

ADVISERS AND AGGREGATION IN FOREIGN POLICY DECISION-MAKING*

Tyler Jost[†] Joshua D. Kertzer[‡] Eric Min[§] Robert Schub[¶]
Last modified: January 7, 2022

Word count: 11993

ABSTRACT: Do advisers affect foreign policy and, if so, how? Recent scholarship on elite decision-making prioritizes leaders and the institutions that surround them, rather than the dispositions of advisers themselves. We argue that despite the hierarchical nature of foreign policy decision-making, advisers' predispositions towards the use of force shape state behavior through participation in deliberations. We test our argument by introducing an original dataset of 2,881 foreign policy deliberations between US presidents and their advisers from 1947 to 1988. Applying a novel machine learning approach to estimate the hawkishness of 1,073 Cold War-era foreign policy decision-makers, we show that adviser-level hawkishness has consistently large effects on foreign policy decisions. Conflictual policy choices grow more likely as hawks increasingly dominate the debate, even when accounting for leader dispositions. These results enrich our understanding of international conflict by demonstrating that advisers' dispositions, which aggregate via deliberation, systematically shape foreign policy.

*This research was funded by the Defense Advanced Research Projects Agency (Grant #W911NF1920162). We're also grateful to the Institute of Quantitative Social Science (IQSS) and the Weatherhead Center for International Affairs (WCFA) at Harvard University for their logistical support, and to Saadi Ali, Perry Arrasmith, Christine Bienes, Florian Bochert, Lexa Brenner, Alexis Carraway, Youmna Chamieh, Erica Colston, Joshua Conde, Vanessa Cordova, Jacinta Crestanello, Charlie Davis, Darius Diamond, Iain De Boer, A.J. Dilts, Jelena Dragicevic, Corbin Duncan, Will Engelmayer, Karla Flores, Kendrick Foster, Angela Fu, John Gehman, John Gleb, Joolean Hall, Joseph Hasbrouck, Sarah Jacob, Joshua Kelley, Joel Kesselbrenner, Samantha Ku, Jacob Kuckelman, Max Kuhelj Bugaric, Johannes Lang, Steven LeGere, Isabel Levin, Rachel Levine, Daiana Lilo, Angel Lira, Clare Lonergan, Parker Mas, Holly McCarthy, Alexandra McClintock, Sean McGuire, Deja McKnight, Scarlett Neely, Yihao Niu, Denise Nzeyimana, Christine Ow, Clay Oxford, Marc Prophet, Kasey Quicksill, Janna Ramadan, Bianca Rodriguez, Noah Schimenti, Isabel Slavinsky, Bruno Snow, Matthew Stafford, Randle Steinbeck, Claire Sukumar, Ollie Sughrue, Hannah Suh, Maya Taha, Sal Vidal, Matthew Von, Eileen Walsh, and Taylor Whitsell for research assistance. Authors listed in alphabetical order.

[†]Assistant Professor, Department of Political Science, Brown University. Email: tyler.jost@brown.edu. Web: <http://www.tylerjost.com>

[‡]Professor of Government, Department of Government, Harvard University. 1737 Cambridge St, Cambridge MA 02138. Email: jkertzer@gov.harvard.edu. Web: <http://people.fas.harvard.edu/~jkertzer/>

[§]Assistant Professor, Department of Political Science, University of California, Los Angeles. Email: eric.min@ucla.edu. Web: <http://www.ericmin.com>

[¶]Assistant Professor, Department of Political Science, University of Nebraska–Lincoln. Email: robert.schub@unl.edu. Web: <http://www.robertschub.com>

The past decade has seen a resurgence of interest in the study of leaders in International Relations (Saunders, 2011; Chiozza and Goemans, 2011; Croco, 2011; Jervis, 2013; Rathbun, 2014; Weeks, 2014; Kertzer, 2016; Renshon, 2017; Whitlark, 2017; Yarhi-Milo, 2018; Holmes, 2018; Horowitz et al., 2018; McManus, 2021). Who assumes the role of president, prime minister, or dictator dramatically shapes the foreign policies a state chooses, in large part because leaders systematically vary from one another: in their experiences before entering office (Horowitz, Stam and Ellis, 2015; Fuhrmann, 2020); in their personalities, leadership styles, and operational codes (Kaarbo, 1998; Hermann and Preston, 1994); and in their traits or dispositions like hawkishness (Kahneman and Renshon, 2007; Yarhi-Milo, Kertzer and Renshon, 2018; Carter and Smith, 2020). Leader characteristics, this body of scholarship has argued, are critically important to understanding when states are prone to interstate conflict.

One of the critiques of the study of leaders in IR, like the study of political psychology more generally, concerns the problem of aggregation: although leaders sometimes act alone, many of the most important decisions in foreign policy are made in small groups (Hafner-Burton et al., 2017). History is replete with images of advisers at their leaders' side at critical moments: Otto von Bismarck during the Franco-Prussian War, Vo Nguyen Giap during Dien Bien Phu, John Foster Dulles during the Berlin Crisis, Moshe Dayan during the Six Days War, P.N. Haksar during the Bangladesh War, and so on.

Do advisers systematically shape foreign policy behavior and, if so, how? Whether in reference to the "best and brightest" in Vietnam, or the "Vulcans" who advocated for the invasion of Iraq, many popular accounts of foreign policy decision-making suggest that critical foreign policy choices often hinge on whether a leader's inner circle is filled with hawks or doves (e.g., Halberstam, 1972; Mann, 2004). Yet much of the academic literature in IR has relatively little to say about the role of adviser dispositions in foreign policy behavior. The new wave of first-image IR scholarship predominantly focuses on leaders rather than those that counsel them. In foreign policy analysis (FPA), there exists a vibrant literature on small group decision-making, but it tends to study advisers situationally rather than dispositionally, focusing on the quality of the group's advisory *process* and how leaders and advisers relate (e.g., 't Hart, Stern and Sundelius, 1997; Schafer and Crichlow, 2010; Mintz and Wayne, 2016; Saunders, 2017), rather than on advisers' predispositions themselves. This is partly for substantive reasons. Foreign policy decision-making is hierarchical:

leader dispositions tend to dominate those of advisers, and leaders can also appoint advisers with similar worldviews or disregard advice incongruent with their prior beliefs. The asymmetry in focus also stems from methodological considerations: thanks to large-scale data-collection efforts (Chiozza and Goemans, 2011; Horowitz, Stam and Ellis, 2015), we have excellent data on leader-level characteristics, but relatively little data on adviser-level ones.

In this article, we develop a theory of foreign policy decision-making that emphasizes adviser characteristics. Unlike recent scholarship that tends to assume that leaders dominate decision-making groups, we argue that a leader’s foreign policy beliefs are less well-defined than commonly assumed. Instead, leaders frequently use deliberations to search for information and recommendations from their advisers. Since advisers’ perspectives are shaped by their predispositions, deliberation provides a pathway for adviser characteristics to affect foreign policy behavior by shaping the information and problem representations available to leaders.

Our contribution is not simply to develop a theoretical logic by which adviser dispositions nudge states towards conflict or cooperation. We also leverage big data and machine learning techniques to offer systematic and high-resolution empirical tests on whether and how advisers shape foreign policy. First, we collected, digitized, and processed the transcripts of 2,881 US foreign policy decision-making meetings from 1947 to 1988. We compiled the records through a combination of in-person collection at seven libraries and archives, as well as from online repositories. These include all available meeting records of the US National Security Council (NSC), as well as 2,088 informal meetings in which presidents discussed foreign policy issues with their advisers. We segmented each transcript by speaker, meaning that our data identify who said what and when they said it. In total, the corpus includes over 100,000 speech segments attributable to a specific president or adviser. We believe this represents the largest segmented corpus of foreign policy deliberations to date in political science.

Second, we introduce a novel machine learning-based approach to measure leader and adviser hawkishness at a distance, based off an original dataset of the biographical characteristics of any individual who participated in our sample of meetings. Our biographical data describe the backgrounds and professional experiences of 1,073 individuals ranging from secretaries of state, to CIA directors, to Pentagon bureaucrats – an adviser-level counterpart to the leader-level datasets that have revolutionized the study of leaders in IR. This innovation allows us to study advisers in US

foreign policy on a far larger scale than has traditionally been possible, and study quantitatively what has traditionally been the preserve of qualitative approaches.

Finally, we manually code over 2,300 decisions reached during these meetings, which ranged from diplomatic cooperation to interstate conflict. These include some of the most important foreign policy decisions in modern American history, such as the decision to blockade Soviet ships during the Cuban Missile Crisis, the decision to bomb North Vietnam during the Gulf of Tonkin crisis, and the decision to engage Mikhail Gorbachev during the Reykjavik Summit.

A series of statistical tests leveraging these data yield two major findings. First, our results show the central importance of advisers in foreign policy decision-making. More hawkish advisory groups are associated with conflictual foreign policies, even after considering several potential pathways for selection effects. Contrary to accounts assuming that decision-making outcomes simply reflect leader characteristics, we find that adviser-level hawkishness appears to have larger and more consistent effects on foreign policy decision-making in the United States than leader-level hawkishness does – and that appointment to and participation in foreign policy groups does not merely mirror the hawkishness of the leader. Second, our results demonstrate how deliberations serve as an important mechanism for adviser characteristics to shape foreign policy outcomes. We show that leaders in the United States consistently solicited information from advisers during meetings, and that advisers used these opportunities to make arguments that emphasize ideas congruent with their dispositions, rather than just their institutional roles. Collectively, the findings suggest that adviser traits matter much more than the existing literature implies.

1 Leaders, Advisers, and Trait Aggregation in Groups

The division between hawks and doves is central to our understanding of why states choose conflict over cooperation (Modigliani, 1972; Herrmann, Tetlock and Visser, 1999; Schultz, 2005; Kahneman and Renshon, 2007). Hawks and doves differ in their beliefs about the nature of international politics, adversary motivations, and the efficacy and appropriateness of using force. Existing research (e.g., Rathbun, 2007; Yarhi-Milo, Kertzer and Renshon, 2018; Lin-Greenberg, 2019) finds that these beliefs shape the foreign policies individuals support. Consequently, knowing an individual’s *general* hawkishness often predicts their propensity to endorse *specific* conflictual

policies.

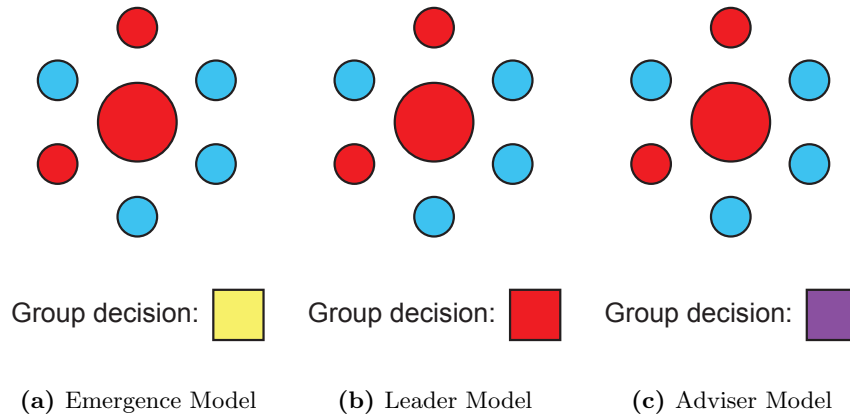
Despite the centrality of hawkishness to our understanding of policy preferences, there is debate about whether and how individual characteristics like hawkishness aggregate to shape state behavior (Hafner-Burton et al., 2017). Most foreign policy decisions occur in settings in which leaders and foreign policy advisers interact. American presidents made key decisions during episodes spanning from the Berlin Crisis to the Persian Gulf War in consultation with advisers ranging from John Foster Dulles to Colin Powell. Prior to the Soviet Union’s invasion of Afghanistan, Leonid Brezhnev conferred with an advisory troika consisting of Defense Minister Dmitry Ustinov, KGB director Yuri Andropov, and Foreign Minister Andrei Gromyko. Even Richard Nixon and Mao Zedong – known for restrictive decision-making styles – routinely consulted with select advisers, such as Henry Kissinger and Zhou Enlai respectively.

Like the leaders they serve, advisers presumably possess stable dispositions including on hawkishness. The central question, however, is how these leader and adviser-level traits aggregate to shape foreign policy outcomes. The existing literature suggests two answers. The first proposes that the dispositions of group members (leaders or advisers) have no bearing on the foreign policy decisions reached – an assumption shared by several disparate theoretical traditions in IR. In the second, individual traits do matter, but because foreign policy is hierarchical, foreign policy decisions simply follow from leaders’ dispositions. We discuss each in turn, before presenting our adviser-centric alternative account.

1.1 The Emergence Model

Several theoretical traditions in IR assume that the dispositions of group members – leaders or advisers – should not aggregate in systematic or predictable ways (Figure 1a). These theories offer markedly distinct justifications for this nonetheless common assumption. First, realist scholarship emphasizes that the structure of the international system creates powerful incentives for states to behave as “unitary actors.” In this view, the domestic features of the state – including the types of individuals constituting decision-making groups – exert little effect on state behavior, particularly compared to structural variables such as polarity, the balance of power, or alliances (e.g., Mearsheimer, 2001). Second, game theoretic scholarship remains skeptical about the challenges of aggregation (Powell, 2017), suggesting that the aggregation of traits in groups is sufficiently complex

Figure 1: Three Models of Aggregation in Groups



Does knowing something about the predispositions of decision-makers in foreign policy tell us anything about the ultimate decisions reached? Emergence models (panel a) argue that even if you could reconstruct the ideal points of leaders' inner circles, group decisions can't be reduced to the traits of the members that comprise them. Leader models (panel b) emphasize the traits just of those at the top. In contrast, our adviser model suggests that adviser predispositions can also shape foreign policy decisions.

so as to warrant a “methodological bet” that they are not worth studying. Third, a body of holist or constructivist scholarship argues that group-level outcomes cannot be reduced to attributes of the members comprising the group (Wendt, 2004). Just as international politics is a complex social system, so too is domestic policymaking (Jervis, 1998). While these intellectual traditions differ dramatically, they share an assumption that when studying state behavior, the individual level of analysis is the wrong place to look: neither leader nor adviser characteristics should neatly map onto foreign policy decisions.

1.2 The Leader Model

A second view of aggregation, evident in much recent work, focuses on the traits of leaders (Figure 1b). Simply put, some leaders are more hawkish than others, and states led by hawkish leaders are more likely to engage in conflictual behavior (Jervis, 2013, p. 165, Carter and Smith, 2020). For example, Yarhi-Milo (2018, p. 82) finds that some American presidents during the Cold War, such as Richard Nixon and Ronald Reagan, tended to exhibit more hawkishness than others, such as John F. Kennedy and Jimmy Carter. Similar studies note the division between hawkish and dovish leaders in other countries as well (e.g., Kennedy, 2011).

The leader model posits that group decisions reflect leader traits for two reasons. First, for-

eign policy decision-making groups tend to be hierarchical: leaders enjoy more substantive and procedural authority than other group members. Leaders – not advisers – make the final decisions regarding whether to wage war, pursue arms, or engage in diplomatic cooperation. Leaders also enjoy considerable discretion in choosing other members of the decision-making group (Horowitz, Stam and Ellis, 2015, pp. 6-7). As Krasner (1972, p. 166) argues, “The President chooses most of the important players and sets the rules.... These individuals must share his values.” Leaders can select and remove advisers, such that adviser dispositions align with the leader’s type. Saunders (2011, p. 26) similarly notes that leaders can shape the decision-making group by “hiring advisers or government officials who share [similar] beliefs.” Byman and Pollack (2001, p. 143) suggest that adviser preferences are often “determined by the leader,” rather than the adviser’s bureaucratic position or worldview. Leaders might also structure the advisory process to afford privileged access to advisers with congruent dispositions, shutting out advisers with dissimilar perspectives.¹ If hawkish leaders surround themselves with hawkish advisers, and dovish leaders with only dovish confidants, the real causal power comes from leader-level traits.

Second, the leader model argues that leader beliefs supersede the information and perspectives that advisers provide. In this view, leaders enter office with fixed preferences and firm beliefs about optimal strategies to achieve them (e.g., George, 1969; Byman and Pollack, 2001; Saunders, 2011). When making decisions under uncertainty, leaders may privilege their own “mental Rolodex” regarding the nature of international politics (Horowitz, Stam and Ellis, 2015, p. 10), placing more emphasis on “vivid, personalized, and emotionally involving” information from first-hand experiences, rather than the abstract intelligence provided by their bureaucratic advisers (Yarhi-Milo, 2014, p. 16). Cognitive barriers, such as the desire for consistency (Jervis, 1976, Chapter 4) and motivated reasoning (Kertzer, Rathbun and Rathbun, 2020), may also cause leaders to discount information from advisers that is incongruent with prior beliefs and worldviews. Hawkish leaders may prioritize information from hawkish advisers and discount information from dovish confidants – and vice versa for dovish leaders. As such, advisers are most “influential” when the information they provide is congruent with what the leader already believes.² But if the only

¹Even when leaders seek heterogeneity in adviser dispositions, they may calibrate the overall level of hawkishness in their advisory team to match their own.

²For a similar argument regarding a leader’s beliefs about fighting for reputation, see Yarhi-Milo (2018, pp. 36-38).

adviser information that matters is that which confirms the leader’s prior beliefs, the effects of adviser dispositions are once again epiphenomenal to those of leaders themselves.

Perhaps because of the presumed importance of leaders, far more of the existing empirical literature focuses on leaders rather than their advisers (Chiozza and Goemans, 2011; Horowitz, Stam and Ellis, 2015; Kertzer, 2016; Carter and Smith, 2020). Qualitative approaches to studying elite decision-making often consider advisers chiefly to illustrate the leader’s importance by providing a counterfactual of what a different individual might have done in the same situation (e.g. Jervis, 2013, p. 149; Yarhi-Milo, 2018, p. 42). This asymmetry in focus is also likely a function of methodological considerations: in quantitative IR, we have excellent datasets of leader-level background characteristics, but as of yet, nothing comparable for adviser-level characteristics. Experimentalists in IR have begun conducting elite experiments (e.g. Yarhi-Milo, Kertzer and Renshon, 2018; Tomz, Weeks and Yarhi-Milo, 2020), but given the challenges associated with hauling leaders into the lab, these studies always study decision-makers as individuals, rather than as accompanied by advisory groups. It is this imbalance we seek to rectify.

1.3 The Adviser Model

In contrast to the emergence and leader-centric views that much of the existing literature espouses, we develop a model of aggregation emphasizing advisers’ dispositions. Advisers, like leaders, possess stable and well-defined foreign policy predispositions (e.g., Holsti and Rosenau, 1984). For example, advisers like Curtis Lemay and Donald Rumsfeld are often identified as more hawkish, while advisers like Cyrus Vance and George Ball are generally considered more dovish. In this adviser-centric aggregation model, as Figure 1c shows, foreign policy decisions are driven by the predispositions of advisers participating in the decision.

Why can individuals who are formally subservient to the leader nevertheless shape foreign policy behavior? The core intuition of the adviser model is that leaders need help making sense of the world. International politics is characterized by complex problems in which the nature, potential solutions, and even the optimal outcome are fundamentally unclear (Sylvan and Voss, 1998). Before making foreign policy decisions, leaders must determine the type of situation they face, what information, interests, and norms are pertinent, and adjudicate between conflicting accounts. Sometimes, even questions as mundane as “What happened?” are not straightforward.

During the Gulf of Tonkin crisis in August 1964, Lyndon Johnson and his advisers struggled for hours to determine whether North Vietnam had conducted a second attack on the USS Maddox.

In international politics, however, the search for credible information is costly. Collecting and processing quality information requires more time and resources than a typical leader possesses. Foreign policy is only one of many issues for which leaders are responsible. Leaders need help filling their gaps in understanding ([Schub, 2021](#)). It is here that advisers are useful to leaders. By looking to advisers, leaders increase their capacity to make sense of the international environment ([Jost, 2021](#)). In this regard, leaders are no different than most human beings who look for credible sources when facing constraints on their capacity to collect, review, and evaluate vast volumes of information.

Deliberations serve as a critical forum where a leader's search for information and the dispositions of advisers intersect. Leaders use deliberations to solicit input from advisers, who offer perspectives that are congruent with their hawkish and dovish predispositions. The complexity of international politics means there are numerous ways to view the problems states face, such as interpreting a signal or assessing the material balance of power. Consider the task of ascribing meaning to an adversary's military mobilization. Hawks in the debate may introduce arguments that focus on threatening aspects of the situation, whereas doves may raise counterarguments that float alternative explanations, downplaying the threat to the state's security ([Yarhi-Milo, Kertzer and Renshon, 2018](#)).³

Which advisers participate in policy deliberations thus shapes the information leaders use to make foreign policy decisions. When making decisions, the perspective upon which the leader draws depends on whether hawkish or dovish advisers dominate the debate. Particularly when they disagree, advisers use arguments and counter-arguments ([Risse, 2000](#)) to compete with one another over the direction they want leaders to take. When hawkish dominate deliberations, for example, it drowns out the arguments of dovish advisers – and vice versa. How the debate unfolds might even affect which parts of a leader's mental map are activated. Whether a leader's heuristics for how to cope with threats is activated, for example (e.g. [Landau-Wells, 2018](#)), depends on

³Note that because advisers' problem representation stems from their disposition, they do not require strategic motivation to frame problems in different ways. However, our model is also compatible with situations in which advisers strategically opt to emphasize certain topics over others in support of conclusions and policy outcomes congruent with their worldview.

whether advisers present a problem as threatening in the first place. In some instances, input by advisers may lead to genuine persuasion, in which leaders revise their beliefs or even preferences based upon the logical merits and evidence marshaled to support it (Johnston, 2014).

Although decision-making groups are hierarchical, there are several reasons why advisers are rarely dispositional mimeographs of the leaders they serve – and thus do not always interpret issues the same way that leaders would had they collected and processed the information themselves. First, adviser appointment may be affected by multiple considerations, such as a candidate’s education, experience, qualifications, personal connections, and fit for the position (Jost and Kertzer, 2021) – not simply their hawkishness. Past scholarship on cabinet appointments in the United States, for example, contains few references to leaders directly considering a candidate’s hawkishness (Brauer, 1986). Second, as in hiring for any job, there may be adverse selection during the appointment process (Feaver, 2003). Even if leaders prioritize adviser hawkishness, they may struggle to accurately gauge which candidates are hawks and doves. Third, leaders often prefer viewpoint diversity in their advisory group to improve the quality of foreign policy debates by considering multiple perspectives (Mintz and Wayne, 2016).

The adviser model suggests three observable implications. First, most generally, we would expect that the characteristics of advisers participating in deliberations affects foreign policy outcomes. Second, our expectation that advisers shape group outcomes via deliberation rests on the assumption that leaders and advisers routinely search for information during deliberations by asking questions, soliciting advice, and requesting clarification. Third, if deliberation provides a forum for adviser predispositions to shape foreign policy outcomes, hawks and doves should raise systematically different types of considerations at these meetings. We evaluate the empirical validity of each contention below.

Our theoretical model differs from existing approaches to the study of advisers in IR, which tends to study advisers situationally rather than dispositionally. The bureaucratic politics literature argues that adviser recommendations are a function of their institutional affiliation, such that where advisers stand depends on where they sit (Allison and Zelikow, 1999). In contrast, we argue that advisers have predispositions about the use of force that may be shaped by their experiences in different institutional positions, but are not reducible to them. Congruent with recent scholarship finding that adviser recommendations do not always align with bureaucratic affiliation (Schub,

2021), our model posits that adviser dispositions dominate their situational incentives to conform to organizational interests. Similarly, a rich literature studies advisory systems from the perspective of institutional design, showing that some decision-making processes (’t Hart, Stern and Sundelius, 1997; Schafer and Crichlow, 2010; Mintz and Wayne, 2016), decision-making units (Hermann, 2001; LeVeck and Narang, 2017), and institutions (Reiter and Stam, 2002; Sechser, 2004; Weeks, 2014; Jost, 2021) yield more accurate assessments, less biased information provision, and more effective policy outcomes. In contrast, while we control for institutional variables in the analyses, our interest is in advisers’ predispositions themselves, rather than the institutional configurations in which they are embedded. Finally, in contrast to Saunders (2018), who focuses on how advisers shape foreign policy decisions through their ability to go public, our findings point to an additional mechanism of adviser influence: deliberations in settings where leaders lack information needed to make decisions, such that they look to advisers to provide perspectives that fill their informational gaps.⁴

2 Research Design and Data

To adjudicate between these three models of decision-making, we systematically collected and analyzed a large set of archival records documenting meetings in the United States from 1947 to 1988 in which the president discussed foreign policy issues with advisers. We first used these archival records to identify the participants in foreign policy decision-making. We then measured participant hawkishness (our explanatory variable) from a distance using a novel methodological approach and recorded whether conflictual or cooperative decisions were reached during the meeting (our outcome variable). All data were collected specifically for this study. We describe each step and provide additional details in the Appendix. Finally, since neither adviser selection nor meeting participation is random, there are important inferential challenges inherent in observational studies of the role of advisers in decision-making. We discuss how we deal with these challenges in our empirical strategy and presentation of results.

