Appendix: Do External Threats Unite or Divide?

This is an appendix for "Do External Threats Unite or Divide? Security Crises, Rivalries, and Polarization in American Foreign Policy" by Rachel Myrick, accepted at *International Organization.*

Appendix A: Study 1

Appendix B: Study 2

Appendix C: Study 3

Appendix A: Study 1

I. Crisis Countries and Keywords

Country	COW Code	ISO3	Total Speeches	SQL Keywords
Afghanistan	700	AFG	23072	afghan
Angola	540	AGO	3048	angola
Cambodia	811	KHM	10175	cambodia, khmer, angkor
China	710	CHN	76264	china, chinese
Congo - Kinshasa	490	COD	3443	congo, zaire
Cuba	40	CUB	37837	cuba
Dominican Republic	42	DOM	3557	dominican
Egypt	651	EGY	12890	egypt
France	220	FRA	71937	france, french, franco
Germany	255	DEU	115299	german, prussia
Grenada	55	GRD	1693	grenada
Guatemala	90	GTM	3988	guatem
Haiti	41	HTI	7901	haiti
Iran	630	IRN	34222	iran, persia
Iraq	645	IRQ	43859	iraq
Israel	666	ISR	31350	israel
Japan	740	JPN	63152	japan
Jordan	663	JOR	8094	jordan
Libya	620	LBY	5556	libya
North Korea	731	PRK	19027	korea†
Panama	95	PAN	26782	panama
Russia	365	RUS	143771	russia, soviet
Spain	230	ESP	16465	spain
Sudan	625	SDN	3516	sudan
Syria	652	SYR	8443	syria
Thailand	800	THA	12190	thai
Turkey	640	TUR	17799	turkey, turkish
United Kingdom	200	GBR	144279	united kingdom, england, the english, britain, british
Vietnam	816	VNM	64171	vietnam
Yugoslavia	345	YUG	15313	yugoslavia, bosnia, serbia

† Speech must include the word “north” after 1953

II. International Crises and Triggering Entities

Crisis No.	Crisis Name	Triggering Entity	Start Date	Added
-1	USS MAINE EXPLOSION	Spain	1898-02-15	Yes
-2	BOXER REBELLION	China	1900-06-17	Yes
-3	UNRESTRICTED SUBMARINE	Germany	1917-01-31	Yes
59	PANAY INCIDENT	Japan	1937-12-12	No
88	PEARL HARBOR	Japan	1941-12-07	No
104	TRIESTE I	Yugoslavia	1945-05-02	No
108	AZERBAIJAN	Russia	1946-03-04	No
111	TURKISH STRAITS	Russia	1946-08-07	No
114	TRUMAN DOCTRINE	United Kingdom	1947-02-21	No
123	BERLIN BLOCKADE	Russia	1948-06-24	No
125	CHINA CIVIL WAR	China	1948-09-23	No
132	KOREAN WAR I	North Korea	1950-06-25	No
133	KOREAN WAR II	China	1950-10-31	No
140	KOREAN WAR III	North Korea	1953-04-16	No
144	GUATEMALA	Guatemala	1954-02-10	No
145	DIEN BIEN PHU	France	1954-03-20	No
146	TAIWAN STRAIT I	China	1954-09-03	No
152	SUEZ NATN.-WAR	Russia	1956-11-05	No
159	SYRIA/TURKEY CONFRNT.	Syria	1957-08-18	No
165	IRAQ/LEB. UPHEAVAL	Iraq	1958-07-14	No
166	TAIWAN STRAIT II	China	1958-08-23	No
168	BERLIN DEADLINE	Russia	1958-11-27	No
180	PATHET LAO OFFENSIVE	Thailand	1961-03-09	No
181	BAY OF PIGS	Cuba	1961-04-15	No
185	BERLIN WALL	Russia	1961-08-13	No
186	VIET CONG ATTACK	Vietnam	1961-09-18	No
193	NAM THA	Thailand	1962-05-06	No
196	CUBAN MISSILES	Russia	1962-10-16	No
206	PANAMA FLAG	Panama	1964-01-10	No
210	GULF OF TONKIN	Vietnam	1964-08-02	No
211	CONGO II	Congo - Kinshasa	1964-09-26	No
213	PLEIKU	Vietnam	1965-02-07	No
215	DOMINICAN INTERVENTN.	Dominican Republic	1965-04-24	No
222	SIX DAY WAR	Russia	1967-06-06	No
224	PUEBLO	North Korea	1968-01-22	No
225	TET OFFENSIVE	Vietnam	1968-02-27	No
230	VIETNAM SPRING OFF.	Vietnam	1969-02-22	No

Crisis No.	Crisis Name	Triggering Entity	Start Date	Added
233	EC-121 SPY PLANE	North Korea	1969-04-15	No
237	INVASION OF CAMBODIA	Vietnam	1970-04-21	No
238	BLACK SEPTEMBER	Jordan	1970-09-15	No
239	CIENFUEGOS SUB. BASE	Russia	1970-09-16	No
246	VIETNAM PORTS MINING	Vietnam	1972-03-30	No
249	CHRISTMAS BOMBING	Vietnam	1972-12-04	No
255	OCTOBER-YOM KIPPUR WAR	Syria	1973-10-12	No
255	OCTOBER-YOM KIPPUR WAR	Egypt	1973-10-12	No
255	OCTOBER-YOM KIPPUR WAR	Israel	1973-10-12	No
259	MAYAGUEZ	Cambodia	1975-05-12	No
260	WAR IN ANGOLA	Angola	1975-09-01	No
274	POPLAR TREE	North Korea	1976-08-17	No
292	SHABA II	Angola	1978-05-14	No
292	SHABA II	Congo - Kinshasa	1978-05-14	No
303	AFGHANISTAN INVASION	Russia	1979-12-24	No
309	US HOSTAGES IN IRAN	Iran	1979-11-04	No
343	INVASION OF GRENADA	Grenada	1983-10-19	No
354	NICARAGUA MIG-21S	Russia	1984-11-06	No
363	GULF OF SYRTE II	Libya	1986-04-05	No
386	LIBYAN JETS	Libya	1988-12-21	No
391	INVASION OF PANAMA	Panama	1989-12-15	No
393	GULF WAR	Iraq	1990-10-30	No
408	N. KOREA NUCLEAR I	North Korea	1993-03-12	No
411	HAITI MIL. REGIME	Haiti	1994-07-17	No
412	IRAQ DEPLOY./KUWAIT	Iraq	1994-10-07	No
419	DESERT STRIKE	Iraq	1996-08-31	No
422	UNSCOM I	Iraq	1997-11-13	No
427	US EMBASSY BOMBINGS	Sudan	1998-08-07	No
427	US EMBASSY BOMBINGS	Afghanistan	1998-08-07	No
429	UNSCOM II	Iraq	1998-10-31	No
430	KOSOVO	Yugoslavia	1999-02-20	No
434	AFGHANISTAN/US	Afghanistan	2001-09-11	No
440	IRAQ REGIME CHANGE	Iraq	2003-01-13	No
441	N. KOREA NUCLEAR II	North Korea	2002-10-04	No
448	IRAN NUCLEAR II	Iran	2006-01-10	No
450	N. KOREA NUCLEAR III	North Korea	2006-05-05	No
459	N. KOREA NUCLEAR IV	North Korea	2009-03-11	No
464	LIBYAN CIVILWAR	Libya	2011-02-22	No
469	N. KOREA NUCLEAR V	North Korea	2013-02-12	No
470	SYRIA CHEMICAL WEAPONS	Syria	2013-08-21	No
476	TURKEY-RUSSIA JET INCIDENT	Turkey	2015-11-24	No

III. Technical Description of Study 1 Method

To measure polarization of Congressional rhetoric, I closely adapt the method outlined by Peterson and Spirling (2018), who use the accuracy of machine learning classifiers as a proxy for polarization. In their application, Peterson and Spirling measure polarization of rhetoric between Labour and Conservative members of parliament (MPs) in British parliamentary debates. The application in this paper is to use machine learning algorithms to predict whether a randomly drawn speech about a given country is from a Republican or a Democratic legislator. Country-sessions in which there is higher predictive accuracy means it is easier to discern the party of a legislator from their speech, thus suggesting greater polarization.

