Advisers and Aggregation in Foreign Policy Decision-Making*

Tyler Jost† Joshua D. Kertzer‡ Eric Min§ Robert Schub¶
Last modified: July 12, 2023

Word count: 13929

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 the counsel advisers provide in deliberations. We test our argument by introducing three original datasets, including an original dataset of 2,685 foreign policy deliberations between US presidents and their advisers from 1947 to 1988. Applying a novel machine learning approach to estimate the hawkishness of 1,134 Cold War-era foreign policy decision-makers, we show that adviser-level hawkishness affects both the counsel advisers provide in deliberations, and the decisions the leader makes: 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 how advisers' dispositions, which aggregate via deliberation, systematically shape the choices leaders make.

*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 (WCFIA) at Harvard University for their logistical support, and to Diana Agdashian, Saadi Ali, Perry Arrasmith, Christine Bienes, Florian Bochert, Lexa Brenner, Alexis Carraway, Youmna Chamieh, Erica Colston, Joshua Conde, Vanessa Cordova, Jacinta Crestanello, Caroline Curran, George Dalianis, 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, Hope Ha, Joolean Hall, Joseph Hasbrouck, Sarah Jacob, Joshua Kelley, Joel Kesselbrenner, Samantha Ku, Jacob Kuckelman, Max Kuhelj Bugaric, Zaki Lakhani, 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, Nic Pantelick, June Park, Marc Prophet, Kasey Quicksill, Janna Ramadan, Bianca Rodriguez, Kay Rollins, 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

‡John Zwaanstra Professor of International Studies and of Government, Department of Government, Harvard University. 1737 Cambridge St, Cambridge MA 02138. Email: jkertzer@gov.harvard.edu. Web: https://jkertzer.sites.fas.harvard.edu

§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, Rutgers University. Email: robert.schub@rutgers.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; 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 (Kertzer et al., 2022). 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 characteristics in shaping foreign policy. The new wave of first-image IR scholarship predominantly focuses on leaders rather than those that counsel them. This is partly for substantive reasons. Foreign policy decision-making is hierarchical, which might lead us to suspect that leader dispositions dominate those of advisers, that leaders tend to appoint advisers with similar worldviews, and that leaders 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.

Moreover, the scholarship that does exist tends to study advisers situationally rather than dispositionally, focusing on the quality of the group's advisory *process* and how leaders and advisers interact, rather than on advisers' predispositions themselves. The bureaucratic politics literature, for example, argues that the recommendations that advisers supply flow from their institutional affiliation; where advisers stand depends on where they sit, rather than stable and enduring predispositions (Allison and Zelikow, 1999). 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 and Preston, 1994; LeVeck and Narang, 2017), and institutions (Reiter and Stam, 2002; Weeks, 2014; Jost, 2023) yield more accurate assessments, less biased information provision, and more effective policy outcomes. Where discussions of dispositions arise in this literature, it tends to be how *leaders'* dispositions affect the advisory structures that leaders establish (e.g. George, 1980), rather than how advisers' dispositions affect how advisers behave.[1] As Preston (2001, 267) directly suggests, we have much still to learn about whether and how adviser dispositions shape foreign policy behavior, if at all.

In this article, we develop a theory of foreign policy decision-making that emphasizes adviser dispositions. The uncertainty, complexity, and ill-defined nature of foreign policy decision-making means that leaders turn to advisers for counsel. The kind of counsel advisers offer—the information they share, the analysis they provide, and the policies they recommend—is shaped by advisers' core dispositions about the desirability and efficacy of the use of force. As a result, despite the decision-making authority that leaders retain, foreign policy choices bear the fingerprints of advisers as much as those of leaders. Against the claims of the bureaucratic politics literature, those fingerprints are not reducible to the adviser's institutional role: *who* occupies the most important positions of government affects where those advisers "stand" and the corresponding counsel they supply. In short, our theoretical contribution is to offer a theory of foreign policy decision-making that shifts the focus to *advisers* rather than advisory systems (e.g., George, 1980), to *bureaucrats* rather than bureaucracy (e.g., Allison and Zelikow, 1999).

[1] When studies do take the important next step of incorporating a personal characteristic of advisers, the relevance of the characteristic still depends on how it relates to the leader's own characteristics. Adviser experience, for instance, is argued to affect adviser behavior conditional on the *leader's* experience (Saunders, 2017).

To test our theory, we leverage big data and machine learning techniques to offer systematic and unusually high-resolution empirical tests on how advisers shape foreign policy. First, we located, collected, digitized, and processed the transcripts of 2,685 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 1,894 informal meetings in which presidents discussed foreign policy issues with their advisers. We segmented each meeting transcript by speaker, meaning that our data identify not only which advisers provided counsel, but what substantive topics they emphasized. We also reviewed each transcript to manually code over 2,500 decisions that leaders made during these meetings, which ranged from diplomatic cooperation to interstate conflict. These include some of the most consequential foreign policy choices in modern American history, such as the decision to blockade Soviet ships during the Cuban Missile Crisis and the decision to engage Mikhail Gorbachev during the Reykjavik Summit. We believe this to be the most comprehensive resource available by which scholars can study the microfoundations of foreign policy decision-making.

Second, we apply a novel machine learning-based approach that estimates (at a distance) leader and adviser dispositions, such as hawkishness, based off an original dataset of the biographical characteristics of every individual who participated in the meetings in our sample. Our biographical data describe the backgrounds and professional experiences of 1,134 individuals ranging from secretaries of state 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 to study quantitatively what has traditionally been the preserve of qualitative approaches. Collectively, this approach enables us to study not only whether dispositions shape the types of counsel advisers provide, but also whether the aggregated dispositions of advisory groups shape the choices that leaders ultimately make.

Analyzing these data yields two major findings demonstrating the central importance of advisers in foreign policy decision-making. First, we find that adviser dispositions shape the counsel that leaders receive when making consequential foreign policy choices. American presidents consistently solicit information from advisers during deliberations—and advisers use these opportunities to offer

information, perspectives, and policy recommendations congruent with their dispositions. Second, we show that hawkish advisory groups are associated with more conflictual foreign policies, even after considering several potential pathways for selection effects. Contrary to accounts assuming that decision-making outcomes simply reflect leader dispositions, we find that adviser-level hawkishness has large and systematic effects on foreign policy decision-making—and that appointment to and participation in foreign policy groups does not merely mirror the hawkishness of the leader. The theory and findings collectively illustrate the formidable influence that foreign policy advisers can wield by supplying the counsel that leaders demand. More broadly, and contrary to accounts suggesting that individual-level traits have little bearing on group decision-making, our results demonstrate that knowing the dispositions of advisers helps to explain the choices of the most powerful state during the second half of the twentieth century.

1 Leaders, Advisers, and Aggregation in Groups

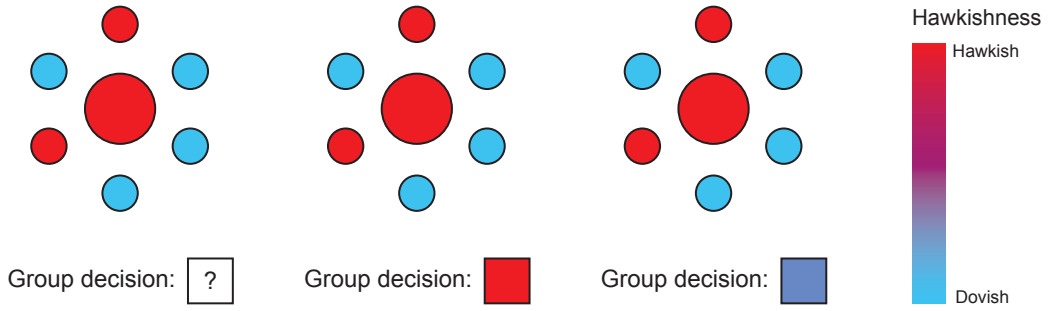
The division between hawks and doves is central to our understanding of why states choose conflict over cooperation (Modigliani, 1972; Herrmann, Tetlock and Visser, 1999; Schultz, 2005; Kahneman and Renshon, 2007). Hawks and doves differ in their beliefs about the nature of international politics, adversary motivations, and the efficacy and appropriateness of using force, shaping the foreign policies that individuals support (Rathbun, 2007; Yarhi-Milo, Kertzer and Renshon, 2018; Lin-Greenberg, 2019). 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 dispositions like hawkishness aggregate to shape state behavior (Hafner-Burton et al., 2017; Kertzer et al., 2022). Most foreign policy decisions occur in group settings in which leaders and 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 domineering decision-making styles—routinely relied on advisers, such as

Henry Kissinger and Zhou Enlai, respectively.

Like the leaders they serve, advisers presumably possess stable dispositions like hawkishness. The central question, however, is whether and 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 dispositional model of advising.

Figure 1: Three Models of Aggregation in Groups



(a) Emergence Model **(b)** Leader Model **(c)** Adviser Model

Note: Large circles represent the leader's level of hawkishness, small circles represent the adviser hawkishness, and the square's color represents the hawkishness of the decision. Emergence models (panel a) argue that group decisions cannot be reduced to the traits of either leaders or advisers. Leader models (panel b) suggest that group decisions reflect the leader's dispositions (large circle). Our adviser model suggests that decisions reflect the dispositions of the advisory group.