⁴The predictions of our model also deviate in one important respect. In Saunders’s fire alarm model, an adviser’s influence depends on the mismatch between their ideal point and their leader’s reputation: Democratic presidents are particularly sensitive to critiques from hawkish advisers, and Republican presidents from dovish advisers. In our model, hawkish and dovish advisers can influence hawkish and dovish leaders equally.

2.1 Identifying Group Participants and Decisions using Archival Records

2.1.1 Archival Record Collection, Processing, and Segmentation

There are a wide number of contexts we could use to study trait aggregation in foreign policy, but we focus here on the United States: as a global hegemon with the largest military budget in the world, the US represents a substantively important case. Crucially, the US maintains an unusually well-kept set of historical records from 1947-1988 of both formal and informal meetings, which we assembled from two types of sources. First, a team of research assistants photographed over a half million pages of records in six presidential libraries, the US National Archives II in College Park, and several other print and digitized resources pertinent to foreign policy decision-making across eight presidential administrations from Harry Truman to Ronald Reagan.⁵ Second, using an automated text scraping protocol, we collected all records included in the *Foreign Relations of the United States (FRUS)* archival database. Using a combination of automated filtering and manual review by research assistants, we extracted all meeting records that met two scope criteria: presidential participation and no participation by foreign leaders (i.e. no diplomatic exchanges).⁶ Note that this excludes lower-level meetings between advisers in which the president did not participate.

As Table 1 summarizes, the collection process yielded 2,881 meeting records. 793 (28% of the sample) of the meetings were formal meetings of the NSC. An additional 2,088 (72% of the sample) were informal meetings.⁷ The inclusion of informal meetings in our sample is particularly important because not all presidents have used the NSC in the same way (Jost, 2021). Some informal meetings featured only the president and a single adviser, while others included dozens of bureaucratic officials. We believe these data constitute the most extensive and complete collection to date of foreign policy meeting records in any country – although we do not claim that our sample encompasses all meetings the president attended. We discuss the potential ways that classified records (particularly for the Carter and Reagan administrations) and the *FRUS* editorial process might affect our meeting sample in Appendix §1.2. The comprehensive nature of these data is essential to avoid drawing inferences based upon a comparatively small section of the complete archival record (Katagiri and Min, 2019; Carson and Min, 2021).⁸

⁵See Appendix §1 for details.

⁶See Appendix §1.2 for details.

⁷Appendix §1.3 disaggregates the sample by formal and informal meetings.

⁸We estimate that our sample includes 81% of the NSC meetings convened during this period. This includes

Table 1: Meeting Records, Speech Acts, and Decisions by Administration

Administration	Records	Speech Acts	Decisions	Conflictual	Cooperative
Truman	237	3,839	103	89	14
Eisenhower	944	30,000	422	331	91
Kennedy	269	15,453	128	104	24
Johnson	546	13,184	95	67	28
Nixon	491	30,672	91	56	35
Ford	255	8,877	56	20	36
Carter	54	1,993	24	13	11
Reagan	85	3,174	63	45	18
<i>Total</i>	2,881	107,192	982	725	257

Research assistants used optical character recognition software to convert the photographs of meeting records into digital text, manually correcting text recognition errors. We then split meeting records into what we call “speech acts” – the uninterrupted words spoken by a single individual during a meeting. Our sample of meetings yielded 107,192 speech acts made by hundreds of meeting participants. We then use these meeting records to generate our outcome measure (the decision reached in each meeting), and use these participant lists to generate our predictor variable (the hawkishness of each participant in the deliberations). We discuss each in turn.

2.1.2 Outcome Variable: Measuring Conflictual and Cooperative Decisions

Our outcome of interest concerns policy choices aimed toward US adversaries. To construct the outcome variable, a team of research assistants identified and classified all substantive decisions made within these meetings – thereby avoiding the truncation bias implicated by studies of decision-making in IR that focus only on major uses of force (Mitchell and Moore, 2002). To qualify as substantive, a decision must have presidential approval and plausibly observable ramifications for US foreign policy. Examples of substantive decisions include authorizing an increase in military spending, accelerating arms testing, moving military personnel or assets, altering strategic priorities, pushing for diplomatic engagement, and crafting language for public statements. Decisions that would not qualify are those that merely note the policy preferences that meeting participants adopted, call for additional study of a topic, or establish a committee to set policy in the future.

For each substantive decision, coders collected contextual background information, classified

approximately 195 meeting records omitted from *FRUS*, as well as about 400 records for which *FRUS* provides only partial excerpts.

the decision into one of several categories, and specified a target of the decision.⁹ The pertinent classification categories were *conflictual acts* which could be verbal or material and span from making threats to using force, and *cooperative acts* which could similarly be verbal or material and span from conveying agreement to providing aid.¹⁰ Targets of the actions were the state or political organization, such as a rebel group, most directly affected by the decision.¹¹ Since the effects of hawkishness are linked to the treatment of adversaries in particular, rather than allies or neutral entities, the analysis focuses on *adversary* targets, with an entity’s classification in this category potentially varying across time depending on the state of bilateral relations.

As the final three columns of Table 1 report, our sample yielded 982 decisions made across formal and informal meetings, 725 of which are conflictual decisions and 257 of which are cooperative decisions toward an adversary.¹² We operationalize the outcome variable of policy decisions in two ways. The first is a meeting’s raw number of conflictual decisions toward adversaries. Because this is a count variable, the corresponding analyses use Poisson regressions.¹³ The second measure accounts for both conflictual and cooperative decisions by subtracting the latter from the former. Positive values indicate a meeting that produces more conflictual decisions than cooperative ones. We use ordinary least squares regressions to analyze this variable. Distributions of these two outcomes are reported in Appendix §2.3.

2.2 Explanatory Variable: Measuring Hawkishness with Biographical Data

To test the hypotheses, we need a measure for leader and adviser hawkishness. A major methodological challenge to the study of elite decision-making is that researchers lack detailed information on the numerous individuals, most of them advisers, in decision-making groups. At present, there is no comprehensive data on adviser characteristics comparable to leader biographical data (e.g., Chiozza and Goemans, 2011; Horowitz, Stam and Ellis, 2015) to enable researchers to study advisers in a nomothetic way. Moreover, even when researchers are able to identify *which* advisers

⁹When the text of the decision itself proved insufficient, as it often did, coders used the full meeting record and contemporaneous policy papers to clarify the decision’s context, nature, and target.

¹⁰Our categories are similar in spirit to the event type categorizations in Goldstein (1992). Minor differences, as well as how we deal with triangular relationships, are discussed in Appendix §2.

¹¹Given the Cold War context in which these meetings took place, “Soviet Bloc” typically served as the target for military spending adjustments, or those strategy changes without an explicitly identified target.

¹²This does not include other cooperative decisions the United States made towards allies or neutral countries, which intuitively constitute the bulk of US cooperation during the Cold War.

¹³Our results also hold using negative binomial models reported in Appendix §4.4.

participate in decision-making, they lack a stable measure to estimate adviser traits and dispositions, such as hawkishness, at a distance. Researchers are often only able to identify an individual as a hawk or dove by observing the position that an adviser takes on a particular issue (e.g., [Feaver, 2003](#)).

We introduce a two-pronged methodological innovation to address this challenge. It pairs systematic coding of the biographic characteristics of presidents and advisers with past surveys administered to real foreign policy elites during the Cold War. This allows us to estimate the hawkishness of elite decision-makers at a distance, without inferring hawkishness from the behavioral outcomes we are using it to explain.

2.2.1 Coding Biographical Characteristics of US Decision-Makers

Ideally, we could administer surveys to presidents and advisers who participated in US foreign policy meetings during the Cold War. Since this is impossible, we adopt a biographical approach, building on work using policymakers' background characteristics as a proxy for their unobservable traits ([Rathbun, 2014](#); [Fuhrmann and Horowitz, 2015](#); [Horowitz, Stam and Ellis, 2015](#); [Kertzer, 2016](#); [Carter and Smith, 2020](#)).

To start, we identify the 1,073 presidents and advisers who spoke at least once during the meetings in our sample. Each segmented speech act discussed above is attributed to one unique speaker. We collected information on the backgrounds and careers of these speakers, ranging from cabinet secretaries, to senior bureaucratic officials (e.g., assistant and deputy secretaries), to mid-level bureaucrats in the State Department, Pentagon, Central Intelligence Agency, and other government agencies.¹⁴ Our codings focused on two aspects of the individual's background. First, we coded their position and the dates on which the position was held. Second, we recorded an array of demographic variables that might affect hawkishness, including gender, birth year, education level, and political party, as well as diplomatic, intelligence, or military experience. Following [Horowitz, Stam and Ellis \(2015\)](#), we also coded combat experience by identifying deployment to a combat theater during a war involving the US.¹⁵ [Table 2](#) provides the coding for Henry Kissinger.

¹⁴Information collection and coding procedures are detailed in [Appendix §3.1](#).

¹⁵We are able to match a speaker name with match a speaker name and position with 106,055 speech acts out of a total of 107,192 speech acts in our dataset – almost 99%. Several hundred of these missing values, however, are explicitly listed as “Unidentified” in the record and are therefore impossible to match.

Table 2: Example Coding for Henry Kissinger

<i>Position #1</i>	National Security Advisor
· <i>Start/End Dates</i>	1/20/1969 to 11/3/1975
· <i>Bureaucratic Affiliation</i>	NSC
<i>Position #2</i>	Secretary of State
· <i>Start/End Dates</i>	9/22/1973 to 1/20/1977
· <i>Bureaucratic Affiliation</i>	State
<i>Gender</i>	Male
<i>Birth Place</i>	Furth, Germany
<i>Date of Birth</i>	5/27/1923
<i>Highest Education</i>	PhD
<i>Prior Experiences</i>	Military: 1943-1946 Diplomatic: 1973-1977 Intelligence: None
<i>Party ID</i>	Republican
<i>President Party ID</i>	Republican

2.2.2 Imputing Decision-Maker Hawkishness with Elite Surveys from the Cold War

While we hope our biographical dataset of American foreign policy decision-makers will be useful to researchers in and of itself, our methodological innovation is to use machine learning approaches to measure adviser hawkishness at a distance. Like [Carter and Smith \(2020\)](#), we incorporate machine learning approaches on biographical codings; our novel contribution is to anchor our measures using information from Cold War-era surveys of American foreign policy elites administered through the Foreign Policy Leadership Project (FPLP) ([Holsti and Rosenau, 1984](#)). Crucially, FPLP surveys included a battery of items measuring respondents’ levels of militant internationalism, a standard measure for hawkishness in the public opinion literature (e.g., [Herrmann, Tetlock and Visser, 1999](#); [Holsti, 2004](#)), as well as demographic questions that mirror those coded in our biographical dataset.

We use this overlap to estimate the hawkishness of meeting participants in our sample based upon the hawkishness of survey respondents in the FPLP with similar biographical characteristics. Our measurement strategy proceeds in three steps. First, we create a measure of hawkishness averaging across a 15-item battery of FPLP questions that tap into respondents’ views towards containing communism using force, prioritizing offensive military action over diplomacy or defensive measures, believing that the American effort in Vietnam was too limited, and so on.¹⁶ Second,

¹⁶See Appendix §3.2 for details.

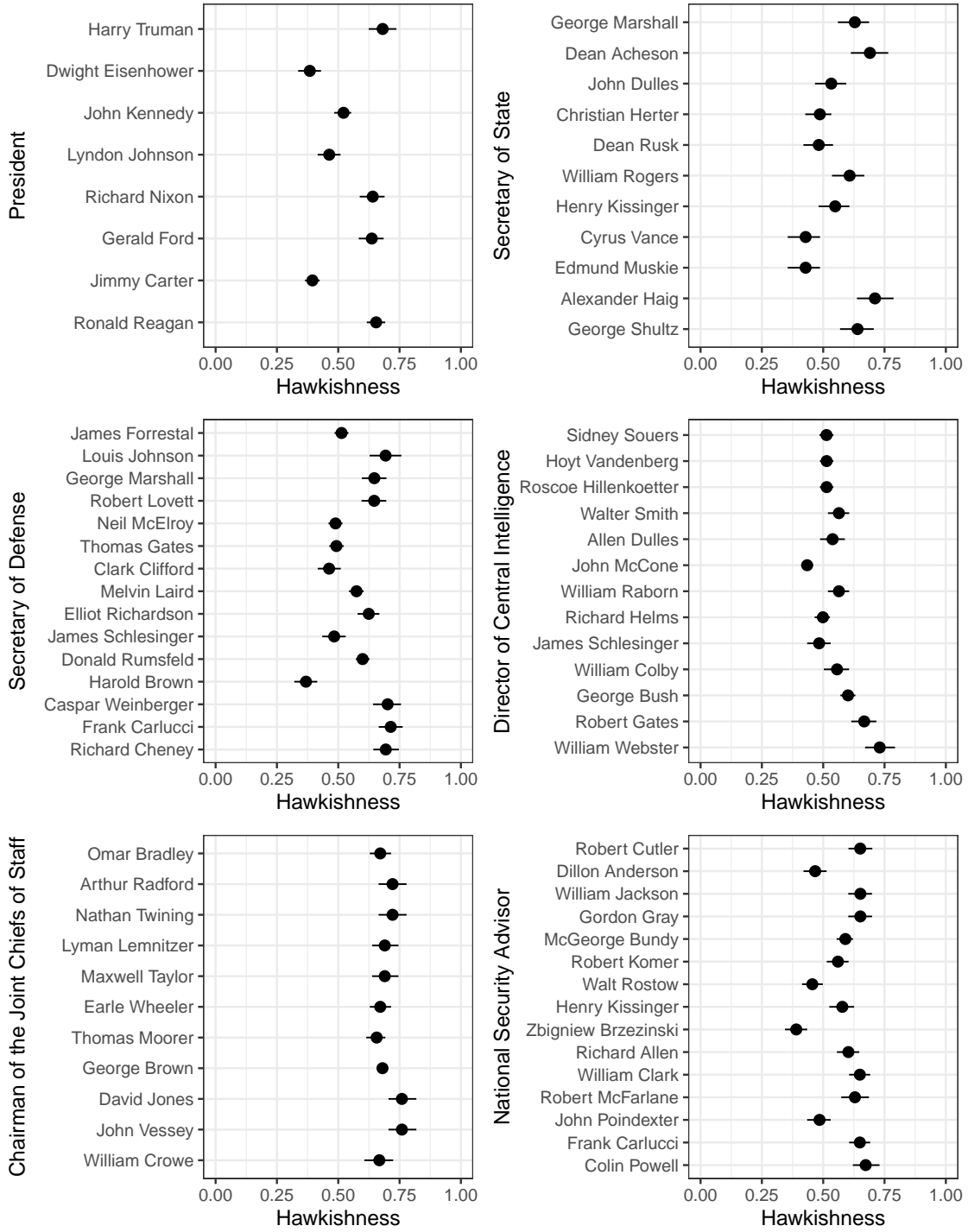
we identify the individual-level characteristics that exist in both the FPLP and our biographical data. These include *gender, birth decade, level of education, military experience, combat experience, diplomatic experience, current military officer, current foreign service officer, and political party*.¹⁷ Third, we harness a series of supervised learning models to adduce the relationship between an individual’s biographical characteristics and their hawkishness in the FPLP. We adopt a boosted linear regression model – a form of ensemble learning where many simple linear models are sequentially trained and reweighted until a final model is established – as our primary method of estimating participant hawkishness. The models are fed the FPLP data, which provides explicit information the model can process to understand the relationship between biographical characteristics and individuals’ level of hawkishness according to their survey responses. The model then uses this information to predict hawkishness on new data, which in our case is the full set of presidents and advisers who participated during meetings in our sample. We use a bootstrapping process through which we randomly sample the FPLP data with replacement 1,000 times. This generates 1,000 predicted hawkishness measures for each individual, and we use the average as an actor’s measure of hawkishness.

In these models, one important predictor of hawkishness is an individual’s party affiliation. However, we know that the Republican and Democratic parties changed their stances in foreign policy issues during the early Cold War (Fordham, 2007). Democrats went from being more hawkish to dovish, while Republicans did the opposite, leading to partisan positions that are more broadly familiar to us today. If we ignored this shift, our measures would underestimate the hawkishness of Truman-era Democrats and overestimate the hawkishness of Eisenhower-era Republicans. To address this issue, we make temporal adjustments to our hawkishness measure that compensate for the shift in party platforms. Specifically, we use longitudinal measures of partisan hawkishness assembled by Jeong (2018) to make time-conditional adjustments to the estimated coefficients for hawkishness of senior meeting participants.¹⁸ This adjustment produces hawkishness measures that

¹⁷As a robustness check in Appendix §4.2, we remove individuals’ current affiliation from our statistical models; our results remain the same.

¹⁸For each individual, we ascertain the administration in which they served, the position that they held, and their partisan affiliation. If the individual is a Democrat or Republican and was a senior official (working at the deputy secretary level or higher), we adjust their hawkishness measure using Jeong’s partisan hawkishness data. Our adjustment used data from the midpoint year of the administration in which the individual served. For individuals with no known party affiliation, we use the raw prediction from the boosted linear model.

Figure 2: Predicted Hawkishness Measures for Senior US Decision-Makers.



Note: Predicted measures of hawkishness with 95% bootstrapped confidence intervals.

Figure 3: Average Speaker Hawkishness in US Foreign Policy Meetings



align more closely with historical assessments of leader and adviser hawkishness.¹⁹

Figure 2 illustrates the predicted hawkishness measures for six senior positions in US foreign policymaking – the president, secretary of state, secretary of defense, the Central Intelligence Agency (CIA) director, Chairman of the Joint Chiefs of Staff (the president’s senior military adviser), and the national security advisor – sorted in chronological order. Chairmen of the Joint Chiefs are generally more hawkish than Secretaries of State, but crucially, some Secretaries of State (like Alexander Haig) are more hawkish than others (like Cyrus Vance). Figure 3 displays group hawkishness at the meeting level over time by calculating the average hawkishness of all meeting participants.

2.2.3 Control Variables: Addressing Adviser Meeting Participation

One challenge with studying the effects of adviser hawkishness on foreign policy decision-making is that adviser participation in meetings is not randomly assigned. The models testing the emergence, leader, and adviser hypotheses include control variables meant to address potential confounding factors.

One threat to inference is that individuals appointed to the advisory team might reflect the leader’s preferences regarding the use of force. To address this issue, we include *administration*

¹⁹As a robustness check, we also use a neural network model that minimizes errors in the predictions of new observations (cross-validation error).

fixed effects across several specifications to account for unobserved invariant components of each administration, particularly those that may have prompted the leader to choose a hawkish advisory team – or use advisers and advisory institutions in systematically different ways (Hermann and Preston, 1994; Jost, 2021). Models with fixed effects study the effect of group composition while holding the leader fixed, which controls for these predilections. We further probe the question of adviser appointment below.

Another threat to inference concerns which advisers are invited to participate in which meetings. Imagine a topic on which the leader is inclined to authorize conflictual decisions. Based on that inclination, the leader could skew meeting invitations toward hawkish advisers. Leader inclination would confound the relationship of interest because it is a common cause of group hawkishness (the explanatory variable for the adviser hypothesis) and policy decisions (the outcome variable). To address this selection process in our models, we manually coded the *agenda topic* for each of the 107,192 speech acts in our sample.²⁰ Using these classifications as a control variable helps minimize potential bias in the meeting’s invitees.

A third potential concern is that advisers may be predisposed to participate when they anticipate positive reactions to their worldview. For instance, hawkish advisers may speak more when the international environment (e.g. recently being attacked) makes the state predisposed to pursue conflictual policies. To address this, we include a set of system- and leader-level control variables that may have motivated hawkish advisers to speak more. Following existing research, these include: a variable measuring the *lagged number of militarized interstate disputes* challenging the US in the last five years, as well as the *national capabilities* to measure military strength and economic health.²¹ Below, we further probe features of the deliberative environment that affect adviser incentives to speak during meetings.²²

Additional control variables track the proportion of meeting participants affiliated with the *Defense Department*, *State Department*, *intelligence agencies*, and *military*, whose bureaucratic interests may skew hawkish or dovish (Allison and Zelikow, 1999) – as well as the (logged) total

²⁰See Appendix §7 for more information on agenda items and their impact on group composition.

²¹These controls follow those in Horowitz, Stam and Ellis (2015), dropping domestic characteristics such as regime type and polity score that do not vary for the United States in our time period. We omit characteristics such as a leader’s age, time in office, and prior military experience because they were attributes used to develop predicted levels of hawkishness for each president.

²²Additionally, Appendix §6.3 shows that advisers do not self-censor when they and the president differ widely in their hawkishness.

years of *military*, *diplomatic*, and *intelligence* experience of meeting participants. Another control captures the *number of attendees* in each meeting (LeVeck and Narang, 2017) and a binary variable indicates whether a meeting was a *formal* meeting of the NSC, as opposed to an informal session outside it. We discuss numerous robustness checks we perform to address the potential for selection effects in Section 3.1.1 below.

3 Results

We present our results in two steps. First, we test our three theorized conceptions of how the dispositions in the group affect foreign policy decisions: an emergence model in which individuals' traits are unconnected to outcomes, a leader model, and an adviser model. Second, we evaluate assumptions and test our hypotheses regarding the mechanisms by which advisers shape group outcomes by exploring the rich textual content of the deliberations in our sample.

3.1 Testing the Three Models of Aggregation

Do the dispositions of group participants help explain the policies ultimately enacted? The emergence model predicts that dispositions of group members should have no systematic effect on policy decisions, whether because structural characteristics of the international system dominate or because group decisions cannot be reduced to individual-level traits. To test this model, we begin with the simplest aggregation procedure: the mean level of hawkishness of all speakers in the meeting.²³ If consequential policy choices are not reducible to the traits of individuals involved in making those choices, as the emergence model suggests, then we ought to observe no relationship between the group's average hawkishness and its policy decisions.

Inconsistent with emergence models, we find that a rise in group hawkishness increases conflictual policy choices toward adversaries. The relationship holds across different specifications as Models 1 and 2 of Table 3 show. Regardless of outcome variable specification or the presence of control variables, the group's average hawkishness consistently has a positive effect on conflictual policy decisions. The left-hand panel of Figure 4 presents the results graphically. Shifting the

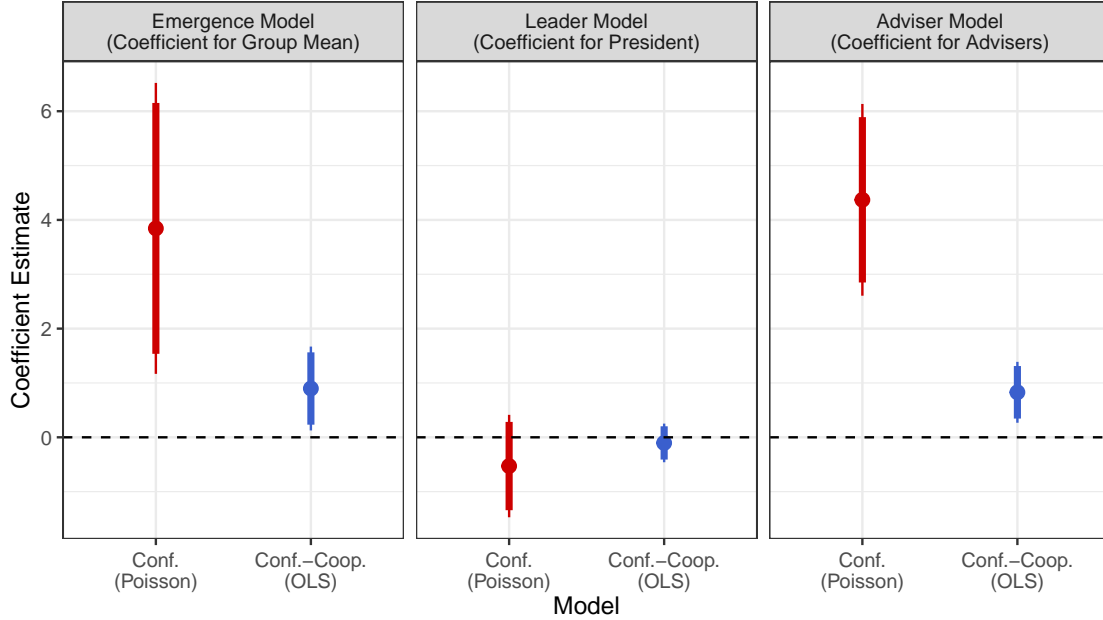
²³Meeting participant here refers to advisers who attend the meeting and speak at least once. We adopt this definition for practical reasons: the full list of meeting attendees is not available for all meetings.