Cleaning Speeches

In order to implement this method, I use the digitized version of the Congressional record compiled by Gentzkow, Shapiro and Taddy (2018). I create a SQL database of Congressional speeches in which any of the 30 countries are mentioned. I merge demographic and political information about each speaker into this database. I then take the following steps to clean the speeches:

1. Remove punctuation
2. Remove non-alpha numeric characters and numbers
3. Change to lowercase
4. Strip white space
5. Remove English stop words (Feinerer, Hornik and Meyer, 2008). These are common words like words like "the," "she," and "is."
6. Remove procedural stop words for Congressional speech (Gentzkow, Shapiro and Taddy, 2017). These are words that appear frequently in Congressional speeches like "chairman," "senator," and "adjourn."
7. Drop speeches that: (1) do not have an identifiable speaker from Congressional speaker map, (2) are not from a Republican or Democrat, (3) are less than 40 characters.
8. Drop country-sessions that do not have a minimum of 50 Republican and 50 Democratic speeches.

Following standard practice in computational text analysis, each speech is then treated as a "bag of words." This means that word order is discarded; each speech can be represented by a vector of terms and the number of time each term appears in the speech, normalized by how frequently the term appears across all speeches (Jurafsky and Martin, 2009; Grimmer and Stewart, 2013). Vocabulary is fixed across all the speeches; words used infrequently (i.e., that do not appear in at least 200 unique speeches) are removed from the data (Gentzkow, Shapiro and Taddy, 2017; Peterson and Spirling, 2018).

Measuring Polarization of Speech

I segment the data by "country-session" (i.e. How polarized is speech about Iran in the 114th Congress?). Within each country-session from the 43rd Congress through the 114th Congress, I use a supervised machine learning method from Peterson and Spirling (2018) to predict the likelihood that a speech was given by a Republican or a Democrat. I run four different machine learning algorithms ("classifiers") and average over a stratified 10 fold cross-validation for each. In stratified 10 fold cross-validation, speeches are randomly divided into 10 equal subsamples or "folds." Within each fold, there are roughly the same number of Republican and Democratic speeches. Nine folds "train" the classifier and one fold "tests" how well the classifier predicts the party of the speaker. This process is repeated 10 times, such that each fold serves as a test set once. The performance of each classifier is averaged over these 10 test sets. For each session, speeches are inversely weighted by party, such that there are roughly equal numbers of Republican and Democratic speeches. The level of polarization within each session is proxied by the average predictive accuracy of the best performing algorithm.[1]

The machine learning algorithms are implemented using the Scikit-learn library in Python. The citations for the four algorithms are:

1. Schmidt, Mark, Nicolas Le Roux, and Francis Bach. 2013. Minimizing finite sums with the stochastic average gradient. https://arxiv.org/abs/1309.2388

2. Bottou, Leon. 2004. Stochastic learning. https://www.semanticscholar.org/paper/Stochastic-Learning-Bottou/7b0db6135b8dd3e2a9efa86163e91c0cd0fdf660

3. Freund, Yoav, and Robert E. Schapire. 1999. Large margin classification using the perceptron algorithm. Machine Learning 37(3):277–296.

4. Crammer, Koby, Ofer Dekel, Joseph Keshet, Shai Shalev-Shwartz, and Yoram Singer. 2006. Online passive-aggressive algorithms. Journal of Machine Learning Research 7(1):551–585.

The result is a country-session dataset involving 30 countries between 1873-2015. Each country-session now has a dependent variable: a measure of the predictive accuracy of the best performing

[1]Using the estimates from any one classifier exclusively does not change the overall findings.

classifier, which captures the level of polarization of rhetoric. Predictive accuracy theoretically ranges between 0.5 and 1, such that 0.5 indicates no distinction between Republican and Democratic speech (i.e., a 50-50 chance of accurately predicting the party of the speaker based on their speech), and 1 indicates perfect distinction. Country-sessions in which the average predictive accuracy of the classifier is closer to 1 are more rhetorically polarized relative to other country-sessions.

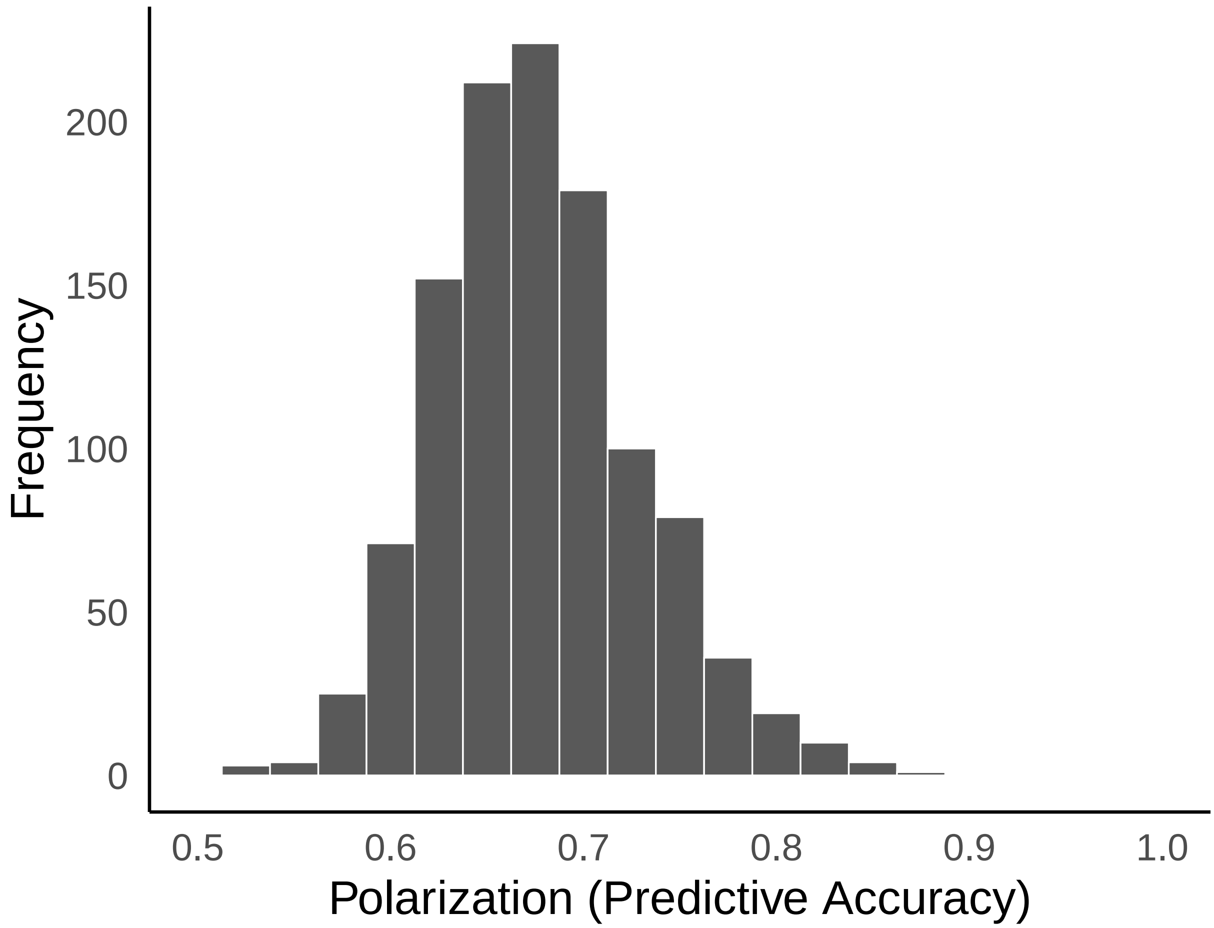
Comparison to Other Methods

There are a variety of different ways to measure the partisanship of speech.[2] I use the methods from Peterson and Spirling (2018) for two main reasons. First, in measuring the partisanship of speech, a supervised machine learning approach has greater construct validity than an unsupervised approach. As Goet (2019) summarizes in a review of different methods for measuring partisanship of speech, "In contrast to their unsupervised siblings, such supervised models attempt to identify which speakers use a vocabulary that is similar to speakers from one versus another party, ensuring that variation in word use is related to a stable construct" (p. 2). Second, in contrast to parametric approaches like the approach used by Gentzkow, Shapiro and Taddy (2017), non-parametric approaches like the one employed by Peterson and Spirling (2018) are more easily interpretable and much less computationally intensive.