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 compared to structural variables, such as polarity, the balance of power, and alliances (e.g., Mearsheimer,

2001). Second, some game theoretic scholarship remains skeptical about the challenges of aggregation, suggesting that the aggregation of traits in groups is sufficiently complex so as to warrant a "methodological bet" that they are not worth studying (Powell, 2017). 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 (Jervis, 1998), so too is domestic policymaking. 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, 165, Carter and Smith, 2020). For example, Yarhi-Milo (2018, 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. Other studies similarly note the division between hawkish and dovish leaders in other countries, such as China and India (e.g., Kennedy, 2011).

The leader model posits that group decisions reflect leader traits for two reasons. First, foreign policy decision-making groups tend to be hierarchical: leaders enjoy more substantive and procedural authority than other group members. Leaders do not just make the final decisions in foreign policy; they also set the rules for decision-making. In this view, leaders might be able to impose their worldview on policy by strategically manipulating adviser appointments or participation in decision-making. As Krasner (1972, 166) argues, "The President chooses most of the important players and sets the rules... These individuals must share his values."[2] Saunders (2011, 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, 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 decision-making process to afford privileged access to advisers

[2]See also Horowitz, Stam and Ellis (2015, 6–7).

with congruent dispositions. Leaders might deliberately manipulate which advisers participate in meetings, steer discussions in directions that suit their preferences, or include a "domesticated dissenter" within meetings to fabricate the appearance of debate when, in fact, their decision has already been made (Saunders, 2015). For example, Lyndon Johnson excluded Vice President Hubert Humphrey from policy deliberations on Vietnam in 1965 after Humphrey expressed opposition to escalation (Burke and Greenstein, 1989). If leaders strictly surround themselves with like-minded advisers, or manipulate the advisory process to ensure the information they receive predominately reflects their worldview, the real causal power comes from the traits of the leader, not those of the advisory group.

Second, the leader model argues that leader beliefs tend to supersede the input 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, 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, 16). Cognitive barriers, such as the desire for consistency (Jervis, 1976) and motivated reasoning (Kertzer, Rathbun and Rathbun, 2020), may also cause hawkish leaders to prioritize input from hawkish advisers and discount it from dovish confidants—and vice versa. As such, advisers are "influential" only when the input they provide is congruent with what the leader already believes.[3] But if the only adviser input 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, 149; Yarhi-Milo, 2018, 42). This asymmetry in focus is also likely a function of methodological considerations: in quantitative IR, we have excellent datasets of leader-level background

[3]For a similar argument regarding a leader's beliefs about fighting for reputation, see Yarhi-Milo (2018, 36–38).

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. Our adviser model is based on three simple intuitions. First, the challenges of foreign policy decision-making causes leaders to turn to advisers for counsel. Second, advisers have stable predispositions about foreign policy. Third, these predispositions affect the nature of the counsel that advisers provide during deliberations—and thus the decisions that leaders are likely to make. We discuss each in turn.

Our model begins with an assumption as well known to scholars of foreign policy analysis as it is to leaders themselves: foreign policy decision-making is hard (Snyder, Bruck and Sapin, 1962; Jervis, 1976; Renshon and Renshon, 2008). International politics is characterized by ill-defined problems in which the nature of the situation, 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, all while facing time constraints, information constraints, irreducible uncertainty, and complex value tradeoffs (George, 1980).

It is here that advisers are useful. Advisers in foreign policy do many things: they offer emotional support and companionship to leaders coping with the stress of decisionmaking (George, 1980, 80, Goldhamer, 1978, 8), they bestow the leader with public legitimacy and political cover (George, 1980, 81; Kenwick and Maxey, 2022), and they coerce leaders using the threat of public protest (Saunders, 2018). Yet advisers do not just provide comfort, cover, and coercion: they also provide counsel, which we can understand as consisting of three distinct functions.[4]

First, advisers engage in *information provision*, giving leaders information they need about the

[4]The three counseling functions we identify—problem representation, information provision, and policy recommendations—map nicely onto Destler's (1972, 10) claim that advisers provide "analysis, information, and advice."

state of the world (Schub, 2022). Historically, advisers served as the king's "eyes and ears and hands and feet" (Aristotle, 1920, Book III, 1287b). Today, they monitor intelligence, diplomatic cables, and news reports. Sometimes, even questions as mundane as "What happened?" are not straightforward, even for highly experienced leaders (Katagiri and Min, 2019). 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*. The sheer volume of information contemporary leaders command—by the mid-1960s, US ambassadors were sending 400,000 words a day by telegraph (Goldhamer, 1978, 58)—means that advisers do not merely provide information to the leaders they serve, but also screen it (Burke and Greenstein, 1989, 6), choosing what pieces of information to relay and what to hold back.

Second, advisers engage in *problem representation*, helping leaders develop a "definition of the situation" they face (Sylvan and Voss, 1998; Snyder, Bruck and Sapin, 1962). This function, which George (1980, 240) called the "diagnostic" function of advising, and Maoz (1990) refers to as "framing," is less about gathering information than it is about interpreting it. Is the conflict in Korea in 1950 a civil war or an act of communist aggression? Is Ho Chi Minh a local nationalist or a Soviet puppet? Should the uptick in violence in Iraq in 2007 be understood as terrorism or insurgency? This is why analogical reasoning is so powerful in foreign policy, since it offers decision-makers schemas they can use to define what values or interests are at stake in a given crisis (Khong, 1992). Much of what advisers do in foreign policy deliberations consists of offering leaders these schemas or perspectives, as reflected in documents like memos written by McGeorge Bundy and George Ball in 1964 and 1965 with titles like "Vietnam: what is our interest there and our object?" and "How valid are the assumptions underlying our Vietnam policies?" In this sense, advisers do not just provide leaders with information, they also provide them with theories.

Third, advisers provide *policy recommendations*, helping leaders select the optimal strategy given the situation they face (Jost, 2023). Burke and Greenstein (1989) call this "reality testing," and George (1980, 243–244) calls it "option assessment," as decision-makers assess the expected consequences of different courses of action. In the Cuban Missile Crisis, for example, advisers deliberated about whether the United States should respond diplomatically, with air strikes, or with a blockade. Just as strategic scripts follow from the images that precede them (Herrmann and Fischerkeller, 1995), the policy recommendations that advisers advocate for are intimately

connected to the problem representations to which they subscribe. Jervis (1976), for instance, noted that most of the debates during the Cold War depended on whether observers viewed US-Soviet relations through the prism of the spiral model or the deterrence model; which model you embraced determined what policies you favored.

Our next assumption posits that advisers have stable and well-defined predispositions that shape the way they view foreign policy. Some, like Curtis LeMay and Donald Rumsfeld, are relatively hawkish, while others, like Cyrus Vance and George Ball, are relatively dovish. The notion that leaders systematically differ from one another in their predispositions—whether operationalized as ideological belief systems, personalities, worldviews, leadership styles, or operational codes—is well established in the foreign policy literature (e.g. George, 1969; Preston, 2001; Walker, Schafer and Young, 2003; Whitlark, 2017; Yarhi-Milo, 2018; Harden, 2021). Much like the leaders they serve (e.g. Horowitz, Stam and Ellis, 2015), adviser traits are forged by early experiences which continue to shape how they behave when in office decades later. Experience planning Allied bombing campaigns against Japan during World War II shaped Robert McNamara's assessments of the feasibility of the bombing campaign against Hanoi during the Vietnam War. Experiences touring American aircraft carriers in the early 1980s colored the views of Chinese admiral Liu Huaqing during Politburo debates in the 1990s concerning Chinese naval modernization.

Finally, we argue that adviser predispositions shape the counsel advisers provide during deliberations, and thus, the decisions that leaders make. Research in political psychology leads us to expect that an adviser's dispositions affects all three of the counseling functions—information provision, problem definition, and policy recommendations. Scholarship on motivated reasoning and confirmation bias argues that our predispositions affect not only the information decision-makers seek out, but also how persuasive they find that information to be (Taber and Lodge, 2006; Kertzer, Rathbun and Rathbun, 2020). Advisers in the George W. Bush administration, for example, were convinced that Iraq had weapons of mass destruction, so they tasked intelligence officers to look for any indication that Iraqi WMDs existed. Predispositions also affect how decision-makers define the situation they face, responding to ill-defined problems by anchoring on their core dispositions. Hawks and doves facing the same strategic setting tend to perceive the situation in fundamentally different ways (Brutger and Kertzer, 2018), suggesting one reason why hawks after the Cold War continued to perceive the same level of international threat even after the Berlin Wall fell (Murray,

2002). Finally, as the discussion of the spiral and deterrence model above shows, predispositions affect the policies we prefer. This assumption lies at the heart of hierarchical models of foreign policy preferences, which envision our more general orientations towards foreign policy (e.g., hawkishness) shaping our preferences towards the use of force in particular circumstances (Hurwitz and Peffley, 1987; Herrmann, Tetlock and Visser, 1999; Rathbun, 2007).