Table 3: Effect of Participant Hawkishness on Foreign Policy Decisions

	<i>Emergence Model</i>		<i>Leader Model</i>		<i>Adviser Model</i>		<i>Advisers + Admin. FEs</i>	
	Conf	Conf. – Coop.	Conf.	Conf. – Coop.	Conf	Conf. – Coop.	Conf.	Conf. – Coop.
	<i>Poisson</i>	<i>OLS</i>	<i>Poisson</i>	<i>OLS</i>	<i>Poisson</i>	<i>OLS</i>	<i>Poisson</i>	<i>OLS</i>
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Mean Hawkishness	3.844*** (1.366)	0.898** (0.394)						
President's Hawkishness			-0.529 (0.481)	-0.104 (0.181)	-1.557*** (0.522)	-0.356* (0.201)		
Advisers' Hawkishness (Acts)					4.370*** (0.900)	0.828*** (0.286)	3.876*** (1.192)	0.932*** (0.352)
No. of Attendees	0.073*** (0.017)	0.035*** (0.007)	0.086*** (0.016)	0.040*** (0.007)	0.108*** (0.016)	0.046*** (0.007)	0.092*** (0.017)	0.042*** (0.007)
Proportion - President	0.920** (0.460)	0.140 (0.118)	1.131*** (0.418)	0.239** (0.106)	1.536*** (0.440)	0.275** (0.118)	1.476*** (0.479)	0.221* (0.127)
Proportion - Military	1.116** (0.555)	0.137 (0.181)	1.544*** (0.457)	0.328** (0.151)	1.254*** (0.475)	0.232 (0.166)	1.417*** (0.535)	0.172 (0.182)
Proportion - Defense	1.766*** (0.340)	0.163 (0.114)	1.204*** (0.327)	0.078 (0.111)	1.883*** (0.348)	0.184 (0.118)	2.107*** (0.360)	0.233* (0.120)
Proportion - Intelligence	1.600*** (0.552)	0.302* (0.161)	1.312** (0.526)	0.243 (0.160)	2.142*** (0.545)	0.367** (0.169)	2.079*** (0.563)	0.372** (0.169)
Proportion - State	0.282 (0.329)	-0.024 (0.076)	0.171 (0.331)	0.004 (0.075)	0.821** (0.353)	0.092 (0.086)	0.763** (0.356)	0.078 (0.087)
Diplomatic Experience	0.111** (0.045)	0.008 (0.015)	0.083* (0.043)	-0.002 (0.014)	0.082* (0.044)	-0.003 (0.014)	0.106** (0.046)	0.004 (0.015)
Intelligence Experience	-0.131** (0.052)	-0.049*** (0.018)	-0.117*** (0.045)	-0.046*** (0.018)	-0.174*** (0.048)	-0.053*** (0.018)	-0.150*** (0.053)	-0.053*** (0.019)
Military Experience	0.034 (0.079)	-0.029 (0.021)	0.071 (0.071)	-0.032* (0.019)	0.016 (0.075)	-0.046** (0.021)	0.001 (0.079)	-0.044** (0.021)
Formal	0.583*** (0.138)	0.143*** (0.051)	0.779*** (0.129)	0.157*** (0.049)	0.736*** (0.133)	0.151*** (0.051)	0.602*** (0.140)	0.146*** (0.053)
5-Year MID Challenges	-0.244 (0.223)	-0.100 (0.075)	0.016 (0.151)	-0.015 (0.047)	-0.039 (0.154)	-0.023 (0.049)	-0.281 (0.228)	-0.125 (0.079)
US CINC	6.134*** (2.316)	2.790*** (0.889)	2.869*** (1.049)	1.485*** (0.341)	1.238 (1.097)	1.129*** (0.372)	5.828** (2.344)	2.574*** (0.933)
Constant	-6.945*** (1.173)	-1.298*** (0.389)	-3.751*** (0.670)	-0.367* (0.212)	-5.401*** (0.775)	-0.609** (0.242)	-7.140*** (1.141)	-1.252*** (0.384)
Administration FEs	✓	✓					✓	✓
Agenda Items	✓	✓	✓	✓	✓	✓	✓	✓
Observations (Meetings)	2,877	2,877	2,880	2,880	2,704	2,704	2,705	2,705

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

Figure 4: Summary of Three Models of Trait Aggregation



Note: Coefficient plots corresponding to main findings, corresponding to Models 1-6 in Table 3. Thicker and smaller bands represent 90% and 95% confidence intervals respectively.

group’s mean hawkishness from its minimum to maximum values while holding other variables constant more than triples the predicted number of conflictual decisions (based on Model 1).

The results demonstrate a clear, consistent, and substantively meaningful relationship between a group’s composition and its decisions. Moreover, including a measure of aggregated group traits (mean hawkishness) improves the statistical model fit compared to a sparser specification without this measure. A likelihood ratio test that compares Model 1 of Table 3 with a null model that omits the measure of mean group hawkishness strongly indicates that accounting for hawkishness yields a significant improvement ($p = 0.005$). A similar exercise using Model 2 produces an analogous result ($p = 0.022$). The findings thus do not support the emergence model.

Are group decisions simply an extension of the leader’s disposition? The leader model predicts that leader hawkishness should systematically affect policy decisions. Whether due to group hierarchy or a leader’s reluctance to hear other views, we should observe a positive relationship between leader hawkishness and conflictual policy choices.²⁴ Results presented for Models 3 and 4 in Table 3 and the middle panel of Figure 4 provide no support for the leader view. Across

²⁴Because each president’s hawkishness in our data does not appreciably change over time, we omit administration fixed effects.

these specifications, presidential hawkishness has a negligible relationship with conflictual decisions toward adversaries.²⁵ Appendix §5 details a range of potential explanations for why leader-level hawkishness fails to be positively associated with conflictual decisions, ranging from strategic interaction to the uniquely institutionalized context of the US. Moreover, replications and extensions of Horowitz, Stam and Ellis (2015), Yarhi-Milo (2018) and Carter and Smith (2020) in Appendix §5.2 show that our leader-level hawkishness results are consistent with recent work finding little evidence that presidential hawkishness predicts American conflict behavior. We also obtain similar results when measuring presidents' hawkishness using measures derived from codings provided by experts of American foreign policy (Appendix §5.1), as well as using a different operationalization of the dependent variable (Appendix §4.5). While other leader-level characteristics matter in American foreign policy (e.g., McDermott, 2007), and while leader hawkishness matters in other national contexts (e.g., Tomz, Weeks and Yarhi-Milo, 2020), leader-level hawkishness appears to have little effect on American conflict behavior, a point the conclusion addresses further.

The adviser model predicts that the hawkishness of participating advisers affects decision outcomes. If advisers exert influence through communicative acts during deliberation, then meetings in which hawkish advisers speak frequently ought to produce more conflictual decisions. For these models, we calculate a weighted average of adviser hawkishness, where each adviser's weight is a function of the proportion of speech acts they contributed to the discussion, reflecting our emphasis on communication as a vehicle for influence.²⁶

As shown by Models 5 and 6 in Table 3 and the right-hand panel of Figure 4, we find strong evidence that adviser hawkishness affects decision outcomes. Meetings in which more hawkish advisers speak more tend to adopt more conflictual policies toward an adversary. This pattern appears across specifications. Holding other variables fixed, shifting the group's weighted hawkishness from the minimum to maximum values nearly triples the expected number of conflictual decisions (based on Model 5's specification). Models 7 and 8 drop the leader measure so that we can include fixed

²⁵We find this result is largely driven by Dwight Eisenhower, who our measure codes as comparatively dovish but nevertheless presided over an administration responsible for a substantial number of conflictual decisions. Yet as robustness checks in Appendix §5.4 show, we fail to find evidence in favor of leader-level hawkishness even when dropping Eisenhower from the sample or rerunning the model with a neural net measure of hawkishness instead.

²⁶Consequently, meetings with identical lists of participants can produce divergent hawkishness levels. To stress the independent effect of advisers, the weighted average score does not include the president. In fact, some specifications control for the president's hawkishness (and thus do not include administration fixed effects) to assure that advisers exert influence rather than merely serving as proxies for the president.

effects to guard against the possibility that results reflect differences between different presidents' management style and preference for formal (i.e. NSC) or informal meetings.

Reagan administration meetings and decisions in 1981 concerning relations with Poland and the USSR illustrate this finding. Hawkish advisers – such as Secretary of State Alexander Haig and Secretary of Defense Caspar Weinberger – dominated an NSC meeting in late March. The meeting scored in the 95th percentile for adviser hawkishness during the Reagan administration. This hawk-dominated discussion produced conflictual policies toward adversaries with an aggressive diplomatic message sent to the Soviets alongside the continuation of a grain embargo. Haig pushed for the strong statement that would “lay out the repercussion for Soviet actions” while Weinberger “noted the continuing nature of the threat.”²⁷ In contrast, when addressing the same issue area six months later, doves drove the discussion. This meeting scored at only the 25th percentile for adviser hawkishness. Countering Weinberger’s skepticism, the relatively dovish NSA Richard Allen argued that the US must push for continued liberalization in Poland and offer food aid to advance those prospects. In accordance with the adviser model, a dovish advisory group nudged the US toward cooperation in the form of \$50 million in food aid.²⁸

To further contextualize the substantive effects, Table 4 presents the predicted number of conflictual decisions toward an adversary for all fully-specified models. These calculations shift each relevant measure of hawkishness from its minimum to maximum value while holding other variables fixed at their means, while also presenting the substantive effects of other contextual variables to provide a relative benchmark. The table shows the dramatic effect of both the mean and weighted mean measures of group hawkishness, which respectively cast doubt on the emergence model and provides evidence consistent with the adviser model. In the leader versus adviser models, it is worth noting that even though the president and advisers have cross-cutting effects on decision-making, the more conflictual nature of hawkish advisers appears to outweigh the effects of the president, lending further support to importance of advisers. Collectively, the results suggest that the leader model is incomplete, and that we must consider the dispositions of advisers in the room.

²⁷NSC Meeting No. 764; 26 March, 1981.

²⁸NSC Meeting No. 779; 15 September, 1981.

Table 4: Predicted Number of Conflictual Decisions

Model	Model	Variable	Min.	Max.
Emergence	Poisson (Model 1)	Mean Hawkishness	0.119	0.416
	OLS (Model 2)	Mean Hawkishness	0.088	0.380
Leader	Poisson (Model 3)	President’s Hawkishness	0.261	0.223
	OLS (Model 4)	President’s Hawkishness	0.233	0.202
Adviser	Poisson (Model 5)	Advisers’ Hawkishness (Acts)	0.091	0.423
		5-Year MID Challenges	0.221	0.196
		US CINC	0.181	0.229
	OLS (Model 6)	Advisers’ Hawkishness (Acts)	0.042	0.335
		5-Year MID Challenges	0.224	0.178
		US CINC	0.099	0.314

Note: Predicted numbers of conflictual decisions toward an adversary across the range of hawkishness, holding all other variables at mean values, formal meetings, and Eisenhower administration.

3.1.1 Selection and Robustness

Three additional questions regarding selection effects merit consideration. First, above we noted that administration fixed effects help identify the effect of adviser dispositions *within* each administration by holding constant unobserved variables, such as the leader’s preferred advisory arrangements. Yet one potential question is whether this methodological choice masks effects that leaders exert through the appointment process. Leaders make appointments for a variety of reasons, including appointee qualifications, personal connections, and public approval. If leaders only appointed advisers who shared their foreign policy worldview (i.e. hawkish presidents only appointed hawks - e.g. [Krasner, 1972](#)), then the advisory environment – and our results – would simply represent an extension of the leader’s disposition. We find no evidence to support this conclusion. Mixed-effect models that include administration random effects find that the intraclass correlation ranges between 0.10 and 0.16: there is approximately six to ten times greater variation in hawkishness within individual administrations than there is between them. We would expect a far lower figure if hawkish leaders simply hired or invited hawkish advisers into meetings.

Second, even if leaders do not merely select mimeographs as their advisers, leaders may still have the ability to decide when meetings take place and which advisers are invited ([Krasner, 1972](#)), in ways that controlling for meeting agendas might fail to capture. To address this concern, we replicate our results using an alternative model specification that uses our adviser hawkishness

scores but ignores our meeting data altogether, instead examining the effect of adviser-level hawkishness on the United States' propensity for initiating militarized interstate disputes in a given month. Our findings, shown in Table 5 below and discussed in Appendix §4.5 in detail, remain the same, despite a different unit of analysis (the time-unit, rather than the meeting-level), and a more restrictive dependent variable (militarized interstate disputes, rather than all foreign policy decisions). Adviser predispositions appear to be significantly associated with propensity for interstate conflict, even controlling for a host of international, domestic, and leader-level variables.

Additionally, one might wonder whether our conceptualization of adviser hawkishness was specific to the Cold War – perhaps making our findings an artifact of the highly competitive US-Soviet relationship. Two factors discredit such an interpretation. First, the militant internationalism measure we use to impute decision-maker hawkishness has been widely used since 1991 (e.g., [Brutger and Kertzer, 2018](#); [Tomz, Weeks and Yarhi-Milo, 2020](#)). In fact, [Murray \(2002\)](#) shows that hawkish beliefs among American decision-makers were surprisingly consistent before and after the Cold War. Second, in Appendix §4.3, we run a robustness check where we drop decisions involving the Soviet Union from our analysis and find that results remain consistent.

3.2 *How Advisers Shape Decisions*

The findings in Section 3.1 indicate that adviser participation in deliberation matters to the ways states behave. Our argument, however, is more specific: adviser participation matters because of the perspective they provide during deliberations. Content analysis of meeting deliberations allows us to empirically probe two aspects of this contention.

3.2.1 *Leaders Search for Information during Deliberations*

The adviser model of aggregation assumes that group members, particularly leaders, should routinely search for information during meetings and that advisers should express opinions even when others disagree. To evaluate whether this assumption of the deliberation mechanism holds, we generated a stratified random sample of 296 formal and informal meetings across presidential administrations (approximately 10% of the full sample).

Drawing upon the study of deliberation elsewhere in political science, we developed a coding scheme (detailed in Appendix §6.1) to identify speech acts that exhibited *information search*. A

Table 5: Effect of NSC Principals' Hawkishness on MID Initiation, Using Monthly Data

	<i>Dependent variable:</i>	
	US MID Initiation	
	(1)	(2)
Advisers' hawkishness	1.309** (0.663)	2.551** (1.232)
President's hawkishness	-0.496 (0.368)	-1.267* (0.753)
War ongoing		-0.065 (0.048)
Deaths per capita in last war (logged)		0.021 (0.033)
Months since last war (logged)		0.060 (0.049)
Victory in last war		0.005 (0.196)
MID challenges to US in last 5 years		-0.005 (0.021)
Average MID outcome in last 5 years		-0.198 (10.114)
Economic recession		-0.170** (0.070)
Unified government		0.111 (0.078)
US material capabilities		-4.277** (2.070)
President's tenure (logged months)		0.020 (0.029)
Constant	0.714*** (0.247)	1.159 (0.839)
Observations	501	501

Note: * $p < 0.1$; ** $p < 0.05$; *** $p < 0.01$. Outcome variable is the number of MIDs initiated in a given month. Advisers' hawkishness reflects average hawkishness score of the most senior advisers in the administration in a given month. See Appendix §4.5 for details.

research assistant reviewed and coded each of the 11,523 speech acts in the random sample. Information search was defined as a speaker requesting another participant to introduce new ideas, information, or recommendations into the discussion. Simply stating one’s own position does not qualify as information search. Rather, speakers must proactively ask *others* for their perspective. A meeting in which participants stated their own position over and over again in slightly different ways would be coded as having no information search.

We find that US foreign policy deliberations exhibited a high level of information search, particularly at the leader level. About one in three presidential speech acts – and over one in eight adviser speech acts – queried for more information from advisers. This finding is difficult to reconcile with leader-centric views that emphasize fixed and unchanging beliefs of leaders. Presidents expend considerable time and effort soliciting input from advisers. As reported in Appendix §6.1, these patterns generally hold across administrations, with some presidents being more open to information than others (Hermann and Preston, 1994).

One question, however, is whether these queries simply led to the identical viewpoints being repeatedly expressed. To explore this contention, we replicated our coding process for speech acts exhibiting *dissent*, defined as a textual indication that the speaker disagreed another meeting participant.²⁹ The data again suggest that leaders and advisers use deliberations to offer conflicting views: 16% of adviser speech acts and 11% of president speech acts featured a dissenting opinion. In total, 61% of meetings featured some form of debate between participants. Collectively, the findings accord with our adviser model emphasizing that advisers matter because leaders seek input – and participation in group meetings allows advisers the opportunity to argue with one another. Appendix §6.3 uses these data to show that expressions of dissent more frequently come from advisers who differ substantially in hawkishness relative to the president, which allays concerns that our results are driven by advisers engaging in self-censorship.

3.2.2 Advisers Provide Perspectives Congruent with their Predispositions

According to the adviser model of aggregation, hawkish and dovish advisers use deliberations to introduce arguments consistent with their predispositions. To evaluate the deliberation mechanism,

²⁹In some cases, advisers directly identified their dissent (e.g., “disagreed” or “objected”), while in other cases dissent could only be discerned in the context of the meeting.

Table 6: Hawkish and Dovish Topics (keyATM)

Topic	Who Emphasizes?	Source	Distinctive Terms
Military violence	hawks	George (1969); Herrmann, Tetlock and Visser (1999)	air, fight , bomb , plane, troop, divis, enem, air_forc, suppli, admir
International threats	hawks	Russett (1990)	risk , paragraph, cours_action, aggress , alli, threat , council, hostil , nation, joint_chief_staff
Military balance	hawks and doves	Herrmann, Tetlock and Visser (1999)	system, bomber , cru, backfir, test, deploy, icbm, submarin , missil , rang
Diplomacy	doves	Russett (1990)	un lear, treati , peac , public, got, someth, thing, help, right, negoti
Adversary interests	doves	Brutger and Kertzer (2018)	recent, leader, elect, visit, parti, econom, attempt, nation, west, communist

we use the rich textual data contained within the collected records, examining the content of speech acts by hawks and doves during the meetings. Drawing on existing research on hawkishness, we identified ex ante five topical categories summarizing the types of content that hawks and doves should stress during deliberations. First, hawks should emphasize that using violence is an effective and appropriate strategy in international politics (George, 1969; Herrmann, Tetlock and Visser, 1999; Weeks, 2014). In contrast, doves typically believe that the use of force is counterproductive and stress the potential costs of confrontation. Second, hawks should be more likely than doves to emphasize the ubiquity of threats or other “competitive elements” between states (Russett, 1990, p. 516). Third, both hawks and doves are attuned to the military balance of capabilities for different reasons. Hawks should emphasize the importance of primacy in material strength and power (Herrmann, Tetlock and Visser, 1999), while doves should instead emphasize that the balance of power limits the potential payoffs to violence. Fourth, doves should tend to ascribe greater promise to diplomacy and the need for international cooperation (Russett, 1990). Finally, doves should tend to emphasize the importance of viewing international disputes from the adversary’s perspective, recognizing that an adversary may face international or domestic constraints that impede a negotiated settlement (Brutger and Kertzer, 2018). The first three columns of Table 6 summarize the topics that existing literature suggests hawks and doves should emphasize.

To test whether hawkish and dovish advisers exhibit different speech patterns on these topics, we employ a semi-supervised text analysis method called the keyword-assisted topic model (keyATM) (Eshima, Imai and Sasaki, 2020). Keyword-assisted topic models have all of the strengths of traditional topic models – an automated content analysis method that represents text as a mixture of semantically interpretable topics. Crucially, keyATM also allows researchers to specify conceptual topics of interest by providing a set of keywords, while remaining agnostic on other topics in the corpus. We thus specify a set of five to seven keywords for each of our five topics, while allowing the model to identify 20 other topics within the corpus. The analysis applies the model to all speech acts (of at least 15 words) in formal NSC meetings. For each qualifying speech act (39,172 in total) the model assigns topic proportions with higher values indicating a closer alignment with that topic.

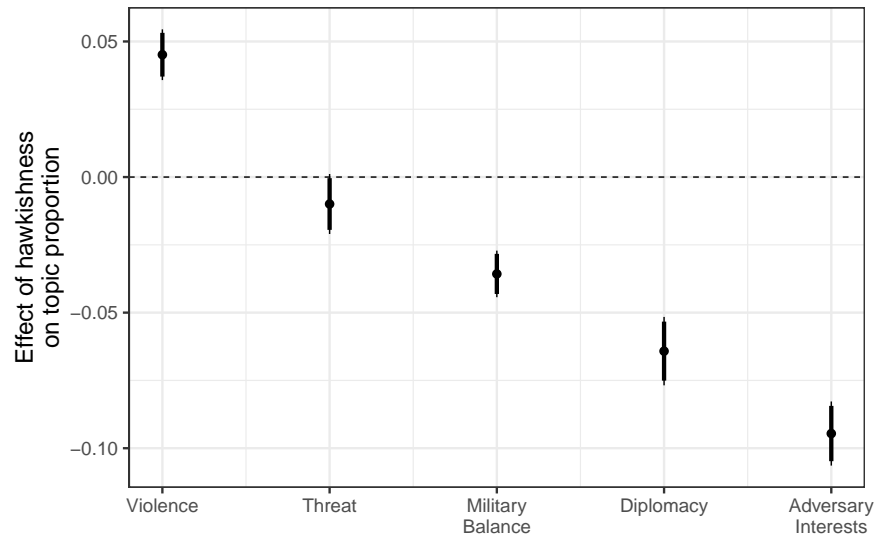
The far-right column of Table 6 demonstrates substantive coherence across the five keyword-assisted topics. It shows the ten word stems that most distinguish each topic from the others, excluding specific individuals (e.g., Khrushchev) or locations (e.g., North Vietnam). Researcher-supplied keywords that proved distinctive are noted in bold text. Topic coherence is generally high, as exemplified by the *military violence* topic, where distinguishing terms are clearly indicative of using force (e.g., “troop”, “fight”). The highest scoring speech act for the topic similarly corroborates the approach: during a February 1968 NSC deliberation, Chairman of the Joint Chiefs of Staff Earle Wheeler stated, “An estimated 50,000 enemy troops participated in the Tet offensive [...] A new attack on Danang is possible. The North Vietnamese Second Division is in the area. More U.S. and South Vietnamese troops are moving in.”³⁰ As Appendix §6.2 details, the other topics display similar measurement validity with only the threat topic being a partial exception.

We find that hawks and doves discuss systematically different considerations, consistent with their underlying predispositions. Figure 5 plots the effect of moving from the least to most hawkish speaker within a single administration on expected topic proportions in a speech act.³¹ The plotted effects are based on OLS specifications that leverage the hawkishness scores from the main analysis above and include administration fixed effects. As shown, hawks and doves differ in the topics they emphasize. Consistent with theoretical expectations, hawks are more likely to address issues

³⁰NSC Meeting No. 581; February 7, 1968.

³¹See Appendix §6.2 for full statistical results.

Figure 5: Effect of speaker hawkishness on topic proportions in speech segments.



Note: Plot shows marginal effect of moving from the least to most hawkish speaker within a fixed administration. Thicker and smaller bands represent 90% and 95% confidence intervals respectively.

related to violence while doves place greater stress on diplomatic possibilities and adversaries' political interests. Beyond statistical significance, these differences are substantively meaningful. An average speech act has an *adversary interests* proportion of 14%. Moving from the least to most hawkish individual reduces the expected proportion by 9% points. Finally, doves are slightly more likely to discuss the *military balance*, while the substantively small relationship between hawkishness and *threat* is null at the 95% confidence level.³²

In sum, analyzing tens of thousands of speech acts in foreign policy meetings offers evidence consistent with the adviser model. Deliberations offer advisers the opportunity to present information and arguments. What topics advisers prioritize during deliberations depend on their disposition. Hawkish advisers make arguments emphasizing military violence, while dovish advisers make arguments emphasizing the limits of power, utility of diplomacy, and adversary perspectives. Inconsistent with the leader model, these dispositional differences in speech pattern occurs *within* the same administration: advisers do not simply provide information congruent with the leader's type. The content of arguments that advisers make during deliberations thus appears to be a key mechanism by which hawkish advisers nudge states toward conflict – and dovish advisers toward

³²The *threat* topic results may stem from the comparatively weak topic coherence, which is discussed in the Appendix §6.2.

cooperation.