[2]See Goet (2019) for an overview of methods.

IV. Distribution of dependent variable (polarization of rhetoric)

Figure 1: Polarization of crisis country rhetoric



V. Description of “High Threat” Crisis Coding

This section describes the coding of “high threat” crises used in Table 3 (Study 1) and Table 5 (Study 2) in the mansucript. I conceptualize “high threat” in three different ways, and then use the coding of international crises involving the United States from the International Crisis Behavior Project v. 12 dataset (Brecher et al., 2017) to classify crises.

(1) Crises that ultimately escalate into military violence

These are crises that involve an outbreak of violence, corresponding to **viol** variable in the ICB2 data (viol > 1).

(2) Crises that involve the USSR/communist threat

These are crises that involve both the United States and the Soviet Union, corresponding to the **suinv** variable in the ICB2 data (suinv > 2).

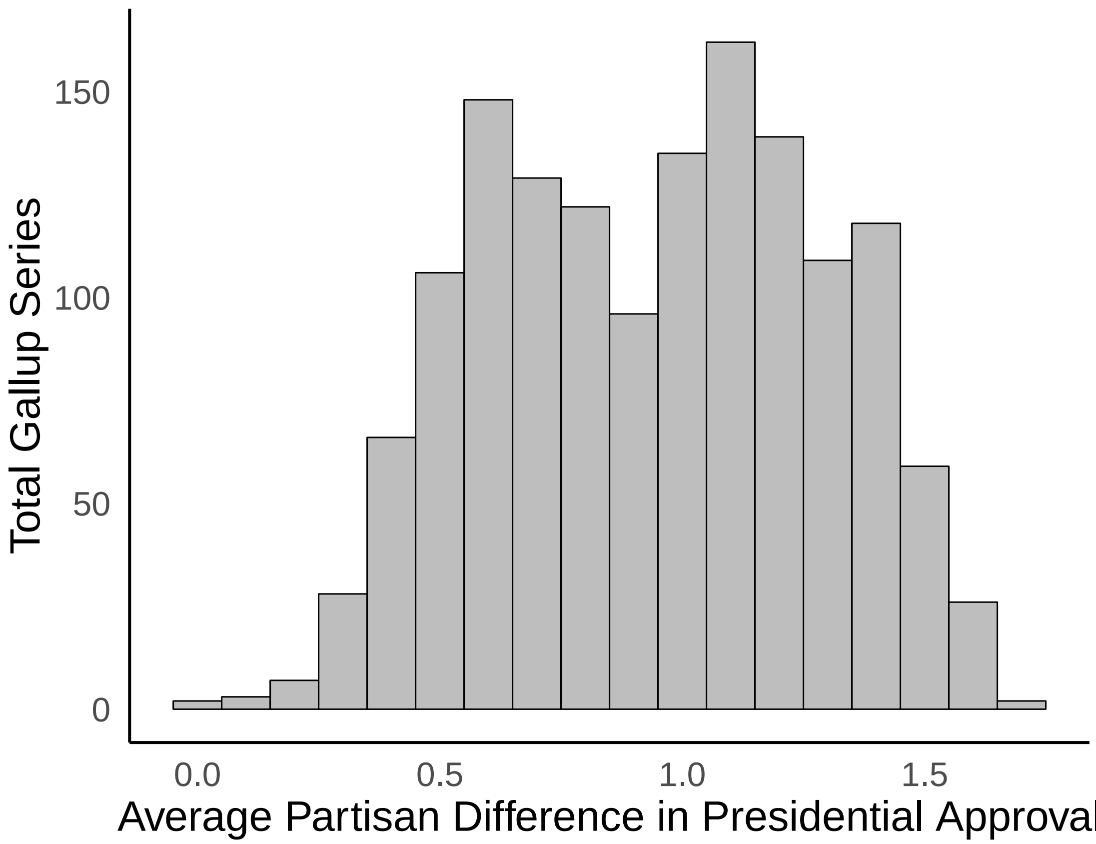
(3) Crises that are perceived as “higher threat” ex-ante by decision-makers

These are crises that have a high perceived gravity of threat prior to knowing the crisis outcome. For example, the Cuban Missile Crisis did not escalate to violence, but it was perceived as a serious threat ex-ante by foreign policymaker. This coding corresponds to the **gravty** variable in the ICB2 data (gravty > 3 and gravty < 7). This classification drops crises where the threat is limited to economic, political, or small-scale military threats.

Appendix B: Study 2

I. Distribution of Dependent Variable

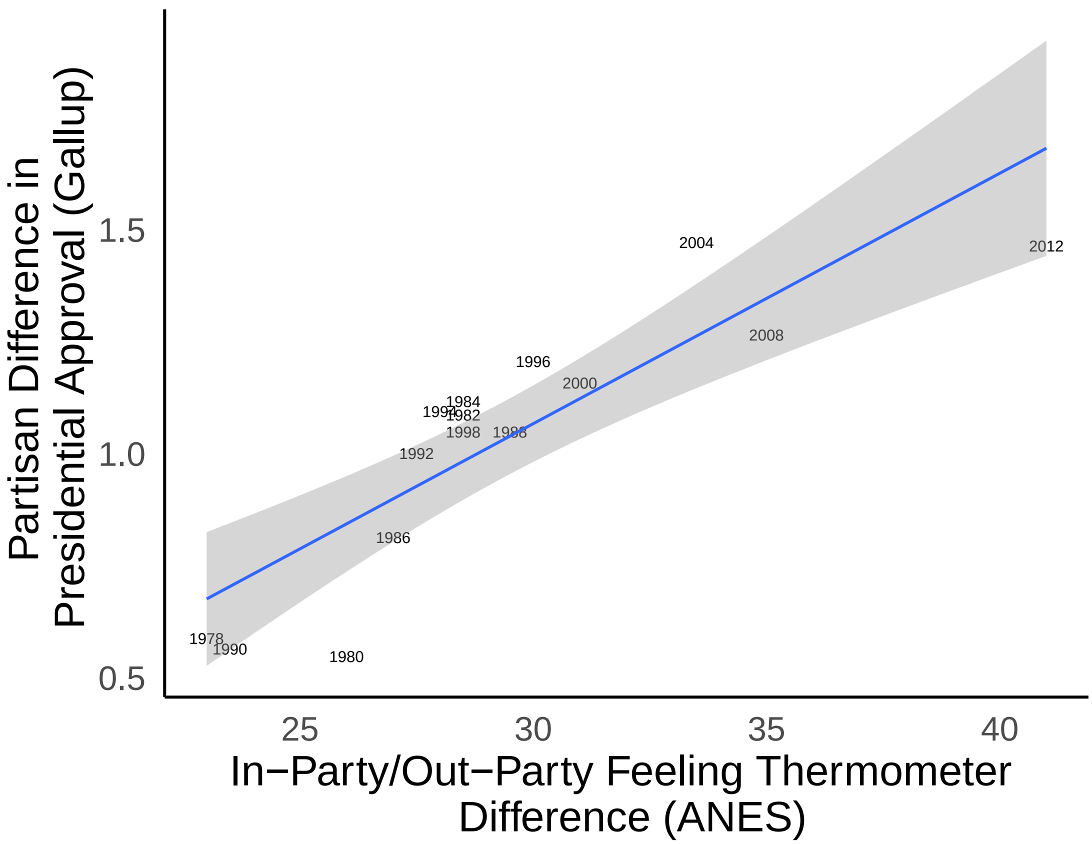
Figure 2: Distribution of partisan difference in presidential approval



II. Relationship between dependent variable and ANES variable

A common measure of affective polarization is the difference in feeling thermometer ratings of one's in-party and out-party from the American National Election Studies (ANES). Survey respondents are asked to rate each party on a scale from 0 to 100, where 100 indicates extremely favorable. This figure shows a strong, positive correlation between the affective polarization measure from the ANES on the X-axis and the dependent variable used in Study 2 (the average partisan difference in presidential approval between in-party and out-party) members) for the corresponding year on the Y-axis. The correlation between the two variables is 0.874.