In sum, our adviser model predicts that foreign policy decisions will at least partially reflect the dispositions of the advisers who participate in the deliberations, as illustrated by Figure 1c. The model suggests a number of observable implications. First, leaders should seek advisers' counsel. Leaders should meet with advisers, and in these meetings, leaders should ask questions, solicit advice, and request clarifications, rather than meetings merely being pro forma opportunities for leaders to keep their subordinates informed. Second, the counsel that advisers supply in these meetings should systematically differ based on their dispositions: hawks and doves should emphasize different pieces of information, or interpret the information in different ways, engaging in arguments and counter-arguments as they compete with one another over the direction they want leaders to take. Third, the dispositions of advisers participating in deliberations should affect the foreign policy decisions reached. The more hawks dominate the discussion, the more conflictual decisions the leader should make.

It is worth noting what is distinctive about our approach. First, it explicitly studies advisers dispositionally, rather than situationally. By drawing our attention to how dispositions affects the information, problem representation and policy recommendations that advisers provide, our dispositional focus not only connects the study of advisers to the study of political psychology more generally, but also leads to substantively different predictions. Unlike the bureaucratic politics literature, whose situational focus assumes that where advisers stand is based on where they sit (Allison and Zelikow, 1999), we argue that advisers' recommendations flow from predispositions that are not reducible to their institutional role. Second, recent scholarship has tended to focus on *coercive* pathways to advisory influence, showing that advisers matter because of the costs they can impose on leaders outside of meetings, such as the threat of leaks or public criticism (e.g. Garrison, 1999; Saunders, 2018, 22–23). In contrast, our model rests on a *counseling* pathway to advisory influence: advisers shape decision-making because they fulfill a leader's psychological and informational needs during deliberations. Our model thus complements this recent wave of research

by identifying an additional pathway by which advisers matter despite ultimately being subservient to the leaders they serve.

Finally, it is important to emphasize that the three models of aggregation are not necessarily mutually exclusive. The leader model might explain certain decisions, while the adviser model explains others. Sometimes leaders may know both what they want to do and how they want to do it, such that no amount of information or deliberation will shape their decision — whereupon we would expect leaders to manipulate the decision-making process to obtain their desired outcome. The value of theorizing an adviser model stems from the fact that there are many circumstances in which leaders are uncertain about which direction to take. Ultimately, the performance of each model is an empirical question we seek to test below.

2 Data

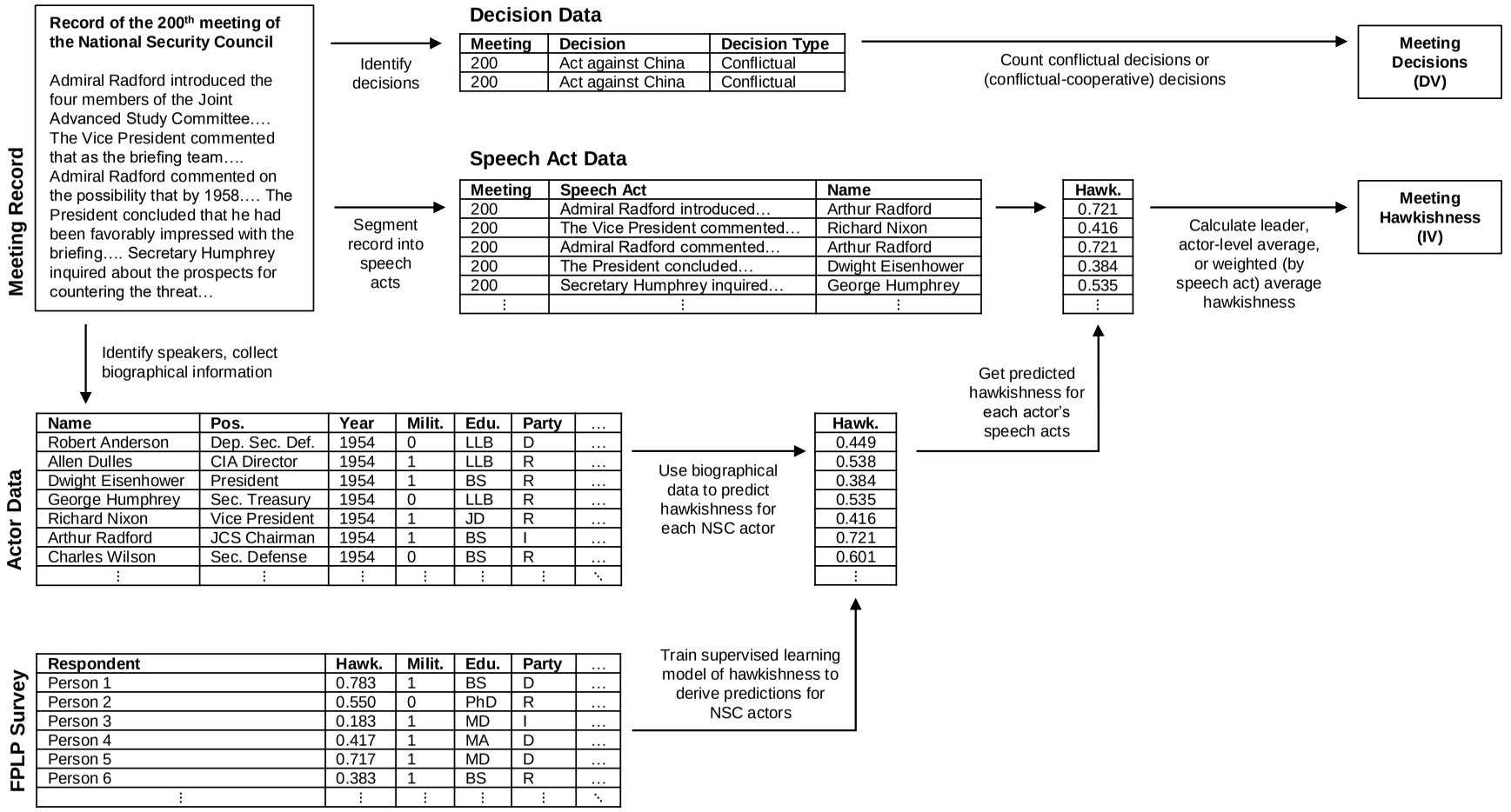
To test our adviser model, we systematically collected and analyzed a large set of archival records documenting meetings in the United States from 1947 to 1988 in which the president discussed foreign policy issues with advisers. We first used these archival records to identify the participants in foreign policy decision-making. We then measured participant hawkishness (our explanatory variable) from a distance using a novel methodological approach. All data were collected specifically for this study. Figure 2 visually summarizes our main datasets and the steps taken to convert them into the measures we ultimately use to test the three models of aggregation. We describe each step below and provide additional details in Appendices §1-4.

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

Figure 2: Diagram of All Primary Datasets, How Key Measures of Interest Were Derived, and the Relationships between Each



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 2,685 meeting records. 791 (29% of the sample) of the meetings were formal meetings of the NSC. An additional 1,894 (71% 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, 2023). 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.[8] The largest single set of records and decisions in our data come from the Eisenhower administration—reflecting the extent to which Eisenhower utilized a highly formalistic advising system in which the NSC played an outsized role. On the one hand, given his formalistic advising structure, Eisenhower's prominence in this data might bias our results in favor of finding evidence that adviser dispositions matter. On the other hand, existing theories might imply that adviser dispositions might be less likely to matter in the context of during this period, as Eisenhower's extensive foreign policy experience prior to his presidency may have improved his ability to monitor adviser information provision better than foreign policy novices like Carter and Reagan (Preston, 2001, 21; Saunders, 2017). 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.

Research assistants used optical character recognition software to convert the photographs of

[5] See Appendix §1 for details.
[6] Appendix §1.2 provides further information.
[7] Appendix §1.3 disaggregates the sample by formal and informal meetings.
[8] Our sample includes approximately 195 meeting records that were omitted from *FRUS*, as well as about 400 records for which *FRUS* provides only partial excerpts. We estimate that our sample includes 81% of all NSC meetings convened during this period.

Table 1: Meeting Records and Speech Acts by Administration

Administration	Records	Speech Acts
Truman	215	3,628
Eisenhower	889	27,880
Kennedy	250	14,996
Johnson	509	12,604
Nixon	453	31,960
Ford	238	8,280
Carter	48	2,142
Reagan	83	3,014
Total	2,685	104,504

meeting records into digital text, manually correcting text recognition errors. We then split each meeting record into what we call "speech acts"—the uninterrupted words spoken by a single individual during a meeting.[9] Our records featured 2,685 meetings containing 104,504 speech acts made by 1,134 unique participants.

2.2 Explanatory Variable: Measuring Hawkishness with Biographical Data

To test our dispositional model of advising, we need a measure of hawkishness for each of the 1,134 people identified in the meeting records. 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 (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 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

[9] See Appendix §2 for examples.

