plot above includes sentiment around the Gulf of Tonkin incident and in the first seven months of 1965. Congressional support for the war was quite high—overwhelmingly in favor, albeit with a substantial minority in strong opposition—through the initial escalations in 1965 and beyond (Gelb & Betts 2016). Once Congress turned against the war by the time of Nixon's inauguration in 1969, we see Nixon continue the fight rather than simply withdraw—but even here there is strong evidence that Nixon's willingness to sustain casualties was conditioned by congressional support or opposition. As Congress turned against the war, Nixon rapidly reduced the number of American soldiers in Southeast Asia and consequently reduced American casualties immensely. Nixon created an uproar when he launched an incursion in Cambodia in 1970, but the action was actually authorized by the Gulf of Tonkin Resolution (Ely 1995) and was aimed at *reducing* American casualties (Kissinger 2011*a*). Notably, when faced with a similar situation in Laos the next year—this time, however, facing a congressional ban on ground troops instead of the arguable authorization he had for Cambodia—Nixon consciously avoided a similar operation. Some of Nixon's most infamous actions—Operations Linebacker I & II—were *air* operations aimed at minimizing U.S. casualties. Lastly, likely the clearest evidence of all that even Nixon was not the Imperial President he claimed to be was his inability to act against North Vietnamese violations of the 1973 Paris Peace Accords (Kissinger 2011*a*). Nixon consciously sought to convey an image of an Imperial President, but it should be recognized he had strong incentives to do so (Jervis 1970). See Chapter 5 for an extensive discussion of the Vietnam War.