4 Conclusion

Foreign policy decisions are made in groups, but whether for theoretical or methodological reasons, we know much more about the role of leader-level characteristics in shaping foreign policy outcomes than adviser-level ones. In this article, we develop an argument linking adviser dispositions to the propensity for interstate conflict. We test our proposition by introducing a new methodological approach that estimates, at a distance, the hawkishness of over 1,000 American presidents and advisers who participated in nearly 3,000 of the most important foreign policy meetings from 1947 to 1988 for which archival records exist. Our theoretical and empirical innovations allow us to move beyond conceptualizing advisers as fungible extensions of leaders and their decision-making processes – and systematically study the ways in which the groups surrounding them matter, particularly for questions of interstate conflict.

Our findings have several implications. First, they cast doubt on a longstanding tradition in IR arguing that the “aggregation problem” renders the study of group member attributes an unfruitful path for inquiry. Our argument instead emphasizes that group deliberation serves as a crucial and under-explored conduit through which individual-level dispositions, such as hawkishness, affect group outcomes. As a result, knowing the dispositions of the advisers who dominate policy debates has substantial explanatory power for state behavior. Aggregation neither washes away individual differences between actors nor forges, *ex machina*, a tabula rasa. Decision-makers’ dispositions remain central to explaining foreign policy decisions even when decision-making occurs in group settings.

In addition, specifying the systematic ways that advisers matter does not imply that leaders are unimportant – and we are careful to note and test ways in which leaders may indirectly shape advisory dynamics. Nevertheless, the findings strongly suggest that leader-centric theories are insufficient to explain foreign policy choices. Specifically, our analysis shows that, when holding the president fixed, shifting from a maximally dovish to a maximally hawkish advisory group essentially triples the expected number of conflictual decisions in a meeting. The deliberative nature of these meetings, which we demonstrate empirically, illuminates an intuitive and compelling reason that

advisers wield such tremendous influence: leaders confront numerous policy challenges on which they are relatively uninformed and hold few preconceived notions. Even when a leader knows *what* they want, there is often cognitive room for hearing disparate perspectives concerning the numerous possible strategies for *how* to get it.³³ We show that when hawks dominate the advisory group, they introduce arguments congruent with a more antagonistic worldview and thereby nudge states toward more conflictual policies.

One important scope condition, however, may concern the nature of the US bureaucracy. The American national security establishment enjoys massive resources, employs thousands of individuals, collects incalculable quantities of information, and ultimately produces outputs that inform, guide, or sway leaders. As a result, American foreign policy advisers may be positioned to persuade presidents (and one another) in ways that bureaucracies elsewhere are not (Jost, 2021). Nevertheless, many countries – particularly democracies – possess advisory systems that mirror the access and participation that the US system affords its national security bureaucracies. Even if our findings only applied to the US, it would still suggest that accounting for the attitudes and beliefs of advisers is of paramount importance to understanding when the world’s current hegemon is prone to conflict – and, perhaps, that policymakers should remain sensitive to when the composition of the leader’s inner circle drifts toward extreme dispositions.

The findings of this study, as well as its limitations, suggest a broad agenda for future research. For one, this study addresses how adviser dispositions shape policy outcomes through the content they contribute to deliberation. Who is able to dominate group discussions and win policy debates is not directly considered. Prospective scholarship can examine the predictors and modes of persuasion in group settings, helping to identify exactly who or what type of communication proves most powerful – as well as build on work in leadership trait analysis to test whether advisor dispositions exert larger effects on foreign policy outcomes when combined with certain leadership styles (Hermann and Preston, 1994). Our study also suggests new directions in the study of bureaucratic politics and the psychology of groups in IR. Future research might explore the causes and consequences of professionalism in the national security bureaucracy, the effects of other adviser dispositions (e.g., risk preference, patience, empathy), and the types of information and signals

³³However, this also implies that leaders might overshadow their advisers when they hold firm beliefs and preferences about certain policy issues (e.g., Saunders, 2011; Yarhi-Milo, 2018). This presents a fruitful avenue for future research.

that certain types of advisers privilege. Looking at other dispositions is also valuable because we find that leader-level hawkishness has a relatively weak effect on US foreign policy decisions, when other research has shown that other leader-level traits have important effects in the US context. Exploring why certain leader-level characteristics are important outside of the US but not inside it, or vice versa, is an important avenue for additional inquiry.

Empirically, our corpus of US meeting records offers a new resource for scholars to study numerous state behaviors at a previously impossible scale, ranging from interstate signaling and threat perception, to alliance formation and crisis escalation. Traditionally, the field of IR has studied these questions either through rich qualitative examination of decision-making with archival documents or through quantitative methods that focus on state behavior rather than decision-making. This corpus offers the field a middle path (Katagiri and Min, 2019): to study state behavior by applying contemporary computational methods to archival documents that span an extended period of time, but to do so in a way that still directly observes the decision-making process. The tools we have harnessed in service of this goal, as well as the data created for and from this endeavor, open numerous doors for further research to demonstrate the value of this approach.

References

- Allison, Graham T. and Philip Zelikow. 1999. *Essence of Decision: Explaining the Cuban Missile Crisis*. 2 ed. New York: Longman.
- Brauer, Carl M. 1986. *Presidential Transitions: Eisenhower Through Reagan*. New York: Oxford University Press.
- Brutger, Ryan and Joshua D. Kertzer. 2018. “A Dispositional Theory of Reputation Costs.” *International Organization* 72(3):693–724.
- Byman, Daniel L. and Kenneth M. Pollack. 2001. “Let Us Now Praise Great Men: Bringing the Statesman Back In.” *International Security* 25(4):107–146.
- Carson, Austin and Eric Min. 2021. “Keeping it Professional: Language and Bureaucracy in the Cold War.”
- Carter, Jeff and Charles E. Smith. 2020. “A Framework for Measuring Leaders’ Willingness to Use Force.” *American Political Science Review* 114(4):1352–1358.

- Chiozza, Giacomo and H.E. Goemans. 2011. *Leaders and International Conflict*. Cambridge, UK: Cambridge University Press.
- Croco, Sarah E. 2011. "The Decider's Dilemma: Leader Culpability, War Outcomes, and Domestic Punishment." *American Political Science Review* 105(3):457–477.
- Dafoe, Allan and Devin Caughey. 2016. "Honor and War: Southern US Presidents and the Effects of Concern for Reputation." *World Politics* 68(2):341–381.
- Eshima, Shusei, Kosuke Imai and Tomoya Sasaki. 2020. "Keyword Assisted Topic Models." arXiv.
- Feaver, Peter D. 2003. *Armed Servants: Agency, Oversight, and Civil-Military Relations*. Cambridge: Harvard University Press.
- Fordham, Benjamin O. 2007. "The Evolution of Republican and Democratic Positions on Cold War Military Spending: A Historical Puzzle." *Social Science History* 31(4):603–636.
- Fuhrmann, Matthew. 2020. "When Do Leaders Free-Ride? Business Experience and Contributions to Collective Defense." *American Journal of Political Science* 64(2):416–431.
- Fuhrmann, Matthew and Michael C. Horowitz. 2015. "When Leaders Matter: Rebel Experience and Nuclear Proliferation." *Journal of Politics* 77(1):72–87.
- George, Alexander L. 1969. "The 'Operational Code': A Neglected Approach to the Study of Political Leaders and Decision-Making." *International Studies Quarterly* 13(2):190–222.
- Goldstein, Joshua S. 1992. "A Conflict-Cooperation Scale for WEIS Events Data." *Journal of Conflict Resolution* 36(2):369–385.
- Hafner-Burton, Emilie M., Stephan Haggard, David A. Lake and David G. Victor. 2017. "The Behavioral Revolution and International Relations." *International Organization* 71(S1):S1–S31.
- Halberstam, David. 1972. *The Best and the Brightest*. New York: Random House.
- Hermann, Margaret G. 2001. "How Decision Units Shape Foreign Policy: A Theoretical Framework." *International Studies Review* 3(2):47–81.
- Hermann, Margaret G. and Thomas Preston. 1994. "Presidents, Advisers, and Foreign Policy: The Effect of Leadership Style on Executive Arrangements." *Political Psychology* 15(1):75–96.
- Herrmann, Richard K., Philip E. Tetlock and Penny S. Visser. 1999. "Mass Public Decisions to Go to War: A Cognitive-Interactionist Framework." *American Political Science Review* 93(3):553–573.
- Holmes, Marcus. 2018. *Face-to-Face Diplomacy: Social Neuroscience and International Relations*.

- Cambridge, UK: Cambridge University Press.
- Holsti, Ole R. 2004. *Public Opinion and American Foreign Policy*. Revised edition ed. Ann Arbor, MI: University of Michigan Press.
- Holsti, Ole R. and James N. Rosenau. 1984. *American Leadership in World Affairs: Vietnam and the Breakdown of Consensus*. Boston: Allen and Unwin.
- Horowitz, Michael C., Allan C. Stam and Cali M. Ellis. 2015. *Why Leaders Fight*. New York: Cambridge University Press.
- Horowitz, Michael C., Philip Potter, Todd S. Sechser and Allan Stam. 2018. "Sizing Up the Adversary: Leader Attributes and Coercion in International Conflict." *Journal of Conflict Resolution* 62(10):2180–2204.
- Jeong, Gyung-Ho. 2018. "Measuring Foreign Policy Positions of Members of the US Congress." *Political Science Research and Methods* 6(1):181–196.
- Jervis, Robert. 1976. *Perception and Misperception in International Politics*. Princeton: Princeton University Press.
- Jervis, Robert. 1998. *System Effects: Complexity in Political and Social Life*. Princeton: Princeton University Press.
- Jervis, Robert. 2013. "Do Leaders Matter and How Would We Know?" *Security Studies* 22(2):153–179.
- Johnston, Alastair Iain. 2014. *Social States*. Princeton: Princeton University Press.
- Jost, Tyler. 2021. "Bureaucracy and War: The Institutional Origins of Miscalculation in International Politics." Book manuscript.
- Jost, Tyler and Joshua D. Kertzer. 2021. "Armies and Influence: Public Deference to Foreign Policy Elites." Working paper.
- Kaarbo, Juliet. 1998. "Power Politics in Foreign Policy: The Influence of Bureaucratic Minorities." *European Journal of International Relations* 4(1):67–97.
- Kahneman, Daniel and Jonathan Renshon. 2007. "Why Hawks Win." *Foreign Policy* pp. 34–38.
- Katagiri, Azusa and Eric Min. 2019. "The Credibility of Public and Private Signals: A Document-Based Approach." *American Political Science Review* 113(1):156–172.
- Kennedy, Andrew. 2011. *The International Ambitions of Mao and Nehru: National Efficacy Beliefs and the Making of Foreign Policy*. New York: Cambridge University Press.

- Kertzer, Joshua D. 2016. *Resolve in International Politics*. Princeton: Princeton University Press.
- Kertzer, Joshua D., Brian C. Rathbun and Nina Srinivasan Rathbun. 2020. "The Price of Peace: Motivated Reasoning and Costly Signaling in International Relations." *International Organization* 74(1):95–118.
- Krasner, Stephen D. 1972. "Are Bureaucracies Important? (Or Allison Wonderland)." *Foreign Policy* (7):159–179.
- Landau-Wells, Marika. 2018. *Dealing with Danger: Threat Perception and Policy Preferences* PhD thesis Massachusetts Institute of Technology.
- LeVeck, Brad L. and Neil Narang. 2017. "The Democratic Peace and the Wisdom of Crowds." *International Studies Quarterly* 61(4):867–880.
- Lin-Greenberg, Erik. 2019. "Backing Up, Not Backing Down: Mitigating Audience Costs through Policy Substitution." *Journal of Peace Research* 56(4):559–574.
- Mann, James. 2004. *Rise of the Vulcans: The History of Bush's War Cabinet*. New York, NY: Penguin.
- McDermott, Rose. 2007. *Presidential Leadership, Illness, and Decision Making*. New York: Cambridge University Press.
- McManus, Roseanne W. 2021. "Crazy Like a Fox? Are Leaders with Reputations for Madness More Successful at International Coercion?" *British Journal of Political Science* 51(1):275–293.
- Mearsheimer, John. 2001. *The Tragedy of Great Power Politics*. W.W. Norton & Company.
- Mintz, Alex and Carly Wayne. 2016. "The Polythink Syndrome and Elite Group Decision-Making." *Political Psychology* 37:3–21.
- 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.
- Modigliani, Andre. 1972. "Hawks and Doves, Isolationism and Political Distrust: An Analysis of Public Opinion on Military Policy." *The American Political Science Review* 66(3):960–978.
- Murray, Shoon Kathleen. 2002. *Anchors against Change: American Opinion Leaders' Beliefs After the Cold War*. University of Michigan Press.
- Powell, Robert. 2017. "Research Bets and Behavioral IR." *International Organization* 71(S1):S265–S277.

- Rathbun, Brian C. 2007. "Hierarchy and Community at Home and Abroad: Evidence of a Common Structure of Domestic and Foreign Policy Beliefs in American Elites." *Journal of Conflict Resolution* 51(3):379–407.
- Rathbun, Brian C. 2014. *Diplomacy's Value: Creating Security in 1920s Europe and the Contemporary Middle East*. Ithaca: Cornell University Press.
- Reiter, Dan and Allan C. Stam. 2002. *Democracies at War*. Princeton: Princeton University Press.
- Renshon, Jonathan. 2017. *Fighting for Status: Hierarchy and Conflict in World Politics*. Princeton: Princeton University Press.
- Risse, Thomas. 2000. "'Let's Argue!': Communicative Action in World Politics." *International Organization* 54(1):1–39.
- Russett, Bruce. 1990. "Doves, Hawks, and US Public Opinion." *Political Science Quarterly* 105(4):515–538.
- Saunders, Elizabeth N. 2011. *Leaders at War: How Presidents Shape Military Interventions*. Ithaca: Cornell University Press.
- Saunders, Elizabeth N. 2017. "No Substitute for Experience: Presidents, Advisers, and Information in Group Decision Making." *International Organization* 71(S1):S219–S247.
- Saunders, Elizabeth N. 2018. "Leaders, Advisers, and the Political Origins of Elite Support for War." *Journal of Conflict Resolution* 62(10):2118–2149.
- Schafer, Mark and Scott Crichlow. 2010. *Groupthink versus High-Quality Decision Making in International Relations*. New York: Columbia University Press.
- Schub, Robert. 2021. "Bureaucracies, Information, and International Crises: What You Know Depends on Where You Sit."
- Schultz, Kenneth A. 2005. "The Politics of Risking Peace: Do Hawks or Doves Deliver the Olive Branch?" *International Organization* 59(1):1–38.
- Sechser, Todd S. 2004. "Are soldiers less war-prone than statesmen?" *Journal of Conflict Resolution* 48(5):746–774.
- Sylvan, Donald A. and James F. Voss, eds. 1998. *Problem Representation in Foreign Policy Decision-Making*. New York: Cambridge University Press.
- 't Hart, Paul, Eric Stern and Bengt Sundelius. 1997. *Beyond Groupthink: Political Group Dynamics and Foreign Policy-Making*. Ann Arbor: University of Michigan Press.

- Tomz, Michael, Jessica Weeks and Keren Yarhi-Milo. 2020. "Public Opinion and Decisions about Military Force in Democracies." *International Organization* 74(1):119–143.
- Weeks, Jessica L.P. 2014. *Dictators at War and Peace*. Ithaca: Cornell University Press.
- Wendt, Alexander. 2004. "The State as Person in International Theory." *Review of International Studies* 30(2):289–316.
- Whitlark, Rachel Elizabeth. 2017. "Nuclear Beliefs: A Leader-Focused Theory of Counter-Proliferation." *Security Studies* 26(4):545–574.
- Yarhi-Milo, Keren. 2014. *Knowing the Adversary: Leaders, Intelligence, and Assessment of Intentions in International Relations*. Princeton: Princeton University Press.
- Yarhi-Milo, Keren. 2018. *Who Fights for Reputation in International Politics? Leaders, Resolve, and the Use of Force*. Princeton, NJ: Princeton University Press.
- Yarhi-Milo, Keren, Joshua D. Kertzer and Jonathan Renshon. 2018. "Tying Hands, Sinking Costs, and Leader Attributes." *Journal of Conflict Resolution* 62(10):2150–2179.

ADVISERS AND AGGREGATION IN FOREIGN POLICY
DECISION-MAKING
Supplementary Appendix
January 7, 2022

Contents

1	Record Collection	1
1.1	Archival Record Collection	1
1.2	FRUS Record Collection	1
1.3	Meeting Statistics	3
2	Methodology for Measuring Decision Outcomes	4
2.1	Coding Approach and Rationale	4
2.2	Coding Typology	4
2.3	Distribution of the Outcome Variable	5
3	Methodology for Measuring Decision-Maker Hawkishness	6
3.1	Biographical Dataset on US Decision-Makers	6
3.2	Foreign Policy Leadership Project Survey	7
3.3	Survey Questions	7
4	Robustness Checks	10
4.1	Disaggregating Results by Meeting Type	10
4.2	Removing Bureaucratic Affiliations	16
4.3	Removing the Soviet Union	16
4.4	Negative Binomial	22
4.5	Time-Unit Replication Analysis	25
5	Discussion of Leader Results	27
5.1	Expert survey results	27
5.2	Consistency with Existing Findings on US Presidential Hawkishness	28
5.2.1	Horowitz and Stam (2014)	31
5.2.2	Carter and Smith (2020)	33
5.3	Leader Constraints	34
5.4	Measurement of Hawkishness	34
5.5	Strategic Behavior	39
5.6	Measurement of the Dependent Variable	39
5.7	Institutional Context	40
6	Probing the Deliberation Mechanism	41
6.1	Information Search and Dissent	41
6.2	Keyword Assisted Topic Model	42
6.3	Issues of Self-Censorship	46
7	Agenda Items and Hawkishness	47

1 Record Collection

1.1 Archival Record Collection

Records of the National Security Council were photographed from multiple sites around the country.

- Truman: Truman Presidential Library (Independence, Missouri)
- Eisenhower: Eisenhower Presidential Library (Abilene, Kansas)
- Kennedy: See note below.
- Johnson: Johnson Presidential Library (Austin, Texas)
- Nixon: Nixon Presidential Library (Yorba Linda, California)
- Ford: Ford Presidential Library (Ann Arbor, Michigan)
- Carter: See note below.
- Reagan: See note below.
- Various supporting documents: National Archives II (College Park, Maryland)

NSC records for the Kennedy and Carter administrations were available for download from each presidential library’s website. It was therefore not necessary to manually photograph document records from these administrations’ meetings.

Records for the Reagan administration come from the widest variety of sources. Approximately 27 records were available for download from the presidential library’s site. Others come from [Saltoun-Ebin \(2014\)](#).

1.2 FRUS Record Collection

All of our informal meeting records come from the *Foreign Relations of the United States* (FRUS) collection. As of October 2020, when we scraped the FRUS from the State Department website, we gathered a total of 139,847 documents. The following process was used to determine which documents would be included in the final analysis:

1. Search for all FRUS documents’ titles for any of the phrases in the list below:
 - memorandum of conference

- memorandum of meeting
 - memorandum of discussion
 - minutes of meeting
 - notes on meeting
 - record
 - conversation
2. Filter down to documents that use the word “president” in the text of document.
 3. Remove documents that include the name of a formal body (such as the National Security Council or Washington Special Acts Group) or mention a foreign leader or diplomat.
 4. Have RAs qualitatively review remaining meetings to remove any irrelevant documents.

This resulted in a final list of 2,232 informal meeting records. The formal NSC meetings collected via archives and informal meetings identified via *FRUS* together constitute an expansive record of US foreign policy deliberations. The set of records, however, is incomplete in the sense that it does not include all meetings in which presidents participated. Two likely causes for records to be missing from our set merit attention: document selection and classification.

First, historians compiling *FRUS* volumes exercise discretion on which documents to include. *FRUS* volumes typically prioritize high-salience events and discussions over the more mundane elements of executive branch functions. The missing mundane documents, while inherently difficult to observe, are unlikely to skew the paper’s core results. Given constraints on the president’s time and attention, the president is less likely to have strong, well-formed views on less salient matters. In such cases, advisers with narrower perspectives and remits are likely to assume a privileged role. Accordingly, we expect that, if anything, decisions surrounding lower salience policy issues would exhibit even more (less) adviser (leader) influence.

Second, materials remain classified both at the time that historians compile *FRUS* volumes and when we collected archival materials. Classification affecting *FRUS* document inclusion is most vividly evident for the controversial initial *FRUS* volume on the 1954 covert action in Guatemala (McAllister et al., 2015). After a publicized dispute, a follow-up volume published decades later filled in the patchy record. The facts that (1) extensive missingness due to classification on

Guatemala prompted such a public outcry and (2) there has not been comparable public complaints for other volumes, suggests that politicization of the declassification process is somewhat limited. The somewhat more difficult issue to address is the ongoing classification of NSC meetings that we were unable to collect from the archives. Missing NSC meetings are overwhelmingly from the Carter and Reagan administrations, the most recent two in our sample. This suggests miss- ingness reflects the extended timeline for declassification as opposed to substantive considerations that might sway the relative influence of presidents versus advisers.

1.3 Meeting Statistics

Table 1 displays summary statistics on all formal meetings of the National Security Council included in our analysis.

Admin.	Records	Mtgs	Speech Acts	Decisions	Conf. Adv.	Coop. Adv.
Truman	125	128	3,121	77	66	11
Eisenhower	337	359	22,198	315	261	54
Kennedy	66	89	12,153	90	75	15
Johnson	72	75	1,459	26	21	5
Nixon	62	90	4,568	30	15	15
Ford	34	42	4,714	9	4	5
Carter	18	41	1,375	18	12	6
Reagan	79	153	3,049	63	45	18
<i>Total</i>	793	977	52,637	628	499	129

Table 1: Coverage of formal NSC records for each administration.

Table 2 displays summary statistics on all informal meetings included in our analysis.

Admin.	Records	Speech Acts	Decisions	Conf. Adv.	Coop. Adv.
Truman	112	718	26	23	3
Eisenhower	607	7,802	107	70	37
Kennedy	203	3,300	38	29	9
Johnson	474	11,725	69	46	23
Nixon	429	26,104	61	41	20
Ford	221	4,163	47	16	31
Carter	36	618	6	1	5
Reagan	6	125	0	0	0
<i>Total</i>	2,088	54,555	354	226	128

Table 2: Coverage of informal meeting records for each administration.

2 Methodology for Measuring Decision Outcomes

2.1 Coding Approach and Rationale

The goal for coding substantive decisions reached during meetings is to specify the target(s) of the decision and whether the decision is cooperative or conflictual. The task is therefore similar to coding event data (e.g., COPDAB or WEIS) but several unique features of our substantive area make existing approaches insufficient. First, decisions reached in a classified setting often do not generate a news report. We must look to the meeting itself, rather than the media, to discern decisions made and attribute them to the proper meeting session. Second, many decisions in NSC and informal meetings pertain to military planning. These include decisions to accelerate or halt arms programs or move forces to a region. While event data coding typically relies on having an explicitly named target of the action, in practice decision-makers do not always name the target. Instead, coders looked for contextual information to discern the primary implicit target of a decision – often the Soviet Union. Third, decisions from meetings frequently involved a triangular relationship in which two actors were targets but in diametrically opposed ways. For instance, a decision to supply arms to Chinese Nationalists is properly coded as a cooperative act toward Chinese Nationalists and conflictual act toward Chinese Communists. Our coding scheme captures these triadic dynamics (Goldstein and Freeman, 1990).

2.2 Coding Typology

We code each decision target as *Adversary*, *Aligned*, or *Non-Aligned*. Examples of each would include the Soviet Union, United Kingdom, and Austria, respectively. The status of other targets varied depending on bilateral relations between it and the United States.