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 all 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.[10] Our codings focused on two aspects of the individual's background. First, we recorded their position and the dates on which the position was held. Second, we gathered 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.[11] Table 2 provides the coding for Henry Kissinger.

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).

[10] Information collection and coding procedures are detailed in Appendix §4.1.

[11] We are able to match a speaker name and position for 102,720 speech acts out of a total of 104,504 in our dataset – over 98%. Several hundred of these missing values, however, are explicitly listed as "Unidentified" in the record and are therefore impossible to match. Our effective match rate is thus over 99%.

Table 2: Example Coding for Henry Kissinger

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

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.[12] Second, 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*.[13] Third, we harness a series of supervised learning models to adduce the relationship between a respondent'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

[12] See Appendix §4.2 for details.

[13] As a robustness check in Appendix §5.5, we remove individuals' current affiliation from our statistical models; our results remain the same.

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. In order to tune the hyperparameters of the boosted model, a five-fold cross-validation process is used to select the model that produces the best out-of-sample predictions. This optimal model is then used 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.[14] We use a bootstrapping process through which we randomly re-sample the FPLP survey 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.[15]

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.[16] This adjustment produces hawkishness measures that align more closely with historical assessments of leader and adviser hawkishness.[17]

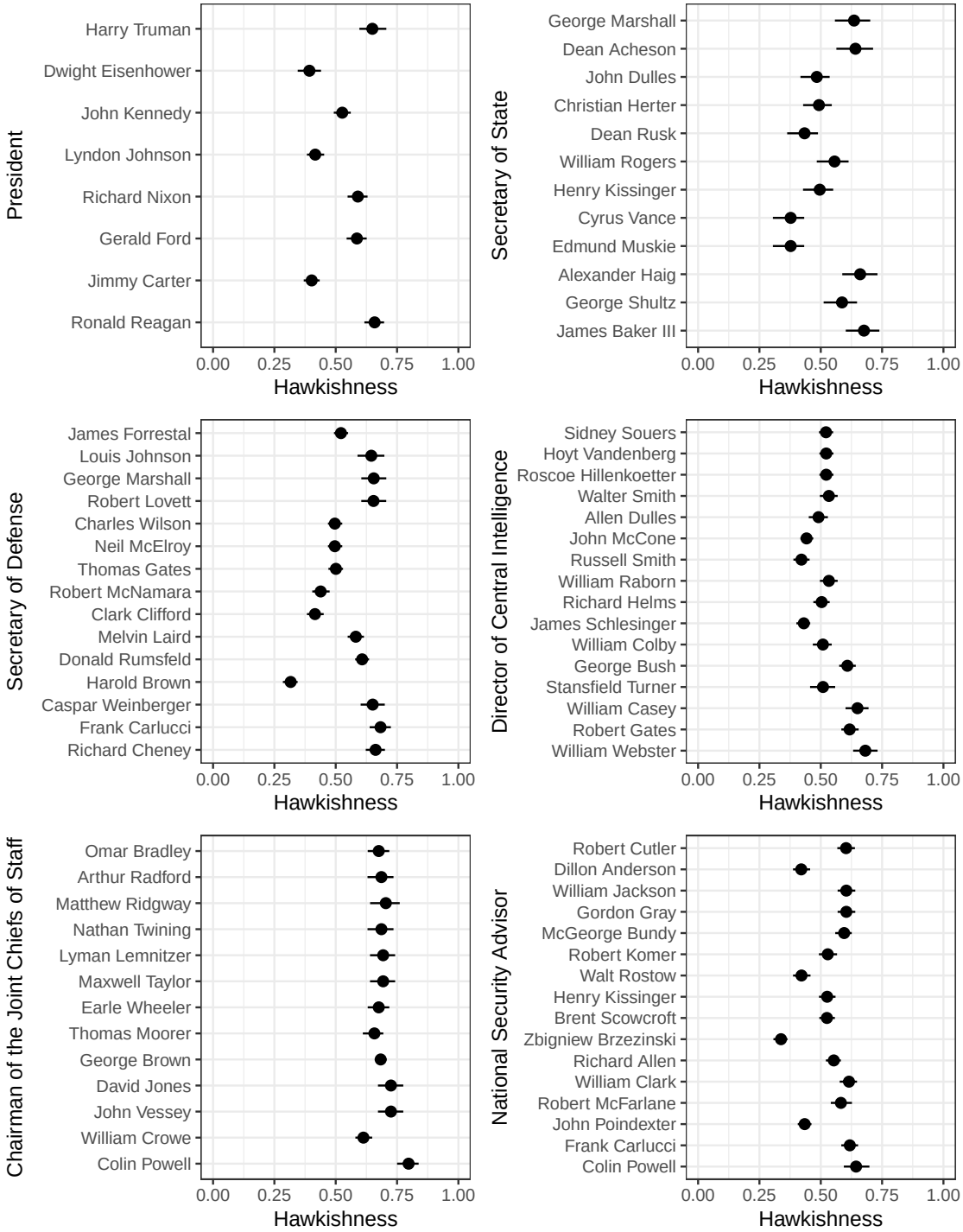
[14] Appendix §4.5 provides comparisons of these attributes between our actor data and the FPLP respondents.

[15] See Appendix §4.3 for more information on model selection and performance. Appendix §4.4 compares the performance of the boosted linear model with a simple linear regression. In Appendices §5.3, §5.8.1, and §8.4, we show that our main results hold using a more computationally intensive process where we propagate the uncertainty inherent to our predicted measures of hawkishness through our statistical analyses. Appendix §5.4 demonstrates that our findings are unchanged by using hawkishness measures produced with a simple linear model.

[16] For 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.

[17] While we have strong reasons to expect the association between partisanship and hawkishness changed markedly over the Cold War, we do not have similarly strong theoretical expectations about heterogeneous effects of the other types of biographical characteristics (age, gender, education, or military background) underpinning our predictive model. For a discussion and supporting evidence, see Appendix §4.6.

Figure 3: Predicted Hawkishness Measures for Senior US Decision-Makers



Note: Hawkishness estimates were calculated through a bootstrapping process that randomly sampled the FPLP survey data with replacement 1,000 times, generating 1,000 predicted hawkishness values for each leader/adviser. Black dots report the mean hawkishness value and bands are 95% confidence intervals across predicted values.

Figure 4: Average Speaker Hawkishness in US Foreign Policy Meetings

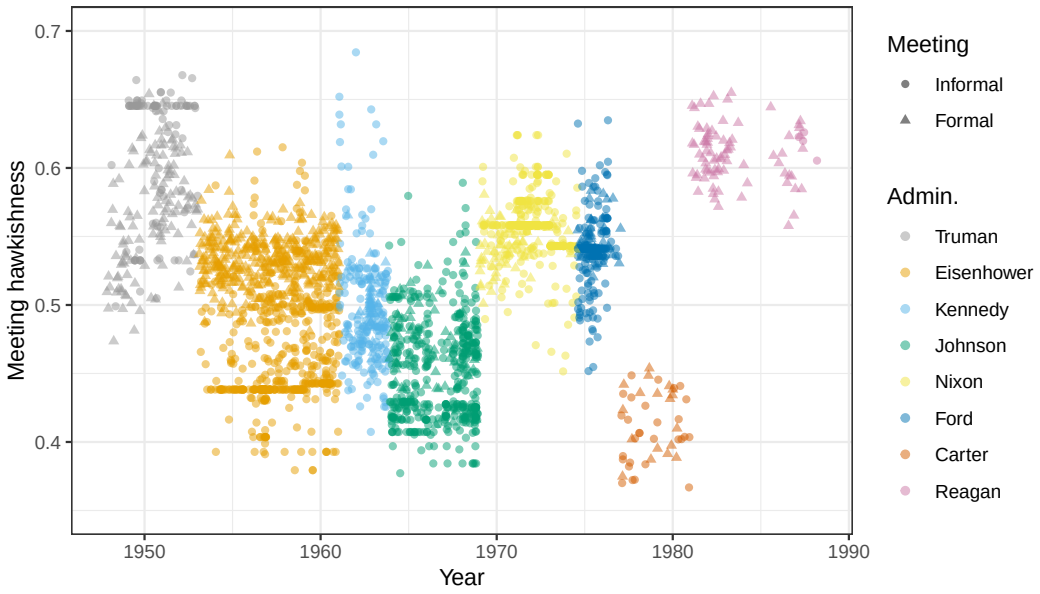


Figure 3 illustrates the predicted hawkishness measures for six senior positions in US foreign policymaking—the president, secretary of state, secretary of defense, the Central Intelligence Agency (CIA) director, Chairman of the Joint Chiefs of Staff (the president's senior military adviser), and the 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 4 displays group hawkishness at the meeting level over time by calculating the average hawkishness of all meeting participants.

3 Results I: Testing the Microfoundations of the Adviser Model

We first take advantage of our rich deliberation data to validate the adviser model's microfoundations. The analysis in this section answers two questions. First, do leaders seek counsel from advisers in foreign policy deliberations? And second, does the nature of the counsel that advisers provide depend on their foreign policy dispositions?