Figure 3: Correlation between dependent variable and commonly used measure of affective polarization from the ANES



III. Robustness Check: Unweighted Results

The polling firm Gallup, which collects presidential approval data used in Study 2, weights the data they collect from each poll to reflect the American adult population. Following **?**, who have assembled this polling data across time, the results presented in Study 2 are also weighted. As a robustness check, I remove the weights and replicate the analyses (Table 4 in manuscript). The substantive conclusion—that crises events are associated with smaller partisan gaps in presidential approval, but that this effect is small and short lived—remains the same.

Table 1: Crises and partisan gaps in presidential approval (Unweighted results)

	(1)	(2)	(3)	(4)	(5)	(6)
Before Crisis	-0.018			-0.021	-0.011	-0.012
	(0.022)			(0.021)	(0.022)	(0.022)
During Crisis	-0.076^{***}			-0.075^{***}	-0.057^{**}	-0.059^{**}
	(0.024)			(0.023)	(0.024)	(0.024)
After Crisis	-0.037			-0.040	-0.036	-0.037
	(0.028)			(0.027)	(0.028)	(0.028)
Change in Disposable Income		0.002		0.002		-0.001
		(0.002)		(0.002)		(0.002)
Unemployment Rate		0.011^{**}		0.012^{**}		0.038^{***}
		(0.005)		(0.005)		(0.005)
Congressional Polarization			1.079^{***}		1.052^{***}	1.506^{***}
			(0.136)		(0.135)	(0.153)
Divided Government			-0.034		-0.034	0.027
			(0.026)		(0.026)	(0.027)
Election Year			0.127^{***}		0.125^{***}	0.126^{***}
			(0.011)		(0.010)	(0.010)
Constant	0.650^{***}	0.570^{***}	-0.016	0.575^{***}	0.007	-0.567^{***}
	(0.021)	(0.030)	(0.083)	(0.030)	(0.082)	(0.119)
President FEs	Yes	Yes	Yes	Yes	Yes	Yes
AIC	-888.61	-878.79	-1066.56	-891.9	-1072.49	-1114.62

Note: $^{*}p<0.1$; $^{**}p<0.05$; $^{***}p<0.01$

IV. Robustness Check: Dropping September 2001

One possible concern is that the public response to September 11th is driving the negative relationship between crisis events and the partisan gap in presidential approval. As a robustness check, I drop polling conducting during September 2001 from the analyses. The magnitude of the coefficients on "During Crisis" reduces by approximately 30 percent relative to the analysis on the full sample (see Table 4 in the manuscript) by removing this one observation. Since the coefficient is still statistically significant at the 95 percent confidence level, the negative relationship is not solely driven by the response to 9/11, although this is clearly an influential data point.

Table 2: International Crises and Partisan Difference in Presidential Approval (Dropping September 2001)

	(1)	(2)	(3)	(4)	(5)	(6)
Before Crisis	-0.014			-0.017	-0.006	-0.007
	(0.022)			(0.022)	(0.024)	(0.024)
During Crisis	-0.061^{***}			-0.059^{***}	-0.042^{**}	-0.044^{**}
	(0.020)			(0.019)	(0.021)	(0.020)
After Crisis	-0.040			-0.043	-0.038	-0.040
	(0.028)			(0.027)	(0.029)	(0.029)
Change in Disposable Income		0.003^{*}		0.003		-0.001
		(0.002)		(0.002)		(0.001)
Unemployment Rate		0.010^{**}		0.011^{**}		0.039^{***}
		(0.005)		(0.005)		(0.005)
Congressional Polarization			1.115^{***}		1.096^{***}	1.555^{***}
			(0.136)		(0.137)	(0.152)
Divided Government			-0.026		-0.026	0.035
			(0.027)		(0.027)	(0.028)
Election Year			0.123^{***}		0.122^{***}	0.122^{***}
			(0.011)		(0.011)	(0.011)
Constant	0.611^{***}	0.536^{***}	-0.083	0.539^{***}	-0.065	-0.648^{***}
	(0.021)	(0.031)	(0.084)	(0.031)	(0.084)	(0.116)
President FEs	Yes	Yes	Yes	Yes	Yes	Yes
AIC	-813.14	-810.5	-982.77	-816.92	-984.14	-1023.97

Note: *p<0.1; **p<0.05; ***p<0.01

V. Analysis of other external threats: Militarized Interstate Disputes

An alternative proxy for heightened threat is to use American engagement in a militarized interstate dispute (Palmer et al., 2015). The"rally 'round the flag" effect posits that engagement in militarized disputes boosts presidential popularity. This robustness check replicates the Study 2 analysis (Table 4 in the manuscript) using MIDs rather than crisis events. While the coefficient on "During MID" is negative and statistically significant, it is difficult to attribute this effect to engagement in the dispute. First, unlike crisis events, which are defined in part by their short time horizon, MIDs often occur over long stretches of time. Second, the initiation of a MID is a strategic interaction that may not be exogenous to presidential approval (Ostrom and Job, 1986; Mitchell and Moore, 2002; Mitchell and Prins, 2004; Mitchell and Thyne, 2010). Across all models, the coefficient on "Before MID" (an indicator for whether a survey was run directly before a MID) is larger in magnitude than the coefficient on "During MID." While this coefficient is not statistically significant (likely because only a handful of surveys were run directly before a MID), it suggests that a smaller partisan gap in presidential approval precedes MIDs. This should make us cautious about claims that MIDs independently reduce the partisan gap in presidential approval.

Table 3: Militarized Interstate Disputes and Partisan Difference in Presidential Approval