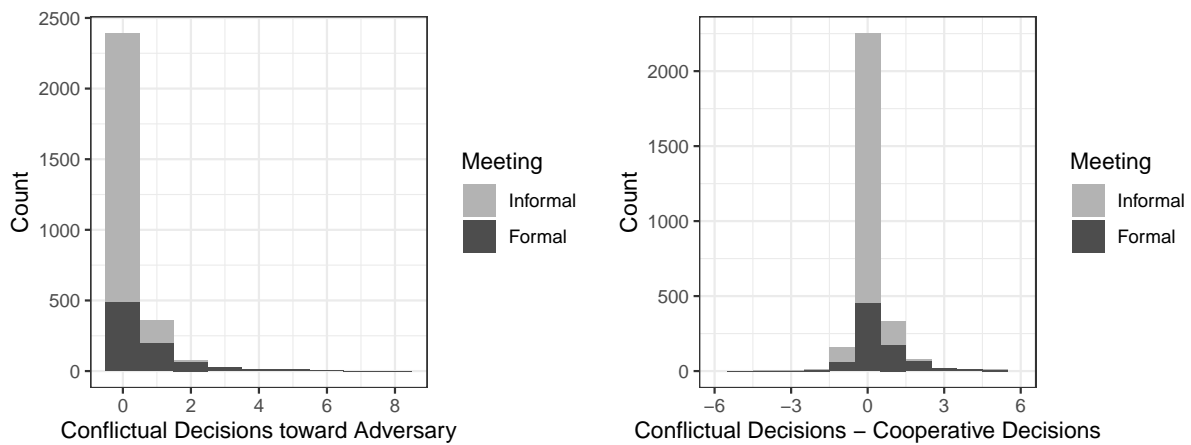
Decisions are ultimately coded as *Cooperative* or *Conflictual*. Given the broad categorization, each category subsumes numerous forms of decisions. *Cooperative* examples include material acts such as providing military aid and verbal acts such as conveying agreement. *Conflictual* examples include material acts such as increasing military spending, imposing sanctions, or even using military force and verbal acts such as making a threat or lodging a protest. Due to common problems associated with efforts to apply distinct scores to each type of action (e.g., protest versus military strike), we employ a count approach that aggregates up to the conflictual versus coop-

erative distinction. Moreover, the large presence of military planning decisions in our data does not have a natural analogue in existing event coding scales which further undercuts the validity of trying to apply those scales to our data.

As one expects, decisions toward adversaries are generally more conflictual than cooperative, though with heterogeneity across administrations.

2.3 Distribution of the Outcome Variable

Figure 1: Distributions of Foreign Policy Decisions



3 Methodology for Measuring Decision-Maker Hawkishness

3.1 Biographical Dataset on US Decision-Makers

Our analysis relies on data regarding more than 1,100 individuals who spoke at least once in a meeting. For each individual, we collected the following biographical information:

- First name
- Middle name
- Last name
- Gender
- Year of birth
- Highest education level
- Years of service in military
- Years of service in diplomacy
- Years of service in intelligence
- Participation in WWII, Korean War, and/or Vietnam Wars
- Political party affiliation

As one can note in the next section, these characteristics align with individual-level information from the Foreign Policy Leadership Project (FPLP) survey, which we use to train a model that is then applied to these biographical data to predict hawkishness for each individual.

We additionally logged data on all positions an individual held in the executive branch of the federal government between 1947 and 1988. This included:

- Position name
- Position start and end date
- Position level (in United States Order of Precedence)
- Position's bureaucratic affiliation

3.2 Foreign Policy Leadership Project Survey

The following attributes of individuals were extracted from 2,282 responses to the 1976 Foreign Policy Leadership Project (FPLP) survey for use in our predictive model:

- Gender
- Birth decade
- Highest level of education
- Military service (binary)
- Military officer (binary)
- Foreign Service Officer (binary)
- Participation in WWII, Korean War, or Vietnam War
- Political party affiliation

As mentioned in the main text, our measure of each respondent’s hawkishness is based on their responses to fifteen specific questions in the survey.¹ These fifteen questions, which were distributed over five separate sections of the survey, are replicated below. Numbers in parentheses represent the item number as recorded in the original dataset containing all responses and do not reflect the actual item number in the survey itself.

3.3 Survey Questions

“Turning to more general considerations, here is a list of possible foreign policy goals that the United States might have. Please indicate how much importance you think should be attached to each goal.” (Very important, Somewhat important, Not important at all, Not sure)

- Containing Communism (#136)

“Somewhat more specifically, please indicate how strongly you agree or disagree with each of the followings statements concerning America’s role in the world.” (Agree Strongly, Agree Somewhat, Disagree Somewhat, Disagree Strongly, No Opinion)

- There is nothing wrong with using the C.I.A. to try to undermine hostile governments. (#155)

¹We obtain these measures from the first wave of the FPLP survey since it is temporally the closest to our study period.

- It is not in our interest to have better relations with the Soviet Union because we are getting less than we are giving to them. (#158)
- The U.S. should take all steps including the use of force to prevent the spread of Communism. (#160)

“This question asks you to indicate your position on certain foreign policy issues, and to state the extent to which your position was shaped by the experience in Vietnam. First indicate how strongly you agree or disagree with each statement by checking one box in each row on the right.” (Agree Strongly, Agree Somewhat, Disagree Somewhat, Disagree Strongly, No Opinion)

- There is considerable validity in the “domino theory” that when one nation falls to communism, others nearby will soon follow a similar path. (#218)
- Any communist victory is a defeat for America’s national interest. (#225)
- The Soviet Union is generally expansionist rather than defensive in its foreign policy goals. (#234)
- Detente permits the USSR to pursue policies that promote rather than restrain conflict (#271)
- Rather than simply countering our opponent’s thrusts, it is necessary to strike at the heart of the opponent’s power. (#277)
- When force is used, military rather than political goals should determine its application. (#280)

“There has been quite a bit of discussion about the consequences of the Vietnam episode. Some of these are listed below. Please indicate your assessment of each statement by checking only one box for each item.” (Agree Strongly, Agree Somewhat, Disagree Somewhat, Disagree Strongly, No Opinion)

- Communist nations have been encouraged to seek triumphs elsewhere as a result of Vietnam. (#309)
- The major assumptions of detente have been proven false by the events in Vietnam. (#310)

“Observers of American foreign policy have identified several factors that may have prevented the United States from achieving its goals in the Vietnam undertaking. In your judgment, how important were the reasons listed below in America’s inability to achieve all of its goals? Please

indicate your assessment by checking only one box in each row. (Very important, Moderately important, Slightly important, Not at all important, Not sure)

- The United States fought with a “no win” approach. (#327)
- The use of American air power was restricted. (#329)
- Insufficient attention was paid to advice from the military. (#332)

4 Robustness Checks

In this section, we conduct a series of robustness checks that uphold and further contextualize the validity of our main findings.

4.1 Disaggregating Results by Meeting Type

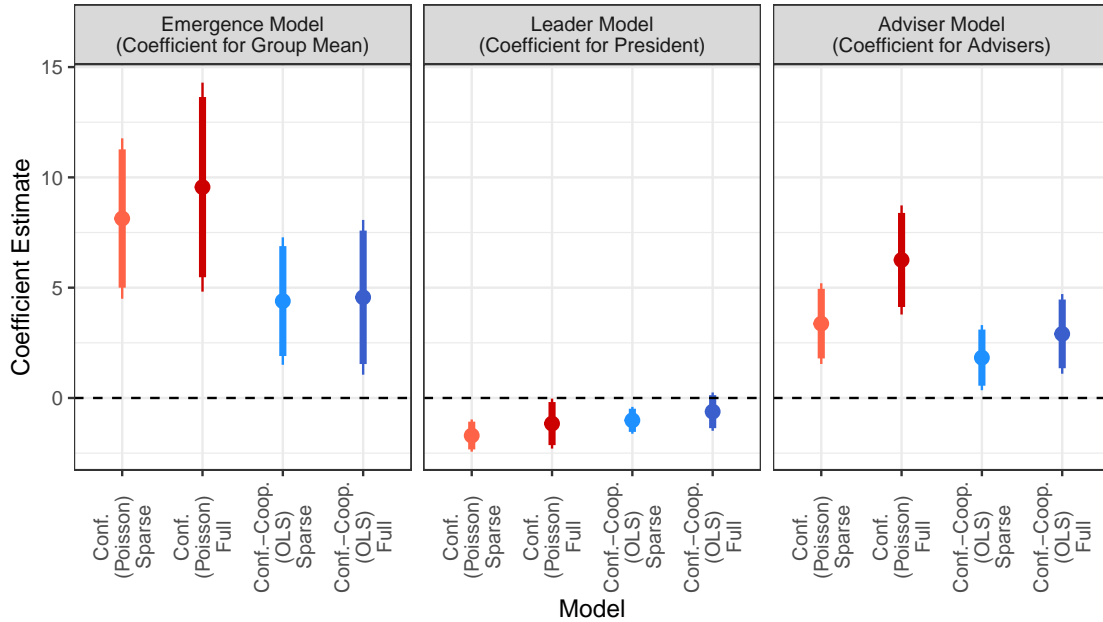
As noted, our main analysis includes both formal NSC meetings as well as informal gatherings, as national security decisions were made in both contexts. Yet we may also believe that differences exist in the dynamics of these separate settings. We can explore potential differences by replicating the previous analysis after disaggregating these two meeting types.

Tables 3 and 4 present full statistical results from regressions that only use data from informal meetings. Figure 2 displays the corresponding coefficient plots. Tables 5 and 6 present full statistical results from regressions that only use data from formal meetings. Figure 3 displays the corresponding coefficient plots.

The results make two distinctions clear. First, the influence of advisers is consistently stronger and more statistically significant in formal meetings compared to informal gatherings. Second, there is no statistically significant relationship between the hawkishness of leaders and conflictual decisions in informal meetings. This contrasts with the negative coefficients found using Poisson models in only formal meetings.

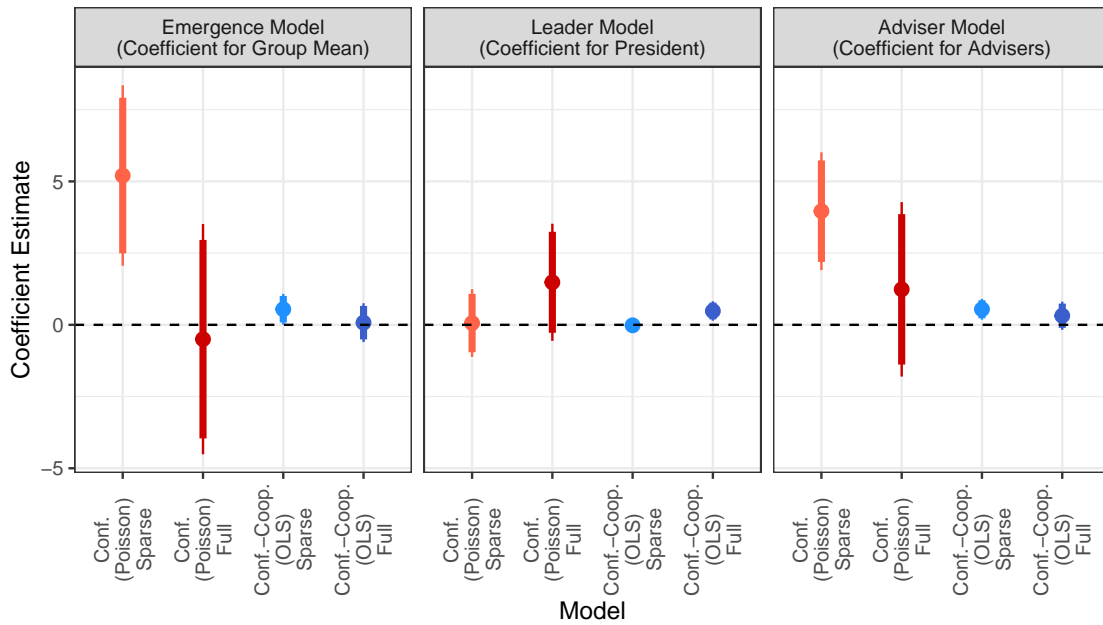
These findings refine our understanding of foreign policy decision-making in critical ways. First, we find that advisers matter more in foreign policy decision-making than work that focuses only on leaders at the expense of advisers might suggest. We also find very little evidence for—and considerable evidence against—the emergent property account, which suggests that group-level properties cannot be reduced to the hawkishness of members that comprise it. We do not rule out the possibility that our analysis omits certain group-level characteristics that may also be shaping policy choices. Nonetheless, we find that hawkishness at the individual level aggregates in foreign policy decision-making groups in sensible ways: the average level of hawkishness in a group is informative and becomes more informative when one takes into account how much people participate in the decision reached, even though our measure of hawkishness is obtained independently of the decisions these dispositions are being used to explain.

Figure 2: Summary of Three Models of Trait Aggregation, Using Formal Meetings



Note: Coefficient plots corresponding to main findings, corresponding to four models on each table. Thicker and smaller bands represent 90% and 95% confidence intervals respectively.

Figure 3: Summary of Three Models of Trait Aggregation, Using Informal Meetings



Note: Coefficient plots corresponding to main findings, corresponding to four models on each table. Thicker and smaller bands represent 90% and 95% confidence intervals respectively.

Table 3: Effect of Mean Participant Hawkishness and President's Hawkishness on Foreign Policy Decisions, Using Formal Meetings

	<i>Emergence Model</i>				<i>Leader Model</i>			
	Conflictual Decisions		Conflictual – Cooperative		Conflictual Decisions		Conflictual – Cooperative	
	<i>Poisson</i>		<i>OLS</i>		<i>Poisson</i>		<i>OLS</i>	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Mean Hawkishness	8.135*** (1.855)	9.558*** (2.417)	4.393*** (1.474)	4.563** (1.790)				
President's Hawkishness		-0.464 (0.293)		-0.275 (0.213)		-0.051 (0.211)		-0.024 (0.147)
No. of Attendees		0.086*** (0.023)		0.058*** (0.019)		0.101*** (0.021)		0.070*** (0.018)
Proportion - President		0.861 (0.888)		0.317 (0.616)		1.313 (0.843)		0.444 (0.606)
Proportion - Military		-0.399 (1.031)		0.081 (0.787)		-0.027 (0.855)		0.333 (0.685)
Proportion - Defense		1.670*** (0.533)		0.894** (0.417)		0.597 (0.509)		0.369 (0.401)
Proportion - Intelligence		1.109 (0.968)		0.849 (0.604)		0.752 (0.905)		0.682 (0.601)
Proportion - State		0.149 (0.689)		0.438 (0.466)		-0.296 (0.669)		0.182 (0.452)
Diplomatic Experience		0.077 (0.062)		0.012 (0.043)		0.075 (0.058)		0.010 (0.041)
Intelligence Experience		-0.094 (0.075)		-0.018 (0.051)		-0.139** (0.055)		-0.049 (0.041)
Military Experience		-0.079 (0.141)		-0.080 (0.101)		0.068 (0.125)		-0.037 (0.091)
5-Year MIDs		6.924** (2.920)		5.350** (2.222)		2.734** (1.365)		2.166** (0.993)
US CINC		0.090 (0.124)		0.030 (0.111)		0.089 (0.123)		0.037 (0.112)
Constant	-5.569*** (1.146)	-9.191*** (1.848)	-2.184** (0.886)	-4.172*** (1.340)	0.380** (0.185)	-2.168** (1.012)	0.981*** (0.163)	-0.413 (0.702)
Administration FEs	✓	✓	✓	✓				
Agenda Items	✓	✓	✓	✓	✓	✓	✓	✓
Observations	793	793	793	793	792	792	792	792

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

Table 4: Effect of Adviser Hawkishness on Foreign Policy Decisions, Using Formal Meetings

	<i>Adviser Model</i>				<i>Advisers + Admin. FEs</i>			
	Conflictual Decisions		Conflictual – Cooperative		Conflictual Decisions		Conflictual – Cooperative	
	<i>Poisson</i>		<i>OLS</i>		<i>Poisson</i>		<i>OLS</i>	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Advisers' Hawkishness (Acts)	3.371*** (0.933)	6.256*** (1.262)	1.828** (0.753)	2.905*** (0.921)	7.554*** (1.493)	7.828*** (1.844)	3.972*** (1.191)	3.759*** (1.369)
President's Hawkishness	-1.975*** (0.373)	-2.593*** (0.665)	-1.201*** (0.318)	-1.388*** (0.502)				
No. of Attendees		0.128*** (0.022)		0.082*** (0.019)		0.101*** (0.023)		0.065*** (0.019)
Proportion - President		1.575* (0.842)		0.514 (0.607)		0.948 (0.904)		0.286 (0.625)
Proportion - Military		-0.070 (0.850)		0.310 (0.682)		0.323 (0.953)		0.416 (0.738)
Proportion - Defense		1.222** (0.523)		0.643 (0.418)		1.726*** (0.542)		0.895** (0.427)
Proportion - Intelligence		1.464* (0.876)		1.004 (0.617)		1.070 (0.957)		0.821 (0.618)
Proportion - State		0.349 (0.674)		0.435 (0.473)		0.208 (0.697)		0.420 (0.484)
Diplomatic Experience		0.061 (0.059)		-0.003 (0.041)		0.063 (0.062)		0.004 (0.043)
Intelligence Experience		-0.230*** (0.060)		-0.080* (0.042)		-0.112 (0.074)		-0.023 (0.051)
Military Experience		-0.120 (0.127)		-0.127 (0.098)		-0.165 (0.139)		-0.129 (0.102)
5-Year MIDs		-0.153 (0.216)		-0.101 (0.149)		-0.390 (0.295)		-0.264 (0.214)
US CINC		1.198 (1.434)		1.546 (1.009)		7.697*** (2.900)		5.632** (2.209)
Constant	-1.429*** (0.535)	-4.049*** (1.124)	0.031 (0.429)	-1.135 (0.768)	-5.214*** (0.929)	-8.286*** (1.656)	-1.918*** (0.714)	-3.631*** (1.198)
Administration FEs					✓	✓	✓	✓
Agenda Items	✓	✓	✓	✓	✓	✓	✓	✓
Observations	788	788	788	788	789	789	789	789

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

Table 5: Effect of Mean Participant Hawkishness and President's Hawkishness on Foreign Policy Decisions, Using Informal Meetings

	<i>Emergence Model</i>				<i>Leader Model</i>			
	Conflictual Decisions		Conflictual – Cooperative		Conflictual Decisions		Conflictual – Cooperative	
	<i>Poisson</i>		<i>OLS</i>		<i>Poisson</i>		<i>OLS</i>	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Mean Hawkishness	5.205*** (1.605)	-0.500 (2.047)	0.543** (0.274)	0.076 (0.346)				
President's Hawkishness		0.228 (0.422)		0.077 (0.066)		-0.002 (0.270)		0.087** (0.044)
No. of Attendees		0.031 (0.035)		0.010 (0.007)		0.011 (0.036)		0.008 (0.006)
Proportion - President		0.815 (0.623)		0.023 (0.093)		0.755 (0.567)		0.051 (0.083)
Proportion - Military		1.896*** (0.716)		0.253* (0.142)		1.567*** (0.565)		0.256** (0.115)
Proportion - Defense		1.885*** (0.459)		0.154* (0.093)		1.823*** (0.457)		0.144 (0.091)
Proportion - Intelligence		1.755** (0.714)		0.248** (0.126)		1.788** (0.703)		0.238* (0.125)
Proportion - State		0.365 (0.432)		-0.013 (0.058)		0.358 (0.435)		-0.015 (0.056)
Diplomatic Experience		0.065 (0.076)		-0.005 (0.012)		0.075 (0.074)		-0.0004 (0.012)
Intelligence Experience		-0.026 (0.095)		-0.010 (0.018)		-0.070 (0.092)		-0.014 (0.018)
Military Experience		0.127 (0.118)		0.005 (0.017)		0.224** (0.110)		0.010 (0.015)
5-Year MIDs		-2.676 (5.017)		-0.737 (0.852)		3.142* (1.867)		0.854*** (0.318)
US CINC		-0.710 (0.844)		0.076 (0.094)		-0.624 (0.836)		0.084 (0.093)
Constant	-5.012*** (1.089)	-2.636 (2.164)	-0.175 (0.185)	0.054 (0.369)	-2.255*** (0.316)	-5.578*** (1.372)	0.056 (0.052)	-0.738*** (0.205)
Administration FEs	✓	✓	✓	✓				
Agenda Items	✓	✓	✓	✓	✓	✓	✓	✓
Observations	2,084	2,084	2,084	2,084	2,088	2,088	2,088	2,088

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

Table 6: Effect of Adviser Hawkishness on Foreign Policy Decisions, Using Informal Meetings

	<i>Adviser Model</i>				<i>Advisers + Admin. FEs</i>			
	Conflictual Decisions		Conflictual – Cooperative		Conflictual Decisions		Conflictual – Cooperative	
	<i>Poisson</i>		<i>OLS</i>		<i>Poisson</i>		<i>OLS</i>	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Advisers' Hawkishness (Acts)	3.962*** (1.047)	1.237 (1.551)	0.543*** (0.186)	0.318 (0.250)	3.667*** (1.199)	-0.083 (1.720)	0.552** (0.220)	0.235 (0.278)
President's Hawkishness	-0.677 (0.624)	0.701 (1.108)	-0.125 (0.109)	0.385** (0.190)				
No. of Attendees		0.043 (0.036)		0.012* (0.007)		0.054 (0.035)		0.014** (0.007)
Proportion - President		1.542** (0.645)		0.079 (0.095)		1.690** (0.672)		0.084 (0.100)
Proportion - Military		1.885*** (0.654)		0.225* (0.131)		2.321*** (0.723)		0.246* (0.140)
Proportion - Defense		2.366*** (0.500)		0.188** (0.095)		2.324*** (0.502)		0.192** (0.096)
Proportion - Intelligence		2.570*** (0.739)		0.299** (0.132)		2.511*** (0.743)		0.306** (0.132)
Proportion - State		1.187** (0.499)		0.045 (0.066)		1.164** (0.497)		0.056 (0.067)
Diplomatic Experience		0.067 (0.076)		-0.001 (0.012)		0.047 (0.079)		-0.009 (0.013)
Intelligence Experience		-0.106 (0.093)		-0.016 (0.018)		-0.071 (0.097)		-0.015 (0.018)
Military Experience		0.206* (0.118)		0.003 (0.016)		0.144 (0.120)		0.003 (0.017)
5-Year MIDs		-0.095 (0.290)		0.094** (0.048)		0.173 (0.445)		0.063 (0.070)
US CINC		1.367 (2.164)		0.594 (0.374)		-6.172 (5.385)		-1.305 (0.912)
Constant	-4.049*** (0.598)	-5.783*** (1.573)	-0.185* (0.101)	-0.840*** (0.233)	-3.865*** (0.817)	-2.414 (2.134)	-0.157 (0.148)	0.109 (0.357)
Administration FEs					✓	✓	✓	✓
Agenda Items	✓	✓	✓	✓	✓	✓	✓	✓
Observations	1,916	1,916	1,916	1,916	1,916	1,916	1,916	1,916

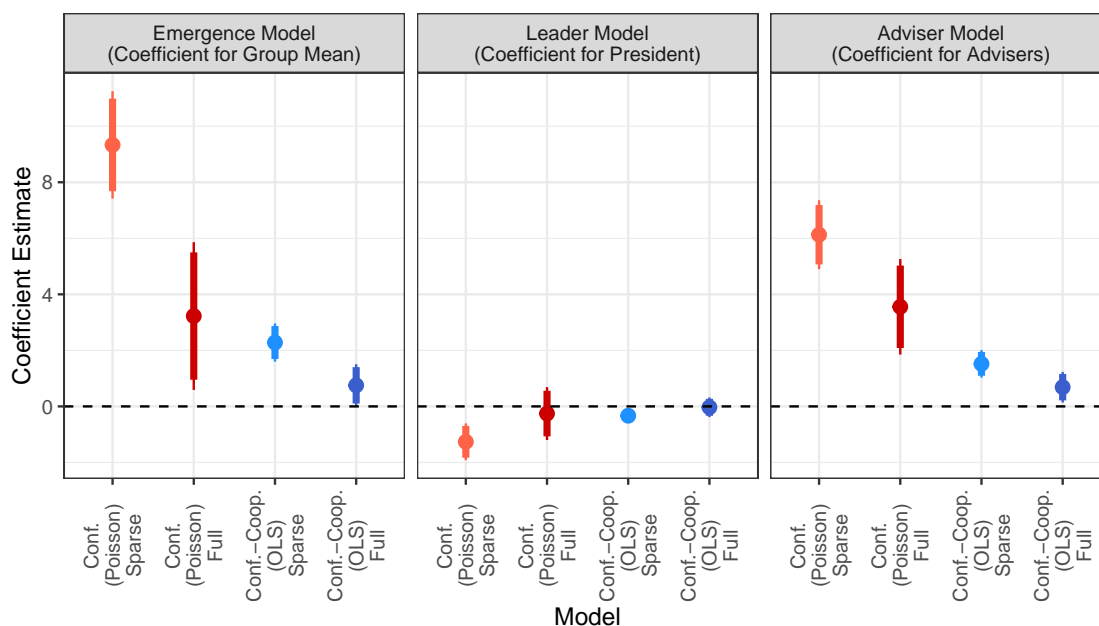
Note: * $p < 0.1$; ** $p < 0.05$; *** $p < 0.01$

4.2 Removing Bureaucratic Affiliations

We remove variables capturing whether a FPLP survey respondent is currently a member of the military or a foreign service officer. Consequently, our predictions of NSC actors' hawkishness also does not take into consideration whether the individual is a member of the military or in the State Department; predictions are made based on all other factors.