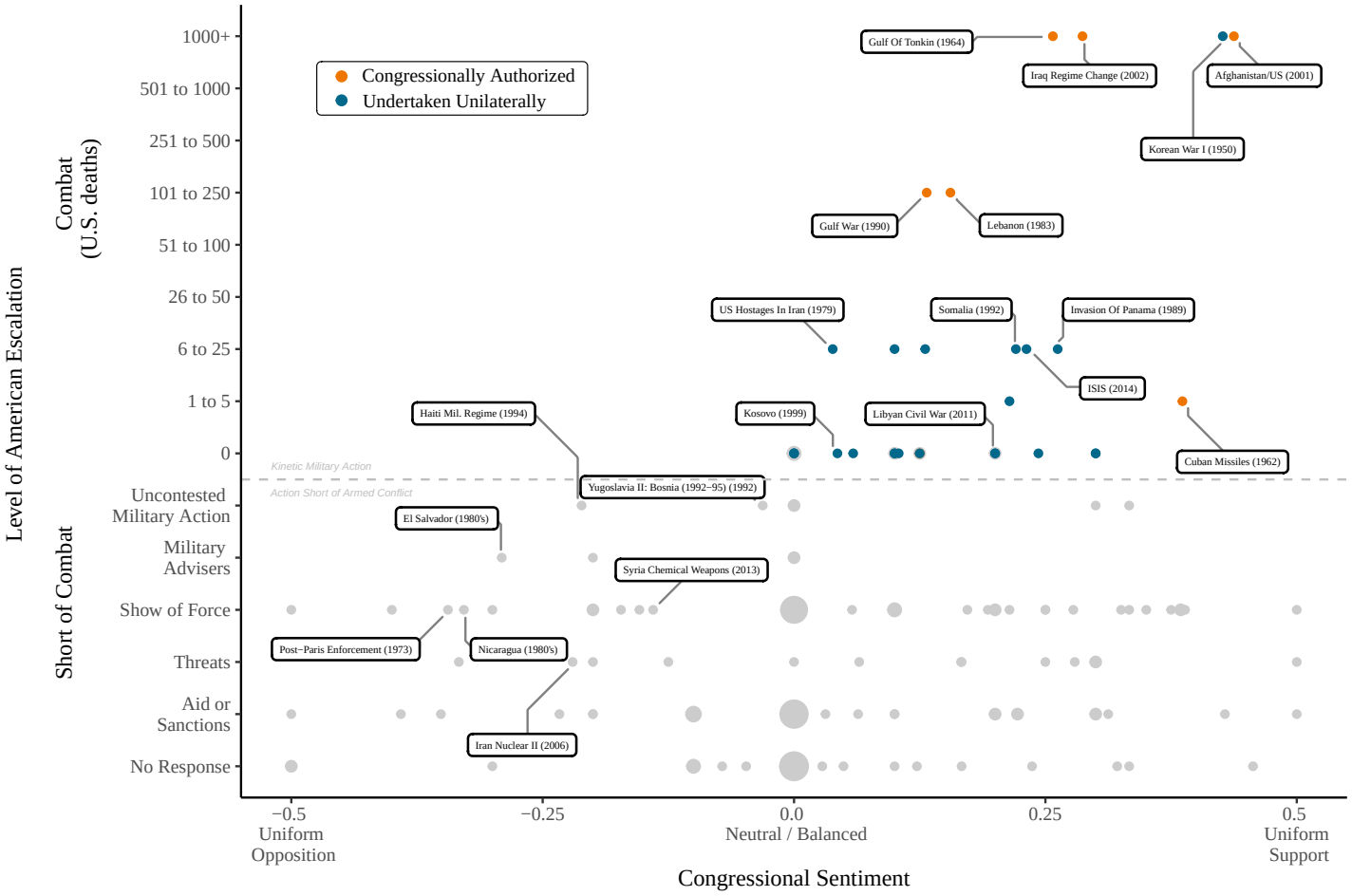


Figure 7: Level of Force Employed by Support in Congress for Use of Military Force in Crisis

troops actually engaged in combat.[36] Consider, in contrast, the administration's quick pullout from Somalia once support evaporated in the legislature after the "Black Hawk Down" incident, or the White House's omission to intervene in the Rwandan genocide due to anticipated resistance from Congress (Clinton 2005).

Conclusion

Democratic control of use of military force decisions is of great concern to academics, politicians, and everyday citizens. Accurately assessing the extent of such control, however, requires an adequate estimate of sentiment in the legislature toward the use of force in individual crises.

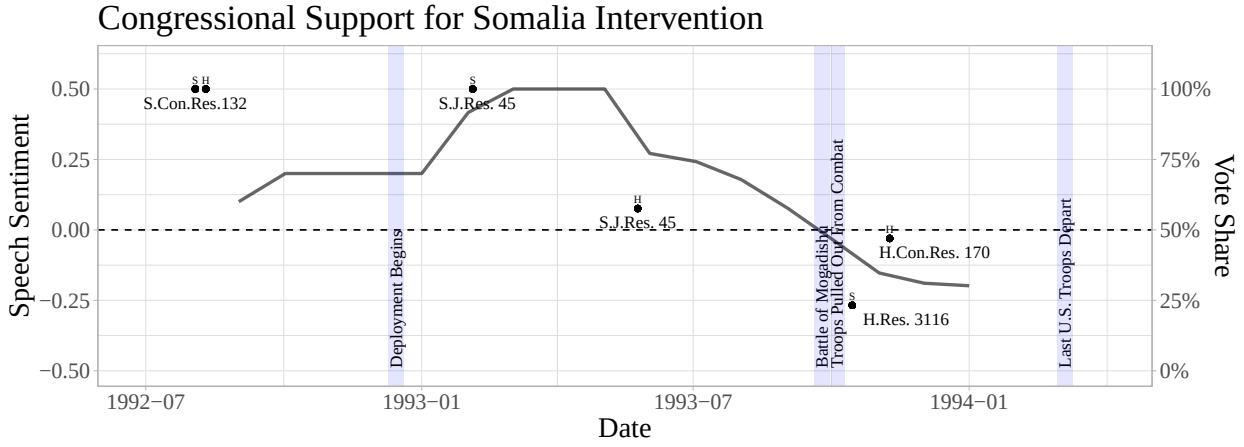
There are, moreover, several other possible uses of such data. Political science, for example, has several prominent theories related to the effect of opposition party signalling to crisis credibility (Schultz 2001, Ramsay 2004), but such theories have been difficult to test without systematic evidence of the position taken by such parties. Such theories would be relatively straightforward to test with this data. Scholars of international relations and American politics, as well, have considered the effect of party polarization on U.S. foreign policy (Schultz 2017, Myrick 2021, Jeong & Quirk 2019)—including a recent special issue of a prominent journal (Friedrichs & Tama 2022)—and yet have often had to rely on the very few use-of-force votes available.