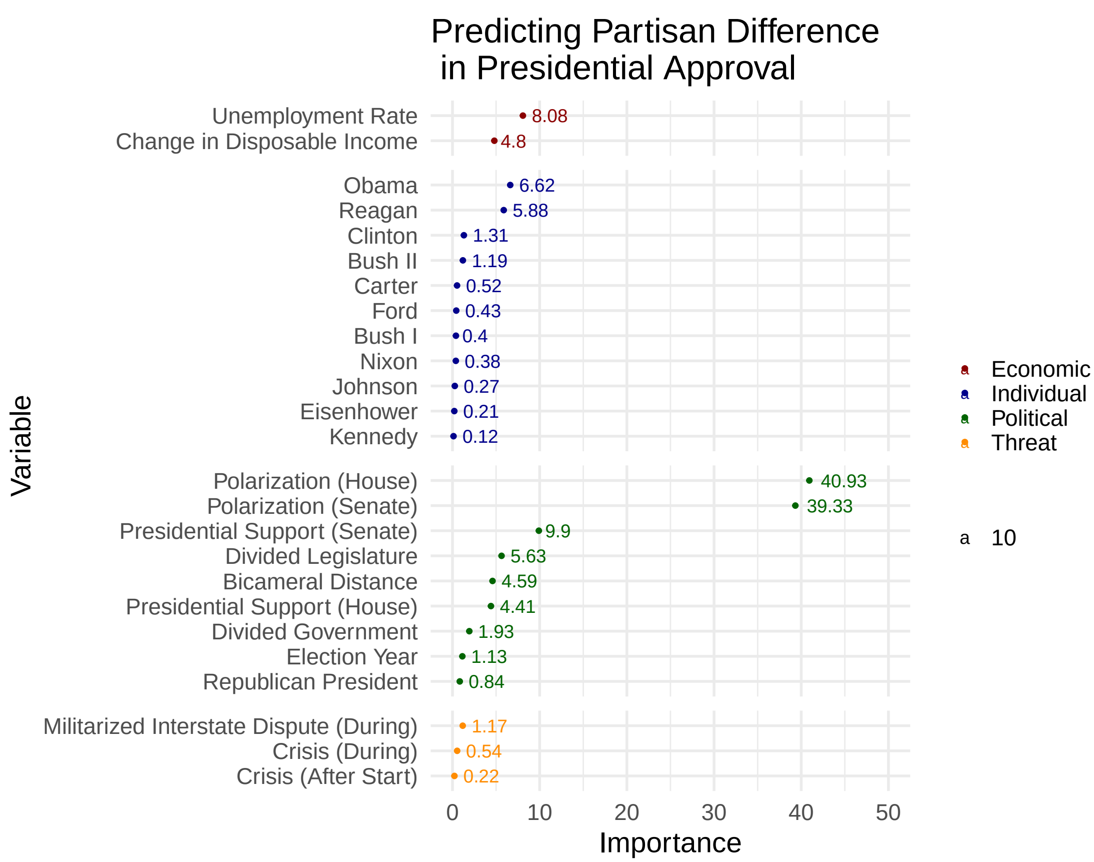
	(1)	(2)	(3)	(4)	(5)	(6)
Before MID	-0.191			-0.193	-0.161	-0.160
	(0.188)			(0.191)	(0.176)	(0.167)
During MID	-0.157^{***}			-0.154^{***}	-0.138^{***}	-0.126^{***}
	(0.032)			(0.032)	(0.033)	(0.034)
Change in Disposable Income		0.002		0.002		-0.001
		(0.002)		(0.002)		(0.002)
Unemployment Rate		0.011^{**}		0.006		0.037^{***}
		(0.005)		(0.005)		(0.005)
Congressional Polarization			1.161^{***}		1.254^{***}	1.692^{***}
			(0.139)		(0.144)	(0.160)
Divided Government			-0.028		-0.031	0.029
			(0.027)		(0.027)	(0.028)
Election Year			0.124^{***}		0.116^{***}	0.117^{***}
			(0.011)		(0.011)	(0.011)
Constant	0.619^{***}	0.534^{***}	-0.109	0.582^{***}	-0.150^{*}	-0.711^{***}
	(0.020)	(0.031)	(0.086)	(0.030)	(0.087)	(0.120)
President FEs	Yes	Yes	Yes	Yes	Yes	Yes
AIC	-803.67	-754.86	-930.18	-803.4	-969.3	-1005.94

Note: *p<0.1; **p<0.05; ***p<0.01

VI. Random Forest

Study 2 looked at standardized coefficients and basic model fit statistics to compare crises to other macro-level characteristics in order to explain variation in the partisan gap in presidential approval. An alternative approach is to use a random forest to more directly compare the predictive importance of variables. A random forest is an ensemble method used to aggregate predictions of individual decision trees. Decision trees partition data into smaller subsets until further partitions no longer enhance the predictive power of the model. After running the random forest, I construct a variable importance plot (**Figure 4**) where higher values indicate better predictors of the dependent variable. The figure confirms that threat variables have very low predictive power in relation many of the macro-level variables included in the model. In general, economic variables and many political variables tend to be much better predictors of partisan differences in presidential approval than threat variables. In particular, polarization among political officials in both the House and Senate is highly predictive of polarization of attitudes towards the president.

Figure 4: Variable importance plot from random forest assessing predictors of partisan difference in presidential approval



Appendix C: Study 3

I. Survey Questionnaire

Part I: Demographic Characteristics & Political Affiliation

After agreeing to participate in the survey, respondents are asked the following demographic questions, consistent with standard language on the American National Election Studies.

sex: *Are you male or female?*

- *Male*
- *Female*

age: *What is your age?*

race: *What racial or ethnic group best describes you?*

- *White*
- *Black or African American*
- *Hispanic or Latino*
- *Asian or Asian American*
- *Native American*
- *Middle Eastern*
- *Mixed Race*
- *Some other race* [Text Entry]

pid1: *Generally speaking, do you think of yourself as a...*

- *Republican*
- *Democrat*
- *Independent / Other*

pid2: [Display only if pid1=="Democrat" or "Republican"] *Would you call yourself a...*

- *Strong [Republican/Democrat]*
- *Not very strong [Republican/Democrat]*

pid2: [Display only if pid1==“Independent / Other”] *Do you think of yourself as closer to the...*

- *Republican Party*
- *Democratic Party*
- *Neither party*

ideo: *In general, do you think of yourself as...*

- *Extremely liberal*
- *Liberal*
- *Slightly liberal*
- *Moderate, middle of the road*
- *Slightly conservative*
- *Extremely conservative*

pres_approval: *Do you approve or disapprove of the way Donald Trump is handling his job as president?*

- *Disapprove Strongly*
- *Disapprove Somewhat*
- *Neither Approve nor Disapprove*
- *Approve Somewhat*
- *Approve Strongly*

state: *In what state do you currently reside?* [Dropdown menu]

educ: *What is the highest level of school you have completed?*

- *Did not graduate from high school*
- *High school graduate*

- *Some college, but no degree (yet)*
- *2-year college degree*
- *4-year college degree*
- *Postgraduate degree (MA, MBA, MD, JD, PhD, etc)*

Part II: Primes

Respondents are randomly assigned to one of five conditions: one Control Group and four treatment groups. Respondents in the Control Group move to the outcomes (Part III), while respondents in any of the Treatment Groups read the following text. There are two treatments, each with two options. The first treatment (T1) is whether the text is presented with a Non-Partisan Frame or a Partisan Frame. The second treatment (T2) is whether or not there is an addition of an ideological statement at the end of the security threat.

The text on the first screen is accompanied with an image of either the cover of the Worldwide Threat Assessment (Non-Partisan Frame) or of President Trump (Partisan Frame). The statements in these primes are based on a publicly issued government report called the 2019 Worldwide Threat Assessment presented annually to Congress by the Director of National Intelligence (DNI). While the statements are attributed to different sources in the treatments, both attributions are factually correct; the intelligence community is a group of non-partisan experts but the DNI is appointed by the Trump administration:

open: *On the next page, you will read statements based on a real, recent report from [T1: non-partisan experts the Trump administration] called the "Worldwide Threat Assessment." Please read this information carefully*

[NEW SCREEN]

prime: *A recent report from [T1: non-partisan experts /the Trump administration] says that the risk of conflict between the United States and China is higher than any time since the end of the Cold War. According to the report, [T1: experts / President Trump and his cabinet officials] say that:*

- *China is aggressively expanding its economic and military influence, as well as its nuclear capabilities.*
- *China is using intelligence services to steal information and spy on U.S. citizens.*

- *China has the ability to launch cyber attacks that can disrupt critical infrastructure — such as electric grids or natural gas pipelines — in the United States*

[T2: NULL / *A coming "ideological battle" between the United States and China will threaten support for democracy and human rights globally.*]

[NEW SCREEN]

Part III: Outcomes

There are four sets of outcomes measured in this experiment. The first set of outcomes uses a feeling thermometer to gauge attitudes as a proxy for affective polarization.

affective: *We'd also like to get your feelings about some groups in American society. Rate the following groups between 0 and 100. Ratings from 50-100 mean that you feel favorably toward the group; ratings from 0-50 degrees mean that you don't feel favorably towards the group and that you don't care too much for that group.*

- *Democrats*
- *Republicans*

[NEW SCREEN]