Tables 7 and 8, as well as Figure 4, indicate that our main results are upheld with this revised measure. This is not unsurprising; the correlation between our original hawkishness measure and that produced without incorporating information on bureaucratic affiliation is 0.966.

Figure 4: Summary of Three Models of Trait Aggregation, Removing Bureaucratic Affiliation from Supervised Learning Model



Note: Coefficient plots corresponding to main findings, corresponding to four models on each table. Thicker and smaller bands represent 90% and 95% confidence intervals respectively.

4.3 Removing the Soviet Union

One may be concerned that our results are primarily driven by Cold War dynamics — in which hawkishness might be understood less as a universal disposition, and more specific to the US-Soviet relationship. We can partially address this by replicating our analysis after removing all decisions involving the Soviet Union. Note that the original data contains 725 conflictual decisions and 257 cooperative decisions. Once we remove any decisions that were coded as involving the Soviet Union,

Table 7: Effect of Mean Participant Hawkishness and President’s Hawkishness on Foreign Policy Decisions, Removing Bureaucratic Affiliation from Supervised Learning Model

	<i>Emergence Model</i>				<i>Leader Model</i>			
	Conflictual Decisions		Conflictual – Cooperative		Conflictual Decisions		Conflictual – Cooperative	
	<i>Poisson</i>		<i>OLS</i>		<i>Poisson</i>		<i>OLS</i>	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Mean Hawkishness	9.334*** (0.978)	3.224** (1.345)	2.279*** (0.348)	0.752* (0.384)				
President’s Hawkishness					-1.264*** (0.335)	-0.255 (0.482)	-0.332*** (0.122)	-0.032 (0.180)
No. of Attendees		0.076*** (0.017)		0.036*** (0.007)		0.085*** (0.016)		0.040*** (0.007)
Proportion - President		0.956** (0.461)		0.153 (0.118)		1.101*** (0.418)		0.229** (0.106)
Proportion - Military		1.793*** (0.471)		0.298* (0.157)		1.531*** (0.456)		0.324** (0.151)
Proportion - Defense		1.743*** (0.339)		0.164 (0.114)		1.186*** (0.327)		0.077 (0.111)
Proportion - Intelligence		1.577*** (0.548)		0.297* (0.161)		1.295** (0.528)		0.242 (0.160)
Proportion - State		0.267 (0.328)		-0.020 (0.076)		0.177 (0.331)		0.006 (0.075)
Diplomatic Experience		0.115** (0.045)		0.008 (0.015)		0.079* (0.043)		-0.002 (0.014)
Intelligence Experience		-0.133** (0.052)		-0.050*** (0.018)		-0.111** (0.045)		-0.045** (0.018)
Military Experience		0.010 (0.078)		-0.035* (0.021)		0.077 (0.071)		-0.029 (0.019)
Formal		0.590*** (0.138)		0.143*** (0.051)		0.764*** (0.130)		0.153*** (0.049)
5-Year MID Challenges		-0.238 (0.223)		-0.101 (0.075)		0.067 (0.150)		-0.003 (0.046)
US CINC		6.351*** (2.317)		2.824*** (0.889)		3.090*** (1.052)		1.526*** (0.342)
Constant	-6.883*** (0.639)	-6.604*** (1.160)	-1.095*** (0.221)	-1.210*** (0.385)	-0.736*** (0.172)	-4.052*** (0.670)	0.334*** (0.065)	-0.441** (0.210)
Administration FEs	✓	✓	✓	✓				
Agenda Items	✓	✓	✓	✓	✓	✓	✓	✓
Observations	2,877	2,877	2,877	2,877	2,880	2,880	2,880	2,880

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

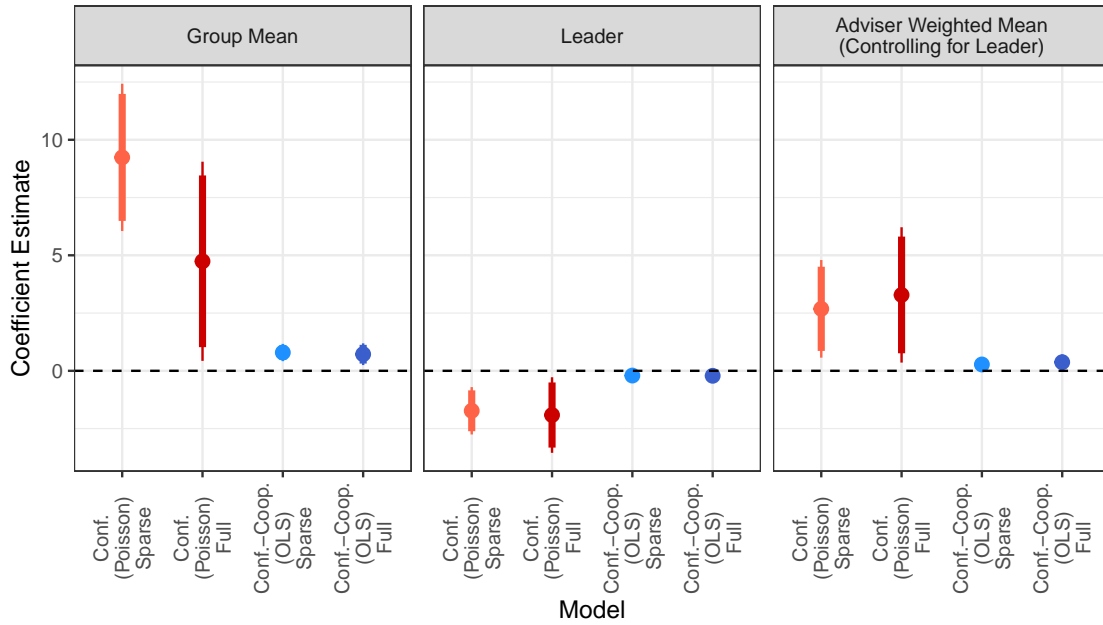
Table 8: Effect of Adviser Hawkishness on Foreign Policy Decisions, Removing Bureaucratic Affiliation from Supervised Learning Model

	<i>Adviser Model</i>				<i>Advisers + Admin. FEs</i>			
	Conflictual Decisions		Conflictual – Cooperative		Conflictual Decisions		Conflictual – Cooperative	
	<i>Poisson</i>		<i>OLS</i>		<i>Poisson</i>		<i>OLS</i>	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Advisers' Hawkishness (Acts)	6.128*** (0.627)	3.555*** (0.870)	1.518*** (0.251)	0.685** (0.279)	4.964*** (0.826)	2.574** (1.154)	1.395*** (0.328)	0.730** (0.337)
President's Hawkishness	-2.488*** (0.345)	-1.248** (0.530)	-0.666*** (0.136)	-0.282 (0.204)				
No. of Attendees		0.105*** (0.016)		0.046*** (0.007)		0.090*** (0.017)		0.042*** (0.007)
Proportion - President		1.446*** (0.440)		0.271** (0.118)		1.477*** (0.485)		0.229* (0.127)
Proportion - Military		2.052*** (0.473)		0.405** (0.157)		2.187*** (0.478)		0.381** (0.160)
Proportion - Defense		1.763*** (0.347)		0.167 (0.118)		2.004*** (0.359)		0.220* (0.120)
Proportion - Intelligence		1.992*** (0.540)		0.346** (0.169)		1.988*** (0.556)		0.357** (0.169)
Proportion - State		0.730** (0.352)		0.089 (0.086)		0.666* (0.353)		0.070 (0.087)
Diplomatic Experience		0.083* (0.044)		-0.003 (0.014)		0.117** (0.046)		0.005 (0.015)
Intelligence Experience		-0.158*** (0.047)		-0.052*** (0.018)		-0.146*** (0.053)		-0.052*** (0.019)
Military Experience		0.015 (0.076)		-0.045** (0.021)		-0.015 (0.079)		-0.048** (0.022)
Formal		0.718*** (0.133)		0.146*** (0.051)		0.589*** (0.140)		0.143*** (0.053)
5-Year MID Challenges		-0.009 (0.153)		-0.018 (0.048)		-0.275 (0.227)		-0.125 (0.079)
US CINC		1.823* (1.089)		1.223*** (0.372)		6.203*** (2.343)		2.627*** (0.933)
Constant	-3.466*** (0.335)	-5.237*** (0.754)	-0.313** (0.131)	-0.599** (0.235)	-3.983*** (0.529)	-6.387*** (1.108)	-0.510** (0.205)	-1.132*** (0.377)
Administration FEs					✓	✓	✓	✓
Agenda Items	✓	✓	✓	✓	✓	✓	✓	✓
Observations	2,704	2,704	2,704	2,704	2,705	2,705	2,705	2,705

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

we are left with 280 conflictual decisions and 87 cooperative decisions. This highlights the reality that many of the NSC's decisions over the period of investigation were indeed about the Soviet Union. Nonetheless, Tables 9 and 10, as well as Figure 5, show that our main substantive findings remain even when only analyzing decisions targeted at all other countries.

Figure 5: Summary of Three Models of Trait Aggregation, Removing Decisions Involving the USSR



Note: Coefficient plots corresponding to main findings, corresponding to four models on each table. Thicker and smaller bands represent 90% and 95% confidence intervals respectively.

Table 9: Effect of Mean Participant Hawkishness and President's Hawkishness on Foreign Policy Decisions, Removing Decisions Involving the USSR

	<i>Emergence Model</i>				<i>Leader Model</i>			
	Conflictual Decisions		Conflictual – Cooperative		Conflictual Decisions		Conflictual – Cooperative	
	<i>Poisson</i>		<i>OLS</i>		<i>Poisson</i>		<i>OLS</i>	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Mean Hawkishness	9.237*** (1.626)	4.741** (2.199)	0.792*** (0.191)	0.719*** (0.240)				
President's Hawkishness					-1.730*** (0.523)	-1.914** (0.834)	-0.205*** (0.070)	-0.215* (0.111)
No. of Attendees		0.028 (0.030)		0.003 (0.004)		0.044 (0.028)		0.008* (0.004)
Proportion - President		0.007 (0.723)		-0.048 (0.072)		0.218 (0.672)		0.025 (0.065)
Proportion - Military		1.283 (0.848)		-0.011 (0.111)		1.870*** (0.685)		0.158* (0.092)
Proportion - Defense		1.934*** (0.538)		0.123* (0.069)		1.308** (0.513)		0.076 (0.068)
Proportion - Intelligence		2.026*** (0.719)		0.171* (0.098)		1.703** (0.706)		0.140 (0.097)
Proportion - State		0.241 (0.493)		-0.032 (0.046)		0.150 (0.491)		-0.017 (0.046)
Diplomatic Experience		0.138* (0.074)		0.005 (0.009)		0.093 (0.069)		-0.0002 (0.009)
Intelligence Experience		-0.034 (0.075)		0.001 (0.011)		0.010 (0.069)		0.006 (0.011)
Military Experience		0.062 (0.121)		0.007 (0.013)		0.086 (0.111)		-0.002 (0.011)
Formal	0.955*** (0.137)	0.199 (0.209)	0.098*** (0.020)	0.036 (0.031)	1.158*** (0.120)	0.379** (0.192)	0.117*** (0.018)	0.045 (0.030)
5-Year MIDs		-1.127*** (0.389)		-0.101** (0.046)		-0.437* (0.243)		-0.038 (0.028)
US CINC		4.951 (4.052)		1.379** (0.542)		0.244 (1.706)		0.247 (0.208)
Constant	-8.670*** (1.043)	-7.420*** (1.921)	-0.483*** (0.120)	-0.789*** (0.237)	-1.940*** (0.272)	-2.639** (1.072)	0.139*** (0.037)	0.092 (0.129)
Administration FEs	✓	✓	✓	✓				
Agenda Items	✓	✓	✓	✓	✓	✓	✓	✓
Observations	2,877	2,877	2,877	2,877	2,880	2,880	2,880	2,880

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

Table 10: Effect of Adviser Hawkishness on Foreign Policy Decisions, Removing Decisions Involving the USSR

	<i>Adviser Model</i>				<i>Advisers + Admin. FEs</i>			
	Conflictual Decisions		Conflictual – Cooperative		Conflictual Decisions		Conflictual – Cooperative	
	<i>Poisson</i>		<i>OLS</i>		<i>Poisson</i>		<i>OLS</i>	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Advisers' Hawkishness (Acts)	2.683** (1.078)	3.284** (1.494)	0.275* (0.142)	0.373** (0.173)	6.141*** (1.315)	2.098 (1.932)	0.581*** (0.175)	0.451** (0.214)
President's Hawkishness	-2.262*** (0.535)	-3.052*** (0.889)	-0.280*** (0.076)	-0.373*** (0.122)				
No. of Attendees		0.062** (0.029)		0.010** (0.004)		0.046 (0.030)		0.007 (0.004)
Proportion - President		0.874 (0.720)		0.056 (0.072)		1.057 (0.770)		0.029 (0.077)
Proportion - Military		1.703** (0.739)		0.105 (0.101)		2.159** (0.848)		0.067 (0.110)
Proportion - Defense		2.085*** (0.562)		0.127* (0.072)		2.360*** (0.583)		0.158** (0.073)
Proportion - Intelligence		2.626*** (0.752)		0.211** (0.103)		2.745*** (0.740)		0.217** (0.103)
Proportion - State		0.941* (0.544)		0.029 (0.052)		0.896 (0.550)		0.025 (0.053)
Diplomatic Experience		0.092 (0.070)		-0.001 (0.009)		0.144* (0.075)		0.005 (0.009)
Intelligence Experience		-0.037 (0.071)		0.001 (0.011)		-0.078 (0.076)		-0.003 (0.011)
Military Experience		0.097 (0.120)		-0.005 (0.012)		0.083 (0.126)		-0.001 (0.013)
Formal	1.073*** (0.128)	0.335* (0.196)	0.112*** (0.019)	0.045 (0.031)	1.139*** (0.134)	0.231 (0.212)	0.118*** (0.020)	0.041 (0.032)
5-Year MIDs		-0.481* (0.251)		-0.047 (0.030)		-1.291*** (0.409)		-0.121** (0.048)
US CINC		-2.096 (1.821)		0.037 (0.225)		3.688 (4.172)		1.257** (0.566)
Constant	-3.139*** (0.607)	-3.774*** (1.268)	0.027 (0.077)	0.014 (0.147)	-6.863*** (0.858)	-6.008*** (1.910)	-0.362*** (0.110)	-0.593** (0.233)
Administration FEs					✓	✓	✓	✓
Agenda Items	✓	✓	✓	✓	✓	✓	✓	✓
Observations	2,704	2,704	2,704	2,704	2,705	2,705	2,705	2,705

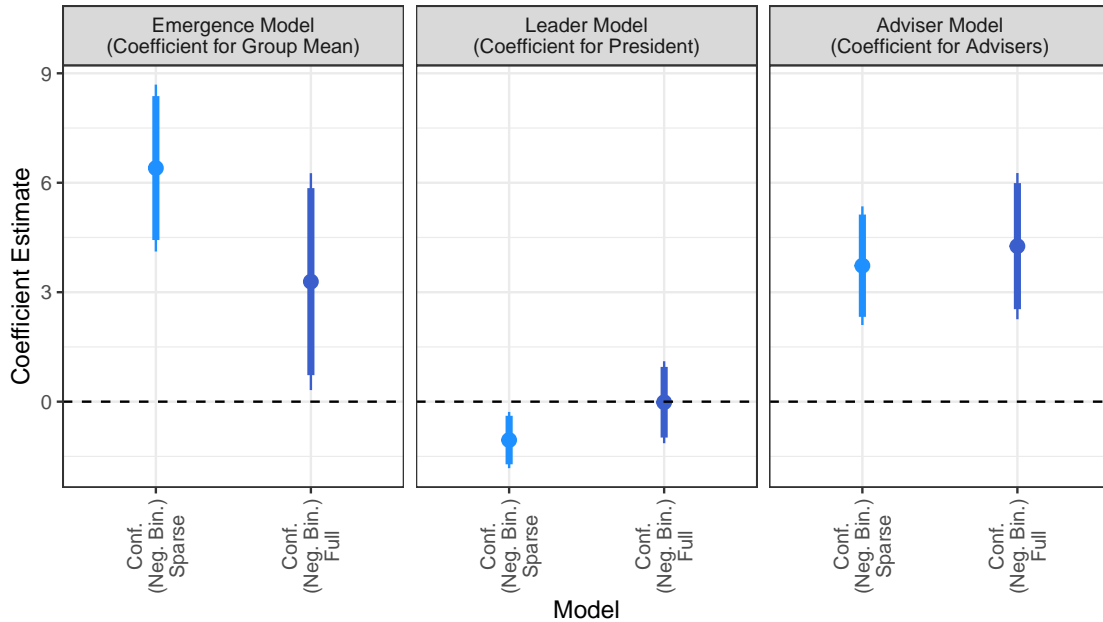
Note: * $p < 0.1$; ** $p < 0.05$; *** $p < 0.01$

4.4 Negative Binomial

One half of the main statistical analysis of conflictual decisions made in meetings relies on Poisson regression models in which the outcome of interest is a count variable. The variable itself does not feature a prominent amount of overdispersion; the mean value across all meetings is 0.252, while the variance is 0.486. However, dispersion tests suggest that a negative binomial model may still be appropriate to account for overdispersion.

Tables 11 and 12 replicate the Poisson models in the main text but instead rely on negative binomial models. Figure 6 illustrates the corresponding coefficient plots. All results are consistent with the findings from simpler Poisson models.

Figure 6: Summary of Three Models of Trait Aggregation, Using Negative Binomial Models



Note: Coefficient plots corresponding to main findings, corresponding to four models on each table. Thicker and smaller bands represent 90% and 95% confidence intervals respectively.

Table 11: Effect of Mean Participant Hawkishness and President's Hawkishness on Foreign Policy Decisions, Using Negative Binomial Models

	<i>Emergence Model</i>		<i>Leader Model</i>	
	Conflictual Decisions			
	(1)	(2)	(3)	(4)
Mean Hawkishness	6.403*** (1.167)	3.290** (1.517)		
President's Hawkishness			-1.052*** (0.393)	-0.015 (0.573)
No. of Attendees		0.074*** (0.021)		0.085*** (0.020)
Proportion - President		0.959* (0.500)		1.086** (0.458)
Proportion - Military		1.291** (0.619)		1.574*** (0.517)
Proportion - Defense		1.849*** (0.380)		1.320*** (0.375)
Proportion - Intelligence		1.570*** (0.608)		1.219** (0.597)
Proportion - State		0.318 (0.355)		0.239 (0.363)
Diplomatic Experience		0.105** (0.052)		0.067 (0.050)
Intelligence Experience		-0.126** (0.059)		-0.111** (0.054)
Military Experience		0.028 (0.088)		0.081 (0.080)
Formal	1.429*** (0.103)	0.568*** (0.156)	1.749*** (0.094)	0.753*** (0.150)
5-Year MIDs		-0.155 (0.256)		0.112 (0.173)
US CINC		5.940** (2.700)		2.892** (1.212)
Constant	-5.884*** (0.728)	-6.689*** (1.331)	-1.690*** (0.209)	-4.286*** (0.775)
Administration FEs	✓	✓		
Agenda Items	✓	✓	✓	✓
Observations	2,877	2,877	2,880	2,880

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

Table 12: Effect of Adviser Hawkishness on Foreign Policy Decisions, Using Negative Binomial Models

	<i>Adviser Model</i>		<i>Advisers + Admin. FEs</i>	
	Conflictual Decisions			
	(1)	(2)	(3)	(4)
Advisers' Hawkishness (Acts)	3.727*** (0.830)	4.263*** (1.022)	5.251*** (1.018)	3.268** (1.334)
President's Hawkishness	-1.541*** (0.400)	-1.199* (0.618)		
No. of Attendees		0.110*** (0.020)		0.096*** (0.021)
Proportion - President		1.579*** (0.487)		1.598*** (0.524)
Proportion - Military		1.337** (0.541)		1.650*** (0.605)
Proportion - Defense		2.074*** (0.399)		2.250*** (0.405)
Proportion - Intelligence		2.133*** (0.618)		2.125*** (0.623)
Proportion - State		0.950** (0.391)		0.880** (0.389)
Diplomatic Experience		0.069 (0.050)		0.103** (0.052)
Intelligence Experience		-0.168*** (0.056)		-0.153** (0.060)
Military Experience		0.029 (0.084)		0.0004 (0.087)
Formal	1.630*** (0.099)	0.698*** (0.152)	1.531*** (0.102)	0.571*** (0.158)
5-Year MIDs		0.025 (0.176)		-0.215 (0.261)
US CINC		1.042 (1.258)		5.356* (2.735)
Constant	-3.479*** (0.463)	-5.747*** (0.884)	-5.208*** (0.647)	-6.815*** (1.297)
Administration FEs			✓	✓
Agenda Items	✓	✓	✓	✓
Observations	2,704	2,704	2,705	2,705

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

4.5 Time-Unit Replication Analysis

One may be concerned that the unit of analysis in our main text – the individual meeting – could introduce issues of selection effects, as the timing of, attendance at, and participation in meetings are likely not random.

To address this concern, it may be appropriate to consider an alternate research design that tracks propensity to engage in conflictual behavior over equally spaced periods of time, using measures of hawkishness that are created based on the characteristics of key actors regardless of their actual participation in decision-making meetings. Evidence of a positive relationship between conflictual behavior and the hawkishness of important decision-making elites using this approach would bolster confidence in our main finding.

We perform this robustness check by creating monthly-level data of conflictual activity and hawkishness. Conflictual activity is measured using new militarized interstate disputes (MIDs) launched by the United States in a given month. We limit our analysis to MIDs that feature a hostility level of 3 or higher.² To capture hawkishness, we calculate the mean hawkishness of the NSC principals (President, Vice President, Secretary of State, Secretary of Defense, Director of Central Intelligence, and the Chairman of the Joint Chiefs) serving in office each month.

Table 13 displays the results of two Poisson models that regress new MIDs initiated by the US on the mean hawkishness of NSC principals at the monthly level. We include several control variables included in analysis of US presidents by [Dafoe and Caughey \(2016\)](#), which are meant to capture structural and political conditions that may influence decisions regarding conflict. Since the original measures in the [Dafoe and Caughey \(2016\)](#) article are at the leader level, we create analogous versions of these measures at the monthly level. We continue to find a positive and statistically significant relationship between conflictual activity and key decision-makers' hawkishness.

²We do not use our conflictual decision variable, since these are explicitly linked to individual meetings.

Table 13: Effect of NSC Principals' Hawkishness on MID Initiation, Using Monthly Data

	<i>Dependent variable:</i>	
	US MID Initiation	
	(1)	(2)
Principals' hawkishness	1.309** (0.663)	2.551** (1.232)
President's hawkishness	-0.496 (0.368)	-1.267* (0.753)
War ongoing		-0.065 (0.048)
Deaths per capita in last war (logged)		0.021 (0.033)
Months since last war (logged)		0.060 (0.049)
Victory in last war		0.005 (0.196)
MID challenges to US in last 5 years		-0.005 (0.021)
Average MID outcome in last 5 years		-0.198 (10.114)
Economic recession		-0.170** (0.070)
Unified government		0.111 (0.078)
US material capabilities		-4.277** (2.070)
President's tenure (logged months)		0.020 (0.029)
Constant	0.714*** (0.247)	1.159 (0.839)
Observations	501	501

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

5 Discussion of Leader Results

Our central proposition in this paper is that predispositions of advisers affect a state’s foreign policy behavior. We believe this finding in itself makes a valuable contribution, regardless of whether leader predispositions exert similar effects on foreign policy behavior in parallel to advisers.

Nevertheless, one of the striking findings reported in the main text is that we find more support for adviser-centric models than leader-centric ones: groups that feature higher deliberation participation by hawkish advisers during group interactions were much more likely to choose conflictual foreign policies, whereas leader-level hawkishness never displayed a significant positive association with the conflictual nature of the group’s decision. In the discussion below, we first show that our results replicate with measures of presidential hawkishness derived from an expert survey, and also obtain similar results using replications and extensions of [Yarhi-Milo \(2018\)](#), [Horowitz, Stam and Ellis \(2015\)](#) and [Carter and Smith \(2020\)](#), all of which display null or negative effects between presidential hawkishness and US conflict initiation. We then present a number of potential explanations for these results.