[36] In the case of Bosnia, the Clinton administration notably refused to deploy ground troops during the actual fighting. Instead, it delayed any deployment until after the Dayton Peace Accords in 1995. Moreover, while House Republicans were mostly against the deployment, it had significant support in the Senate—most notably from soon-to-be Republican presidential candidate Bob Dole (Hendrickson 2002).

Appendix

Dynamic Congress Sentiment

Because the speech data it coded at the individual speech level, it is also possible to track changing sentiment in Congress *within* a crisis. For example, the plot shows congressional support by month for the use of U.S. military force in Somalia in the early 1990's. These dynamic congressional support scores are calculates the same way as described above, but with the underlying speeches limited to those in a specific temporal subset of the crisis. Here, the score is shown in monthly intervals, inside a six-moth window. Thus, in any given month only those speeches made in the past six months are considered. This allows for a "rolling average" of congressional sentiment during the crisis.



Plotted above is not only this dynamic congressional sentiment but also key votes taken by legislators. Both houses unanimously passed S. Con. Res. 132 in 1992, which "call[ed] upon the international community, through the United Nations...to immediately expand its relief efforts in Somalia" and "urge[d] the President to work with the United Nations Security Council to deploy these security guards immediately." A departing President Bush deployed the first troops shortly before Clinton's inauguration, and soon after the new president came into office both Houses of Congress passed versions of S.J. Res. 45 authorizing the use of force in Somalia, although the two versions were never reconciled. By the time the time the Battle of Mogadishu—"Black Hawk Down"—occurred in October 1993, however, congressional support for the operation had decreased

enormously. In the aftermath of the disaster, Congress effectively legislated an American pullout from the war torn country. This rapid change in congressional sentiment within the crisis is clearly seen in the dynamic congressional support score, as initial support for the intervention decreased substantially in late 1993.

References

Acheson, Dean. 1969. *Present at the Creation: My Years in the State Department.* New York: W.W. Norton.

Bendix, William & Gyung-Ho Jeong. 2022. "Beyond party: ideological convictions and foreign policy conflicts in the US congress." *International Politics* .
URL: *https://link.springer.com/10.1057/s41311-021-00370-9*

Blomstedt, Larry. 2016. *Truman, Congress, and Korea: The Politics of America's First Undeclared War.* University Press of Kentucky.
URL: *https://www.jstor.org/stable/j.ctt189ttn6*

Burns, Sarah. 2019. *The politics of war powers: the theory & history of Presidential unilateralism.* American political thought Lawrence: University Press of Kansas.

Bush, George H.W. & Brent Scowcroft. 1999. *A World Transformed.* 1. vintage books ed ed. New York: Vintage Books. OCLC: 247367184.

Bush, George W. 2011. *Decision points.* New York: Broadway Paperbacks. OCLC: 760311339.

Böller, Florian. 2021. "Vandenberg Vanished: US Congress and the Politicisation of Military Interventions." *Parliamentary Affairs* p. gsab023.
URL: *https://academic.oup.com/pa/advance-article/doi/10.1093/pa/gsab023/6199161*

Christenson, Dino & Douglas L. Kriner. 2020. *The Myth of the Imperial Presidency: How Public Opinion Checks the Unilateral Executive.* Chicago ; London: The University of Chicago Press.

Christenson, Dino P. & Douglas L. Kriner. 2017. "Constitutional Qualms or Politics as Usual? The Factors Shaping Public Support for Unilateral Action: CONSTITUTIONAL QUALMS OR POLITICS AS USUAL?" *American Journal of Political Science* 61(2):335–349.
URL: *http://doi.wiley.com/10.1111/ajps.12262*

Clinton, Bill. 2005. *My life.* 1st vintage books ed ed. New York: Vintage Books. OCLC: ocm60503998.