3.1 Leaders Seek Counsel during Deliberations

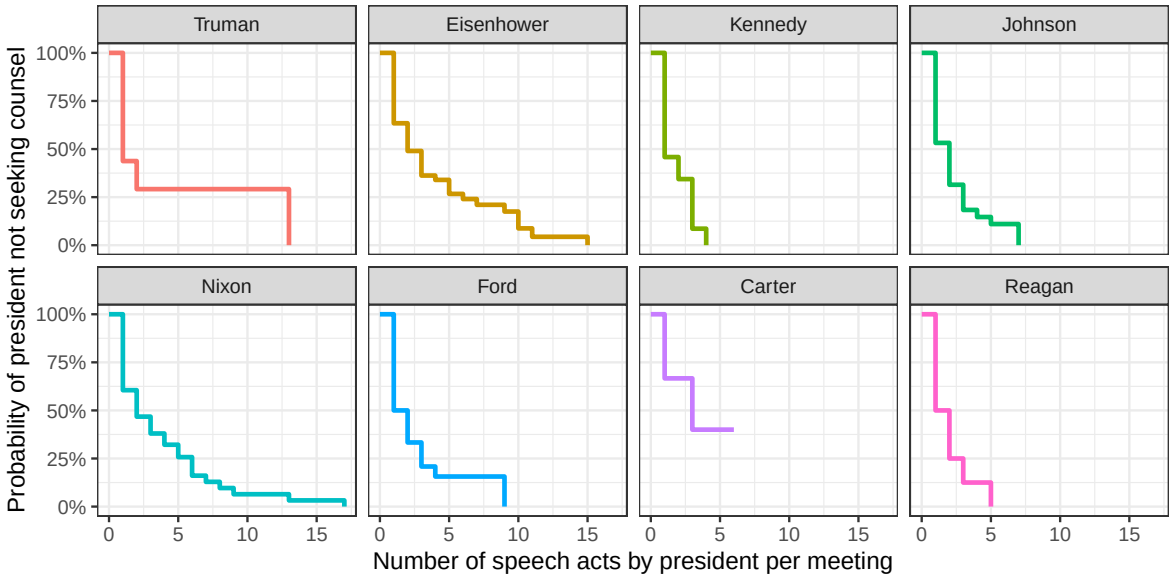
The first assumption of our adviser model is that leaders seek counsel from their advisers. The 2,685 meeting records in our data demonstrate that leaders routinely met with advisers. Our adviser model assumes, however, that these meetings should be genuinely deliberative: leaders should routinely seek counsel from their advisers, rather than merely informing them of decisions that have already been made. Deliberation also requires dissent: advisers should be willing to express opinions even when others disagree.

To evaluate whether these assumptions hold, we examine the timing and frequency of such deliberation in meetings. Since automated methods are unlikely to capture the subtleties of deliberation we wish to identify, we generated a stratified random sample of 270 formal and informal meetings across presidential administrations (approximately 10% of the full sample), for which we coded our concepts of interest manually. Drawing upon the study of deliberation elsewhere in political science, we developed a coding scheme (detailed in Appendix §8.1) to identify speech acts that exhibited *seeking counsel*. A research assistant reviewed and coded each of the 12,006 speech acts in the random sample. Seeking counsel was defined as a speaker requesting another participant to introduce new information, ideas, or recommendations into the discussion. Simply stating one's own position does not qualify as seeking counsel. Rather, speakers must have proactively asked *others* for their perspective. A meeting in which participants reiterated their own position over and over again in slightly different ways would be coded as having no search for counsel.

We find that US foreign policy deliberations exhibited a high level of seeking counsel, particularly by leaders. About one in three presidential speech acts—and over one in ten adviser speech acts—queried for more information from advisers. Figure 5 formalizes this intuition through a simple Cox model, in which administration is regressed on the duration of time before the leader seeks counsel. The model highlights that even the least inquisitive presidents still quickly sought counsel in their deliberations. While some of these queries might be performative, it is clear that presidents expend considerable time and effort soliciting input from advisers. This finding is difficult to reconcile with the leader model's emphasis on fixed and unchanging leader beliefs.

One question, however, is whether these queries simply led to the identical viewpoints being expressed ad nauseam. If leaders manipulated deliberations to ensure that only pre-approved

Figure 5: Leader Search for Counsel during Meetings



Note: Each panel depicts a survival curve reflecting likelihood of a president seeking counsel as a function of speech acts by the president. The sharper the decline, the *faster* a president sought counsel. See Appendix §8.1 for coding details.

viewpoints would be voiced, as the leader model contends, we would expect little disagreement among participants. To explore this contention, we replicated our coding process for speech acts exhibiting *dissent*, defined as a textual indication that the speaker disagreed with another meeting participant.[18] The data again suggest that leaders and advisers use deliberations to offer conflicting views: 14% of adviser speech acts and 10% of president speech acts in our random sample featured a dissenting opinion, and 64% of meetings featured some form of debate between participants. Appendix §8.6 uses these data to show that the probability of meeting attendees speaking is not a function of their dispositional distance from the president. The prevalence of debate suggests that, on average, leaders do not consistently manipulate deliberations to ensure conformity with their own worldview. We also find that expressions of dissent more frequently come from advisers who *differ* substantially in hawkishness relative to the president, which helps to allay concerns that self-censorship dominated the meetings. Collectively, the findings suggest that leaders seek input from advisers, and that participation allows advisers the opportunity to deliberate.

[18] In 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. This further motivated our choice to manually code the concept on a subsample of meetings rather than using automated methods on our whole dataset.

3.2 Advisers Provide Counsel Congruent with their Dispositions

The second assumption of our adviser model is that the counsel that advisers supply during deliberations depends on their predispositions. As noted in our theory section, we can think of counsel as consisting of information provision, problem representation (or analysis), and policy recommendations. To get at information provision and analysis, we use the rich textual data contained within the collected records, examining the content of speech acts by hawks and doves during the meetings.[19] Drawing on existing research on hawkishness, we identified ex ante five conceptual categories summarizing the types of considerations 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). 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 may be attuned to the military balance of capabilities, albeit for different reasons: hawks might emphasize the importance of primacy in material strength and power (Herrmann, Tetlock and Visser, 1999), while doves might 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 3 summarize the concepts 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 straightforward dictionary approach. We specify a set of nine to fourteen keywords, some of which are listed in Table 3, that capture each of our conceptual topics.[20] Using this list, we calculate the proportion of words associated with an individual adviser in a specific meeting that overlaps with these keywords (if at least 50 words in total). A total of 11,609 adviser-meeting observations, representing 100,089 speech acts (96% of our data), are processed in this manner.

[19] We group information provision and problem representation together given difficulties in empirically disentangling them using the automated content approaches we employ.

[20] Appendix §8.2 features full lists of terms for each topic.

Table 3: Sample of Hawkish and Dovish Terms

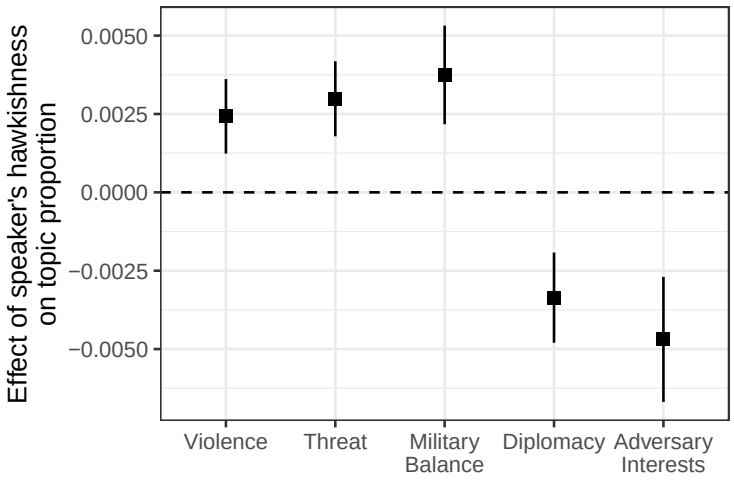
Topic	Who Emphasizes?	Source	Distinctive Terms
Military violence	hawks	George (1969); Herrmann, Tetlock and Visser (1999)	attack, bomb, escalate, invade, strike
International threats	hawks	Russett (1990)	hostile, threat, war, aggression, enemy
Military balance	hawks and doves	Herrmann, Tetlock and Visser (1999)	capability, missile, nuclear, silo, tank
Diplomacy	doves	Russett (1990)	diplomatic, negotiate, summit, talk, treaty
Adversary interests	doves	Brutger and Kertzer (2018)	china, communist, soviet, ussr, vietnam

Several examples suggest that these proportion measures effectively identify text related to our concepts of interest. For example, the highest scoring text for diplomacy comes from a meeting in June 1976, where Chairman of the Council of Economic Advisers Alan Greenspan reported on a recent Puerto Rico economic summit: "It was an extraordinary meeting, especially in the context of other meetings I have attended. There was a real intellectual grappling with major philosophical issues... We may have developed a new form of international institution. We have broken down the formality and protocol of summit meetings so that true dialogue can take place." The text most indicative of military violence comes an October 1958 meeting, in which Chief of Staff of the Air Force Thomas White noted that "our problem was that we must assume that the Soviets will strike first. If they do we cannot stop them by our Distant Early Warning lines. We must therefore find the number of bombers which it is logical for us to maintain in order to strike back after the initial Soviet attack."