5.1 Expert survey results

To show the leader-level results are not an artifact of our measure of hawkishness, we also fielded an expert survey, where we asked 14 prominent political scientists and historians who had published work on Cold War-era US foreign policy to score Truman, Eisenhower, Kennedy, Johnson, Nixon, and Ford in terms of their hawkishness on a 1-to-4 scale, where 1 represents the least hawkish, while 4 represents the most hawkish.³ We took the responses of all 14 experts and calculated an average score for each president, to produce an alternative measure of presidential hawkishness we can use to replicate the results in the main text.⁴

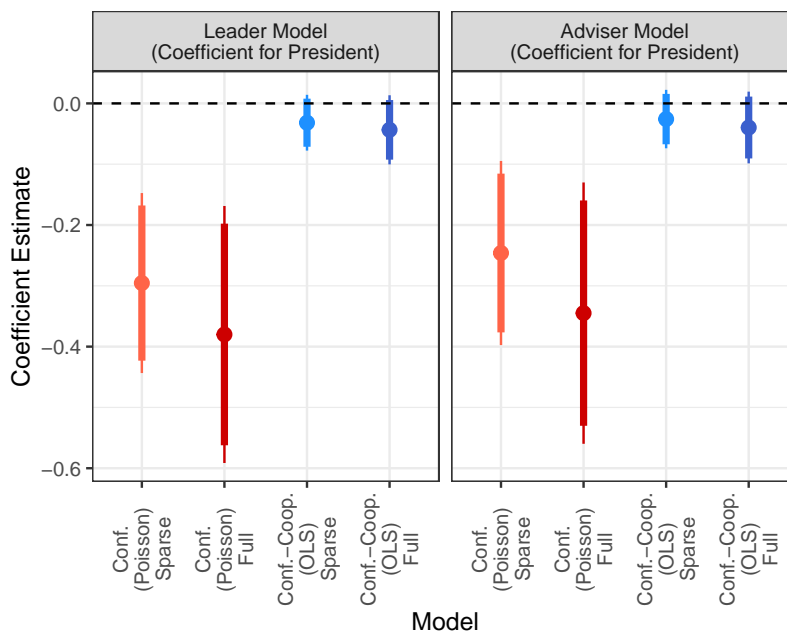
Table 14 replicates the analysis from the paper, but utilizing our expert codings for presidential hawkishness instead. Across all relevant models, we continue to see either statistically significant

³For other examples of expert surveys in political science, see [Braumoeller \(2013\)](#) and [Yarhi-Milo \(2018\)](#). Our expert survey, which contains no personally identifying information, was declared exempt by the Institutional Review Board of [blinded for peer review]. Respondents, who were political scientists and historians who had published work on Cold War-era US foreign policy, were assured anonymity and confidentiality of their responses.

⁴One expert did not provide a score for President Ford. Ford’s average score is therefore based on 13 responses. Note that we do not have expert codings of Carter or Reagan, which results in the omission of 139 meetings, or 5% of all meetings in the main paper.

and negative coefficients for president’s hawkishness or coefficients without any statistical significance. Figure 7 visually reinforces this point.

Figure 7: Summary of Two Models of Trait Aggregation, Using Expert Codings



Note: Coefficient plots corresponding to main findings, corresponding to four models on each table. Thicker and smaller bands represent 90% and 95% confidence intervals respectively.

5.2 Consistency with Existing Findings on US Presidential Hawkishness

Importantly, our findings regarding leader hawkishness are consistent with previous studies of American foreign policy. By this, we do not mean to suggest that the existing literature has found the leader predispositions *in general* do not matter. The main text provides an overview of the many findings that support various ways that leaders might affect policy outcomes. However, most of the literature that directly examines American foreign policy does not specifically examine the president’s *hawkishness*, but rather other traits like honor culture (Dafoe and Caughey, 2016), personality traits (Gallagher and Allen, 2014; Harden, Forthcoming), self-monitoring (Yarhi-Milo, 2018), or general beliefs about intervention or nuclear weapons (Saunders, 2011; Whitlark, 2017). Our findings regarding leader hawkishness do not, of course, undermine the validity of these other pathways by which leaders might matter.

Nonetheless, while a growing literature focuses on the importance of leader-level characteristics

Table 14: Effect of President's Hawkishness on Foreign Policy Decisions, Using Expert Codings

	<i>Leader Model</i>				<i>Adviser Model</i>			
	Conflictual Decisions		Confictual – Cooperative		Conflictual Decisions		Confictual – Cooperative	
	<i>Poisson</i>	<i>OLS</i>	<i>Poisson</i>	<i>OLS</i>	<i>Poisson</i>	<i>OLS</i>	<i>Poisson</i>	<i>OLS</i>
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
President's Hawkishness	-0.295*** (0.076)	-0.380*** (0.108)	-0.032 (0.023)	-0.043 (0.029)	-0.246*** (0.077)	-0.345*** (0.110)	-0.026 (0.025)	-0.040 (0.030)
Advisers' Hawkishness (Acts)					3.429*** (0.761)	2.784*** (1.051)	0.723*** (0.241)	0.686*** (0.315)
No. of Attendees		0.089*** (0.017)		0.039*** (0.007)		0.102*** (0.017)		0.044*** (0.007)
Proportion - President		1.616*** (0.443)		0.287*** (0.109)		1.848*** (0.478)		0.285*** (0.125)
Proportion - Military		1.910*** (0.465)		0.321** (0.151)		1.629*** (0.516)		0.223 (0.173)
Proportion - Defense		1.314*** (0.330)		0.075 (0.111)		1.773*** (0.353)		0.167 (0.118)
Proportion - Intelligence		1.466*** (0.532)		0.287* (0.162)		2.037*** (0.563)		0.386*** (0.171)
Proportion - State		0.078 (0.335)		-0.019 (0.076)		0.629* (0.365)		0.077 (0.088)
Diplomatic Experience		0.127*** (0.046)		0.003 (0.014)		0.110** (0.048)		-0.002 (0.015)
Intelligence Experience		-0.086* (0.048)		-0.040** (0.018)		-0.097** (0.049)		-0.043** (0.019)
Military Experience		-0.041 (0.074)		-0.035* (0.019)		-0.023 (0.078)		-0.040* (0.021)
Formal	1.699*** (0.083)	0.534*** (0.137)	0.428*** (0.032)	0.137*** (0.051)	1.625*** (0.086)	0.494*** (0.139)	0.415*** (0.033)	0.128*** (0.053)
5-Year MID Challenges		-0.018 (0.127)		-0.025 (0.038)		0.075 (0.138)		0.010 (0.044)
US CINC		5.794*** (1.133)		1.793*** (0.361)		4.653*** (1.199)		1.537*** (0.395)
Constant	-1.464*** (0.197)	-3.404*** (0.544)	0.131** (0.062)	-0.331** (0.136)	-3.471*** (0.493)	-5.301*** (0.847)	-0.279* (0.151)	-0.764*** (0.231)
Administration FEs	✓	✓	✓	✓	✓	✓	✓	✓
Agenda Items								
Observations	2,704	2,704	2,704	2,704	2,705	2,705	2,705	2,705

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

in IR more generally, the findings of the leader literature on the effects of leader-level hawkishness on foreign policy behavior in the United States is mixed. Much of the existing quantitative research on hawkishness in foreign policy is at the party- rather than leader-level (e.g. Palmer, London and Regan, 2004; Koch and Cranmer, 2007; Foster, 2008; Clare, 2014; Williams, 2014; Heffington, 2018; Bertoli, Dafoe and Trager, 2019). One of the few exceptions is Yarhi-Milo (2018): while her analysis is primarily concerned with whether high/low self-monitoring leaders are more/less prone to militarized disputes, it also includes a control variable for presidential hawkishness, estimated using a WordScore analysis of presidential speeches.

Interestingly, while Yarhi-Milo’s measurement strategy is different from ours, the models yield results quite similar to our own.⁵ Specifically, Yarhi-Milo finds a *negative* and weakly significant ($p < 0.1$) relationship between presidential hawkishness and militarized dispute involvement and a negative and insignificant statistical relationship between presidential hawkishness and militarized dispute initiation. That is, both the direction and strength of presidential hawkishness as a predictor of conflictual behavior in Yarhi-Milo (2018) match our findings. Yarhi-Milo also includes a second proxy for hawkishness by coding the president’s party affiliation (Republicans as hawks; Democrats as doves). This approach finds that neither the relationship between presidential hawkishness and militarized disputes involvement nor the relationship between presidential hawkishness and militarized dispute initiation are statistically significant.

One might wonder whether congruence of these results is simply a coincidence. To explore this possibility, we replicate and extend two foundational cross-national studies that have shown how leader dispositions affect interstate conflict – Horowitz and Stam (2014) and Carter and Smith (2020) – but subset them to the United States. Our approach here is akin to Potter (2007) and Johnston (2012), who note that patterns that hold in the aggregate in cross-national models may not hold in specific subsets of states. We choose these studies because their approach is similar in spirit to our own: estimating hawkishness at a distance by coding biographical characteristics of an individual’s background and experience.⁶ We discuss each in turn.

⁵Yarhi-Milo’s results are reported in Table 4.5 of the original book (pp. 89-90).

⁶We chose not to replicate existing studies that were not directly concerned with hawkishness (e.g. Croco, 2011; Colgan and Weeks, 2015; Dafoe and Caughey, 2016; McManus, 2019; Fuhrmann, 2020) or not directly concerned with conflictual foreign policy (e.g. Fuhrmann and Horowitz, 2015).

5.2.1 Horowitz and Stam (2014)

The first study we replicate is [Horowitz and Stam \(2014\)](#). Built upon the Archigos leader dataset ([Goemans, Gleditsch and Chiozza, 2009](#)), Horowitz and Stam code numerous leader characteristics, including military experience – both with and without exposure to conflict – which they argue affects a leader’s dispositions. Specifically, they posit that leaders with non-combat military experience should be more likely to be hawkish but that leader exposure to combat may temper this propensity to support interstate violence. In the full cross-national analysis, Horowitz and Stam find that leaders with military experience but without accompanying combat experience (i.e. more hawkish leaders) are indeed more likely to initiate militarized interstate disputes (MIDs) and interstate wars, while those with military and combat experience (i.e. less hawkish leaders) are not.

The original dataset has 11,525 leader-year observations, of which 2,175 involve MID initiation. The US-only subset has 180 president-year observations, of which 109 feature MID initiation. The dataset used to analyze wars has 11,807 leader-year observations, of which 114 experience war initiation. The US-only subset has 128 president-year observations and five war initiations.

Table 15 replicates Models 1 and 2 as reported on Table 2 of the original article ([Horowitz and Stam, 2014](#), 543-544). Models 1 and 3 in this table are bivariate versions of the analysis and are provided for informational purposes only. Models 2 and 4 in this present table are direct replications of the “MID model” and “war model” by Horowitz and Stam. Note that several variables in the original analyses are dropped because of lack of variation in the US-only data.

The fully-specified models yield findings generally consistent with our own leader models. In Model 2, there is a *negative* and statistically insignificant relationship between leaders with non-combat military experience (i.e. hawkish presidents) and MIDs. With respect to war initiation, Model 4 produces a similar and statistically significant result: presidents who have served in the military but had no combat experience are less likely to initiate conflicts. In short, the US-specific coefficients in both Model 2 and Model 4 are in the opposite direction of Horowitz and Stam’s cross-national results but consistent with our own. However, the war initiation finding should be interpreted with caution given that only five war initiations exist in the data. Of five US presidents that had military experience but no combat experience (Chester A. Arthur, Richard Nixon, Jimmy Carter, Ronald Reagan, and George W. Bush), only one – George W. Bush – initiated a war during

his time in office.

Table 15: Replication of Models in Table 1 of Horowitz and Stam (2014), Only US Presidents

	<i>Dependent variable:</i>			
	MID initiation		War initiation	
	(1)	(2)	(3)	(4)
Military service, no combat	1.161 (0.756)	-0.127 (0.613)	0.907 (1.482)	-20.827*** (1.715)
Military service, combat	0.419 (0.598)	0.132 (0.583)	1.138 (1.195)	0.285 (1.104)
Leader age		0.016 (0.035)		0.033 (0.073)
Material capabilities		-10.176*** (3.741)		-26.553 (16.261)
Tau B with system leader		0.781 (0.486)		-2.093 (1.421)
Time in office		0.144 (0.230)		0.922* (0.515)
Five-year MID challenge lag		-0.134 (0.450)		-16.645*** (1.408)
Constant	0.056 (0.497)	1.010 (1.935)	-3.951*** (0.929)	0.583 (5.421)
Observations	180	178	128	126
Log Likelihood	-117.295	-108.041	-20.557	-12.247
Akaike Inf. Crit.	240.590	238.083	47.114	46.493
Peace Year Splines	✓	✓	✓	✓
Clustered SEs (leaders)	✓	✓	✓	✓
Only US Leaders	✓	✓	✓	✓

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

5.2.2 Carter and Smith (2020)

The second study we extended is [Carter and Smith \(2020\)](#), which creates a measure of leaders' latent willingness to use force. Using a Bayesian latent variable framework, Carter and Smith create four models that each produce an estimated of hawkishness for all state leaders between 1875 and 2004. These four measures are referred to as M1, M2, M3, and M4. In the original analysis, Carter and Smith showed that the four measures of leaders' latent hawkishness outperform military experience (as in Horowitz and Stam) in predicting the the initiation of ICB crises, militarized interstate disputes (MIDs), and severe MIDs.⁷

For our purposes, however, we are interested in whether these four measures of latent hawkishness predict conflictual foreign policy when subset to the United States. Since the original text in Carter and Smith does not directly address this question, we conduct two analyses similar in spirit to our replication of Horowitz and Stam in Section 5.2.1 above. Specifically, we subset the four measures of latent hawkishness to US presidents, combined these measures with the control variables used by Horowitz and Stam, and then examine whether they predict the initiation of MIDs, severe MIDs, and interstate war.

The results are again generally consistent with the findings in our leader models. First, [Table 16](#) shows that two of the latent hawkishness measures display a *negative* and statistically significant relationship with MID initiation, whereas two are positive and not statistically significant. Second, [Table 17](#) reports an analysis of severe MID initiation. The coefficients of all four measures of latent hawkishness are negative and not statistically significant. Third, [Table 18](#) shows that there are negative and statistically significant relationships between each latent hawkishness measure and war initiation.

Finally, [Table 19](#) replicates the original analysis in [Table 2](#) of [Carter and Smith \(2020\)](#) – which compares the predictive performance of the models based on M1, M2, M3, and M4 data with a model based on prior military service – but only using observations from the United States. Compared to the original table, the US-centric models do not perform very differently from one

⁷Each of these models is compared to a baseline in which the universe of ICB crises, MIDs, and severe MIDs are predicted using the leader's prior military service. Carter and Smith suggest that M2 exhibits the highest performance in predicting both ICB crises and severe MIDs. That said, [Table 2](#) in the original text (p. 1357) presents Vuong statistics which show that almost every single model based on the latent measures outperforms the prior military experience model.

another. Only two models (using M3 and M4 to predict ICB initiation) perform better than the prior military model, and only at the 90% level.

In sum, similar to our replication of the analysis in Horowitz and Stam, we find that latent presidential hawkishness does not predict US conflict behavior in ways consistent with cross-national models. Instead, the results are consistent with our finding that presidential hawkishness exhibits only a modest – and, if anything, inverse – relationship with conflictual behavior in American foreign policy. These findings thus suggest that the weak relationship we obtain between presidential hawkishness and US foreign policy behavior is unlikely to be an artifact of our measures, and actually reflects a broader pattern in the existing literature that appears to have gone unappreciated. In the discussion below, we present a number of potential explanations for this finding.

5.3 Leader Constraints

One set of explanations might emphasize the usual constraints that might “box” leaders in – either from above by the pressures of the international system, or from below by the vice grip of bureaucratic politics (Jervis, 2013). In the American context, the US national security bureaucracy may exert unusually high influence over foreign policy (Jost, 2021), constraining American presidents to an extent that is less the case in other contexts. Yet other research has found evidence of the importance of leader-level factors in American foreign policy (e.g. Larson, 1985; Saunders, 2011; Yarhi-Milo, 2018), suggesting that the usual constraints of the international system and bureaucratic politics are insufficient by themselves to explain our pattern of results.

5.4 Measurement of Hawkishness

A second set of potential explanations stem from our measurement strategy for the explanatory variable. For instance, one explanation relates to the difficulties of accurately measuring hawkishness at a distance – though it is unclear why this would plague leader-level hawkishness measures but not adviser-level hawkishness ones, which display results consistent with theoretical expectations. Given that there are only eight presidents but hundreds of advisers, it is possible that measurement error is more pronounced in the leader models than adviser-centric models.

Yet as noted in section 5.1, we replicate our results using the expert survey, the results of which (in Appendix 5.1) similarly show a weak or negative effect for presidential hawkishness. As noted in

Table 16: Extension of Carter and Smith (2020): MID Initiation, Only US Presidents

	<i>Dependent variable:</i>				
	MID initiation				
	(1)	(2)	(3)	(4)	(5)
Military	-0.057 (0.617)				
M1		0.628 (0.731)			
M2			0.241 (0.632)		
M3				-1.191** (0.585)	
M4					-1.196** (0.568)
Prior war win	0.243 (0.569)	-0.356 (0.731)	-0.021 (0.672)	0.441 (0.363)	0.448 (0.363)
Leader age	0.014 (0.030)	0.004 (0.029)	0.007 (0.032)	0.033 (0.029)	0.037 (0.030)
Material capabilities	-10.212*** (3.443)	-9.218*** (3.219)	-9.710*** (3.215)	-8.134** (3.191)	-8.160** (3.180)
Tau B with system leader	0.791** (0.366)	0.707* (0.367)	0.762** (0.361)	0.822** (0.365)	0.806** (0.366)
Time in office	0.128 (0.210)	0.133 (0.208)	0.142 (0.209)	0.164 (0.214)	0.151 (0.214)
Five-year MID challenge lag	-0.103 (0.429)	-0.128 (0.430)	-0.104 (0.429)	-0.061 (0.436)	-0.078 (0.436)
Peace year splines 1	1.061 (1.490)	1.701 (1.647)	1.446 (1.761)	0.700 (1.502)	0.472 (1.521)
Constant	1.005 (1.512)	1.578 (1.670)	1.304 (1.782)	0.757 (1.521)	0.532 (1.537)
Observations	178	178	178	178	178
Log Likelihood	-108.052	-107.811	-108.038	-106.036	-105.891
Akaike Inf. Crit.	238.103	237.622	238.077	234.071	233.781
Peace Year Splines	✓	✓	✓	✓	✓
Only US Presidents	✓	✓	✓	✓	✓

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

Table 17: Extension of Carter and Smith (2020): Severe MID Initiation, Only US Presidents

	<i>Dependent variable:</i>				
	Severe MID initiation				
	(1)	(2)	(3)	(4)	(5)
Military	-0.052 (0.930)				
M1		-0.548 (1.014)			
M2			-0.637 (0.926)		
M3				-0.910 (0.899)	
M4					-0.913 (0.870)
Prior war win	0.548 (0.787)	0.976 (1.001)	1.073 (0.970)	0.740 (0.567)	0.748 (0.567)
Leader age	-0.024 (0.044)	-0.016 (0.043)	-0.006 (0.048)	-0.013 (0.042)	-0.008 (0.044)
Material capabilities	1.629 (4.443)	1.128 (4.256)	0.974 (4.252)	3.435 (4.555)	3.411 (4.531)
Tau B with system leader	0.044 (0.515)	0.095 (0.517)	0.071 (0.508)	0.096 (0.522)	0.078 (0.519)
Time in office	0.012 (0.307)	0.012 (0.299)	-0.020 (0.305)	0.050 (0.306)	0.035 (0.306)
Five-year MID challenge lag	1.983* (1.094)	1.989* (1.091)	1.959* (1.090)	2.046* (1.104)	2.026* (1.102)
Constant	-2.995 (2.251)	-3.569 (2.482)	-4.062 (2.759)	-3.276 (2.234)	-3.488 (2.278)
Observations	178	178	178	178	178
Log Likelihood	-57.164	-57.014	-56.915	-56.637	-56.594
Akaike Inf. Crit.	130.328	130.029	129.830	129.273	129.188
Peace Year Splines	<i>No</i>	<i>No</i>	<i>No</i>	<i>No</i>	<i>No</i>
Only US Presidents	✓	✓	✓	✓	✓

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

Table 18: Extension of Carter and Smith (2020): Interstate War, Only US Presidents

	<i>Dependent variable:</i>				
	initiation				
	(1)	(2)	(3)	(4)	(5)
Military	-17.773*** (1.694)				
M1		-276.014*** (93.262)			
M2			-23.892** (9.846)		
M3				-10.699*** (2.651)	
M4					-11.153*** (2.459)
Prior war win	18.719*** (1.081)	267.652*** (89.740)	27.308** (10.648)	4.272*** (1.592)	4.713*** (1.656)
Leader age	0.033 (0.044)	-0.087 (0.068)	0.307*** (0.109)	0.188*** (0.058)	0.256*** (0.065)
Material capabilities	-13.930 (11.909)	-42.495*** (15.271)	-6.438 (13.621)	-13.965 (14.728)	-13.934 (15.902)
Tau B with system leader	0.523 (1.226)	1.452 (0.928)	-0.647 (1.178)	2.267 (1.390)	2.167* (1.264)
Time in office	0.296 (0.559)	1.996*** (0.566)	0.355 (0.701)	0.689 (0.568)	0.534 (0.605)
Five-year MID challenge lag	-16.581*** (0.843)	-16.917*** (0.617)	-17.538*** (0.865)	-17.557*** (0.764)	-17.438*** (0.685)
Constant	-3.474 (2.603)	-134.039*** (44.404)	-34.495*** (11.699)	-6.525*** (2.442)	-10.307*** (2.850)
Observations	126	126	126	126	126
Log Likelihood	-14.554	-10.568	-12.209	-9.776	-9.492
Akaike Inf. Crit.	45.108	37.135	40.419	35.552	34.985
Peace Year Splines	<i>No</i>	<i>No</i>	<i>No</i>	<i>No</i>	<i>No</i>
Clustered SEs (leaders)	✓	✓	✓	✓	✓
Only US Presidents	✓	✓	✓	✓	✓

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

Table 19: Replication of Table 2 in Carter and Smith (2020) using only observations from the United States.

Panel A: ICB initiation				
	Military	M1	M2	M3
Military	-	-	-	-
M1	0.78			
M2	0.99	-0.03		
M3	1.37 [†]	1.16	1.11	
M4	1.41 [†]	1.15	1.13	-0.40
Panel B: MID initiation				
	Military	M1	M2	M3
Military	-	-	-	-
M1	0.25			
M2	0.79	1.03		
M3	0.14	0.08	-0.04	
M4	0.17	0.11	-0.01	0.62
Panel C: Severe MID initiation				
	Military	M1	M2	M3
Military	-	-	-	-
M1	0.33			
M2	0.59	0.54		
M3	0.41	0.32	0.25	
M4	0.40	0.31	0.24	-0.19

Note: [†] $p < 0.1$; * $p < 0.05$; ** $p < 0.01$

the main text, our negative results for leader hawkishness is largely driven by Dwight Eisenhower, who our measure codes as relatively dovish but whose administration nonetheless made a substantial number of conflictual decisions. We therefore carried out a number of supplementary tests. First, we administered the expert survey described above to rate the hawkishness of the presidents in our sample. Consistent with our boosted GLM measure, the average expert rating characterizes Eisenhower as comparatively dovish (two on a four-point scale). Second, given the high variance of Eisenhower’s hawkishness score in the expert survey, we performed an additional robustness check in which we removed Eisenhower from our analysis. When we do so, the estimated coefficient of leader hawkishness loses statistical significance and there is no meaningful relationship associated with the leader model. Third, we reran the leader model with the neural net measure of presidential hawkishness (which codes Eisenhower as comparatively hawkish) and found similarly null results. Collectively, these robustness checks suggest that our weak leader-level findings are not an artifact of our hawkishness measure.

5.5 Strategic Behavior

A third set of explanations are rooted in the strategic choice literature, which emphasizes the challenges in studying strategic interactions with observational data (Signorino, 1999). If leader hawkishness is observable, adversaries should take it into account and adjust their behavior accordingly (Smith, 1996; Foster, 2008; Clark, Fordham and Nordstrom, 2011). As a result, we may be more likely to observe conflict when dovish leaders are in power than when hawkish ones, because the former can be strategically targeted, whereas adversaries will avoid provocative acts when the latter are in charge. Yet if this is the case, it is unclear why foreign observers take leader hawkishness into account but not adviser hawkishness, which at least for top-level advisers, should be at least partially observable to adversary decision-makers. US foreign policy analysts routinely look at the advisers leaders surround themselves with to gain insights into their potential foreign policy decisions – the foundational line of inquiry for the Kremlinologists who studied the elite politics of the Soviet Union during the Cold War.