Cox, Henry Bartholomew, ed. 1984. *War, Foreign Affairs and Constitutional Power.* Vol. 2 Cambridge, Mass: Ballinger.

Ely, John Hart. 1995. *War and Responsibility: Constitutional Lessons of Vietnam and Its Aftermath.* Princeton Paperbacks 4. print., 1. paperback print ed. Princeton, NJ: Princeton Univ. Press. OCLC: 832552066.

Esper, Mark T. 2022. *A sacred oath: memoirs of a Secretary of Defense during extraordinary times.* OCLC: 1313801198.

Fearon, James D. 1994. "Domestic Political Audiences and the Escalation of International Disputes." *American Political Science Review* 88(3):577–592.

Fearon, James D. 1995. "Rationalist Explanations for War." *International Organization* 49(3):379–414.

Fisher, Louis. 2013. *Presidential war power*. Third edition, revised ed. Lawrence, Kansas: University Press of Kansas.

Ford, Gerald R. 1979. *A time to heal: the autobiography of Gerald R. Ford*. 1st ed ed. New York: Harper & Row.

Frederick, Bryan, Jennifer Kavanagh, Stephanie Pezard, Alexandra Stark, Nathan Chandler, James Hoobler & Jooeun Kim. 2021. *Assessing Trade-Offs in U.S. Military Intervention Decisions: Whether, When, and with What Size Force to Intervene*. Santa Monica, CA: RAND Corporation.

Friedrichs, Gordon M. & Jordan Tama. 2022. "Polarization and US foreign policy: key debates and new findings." *International Politics* 59(5):767–785.
URL: *https://link.springer.com/10.1057/s41311-022-00381-0*

Gates, Robert Michael. 2015. *Duty: memoirs of a Secretary at war*. First vintage books edition ed. New York: Vintage Books, a division of Penguin Random House LLC.

Gelb, Leslie H. & Richard K. Betts. 2016. *The irony of Vietnam: the system worked*. Second edition ed. Washington, D.C: Brookings Institution Press.

Gibbons, William Conrad. 2014. *The U.S. Government and the Vietnam War: Executive and Legislative Roles and Relationships, Part III: 1965-1966*. Princeton: Princeton University Press.
URL: *https://muse.jhu.edu/pub/267/monograph/book/34425*

Gowa, Joanne. 1998. "Politics at the Water's Edge: Parties, Voters, and the Use of Force Abroad." *International Organization* 52(2):307–324.

Hendrickson, Ryan C. 2002. *The Clinton wars: the constitution, Congress, and war powers*. 1st ed ed. Nashville: Vanderbilt University Press.

Howell, William G. & Jon C. Pevehouse. 2005. "Presidents, Congress, and the Use of Force." *International Organization* 59(01).
URL: *http://www.journals.cambridge.org/abstract*$_S$0020818305050034

Howell, William G. & Jon C. Pevehouse. 2007. *While Dangers Gather: Congressional Checks on Presidential War Powers*. Princeton, N.J.: Princeton University Press.

Jeong, Gyung-Ho. 2018. "Measuring Foreign Policy Positions of Members of the US Congress." *Political Science Research and Methods* 6(1):181–196.
URL: *https://www.cambridge.org/core/product/identifier/S2049847016000030/type/journal*$_a$*rticle*

Jeong, Gyung-Ho & Paul J. Quirk. 2019. "Division at the Water's Edge: The Polarization of Foreign Policy." *American Politics Research* 47(1):58–87.
URL: *http://journals.sagepub.com/doi/10.1177/1532673X17719721*

Jervis, Robert. 1970. *The Logic of Images in International Relations*. Princeton, NJ: Princeton University Press.

Kertzer, Joshua D., Deborah Jordan Brooks & Stephen G. Brooks. 2021. "Do Partisan Types Stop at the Water's Edge?" *The Journal of Politics* 83(4):1764–1782.
URL: *https://www.journals.uchicago.edu/doi/10.1086/711408*

Kissinger, Henry. 2011*a*. *White House years*. First simon & schuster trade paperback ed ed. New York: Simon & Schuster trade paperbacks.