We find that hawks and doves discuss systematically different considerations, consistent with their underlying predispositions. Figure 6 plots the effect of moving from the least to most hawkish speaker within a single administration on expected topic proportions in a speech act.[21] The plotted effects are based on OLS specifications that leverage the hawkishness scores described above, while also including administration fixed effects and controls for whether the meeting was formal or

[21] See Appendix §8.2 for full statistical results.

Figure 6: Effect of Adviser Hawkishness on Topic Proportions during Meetings



Note: Plot shows marginal effect of moving from a meeting's least to most hawkish adviser within a fixed administration. Bands represent 95% confidence intervals. Square points indicate 95% statistical significance.

whether the meeting record was a transcript. As shown, hawks and doves differ in the topics they emphasize. Consistent with theoretical expectations, hawks are more likely to address issues related to violence and threat, while doves place greater stress on diplomatic possibilities and adversary interests. Hawks' discussions of military balance appear to overwhelm those of doves. Beyond statistical significance, these differences are substantively meaningful, despite the seemingly small magnitude of the estimated coefficients. An average adviser-meeting observation has a *diplomacy* proportion of 0.0045. Moving from the least to most hawkish individual reduces the expected proportion by 0.0034, or about an 76% decrease.[22]

To explore our contention that predispositions shape the recommendations advisers provide, we drew another random sample of 425 meetings (approximately 16% of all meetings in our dataset) and manually coded the 20,451 speech acts in them for whether each speaker was calling for a highly hawkish policy: the use of coercive force against, or conflict with, an adversary. The analysis confirms the straightforward intuition that hawks tend to be more likely to make conflictual policy recommendations than doves do. Individuals in the 80th percentile of hawkishness are 72% more

[22]In Appendix §8.5, we replicate this analysis using a keyword-assisted topic model (keyATM). We find substantively similar relationships, with the exception of military balance, which is not associated with either hawks or doves using the keyATM approach. While one possibility is that these effects are an artifact of skillful leaders who steer deliberations toward their worldview, additional analysis in Appendix §8.3 suggest that advisers are in fact less likely to emphasize hawkish topics under hawkish leaders. On average, the topics that advisers emphasize do not necessarily conform to a leader's worldview, even if leaders might prefer they would.

likely to recommend conflictual policies than individuals in the 20th percentile of hawkishness are. Advisers may have reasons to go "against type" to improve their persuasive power—as hawks advocating against conflict may have more appeal than doves (Saunders, 2018; Mattes and Weeks, 2019)—but on average they do not.

In sum, analyzing tens of thousands of speech acts in foreign policy meetings offers evidence consistent with the microfoundations of the adviser model. Deliberations give advisers the opportunity to offer counsel, and the considerations and recommendations advisers raise depend on their disposition. Hawkish advisers raise considerations emphasizing military violence, while dovish advisers make arguments emphasizing the utility of diplomacy and adversary perspectives. These dispositional differences also hold in terms of the policy recommendations for which hawks and doves advocate. Next, we turn to the question of foreign policy decisions themselves.

4 Results II: Testing the Three Models of Aggregation

The above analysis shows how adviser dispositions affect the counsel they provide, but does not test how these dispositions aggregate to affect decision-making. We now turn to our entire meeting record data to address this question. In order to fully test our adviser model—and compare it to the emergence and leader models—we introduce an additional dataset, which studies the decisions made in each of our deliberation meetings, which we use for our main analysis.

4.1 Outcome Variable: Conflictual Decisions

Given our interest in how hawkishness as a leader- and adviser-level disposition aggregates in foreign policy decision-making, our central outcome of interest concerns policy choices aimed toward US adversaries in each meeting. 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

Table 4: Decisions Against Adversaries by Administration

Administration	Decisions	Conflictual	Cooperative
Truman	102	88	14
Eisenhower	417	329	88
Kennedy	124	100	24
Johnson	951	65	26
Nixon	75	43	32
Ford	56	20	36
Carter	22	12	10
Reagan	63	45	18
Total	950	702	248

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, a team of coders collected contextual background information, classified the decision into one of several categories, and specified a target of the decision.[23] 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.[24] Targets of the actions were the state or political organization, such as a rebel group, most directly affected by the decision.[25] 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 Table 4 shows, our sample yielded 950 decisions towards adversaries made across formal and informal meetings, 702 of which are conflictual decisions and 248 of which are cooperative.[26] We use these data to produce two measures. The first is a meeting's raw number of conflictual decisions toward adversaries. Because this is a count variable, the corresponding analyses use

[23] When 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.

[24] Our 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 §3.

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

[26] This 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.

Poisson regressions.[27] 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 §3.3.

4.1.1 Control Variables

One challenge with studying the effects of adviser hawkishness on foreign policy decision-making is that adviser participation is not randomly assigned. To address these potential selection effects, we employ a two-pronged empirical strategy, beginning with a set of control variables meant to address potential confounding factors for our main analysis, and then proceeding to an additional set of robustness tests we describe in Section 4.1.3 below.[28]

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 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, 2023). Models with fixed effects study the effect of group composition while holding the leader constant, 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. Although which advisers attend formal meeting is partly routinized, 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 104,504 speech acts in our sample.[29] Using these classifications as a control variable helps minimize potential bias in the

[27] Our results also hold using negative binomial models reported in Appendix §5.7.

[28] While these controls are relatively blunt instruments, they help to address the challenges associated with studying the effects of hawkishness on foreign policy decision-making. We turn to an alternative approach that bypasses many of these concerns in section 4.1.3 below.

[29] See Appendix §9 for more information on agenda items and their impact on group composition.

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. While this concern is allayed somewhat by the self-censorship and dissent results presented above, we also 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.[30]

Additional control variables track the proportion of meeting participants affiliated with the *Defense Department*, and *intelligence agencies*, *military*, and *State Department*, as bureaucratic interests may skew hawkish or dovish (Allison and Zelikow, 1999) – as well as the (logged) total 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 indicating whether a meeting was a *formal* meeting of the NSC, as opposed to an informal session outside it.

4.1.2 Results

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.[31] 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.

[30] These 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.

[31] Meeting 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 attendees is not available for all meetings.

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 5 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.[32] The left-hand panel of Figure 7 presents the results graphically. Shifting the 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 5 with a null model that omits the measure of mean group hawkishness strongly indicates that accounting for hawkishness yields a significant improvement ($p = 0.007$). A similar exercise using Model 2 produces an analogous result ($p = 0.025$). The findings thus do not support the emergence model.

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.[33] Results presented for Models 3 and 4 in Table 5 and the middle panel of Figure 7 suggest that presidential hawkishness does not have a systematic relationship with conflictual decisions toward adversaries.[34] Appendix §7 details a range of potential explanations for why leader-level hawkishness may not 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) and Carter and Smith (2020) in Appendix §7.2 similarly find little evidence that presidential hawkishness predicts American conflict behavior, consistent with Yarhi-Milo's (2018) recent work. We also obtain similar results when measuring presidents' hawkishness using measures derived from codings provided by experts of American foreign policy (Appendix §7.1), as well as using

[32] Full results, including bivariate specifications, are provided in Appendix §5.1.

[33] Because each president's hawkishness in our data does not appreciably change over time, we omit administration fixed effects.

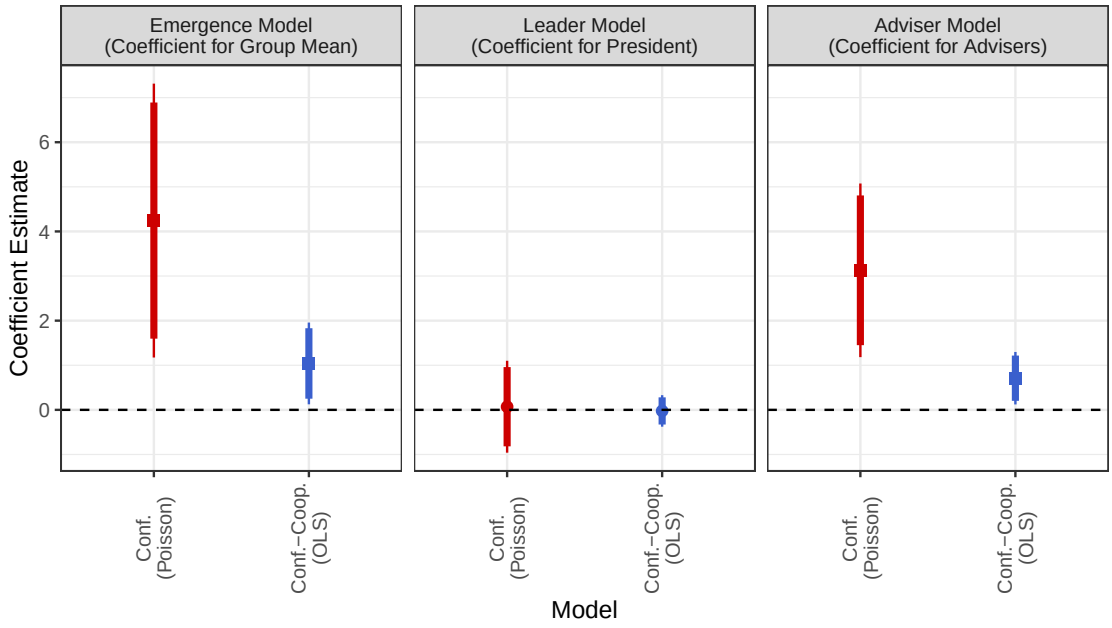
[34] We find this result is partly driven by Dwight Eisenhower, who our measure codes as comparatively dovish but nevertheless presided over an administration responsible for a substantial number of conflictual decisions. Yet as robustness checks in Appendix §7.4 show, we do not find evidence in favor of leader-level hawkishness even when dropping Eisenhower from the sample or rerunning the model with a neural net measure of hawkishness instead.