Another explanation rooted in strategic choice points to the importance of leaders' incentive structures. A number of theoretical models in IR argue that leaders face political incentives to go “against type” (e.g. Schultz, 2005; Kreps, Saunders and Schultz, 2018; Mattes and Weeks, 2019; Kertzer and Brooks, 2021), such that leaders' domestic or international incentives may trump their foreign policy dispositions (Saunders, 2022). If hawks are sometimes more likely to “bring the olive branch,” this might exert a countervailing effect in the opposite direction of the leader's hawkishness. At the same time, the bureaucratic politics literature suggests that the bureaucratic incentives of advisers should typically trump predispositions like hawkishness – Miles' Law, which holds that where you stand depends on where you sit (Marsh, 2014). As such, it is unclear why leader incentives override predispositions but adviser incentives do not, suggesting that these strategic choice explanations are at best incomplete.

5.6 Measurement of the Dependent Variable

Another potential explanation focuses on the operationalization of our outcome measure, which encompasses a wider range of foreign policy behaviors than those in much of the existing literature on leaders, many of which focus strictly on high-level conflictual decisions in militarized interstate

disputes. Our analysis instead includes a broader range of conflictual and cooperative policies national security groups might choose. As discussed in the main text, this includes other important behaviors such as military threats, troop deployments, military spending, economic restrictions, withholding aid, arms control, or diplomatic engagement. It is possible that leaders are more influential in deciding to use military force, whereas advisers are influential across a broad range of national security decisions. If true, this would suggest an important – and overlooked – scope condition on the the study of leaders that political scientists should be careful to emphasize, at least within American foreign policy. It would also reinforce the importance of concerns about truncation bias in IR (Mitchell and Moore, 2002). At the same time, however, analysis in Appendix §4.5 shows our results are robust to a more restrictive unit of analysis that only focuses on militarized interstate disputes, suggesting our fine-grained measure of the dependent variable is unlikely to explain the relatively weak effects of leader-level hawkishness here.

5.7 Institutional Context

A fifth explanation, which we find the most persuasive, suggests that whether leader hawkishness matters depends on the institutional setting. Specifically, leaders may *less* influential in formal, institutionalized settings – such as a meeting of the National Security Council (NSC). In the US system, this might be because the National Security Act of 1947 mandates certain cabinet members to serve on the NSC, although presidents have historically used policies and directives to shape the other attendees. Advisers might also be better positioned to persuade leaders in formal meetings with an established agenda for which they can mobilize bureaucratic resources to prepare – although past research also emphasizes that advisers may provide more candid information in informal settings. In contrast, leaders might be *more* influential in informal settings – such as a small, ad hoc meetings – over which they exert more control.

Importantly, supplementary analysis in Section 4.1 shows that our leader-level results are stronger in informal meetings than formal ones. This means that leader effects might be partially masked in the analysis presented in the main text, which pools formal and informal meetings. If true, this suggests another important scope condition to the study of leaders in political science. Finally, we note that adviser hawkishness remains consistent in both formal and informal settings, which suggests that our principal finding regarding the central important of adviser predispositions

holds regardless of the institutional setting.

6 Probing the Deliberation Mechanism

6.1 Information Search and Dissent

Drawing upon coding methodologies from studying deliberation elsewhere in political science (e.g., Parthasarathy, Rao and Palaniswamy, 2019), we developed a coding scheme to identify speech acts that exhibited *information search*, defined as a textual indication that the speaker requested more information from another meeting participant. These might be instances in which the speaker sought clarification regarding an idea another speaker already expressed or asked another participant to introduce new ideas, information, or recommendations into the discussion. Examples of textual indicators of information search category included: “inquired,” “asked,” “request for advice,” “raised the question,” and “called on.” For instance, while deliberating armed conflict in the Congo during an NSC Executive Committee meeting on December 17, 1962, National Security Advisor McGeorge Bundy queried Chairman of the Joint Chiefs of Staff Maxwell Taylor about bargaining leverage afforded by different military strategies – while President Kennedy called on UN Ambassador Adlai Stevenson for more information.

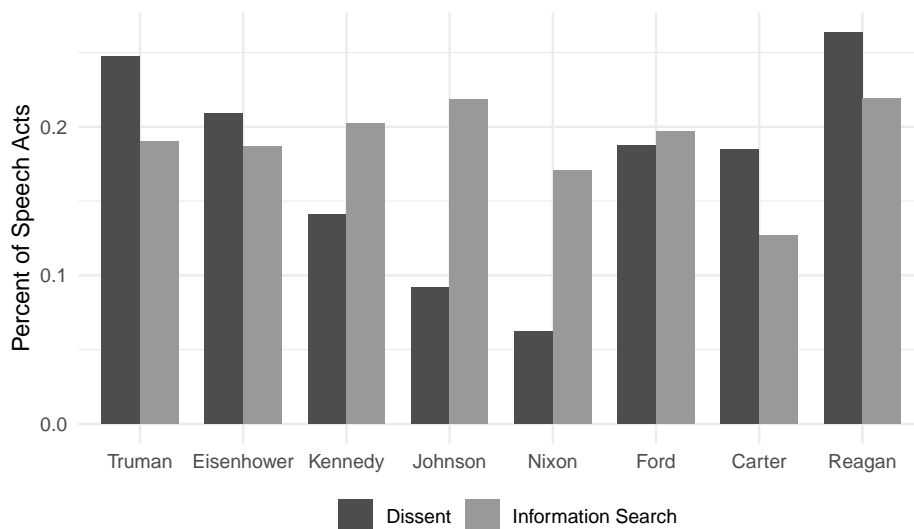
The second characteristic coded by the research assistant was *dissent*, defined as a textual indication that the speaker is disagreeing with an idea another meeting participants has expressed. In some cases, advisers directly identified their dissent by using terms such as “disagreed” or “objected.” For example, in a January 1958 meeting, National Security Advisor Gordon Gray stated that he “did not agree” with data provided by the Treasury Department regarding homeland defense. In many other cases, however, dissent could only be discerned in the context of the meeting – but was nevertheless evident because a meeting participant offered an argument that contradicted those provided by others.⁸

Figure A1 plots deliberation patterns by administration. While information search and dissent are generally consistent across administrations, the Johnson and Nixon administration exhibit comparatively lower levels of dissent. This descriptive finding is consistent with past work on presidential decision-making, which document how Johnson (Logevall, 1999, Chapter 12) and Nixon

⁸This context-specificity is why we use human coding for this analysis rather than an automated content approach.

(McDermott, 2007, Chapter 6) restricted policy debates between advisers.

Figure A1: Deliberation Patterns across Administrations



6.2 Keyword Assisted Topic Model

The keyword assisted topic model (keyATM) is a topic model in which the researcher can attempt to identify the prevalence of specific topics by supplying keywords that are informed by qualitative knowledge of the issue at hand. See [Eshima, Imai and Sasaki \(2020\)](#) for more details. In our case, we sought to identify times when speech acts raised the topics of violence, threat, military balance, diplomacy, and the adversary’s interests. [Table A1](#) enumerates all keywords provided to search for these five topics.

Table A1: Supplied Keywords for KeyATM Topics

<u>Topic</u>	<u>Keywords</u>
Violence	attack, bomb, fight, invad, strike
Threat	hostil, threat, risk, aggress, fall, loss, belliger
Military Balance	missil, bomber, fighter, carrier, submarin
Diplomacy	diplomaci, diplomat, peac, talk, treati, negoti
Adversary Interests	communist, china, vietnam, soviet, ussr

Our keyATM model sought to identify these five topics but also allowed for 20 additional topics that the model would independently and inductively identify. [Table A2](#) lists the most frequent terms

that were associated with the 25 topics in the model. Words with “[X]” beside them represent terms that were included in that category in Table 6.2, while words with “[5]” represent terms that were associated with the Adversary Interests topic but also showed up as common terms for other topics.

Finally, Table A3 displays results for regressions that use speech acts as the unit of analysis. We regress the prevalence of the five key topics on the hawkishness of the individual responsible for the associated speech act.

Table A2: Most Frequent Terms for All keyATM Topics

<u>Topic</u>	<u>Most Frequent Terms</u>
Threat	unit_state, presid, secretari_dull, point, forc, war, militari, said, propos, problem
Military Balance	missil [X], soviet [5], will, can, limit, program, mirv, option, number, one
Diplomacy	will, go, think, presid, get, want, can, us, say, said
Violence	general, attack [X], forc, said, will, now, use, one, militari, presid
Adversary Interests	soviet [X], govern, situat, presid, communist [X], unit_state, mr_dull, said, countri, point
Other_1	meet, mr, presid, govern, discuss, said, propos, berlin, governor, confer
Other_2	jordan, isra, israel, arab, syria, iraq, oil, king, saudi, egypt
Other_3	program, budget, present, militari, mr, fy, forc, point, secretari, said
Other_4	congo, canal, agent, un, use, chemic, biolog, lumumba, broadcast, station
Other_5	will, need, increas, year, can, secretari, econom, food, product, countri
Other_6	paragraph, propos, polici, council, point, presid, said, suggest, nsc, mr
Other_7	program, mr, propos, said, paragraph, presid, point, report, problem, nsc
Other_8	forc, nato, capabl, program, militari, defens, increas, plan, cost, nuclear
Other_9	dr, test, program, disarm, soviet [5], governor_stassen, satellit, propos, said, space
Other_10	program, countri, assist, fund, point, million, secretari_humphrey, mr, money, problem
Other_11	think, unclear, go, cuba, note_note, say, ship, well, get, now
Other_12	council, report, presid, nsc, mr, propos, recommend, paragraph, general, present
Other_13	oil, secretari, mr, said, trade, list, british, problem, control, cuba
Other_14	lao, govern, mr_dull, general, situat, cuba, french, castro, forc, phoumi
Other_15	mr, said, report, council, secretari, presid, propos, paragraph, present, nsc
Other_16	will, el, salvador, presid, admir, prime_minist, shultz, nicaragua, cuba, de
Other_17	nato, forc, problem, countri, pakistan, assist, africa, european, europ, african
Other_18	peopl, servic, said, get, problem, general, presid, inform, one, program
Other_19	mr, said, vessel, port, requir, ship, oper, program, mat, devic
Other_20	program, presid, shelter, said, governor, build, govern, stockpil, peopl, state

Notes: [X] indicates a supplied keyword for that topic; [#] indicates a supplied keyword for the #th topic.

Table A3: Hawkishness and Speech Act Content

	Violence	Threat	Mil. Balance	Diplomacy	Adv. Interests
Speaker Hawkishness	0.12*** (0.01)	-0.03* (0.01)	-0.09*** (0.01)	-0.17*** (0.02)	-0.25*** (0.02)
Truman	0.06*** (0.01)	0.09*** (0.01)	-0.09*** (0.01)	-0.50*** (0.01)	0.08*** (0.01)
Eisenhower	-0.01 (0.01)	0.16*** (0.01)	-0.07*** (0.01)	-0.53*** (0.01)	0.14*** (0.01)
Kennedy	0.02** (0.01)	0.001 (0.01)	-0.06*** (0.01)	-0.16*** (0.01)	-0.02** (0.01)
Johnson	0.27*** (0.01)	0.004 (0.01)	-0.09*** (0.01)	-0.24*** (0.01)	0.05*** (0.01)
Nixon	0.09*** (0.01)	0.003 (0.01)	0.10*** (0.01)	-0.08*** (0.01)	0.01 (0.01)
Ford	0.06*** (0.01)	-0.01 (0.01)	0.38*** (0.01)	-0.28*** (0.01)	-0.02** (0.01)
Reagan	-0.04*** (0.01)	0.01 (0.01)	0.01 (0.01)	-0.02** (0.01)	0.04*** (0.01)
Constant	-0.005 (0.01)	0.04*** (0.01)	0.15*** (0.01)	0.68*** (0.01)	0.19*** (0.01)
Observations	37,333	37,333	37,333	37,333	37,333

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

Notes: OLS regression with the speech act in formal NSC meetings as the unit of analysis. Hawkishness scores are from the boosted linear approach and outcome scores are topic proportions from the keyATM analysis.

6.3 Issues of Self-Censorship

One may be concerned that our results are at least in part a consequence of self-censorship. It is plausible that advisers who have dispositions at odds with the president would be less willing to voice their views during meetings. This, in turn, would lead meetings to hew toward policies that are preferred by the president and supported by like-minded advisers.

We can address this concern using our data on deliberation. For each speech act in this hand-coded data, we measure the absolute distance between the hawkishness of the speaker and the hawkishness of the president, conditional on the president’s attendance. This quantity gauges how far apart the adviser is from the president, regardless of direction. We then analyze the relationship between whether speech acts (made by advisers) express dissent and the absolute distance in hawkishness between the speaker and the president.

Figure A2 displays the distribution of these distances for speech acts that either do or not do express dissent across 250 randomly selected meetings. We see evidence that speech acts expressing dissent tend to come from advisers who are relatively further apart in hawkishness from the president. A t -test indicates that this difference is highly statistically significant ($p \ll 0.01$). This result contradicts what we would expect to see if advisers engaged in self-censorship.

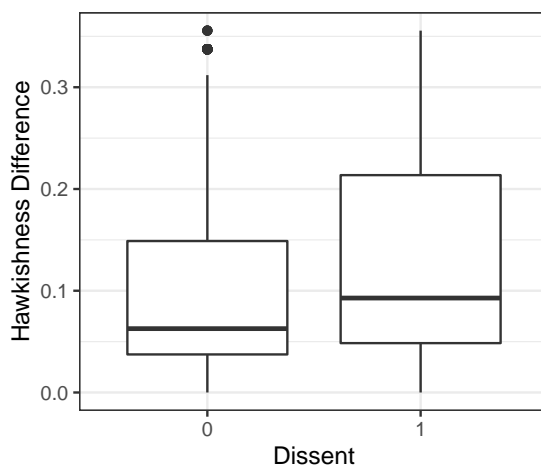


Figure A2: Relationship between speech acts expressing dissent and the speaker’s distance from the president in terms of hawkishness.

7 Agenda Items and Hawkishness

Table A4 displays the distribution of agenda items that were mentioned in at least one speech act in each meeting. Table A5 presents the results of OLS regressions that analyze the relationship between the topics discussed in each meeting and the average hawkishness of participants in the meeting. Model 1 combines all meetings, while Models 2 and 3 split the data into formal and informal meetings respectively. All models include year fixed effects. Several topics appear to have a meaningful association with higher average hawkishness of individuals in a meeting. We therefore control for agenda items in our main analyses.

Table A4: Distribution of agenda items across all meetings.

Agenda Item	No	Yes
Economy	2,639	242
Europe	2,484	397
Asia	2,394	487
Defense	2,615	266
China	2,780	101
Intelligence	2,805	76
Middle East	2,481	400
Strategic Forces	2,678	203
Organization	2,810	71
USSR	2,652	229
Policy	2,728	153
Americas	2,609	272
International Institutions	2,843	38
Diplomacy	2,769	112
Vietnam	2,467	414
Africa	2,777	104
Arms Control	2,736	145
North Africa	2,880	1
Latin America	2,865	16
Other	2,851	30

Table A5: OLS regressions on the relationship between meeting topics and average hawkishness of meeting participants.

	<i>Dependent variable:</i>		
	Mean Hawkishness		
	All	Formal	Informal
	(1)	(2)	(3)
Economy	0.007*** (0.003)	-0.005** (0.002)	-0.004 (0.008)
Europe	0.006*** (0.002)	0.002 (0.002)	-0.003 (0.007)
Asia	0.013*** (0.002)	0.007*** (0.002)	0.003 (0.007)
Defense	0.024*** (0.003)	0.004* (0.002)	0.019** (0.008)
China	0.008* (0.004)	-0.008** (0.004)	-0.001 (0.009)
Other	0.005 (0.007)	-0.002 (0.005)	0.016 (0.021)
Intelligence	0.004 (0.005)	-0.003 (0.004)	0.002 (0.009)
Middle East	0.006*** (0.002)	-0.003 (0.002)	-0.0003 (0.007)
Strategic Force	0.017*** (0.003)	-0.0001 (0.002)	0.008 (0.008)
Organization	0.015*** (0.005)	-0.015*** (0.004)	0.020** (0.009)
USSR	0.012*** (0.003)	-0.001 (0.003)	0.006 (0.007)
Policy	0.047*** (0.004)	0.002 (0.003)	
Americas	0.008*** (0.003)	0.001 (0.003)	0.001 (0.008)
International Institutions	0.012* (0.006)	-0.002 (0.006)	0.007 (0.011)
Diplomacy	0.003 (0.004)	-0.004 (0.004)	-0.004 (0.008)
Vietnam	0.018*** (0.003)	0.003 (0.004)	0.010 (0.007)
Africa	0.009** (0.004)	-0.0002 (0.003)	-0.006 (0.008)
Arms Control	0.011*** (0.004)	-0.002 (0.004)	0.007 (0.008)
North Africa	-0.028 (0.040)	-0.019 (0.023)	
Latin America	0.002 (0.010)	0.0001 (0.016)	-0.008 (0.013)
Constant	0.500*** (0.023)	0.524*** (0.013)	0.560*** (0.024)
Year FEs	✓	✓	✓
Observations	2,877	793	2,084

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

References

- Bertoli, Andrew, Allan Dafoe and Robert F. Trager. 2019. "Is There a War Party? Party Change, the Left–Right Divide, and International Conflict." *Journal of Conflict Resolution* 63(4):950–975.
- Braumoeller, Bear F. 2013. *The great powers and the international system: systemic theory in empirical perspective*. Cambridge University Press.
- Carter, Jeff and Charles E. Smith. 2020. "A Framework for Measuring Leaders' Willingness to Use Force." *American Political Science Review* 114(4):1352–1358.
- Clare, Joe. 2014. "Hawks, Doves, and International Cooperation." *Journal of Conflict Resolution* 58(7):1311–1337.
- Clark, David H., Benjamin O. Fordham and Timothy Nordstrom. 2011. "Preying on the Misfortune of Others: When Do States Exploit Their Opponents' Domestic Troubles?" *Journal of Politics* 73(1):248–264.
- Colgan, Jeff D. and Jessica L.P. Weeks. 2015. "Revolution, Personalist Dictatorships, and International Conflict." *International Organization* 69(1):163–194.
- Croco, Sarah E. 2011. "The Decider's Dilemma: Leader Culpability, War Outcomes, and Domestic Punishment." *American Political Science Review* 105(3):457–477.
- Dafoe, Allan and Devin Caughey. 2016. "Honor and War: Southern US Presidents and the Effects of Concern for Reputation." *World Politics* 68(2):341–381.
- Eshima, Shusei, Kosuke Imai and Tomoya Sasaki. 2020. "Keyword Assisted Topic Models." Working paper, <http://arxiv.org/abs/2004.05964>.
- Foster, Dennis M. 2008. "'Comfort to Our Adversaries'? Partisan Ideology, Domestic Vulnerability, and Strategic Targeting." *Foreign Policy Analysis* 4(4):419–436.
- Fuhrmann, Matthew. 2020. "When Do Leaders Free-Ride? Business Experience and Contributions to Collective Defense." *American Journal of Political Science* 64(2):416–431.
- Fuhrmann, Matthew and Michael C. Horowitz. 2015. "When Leaders Matter: Rebel Experience and Nuclear Proliferation." *Journal of Politics* 77(1):72–87.
- Gallagher, Maryann E. and Susan H. Allen. 2014. "Presidential Personality: Not Just a Nuisance." *Foreign Policy Analysis* 10(1):1–21.
- Goemans, Henk E., Kristian S. Gleditsch and Giacomo Chiozza. 2009. "Introducing Archigos: A Dataset of Political Leaders." *Journal of Peace Research* 46(2):269–283.
- Goldstein, Joshua S. and John R. Freeman. 1990. *Three-Way Street: Strategic Reciprocity in World Politics*. Chicago, IL: University of Chicago Press.
- Harden, John P. Forthcoming. "All the World's a Stage: US Presidential Narcissism and International Conflict." *International Studies Quarterly*.
- Heffington, Colton. 2018. "Do Hawks and Doves deliver? The Words and Deeds of Foreign Policy in Democracies." *Foreign Policy Analysis* 14(1):64–85.

- Horowitz, Michael C. and Allan C. Stam. 2014. "How Prior Military Experience Influences the Future Militarized Behavior of Leaders." *International Organization* 68(3):527–559.
- Horowitz, Michael C., Allan C. Stam and Cali M. Ellis. 2015. *Why Leaders Fight*. New York: Cambridge University Press.
- Jervis, Robert. 2013. "Do Leaders Matter and How Would We Know?" *Security Studies* 22(2):153–179.
- Johnston, Alastair Iain. 2012. "What (If Anything) Does East Asia Tell Us About International Relations Theory?" *Annual Review of Political Science* 15:53–78.
- Jost, Tyler. 2021. "Bureaucracy and War: The Institutional Origins of Miscalculation in International Politics." Book manuscript.
- Kertzer, Joshua D., Deborah Jordan Brooks and Stephen G. Brooks. 2021. "Do Partisan Types Stop at the Water's Edge?" *Journal of Politics* 83(4):1764–1782.
- Koch, Michael T. and Skyler Cranmer. 2007. "Testing the "Dick Cheney" Hypothesis: Do Governments of the Left Attract More Terrorism than Governments of the Right?" *Conflict Management and Peace Science* 24(4):311–326.
- Kreps, Sarah E., Elizabeth N. Saunders and Kenneth A. Schultz. 2018. "The Ratification Premium: Hawks, Doves, and Arms Control." *World Politics* 70(4):479–514.
- Larson, Deborah Welch. 1985. *Origins of Containment: A Psychological Explanation*. Princeton, NJ: Princeton University Press.
- Logevall, Fredrik. 1999. *Choosing War*. Berkeley: University of California Press.
- Marsh, Kevin. 2014. "Obama's Surge: A Bureaucratic Politics Analysis of the Decision to Order a Troop Surge in the Afghanistan War." *Foreign Policy Analysis* 10(3):265–288.
- Mattes, Michaela and Jessica L. P. Weeks. 2019. "Hawks, Doves and Peace: An Experimental Approach." *American Journal of Political Science* 63(1):53–66.
- McAllister, William B., Joshua Botts, Peter Cozzens and Aaron W. Marrs. 2015. *Toward "Thorough, Accurate, and Reliable": A History of the Foreign Relations of the United States Series*. Washington, DC: Office of the Historian, U.S. Department of State.
- McDermott, Rose. 2007. *Presidential Leadership, Illness, and Decision Making*. New York: Cambridge University Press.
- McManus, Roseanne. 2019. "Revisiting the Madman Theory: Evaluating the Impact of Different Forms of Perceived Madness in Coercive Bargaining." *Security Studies* 28(5):976–1009.
- 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.
- Palmer, Glenn, Tamar R. London and Patrick M. Regan. 2004. "What's Stopping You?: The Sources of Political Constraints on International Conflict Behavior in Parliamentary Democracies." *International Interactions* 30(1):1–24.

- Parthasarathy, Ramya, Vijayendra Rao and Nethra Palaniswamy. 2019. "Deliberative Democracy in an Unequal World: A Text-As-Data Study of South India's Village Assemblies." *American Political Science Review* 113(3):623–640.
- Potter, Philip B.K. 2007. "Does Experience Matter? American Presidential Experience, Age, and International Conflict." *Journal of Conflict Resolution* 51(3):351–378.
- Saltoun-Ebin, Jason. 2014. *The Reagan Files: Inside the National Security Council*. Lexington, KY: Seabec Books.
- Saunders, Elizabeth N. 2011. *Leaders at War: How Presidents Shape Military Interventions*. Ithaca: Cornell University Press.
- Saunders, Elizabeth N. 2022. "Elites in the Making and Breaking of Foreign Policy." *Annual Review of Political Science* Forthcoming.
- Schultz, Kenneth A. 2005. "The Politics of Risking Peace: Do Hawks or Doves Deliver the Olive Branch?" *International Organization* 59(1):1–38.
- Signorino, Curtis S. 1999. "Strategic Interaction and the Statistical Analysis of International Conflict." *American Political Science Review* 93(2):279–297.
- Smith, Alastair. 1996. "Diversionary Foreign Policy in Democratic Systems." *International Studies Quarterly* 40(1):133–153.
- Whitlark, Rachel Elizabeth. 2017. "Nuclear Beliefs: A Leader-Focused Theory of Counter-Proliferation." *Security Studies* 26(4):545–574.
- Williams, Laron K. 2014. "Hawks, Doves, and Opportunistic Opposition Parties." *Journal of Peace Research* 51(1):111–125.
- Yarhi-Milo, Keren. 2018. *Who Fights for Reputation in International Politics? Leaders, Resolve, and the Use of Force*. Princeton, NJ: Princeton University Press.