Kissinger, Henry. 2011*b*. *Years of upheaval*. First simon & schuster trade paperback ed ed. New York: Simon & Schuster trade paperbacks.

Kriner, Douglas L. 2010. *After the Rubicon: Congress, presidents, and the politics of waging war*. Chicago series on international & domestic institutions Chicago: University of Chicago Press.

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

Lindsay, James M. 1994. *Congress and the politics of U.S. foreign policy*. Baltimore: Johns Hopkins University Press.

Lunch, William L. & Peter W. Sperlich. 1979. "American Public Opinion and the War in Vietnam." *The Western Political Quarterly* 32(1):21–44. Publisher: [University of Utah, Sage Publications, Inc., Western Political Science Association].
URL: *https://www.jstor.org/stable/447561*

McManus, Roseanne W. 2017. *Statements of Resolve: Achieving Coercive Credibility in International Conflict*. Cambridge: Cambridge University Press.
URL: *http://ebooks.cambridge.org/ref/id/CBO9781316756263*

Myrick, Rachel. 2021. "Do External Threats Unite or Divide? Security Crises, Rivalries, and Polarization in American Foreign Policy." *International Organization* 75(4):921–958. Publisher: Cambridge University Press.
URL: *https://www.cambridge.org/core/journals/international-organization/article/abs/do-external-threats-unite-or-divide-security-crises-rivalries-and-polarization-in-american-foreign-policy/85C21E7EBD9EB38EDDB95F68A0458A88*

Nixon, Richard M. 1986. *No more Vietnams*. 1. avon pr ed. New York: Avon.

Ostermann, Falk & Wolfgang Wagner. 2023. "Introducing the parliamentary deployment votes database." *Journal of Peace Research* p. 002234332211166.
URL: *http://journals.sagepub.com/doi/10.1177/00223433221116655*

Ramsay, Kristopher W. 2004. "Politics at the Water's Edge: Crisis Bargaining and Electoral Competition." *Journal of Conflict Resolution* 48(4):459–486.

Rusk, Dean, Richard Rusk & Daniel S Papp. 1991. *As I saw it*. New York: Penguin. OCLC: 1087603337.

Schlesinger, Arthur M. 1973. *The Imperial Presidency*. Boston: Houghton Mifflin.

Schub, Robert. 2022. "Informing the Leader: Bureaucracies and International Crises." *American Political Science Review* pp. 1–17.
URL: *https://www.cambridge.org/core/product/identifier/S0003055422000168/type/journal$_a$rticle*

Schultz, Kenneth A. 1998. "Domestic Opposition and Signaling in International Crises." *American Political Science Review* 94(4):829–844.

Schultz, Kenneth A. 2001. *Democracy and Coercive Diplomacy*. Cambridge: Cambridge University Press.

Schultz, Kenneth A. 2003. Tying Hands and Washing Hands: The U.S. Congress and Multilateral Humanitarian Intervention. In *Locating the Proper Authorities: The Interaction of Domestic and International Institutions*, ed. Daniel W. Drezner. University of Michigan Press. pp. 105–142.

Schultz, Kenneth A. 2017. "Perils of Polarization for U.S. Foreign Policy." *The Washington Quarterly* 40(4):7–28.
URL: *https://www.tandfonline.com/doi/full/10.1080/0163660X.2017.1406705*

Sofaer, Abraham D. 1976. *War, Foreign affairs, and Constitutional Power*. Cambridge, Mass: Ballinger Pub. Co.

Stevenson, Charles A. 2007. *Congress at war: the politics of conflict since 1789*. 2007 ed ed. Washington, D.C: National Defense University Press : Potomac Books. OCLC: ocn166272954.

Wildavsky, Aaron. 1966. "The Two Presidencies." 4.

Wood, B. Dan. 2012. *Presidential saber rattling: causes and consequences*. New York: Cambridge University Press.

Woodward, Bob. 2010. *Obama's Wars*. 1st simon & schuster hardcover ed ed. New York: Simon & Schuster.