The second outcome asks a standard "Frenemy" question from YouGov.

china_frenemy: *We are now interested in your views towards China. Do you consider China to be a **friend** or an **enemy** of the United States?*

- *Ally*
- *Friendly*
- *Unfriendly*
- *Enemy*
- *Not Sure*

[NEW SCREEN]

The third set asks whether China poses threats to the United States or opportunities for cooperation. The outcomes are measured using a 7-pt Likert scale where "1" indicates "Strongly Disagree" and "7" indicates "Strongly Agree."

china_threat: *Do you agree or disagree with the following statement:* ***China poses a threat to the United States****?*

[NEW SCREEN]

china_opp: *Do you agree or disagree with the following statement:* ***China poses an opportunity for cooperation with the United States.***

[NEW SCREEN]

The fourth set asks about specific instruments the U.S. should use to respond to China. The outcomes are measured using a 5-pt Likert scale where "1" indicates "Very Unacceptable" and "7" indicates "Very Acceptable." :

tool_text: *In your opinion, how acceptable or unacceptable is it for the United States government to take the following actions?*

tool_dip: *Engage in* ***diplomacy*** *(directly talk with foreign leaders) with China*

tool_sanction: *Impose* ***economic sanctions*** *(financial or trade restrictions designed to hurt a country's economy) against China*

tool_covert: *Use* ***covert action*** *to secretly influence China's politics*

tool_milthreat: ***Threaten military force*** *against China*

tool_milforce: ***Use military force*** *against China*

[END OF SURVEY]

II. Sampling Strategy

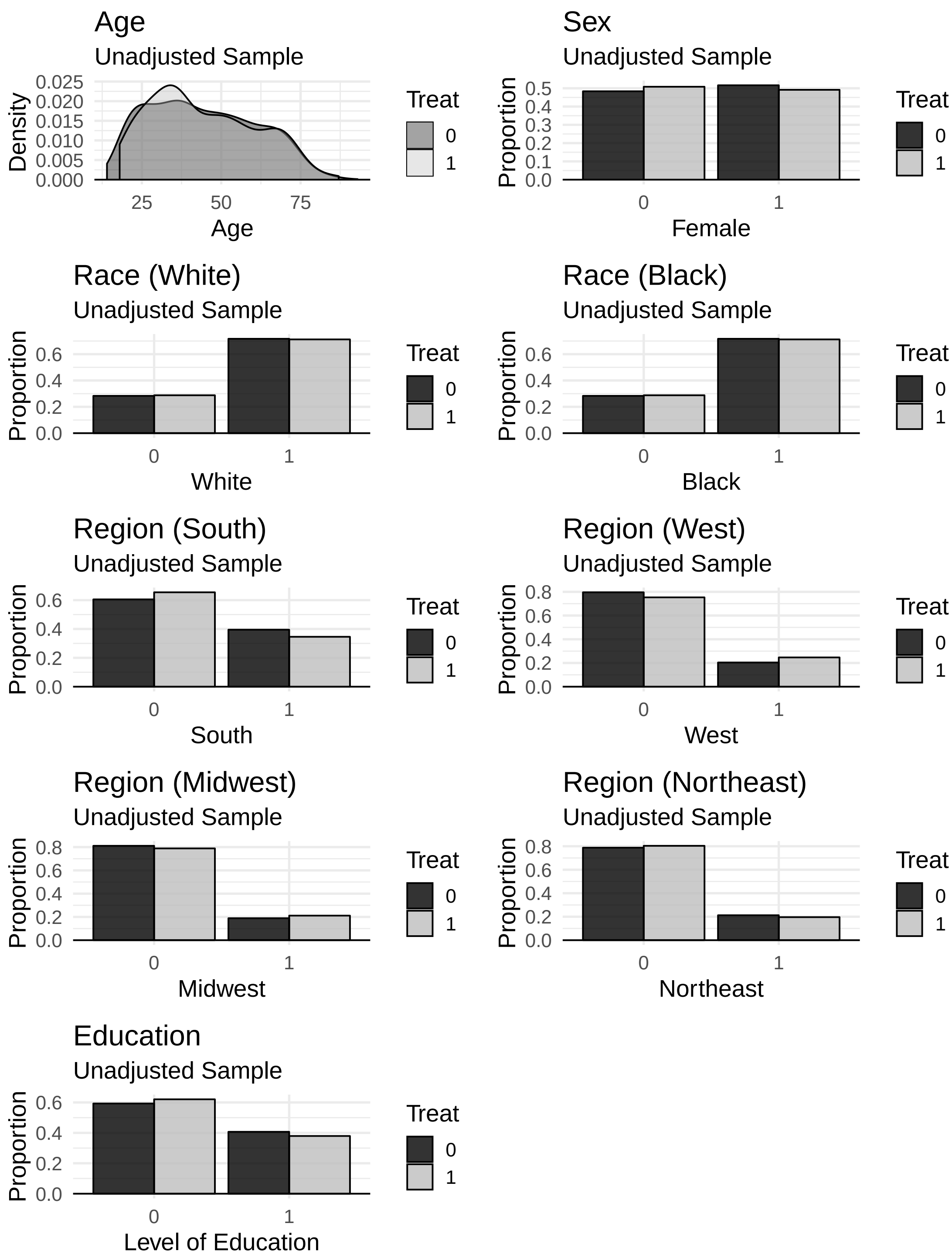
A pilot survey for this study with three treatment conditions was fielded in May 2019 to a sample of 1000 U.S. adults. A full survey with all five treatment conditions was fielded in June 2019 to a sample of 2500 U.S. adults. Both surveys were reviewed by the Stanford University Institutional Review Board (protocol 50970) and fielded through the Lucid for Academics Marketplace. Lucid is a professional survey firm that maintains a marketplace used to recruit survey respondents for academic research. Lucid targets specific demographic quotas by age, sex, ethnicity, race, and region. The target and actual demographic characteristics of the samples drawn from these experiments are displayed in the table below.

Table 4: Targeted and Actual Demographic Characteristics of Pooled Sample

Quota Name	Target	Pilot	Full
Age (18-24)	0.13	0.13	0.14
Age (25-44)	0.41	0.51	0.38
Age (45-64)	0.3	0.24	0.32
Age (65+)	0.16	0.12	0.16
Male	0.5	0.52	0.49
Female	0.5	0.48	0.51
Hispanic	0.11	0.07	0.08
Black	0.12	0.17	0.12
White	0.7	0.69	0.72
Midwest	0.22	0.19	0.21
Northeast	0.22	0.24	0.19
South	0.37	0.40	0.38
West	0.23	0.17	0.21

III. Balance Checks

Figure 5: Checking Covariate Balance in Treated vs. Control Groups



IV. Robustness Check: Replication of Main Analyses with Pretest Results

Figure 6: Does China pose a threat to the United States (Pretest Results)?