Table 5: Effect of Participant Hawkishness on Foreign Policy Decisions

	Emergence Model		*Leader Model*		*Adviser Model*		*Advisers + Admin. FEs*	
	Conf	Conf. − Coop.	Conf.	Conf. − Coop.	Conf	Conf. − Coop.	Conf.	Conf. − Coop.
	Poisson	*OLS*	*Poisson*	*OLS*	*Poisson*	*OLS*	*Poisson*	*OLS*
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Mean Hawkishness	4.243***	1.040**						
	(1.566)	(0.467)						
Advisers' Hawkishness (Acts)					3.128***	0.710**	2.893**	0.708**
					(0.993)	(0.300)	(1.198)	(0.345)
President's Hawkishness			0.071	−0.024	−0.835	−0.138		
			(0.526)	(0.180)	(0.600)	(0.224)		
No. of Attendees	0.069***	0.038***	0.075***	0.042***	0.090***	0.042***	0.078***	0.041***
	(0.017)	(0.007)	(0.016)	(0.006)	(0.017)	(0.007)	(0.017)	(0.007)
Proportion - Defense	2.081***	0.297**	1.328***	0.198*	1.592***	0.182	2.052***	0.293**
	(0.364)	(0.125)	(0.327)	(0.106)	(0.340)	(0.121)	(0.367)	(0.127)
Proportion - Intelligence	1.580***	0.260	1.310**	0.202	1.597***	0.247	1.636***	0.268
	(0.604)	(0.167)	(0.575)	(0.148)	(0.587)	(0.168)	(0.606)	(0.170)
Proportion - Military	0.936*	0.192	1.333***	0.327**	0.932**	0.193	1.173**	0.237
	(0.561)	(0.185)	(0.452)	(0.137)	(0.465)	(0.162)	(0.526)	(0.180)
Proportion - State	0.242	−0.016	0.047	0.031	0.278	−0.011	0.233	−0.013
	(0.359)	(0.090)	(0.350)	(0.078)	(0.360)	(0.090)	(0.364)	(0.093)
Diplomatic Experience	0.123***	0.004	0.085*	0.003	0.071	−0.009	0.119**	0.003
	(0.047)	(0.016)	(0.045)	(0.013)	(0.045)	(0.015)	(0.047)	(0.016)
Intelligence Experience	−0.122**	−0.044**	−0.089*	−0.048***	−0.112**	−0.039**	−0.120**	−0.043**
	(0.053)	(0.018)	(0.046)	(0.016)	(0.047)	(0.018)	(0.053)	(0.019)
Military Experience	0.041	−0.050*	0.133	−0.010	0.056	−0.045*	−0.001	−0.062**
	(0.092)	(0.026)	(0.082)	(0.021)	(0.085)	(0.025)	(0.092)	(0.027)
5-Year MID Challenges	−0.267	−0.110	0.083	−0.047	0.036	0.001	−0.260	−0.116
	(0.232)	(0.078)	(0.148)	(0.041)	(0.150)	(0.046)	(0.233)	(0.080)
US CINC	6.427***	2.831***	2.759***	1.026***	2.072*	1.337***	6.769***	2.909***
	(2.356)	(0.919)	(1.069)	(0.318)	(1.099)	(0.366)	(2.360)	(0.933)
Formal	0.622***	0.139***	0.783***	0.205***	0.763***	0.148***	0.643***	0.145***
	(0.136)	(0.051)	(0.129)	(0.044)	(0.130)	(0.050)	(0.137)	(0.052)
Constant	−7.219***	−1.320***	−4.314***	−0.298	−5.092***	−0.664***	−6.474***	−1.109***
	(1.270)	(0.420)	(0.669)	(0.193)	(0.727)	(0.229)	(1.153)	(0.384)
Administration FEs	✓	✓					✓	✓
Agenda Items	✓	✓	✓	✓	✓	✓	✓	✓
Observations	2,685	2,685	2,685	2,685	2,650	2,650	2,650	2,650

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

Figure 7: Summary of Three Models of Trait Aggregation



Note: Coefficient plots corresponding to main findings, corresponding to Models 1-6 in Table 5. Thicker and smaller bands represent 90% and 95% confidence intervals respectively. Square points indicate 95% statistical significance.

a different operationalization of the dependent variable (Table 7 below). As with most studies of presidential decision-making, a sample of eight leaders limits the conclusiveness of these results, as the idiosyncratic nature of a single president could be exerting an outsized effect on the results. Within these confines, however, our analysis offers little support for the leader model.

The adviser model predicts that the hawkishness of advisers affects decision outcomes in deliberations. If advisers exert influence through the counsel they provide, 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.[35]

As shown by Models 5 and 6 in Table 5 and the right-hand panel of Figure 7, 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

[35] Consequently, 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.

Table 6: Predicted Number of Conflictual Decisions

Model	Model	Variable	Min.	Max.
Emergence	Poisson (Model 1)	Mean Hawkishness	0.101	0.482
	OLS (Model 2)	Mean Hawkishness	0.037	0.420
Leader	Poisson (Model 3)	President's Hawkishness	0.252	0.257
	OLS (Model 4)	President's Hawkishness	0.365	0.358
Adviser	Poisson (Model 5)	Advisers' Hawkishness (Acts)	0.125	0.406
		5-Year MID Challenges	0.221	0.237
		US CINC	0.195	0.289
	OLS (Model 6)	Advisers' Hawkishness (Acts)	0.073	0.341
		5-Year MID Challenges	0.213	0.215
		US CINC	0.103	0.356

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.

appears across specifications. Holding other variables fixed, shifting the group's weighted hawkishness from the minimum to maximum values more than 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 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.

To further contextualize the substantive effects, Table 6 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.

4.1.3 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 leader-level differences in 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), 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 0.043 and 0.179: there is approximately 5.6 to 23 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.

One reason why advisers are not simply dispositional mimeographs of the leaders they serve is that adviser appointment can be affected by multiple considerations, such as a candidate's education, experience, qualifications, personal connections, and fit for the position (Jost and Kertzer, 2023)—not just their hawkishness. Moreover, many leaders prefer viewpoint diversity in their advisory group either to improve the quality of foreign policy debates by considering multiple perspectives (Mintz and Wayne, 2016) or to bolster their domestic credibility (Saunders, 2018). For every example of hawkish leaders like Ronald Reagan selecting dispositionally similar advisers, like Caspar Weinberger, there are also examples of leaders like Barack Obama selecting dissimilar advisers, like Hillary Clinton. We find further evidence in support of this contention in Appendix §6, where we build a dataset of "runner ups" considered for high-level national security positions since the Truman administration, showing that individuals ultimately appointed were not systematically closer in hawkishness to the president than those who were considered but not selected.

Second, even if leaders do not merely select mimeographs as their advisers, leaders may still have the ability to decide when these 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 ignores our meeting data altogether, instead examining the effect of adviser-level hawkishness (limited to NSC principals) on the United States' propensity for being involved in militarized interstate disputes in a given month. Our findings, shown in Table 7 below, 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). This strongly suggests that, even if leaders attempt to manipulate the advisory group or fabricate deliberation in ways that accord with their worldview, advisers are still able to sway foreign policy decisions in aggregate. Moreover, Appendix §5.9 replicates our meeting-level analysis for formal gatherings, only using characteristics of NSC principals, who are obligated to have a presence at every meeting. The results remain consistent. Adviser predispositions appear to be significantly associated with 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 §5.6, we run a robustness check where we drop decisions involving the Soviet Union from our analysis and find that results remain broadly consistent.