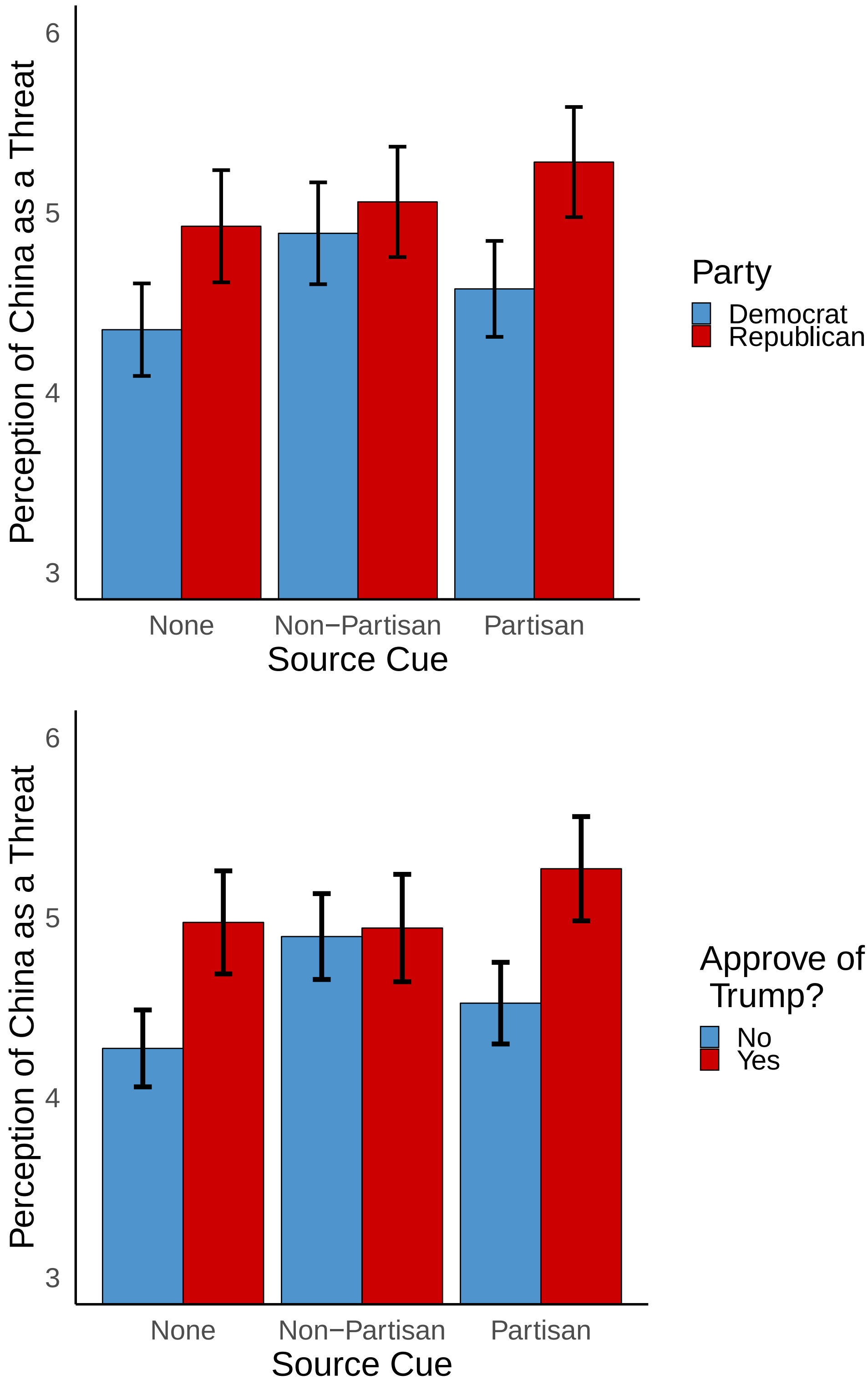


Table 5: How Acceptable is to Threaten or Use Military Force Against China? (Pretest Results)

	Dependent variable:					
	Threaten Military Force			Use Military Force		
	(1)	(2)	(3)	(4)	(5)	(6)
Partisan Cue	0.193^{**}	0.032	0.317^{***}	0.182^{*}	0.059	0.268^{**}
	(0.092)	(0.109)	(0.110)	(0.096)	(0.114)	(0.115)
Non-Partisan Cue	0.189^{**}	0.183^{**}	0.220^{**}	0.149	0.142	0.181^{**}
	(0.092)	(0.088)	(0.086)	(0.097)	(0.092)	(0.090)
Approve of Trump		0.674^{***}			0.753^{***}	
		(0.089)			(0.093)	
Partisan Cue x Approve of Trump		0.313^{**}			0.219	
		(0.154)			(0.161)	
Disapprove of Trump			-0.811^{***}			-0.850^{***}
			(0.086)			(0.090)
Partisan Cue x Disapprove of Trump			-0.266^{*}			-0.184
			(0.150)			(0.157)
Constant	2.350^{***}	2.082^{***}	2.721^{***}	2.369^{***}	2.070^{***}	2.759^{***}
	(0.064)	(0.070)	(0.071)	(0.067)	(0.073)	(0.075)
Observations	1,033	1,033	1,033	1,033	1,033	1,033
Note:				$^{*}p<0.1$; $^{**}p<0.05$; $^{***}p<0.01$		

Table 6: Do External Threats Decrease Affective Polarization? (Pretest Results)

	Dependent variable:			
	In-Party/Out-Party Difference		Out-Party Favorability	
	(1)	(2)	(3)	(4)
Non-Partisan Cue	−4.414	−3.436	1.355	0.977
	(3.060)	(3.010)	(2.183)	(2.147)
Partisan Cue	−2.322	−1.652	1.144	0.728
	(3.055)	(3.000)	(2.178)	(2.139)
Constant	52.826***	30.026***	27.553***	42.582***
	(2.123)	(5.204)	(1.514)	(3.712)
Controls?	No	Yes	No	Yes
Observations	839	839	839	839
Note:				*p<0.1; **p<0.05; ***p<0.01

V. Robustness Check: Impact of Threat Primes on Affective Polarization

Table 7 and **Table 8** look at the robustness of the results of **Table 6** in the mansucript. **Table 7** examines whether the null results in the paper are robust to an alternative way to proxy for affective polarization: using only a measure of out-group favorability. We might be concerned, for example, that the null results are due to the fact that the threat prime increased favorability towards both the in-party *and* the out-party. To check that this is not the case, the dependent variable in **Table 7** is the out-party rating *only*. The Identity Hypothesis (Hypothesis 2B) would anticipate that receiving the threat prime would lead to more positive feelings towards one's out-party. However, the table shows that there are no statistically significant differences between the treatment groups and the control group.

Table 7: Do External Threats Increase Out-Party Favorability? (Full Sample, Robustness Check)

	DV: Out-Party Favorability			
	(1)	(2)	(3)	(4)
Non-Partisan Cue	0.440	−0.039		
	(1.542)	(1.529)		
Partisan Cue	−0.095	−0.236		
	(1.531)	(1.516)		
Ideological Frame			−0.029	−0.353
			(1.531)	(1.517)
Non-Ideological Frame			0.371	0.081
			(1.542)	(1.527)
Constant	26.439***	34.780***	26.439***	34.774***
	(1.231)	(2.445)	(1.231)	(2.445)
Controls?	No	Yes	No	Yes
Observations	2,019	2,019	2,019	2,019
Note:			*p<0.1; **p<0.05; ***p<0.01	

The robustness check in **Table 7** interacts both treatment conditions—the source cue and frame of the report—to see if there are any impacts on affective polarization. In Models 1 and 2, the dependent variable is affective polarization as measured in **Table 6** in the paper (the difference between in-party and out-party ratings on a feeling thermometer). In Models 3 and 4, the dependent variable is out-party rating. The null results in all models emphasize that there are no statistically significant differences between the treatment group and the control group, failing to lend support to the Identity Hypothesis (H2B).

Table 8: Do External Threats Decrease Affective Polarization? (Full Sample, Robustness Check)

	Dependent variable:			
	In-Party/Out-Party Difference		Out-Party Favorability	
	(1)	(2)	(3)	(4)
Non-Partisan, Ideological Threat	-1.837	-1.374	0.808	0.315
	(2.592)	(2.576)	(1.793)	(1.778)
Non-Partisan, Non-Ideological Threat	-2.398	-1.818	0.059	-0.404
	(2.616)	(2.596)	(1.809)	(1.792)
Partisan, Ideological Threat	-1.311	-1.170	-0.826	-0.982
	(2.559)	(2.538)	(1.770)	(1.752)
Partisan, Non-Ideological Threat	-3.501	-3.371	0.673	0.552
	(2.592)	(2.572)	(1.793)	(1.775)
Constant	53.064^{***}	42.425^{***}	26.439^{***}	34.794^{***}
	(1.781)	(3.545)	(1.232)	(2.447)
Controls?	No	Yes	No	Yes
Observations	2,019	2,019	2,019	2,019

Note: *p<0.1; **p<0.05; ***p<0.01

VI. Full Survey Results: Additional Dependent Variables

Table 9: Is China a Friend or Enemy of the U.S.?

	Dependent variable:			
	Enemy (1) to Ally (4)			
	(1)	(2)	(3)	(4)
Partisan Cue	-0.144^{***}	-0.131^{***}	-0.051	-0.233^{***}
	(0.046)	(0.044)	(0.054)	(0.058)
Non-Partisan Cue	-0.264^{***}	-0.272^{***}	-0.259^{***}	-0.264^{***}
	(0.046)	(0.044)	(0.046)	(0.046)
Approve of Trump			-0.065	
			(0.045)	
Partisan Cue x Approve of Trump			-0.218^{***}	
			(0.071)	
Disapprove of Trump				-0.006
				(0.044)
Partisan Cue x Disapprove of Trump				0.183^{***}
				(0.070)
Constant	2.523^{***}	3.195^{***}	2.547^{***}	2.527^{***}
	(0.037)	(0.071)	(0.040)	(0.044)
Controls?	No	Yes	No	Yes
Observations	2,452	2,452	2,452	2,452
Note:		*p<0.1; **p<0.05; ***p<0.01		

Table 10: Does China Pose a Threat or an Opportunity to the United States?

	Dependent variable:					
	China Threat Scale (1-7)			China Opportunity Scale (1-7)		
	(1)	(2)	(3)	(4)	(5)	(6)
Partisan Cue	0.381***	0.080	0.690***	−0.091	0.122	−0.331***
	(0.090)	(0.104)	(0.113)	(0.082)	(0.095)	(0.102)
Non-Partisan Cue	0.681***	0.665***	0.679***	−0.234***	−0.239***	−0.231***
	(0.091)	(0.089)	(0.090)	(0.082)	(0.082)	(0.081)
Approve of Trump		0.243***			0.070	
		(0.087)			(0.080)	
Partisan Cue x Approve of Trump		0.704***			−0.526***	
		(0.139)			(0.127)	
Disapprove of Trump			−0.044			0.072
			(0.086)			(0.078)
Partisan Cue x Disapprove of Trump			−0.644***			0.503***
			(0.138)			(0.125)
Constant	4.385***	4.297***	4.408***	4.699***	4.674***	4.661***
	(0.073)	(0.078)	(0.085)	(0.066)	(0.071)	(0.077)
Observations	2,452	2,452	2,452	2,452	2,452	2,452

Note: *p<0.1; **p<0.05; ***p<0.01

Table 11: How Acceptable is to Threaten or Use Military Force Against China?

	Dependent variable:					
	Threaten Military Force			Use Military Force		
	(1)	(2)	(3)	(4)	(5)	(6)
Partisan Cue	0.133^{**}	-0.028	0.245^{***}	0.118^{*}	-0.032	0.200^{**}
	(0.066)	(0.075)	(0.080)	(0.067)	(0.076)	(0.080)
Non-Partisan Cue	0.165^{**}	0.134^{**}	0.138^{**}	0.200^{***}	0.166^{**}	0.169^{***}
	(0.066)	(0.064)	(0.064)	(0.067)	(0.065)	(0.064)
Approve of Trump		0.464^{***}			0.511^{***}	
		(0.063)			(0.063)	
Partisan Cue x Approve of Trump		0.335^{***}			0.305^{***}	
		(0.100)			(0.101)	
Disapprove of Trump			-0.562^{***}			-0.646^{***}
			(0.061)			(0.061)
Partisan Cue x Disapprove of Trump			-0.286^{***}			-0.231^{**}
			(0.098)			(0.098)
Constant	2.310^{***}	2.143^{***}	2.608^{***}	2.247^{***}	2.062^{***}	2.588^{***}
	(0.053)	(0.056)	(0.060)	(0.053)	(0.056)	(0.060)
Observations	2,452	2,452	2,452	2,452	2,452	2,452
Note:				$^{*}p<0.1$; $^{**}p<0.05$; $^{***}p<0.01$		

Table 12: How Acceptable is it for the U.S. to use the following tools towards/against China?

	Dependent variables (all on 1-5 scale)					
	Diplomacy	Sanction	Assist	Covert	Mil Threats	Mil Force
	(1)	(2)	(3)	(4)	(5)	(6)
Non-Partisan, Ideological	0.141**	0.183**	−0.012	0.234***	0.141*	0.156**
	(0.066)	(0.079)	(0.075)	(0.078)	(0.077)	(0.078)
Non-Partisan, Non-Ideological	0.004	0.213***	−0.071	0.240***	0.190**	0.245***
	(0.066)	(0.079)	(0.075)	(0.078)	(0.077)	(0.078)
Partisan, Ideological	0.051	0.054	−0.084	0.159**	0.127*	0.121
	(0.066)	(0.078)	(0.074)	(0.077)	(0.077)	(0.077)
Non-Partisan, Non-Ideological	0.111*	0.162**	−0.027	0.102	0.139*	0.115
	(0.067)	(0.079)	(0.075)	(0.078)	(0.078)	(0.078)
Constant	4.030***	2.923***	2.342***	2.445***	2.310***	2.247***
	(0.045)	(0.054)	(0.051)	(0.053)	(0.053)	(0.053)
Observations	2,452	2,452	2,452	2,452	2,452	2,452
R^2	0.003	0.004	0.001	0.005	0.003	0.004
Adjusted R^2	0.001	0.003	−0.001	0.004	0.001	0.003
Note:						*p<0.1; **p<0.05; ***p<0.01

References

Brecher, Michael, Jonathan Wilkenfeld, Kyle Beardsley, Patrick James and David Quinn. 2017. "International Crisis Behavior Data, Version 12.".
URL: *https://sites.duke.edu/icbdata/data-collections/*

Feinerer, Ingo, Kurt Hornik and David Meyer. 2008. "Text Mining Infrastructure in R." *Journal of Statistical Software* 25(5):1–54.

Gentzkow, Matthew, Jesse M. Shapiro and Matt Taddy. 2017. "Measuring Polarization in High-Dimensional Data: Method and Application to Congressional Speech." Working Paper.

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

Goet, Niels D. 2019. "Measuring Polarization with Text Analysis: Evidence from the UK House of Commons, 1811-2015." *Political Analysis* 27(4):518–539.

Grimmer, Justin and Brandon M. Stewart. 2013. "Text as Data: The Promise and Pitfalls of Automatic Content Analysis Methods for Political Texts." *Political Analysis* 21(3):267–297.

Jurafsky, Dan and James Martin. 2009. *Speech and natural language processing: An introduction to natural language processing, computational linguistics, and speech recognition.* Prentice Hall.

Mitchell, Sara McLaughlin and Brandon C. Prins. 2004. "Rivalry and Diversionary Uses of Force." *Journal of Conflict Resolution* 48(6):937–961.

Mitchell, Sara McLaughlin and Clayton T. Thyne. 2010. "Contentious Issues as Opportunities for Diversionary Behavior." *Conflict Management and Peace Science* 27(5):461–485.

Mitchell, Sara McLaughlin and Will H. Moore. 2002. "Presidential Uses of Force during the Cold War: Aggregation, Truncation, and Temporal Dynamics." *American Journal of Political Science* 46(2):438–452.

Ostrom, Charles W. and Brian L. Job. 1986. "The President and the Political Use of Force." *American Political Science Review* 80(2):541–566.

Palmer, Glenn, Vito D'Orazio, Michael Kenwick and Matthew Lane. 2015. "The MID4 dataset, 2002-2010: Procedures, Coding Rules, and Description." *Conflict Management and Peace Science* 32(2):222–242.

Peterson, Andrew and Arthur Spirling. 2018. "Classification Accuracy as a Substantive Quantity of Interest: Measuring Polarization in Westminster Systems." *Political Analysis* 26:120–128.