Finally, given space constraints, our analysis here focuses on establishing that adviser dispositions affect the counsel they provide leaders in deliberations, and the decision the leader makes—rather than the follow-up question of when leaders are more likely to heed advisers' counsel, a question we explore in other research. Nonetheless, important recent work by Saunders (2017, 2018) suggests some possible scope conditions to our findings, such that leaders may be less likely to be swayed by their advisers when leaders are more experienced and when dovish leaders are paired with dovish advisers. In Appendix §5.11, we use our decision-making dataset to test both propositions. We find that, at least when it comes to adviser hawkishness, neither leader-adviser gaps in experience nor leader-adviser gaps in predisposition significantly moderate the effects of adviser traits. Across the Cold War, hawkish advisers in the United States were no less able to push

Table 7: Effect of NSC Principals' Hawkishness on MIDs, Using Monthly Data

	Dependent variable:	
	Onset of MIDs involving US	
	(1)	(2)
Advisers' hawkishness	6.816**	11.198**
	(2.734)	(4.629)
President's hawkishness	−3.510**	−3.795
	(1.593)	(2.832)
War ongoing		−0.377**
		(0.153)
Deaths per capita in last war (logged)		0.163
		(0.111)
Months since last war (logged)		0.294
		(0.198)
Victory in last war		−1.070
		(0.760)
MID challenges to US in last 5 years		−0.035
		(0.049)
Average MID outcome in last 5 years		−55.273
		(33.947)
Economic recession		−0.194
		(0.216)
Unified government		0.404
		(0.258)
US material capabilities		−5.518
		(6.850)
President's tenure (logged months)		−0.016
		(0.090)
Constant	−2.989***	−3.126
	(0.836)	(2.942)
Observations	501	501

Note: $^{*}p < 0.1$; $^{**}p < 0.05$; $^{***}p < 0.01$. Outcome variable is the number of US MIDs (display of force or above) beginning in a given month. Advisers' hawkishness reflects average hawkishness score of the most senior advisers in the administration in a given month. See Appendix §5.8 for details.

decision-making in a hawkish direction under experienced leaders than inexperienced ones; similarly, hawkish and dovish advisers appear to influence hawkish and dovish leaders alike. Second, one of the virtues of our main analysis is that it avoids aggregation bias by considering the universe of substantive decisions being made at the meetings, but this raises questions about whether our results are the artifact of lower stakes decisions rather than the high stakes decisions made in crises. To ensure this is not the case, Appendix §5.10 shows that the effects of adviser dispositions remain significant both in and out of international crises featuring decision-making on high-stakes issues.

5 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 dispositions in shaping foreign policy outcomes than adviser-level ones. In this article, we develop an argument linking adviser dispositions to consequential foreign policy choices about peace and 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 over 2,500 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 systematically study the ways in which the advisers surrounding leaders matter, particularly for questions of interstate conflict.

The theory and findings suggest that leaders by themselves may be insufficient to explain many foreign policy decisions. While we emphasize that the leader and adviser models are complementary, we show that leaders consistently turn to advisers for counsel during consequential foreign policy meetings, that adviser dispositions shape the type of counsel that leaders receive, and that shifting from a maximally dovish to a maximally hawkish advisory group triples the expected number of conflictual decisions in a meeting. These dynamics illuminate an intuitive and compelling reason that advisers wield such influence: even experienced leaders confront numerous policy challenges on which they are relatively uninformed and hold few preconceived notions. And even when a leader knows *what* they want, there is often room for hearing disparate perspectives concerning the numerous possible strategies for *how* to get it. Advisers supply the information, analysis, and

recommendations that leaders demand.[36]

More broadly, our findings 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 deliberation serves as a crucial and under-explored conduit through which individual-level dispositions, such as hawkishness, affect foreign policy outcomes. Aggregation does not forge, *ex machina*, a tabula rasa within the group. Knowing the dispositions of the advisers who dominate policy debates has substantial explanatory power for state behavior.

The basic logic of our adviser model suggests that leaders depend on advisers for psychological and informational reasons, which suggests broad applicability across countries beyond the United States. Yet three core elements of the model also imply corresponding scope conditions. First, because advisers must be able to provide counsel congruent with their disposition without fear of leader retribution, our model may be less applicable in authoritarian regimes, particularly personalist ones (Weeks, 2014), in which leaders can severely and arbitrarily punish advisers who disagree. Second, because advisers must have access to the leader, the adviser model may offer limited insight into countries with bureaucratic institutions designed to exclude advisers from decision-making, such as China during Deng Xiaoping's early years (Jost, 2023). Third, since leaders must be at least somewhat receptive to the counsel that advisers provide, the adviser model may not apply under leaders who are extremely closed-minded either for situational (Reiter and Stam, 2002) or dispositional (Preston, 2001) reasons.

Our adviser model suggests a broad agenda for future research. Most broadly, it calls for more scholarly attention to be paid to leaders, advisers, and the institutions that connect them. Yet this study intentionally sets aside several questions that subsequent scholarship could explore. First, it does not differentiate between cases where advisers are successful in shaping decision-making because leaders rely on their counsel to form beliefs on issues they have not fully considered, versus cases where advisers successfully persuade leaders to change their views. Future scholarship could leverage the data we introduce to tease apart these two mechanisms.

Second, while we provide evidence suggesting that what happens during deliberations matters to a leader's decision, it is also possible that adviser influence depends on expending effort by

[36]However, this also implies that leaders might overshadow their advisers when they hold firm beliefs and preferences about certain policy issues (e.g., Preston, 2001; Saunders, 2011; Yarhi-Milo, 2018). This presents a fruitful avenue for future research.

setting the agenda and building bureaucratic coalitions *prior to* deliberations. It is also possible that some advisers may be more influential than others—and that past deliberations shape future ones in intriguing, path-dependent ways.

Third, while our primary aim in this manuscript is to show how advisers matter systematically for outcomes of broad concern to the field of IR, these dynamics are clearly the final set of steps in a long causal chain. A more comprehensive approach would systematically study the antecedent stages prior to entering the "room where it happens"—from institutional design, to adviser appointment, to adviser attendance, to adviser behavior, to the decision.

Finally, our method of inquiry suggests how scholars might connect the study of decision-making with the study of numerous other international behaviors, 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 archival documents or through quantitative methods that focus on state behavior rather than decision-making. The approach developed here offers a middle path: to study state behavior by quantitatively analyzing archival documents that span an extended period of time (Min, 2022), but in a way that still directly observes the decision-making process.

References

Allison, Graham T. and Philip Zelikow. 1999. *Essence of Decision: Explaining the Cuban Missile Crisis*. 2 ed. New York: Longman.

Aristotle. 1920. *Politics*. Oxford: Clarendon Press.

Brutger, Ryan and Joshua D. Kertzer. 2018. "A Dispositional Theory of Reputation Costs." *International Organization* 72(3):693–724.

Burke, John P. and Fred L. Greenstein. 1989. *How Presidents Test Reality: Decisions on Vietnam, 1954 and 1965*. Russell Sage Foundation.

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.

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.

Destler, I.M. 1972. *Presidents, Bureaucrats, and Foreign Policy: The Politics of Organizational Reform*. Princeton University Press.

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.

Garrison, Jean A. 1999. *Games Advisors Play*. College Station: Texas A&M University Press.

George, Alexander. 1980. *Presidential Decision Making in Foreign Policy*. Boulder: Westview Press.

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.

Goldhamer, Herbert. 1978. *The Adviser*. New York: Elsevier.

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.

Harden, John P. 2021. "All the World's a Stage: US Presidential Narcissism and International Conflict." *International Studies Quarterly* 65(3):825–837.

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. and Michael P. Fischerkeller. 1995. "Beyond the Enemy Image and Spiral Model: Cognitive-Strategic Research After the Cold War." *International Organization* 49(3):415–50.

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.

Hurwitz, Jon and Mark Peffley. 1987. "How Are Foreign Policy Attitudes Structured?" *American Political Science Review* 81(4):1099–1120.

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.

Jost, Tyler. 2023. *Bureaucracy and War: The Institutional Origins of Miscalculation in International Politics*. New York: Cambridge University Press.

Jost, Tyler and Joshua D. Kertzer. 2023. "Armies and Influence: Public Deference to Foreign Policy Elites." *Journal of Conflict Resolution* Forthcoming.

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.

Kenwick, Michael R and Sarah Maxey. 2022. "You and Whose Army? How Civilian Leaders Leverage the Military's Prestige to Shape Public Opinion." *The Journal of Politics* 84(4).

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.

Kertzer, Joshua D., Marcus Holmes, Brad L. LeVeck and Carly Wayne. 2022. "Hawkish Biases and Group Decision Making." *International Organization* 76(3):513–548.

Khong, Yuen Foong. 1992. *Analogies at War: Korea, Munich, Dien Bien Phu, and the Vietnam Decisions of 1965*. Princeton: Princeton University Press.

Krasner, Stephen D. 1972. "Are Bureaucracies Important? (Or Allison Wonderland)." *Foreign Policy* (7):159–179.

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.

Maoz, Zeev. 1990. "Framing the National Interest: The Manipulation of Foreign Policy Decisions in Group Settings." *World Politics* 43(1):77–110.

Mattes, Michaela and Jessica L. P. Weeks. 2019. "Hawks, Doves and Peace: An Experimental Approach." *American Journal of Political Science* 63(1):53–66.

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.

Min, Eric. 2022. "Painful Words: The Effect of Battlefield Activity on Conflict Negotiation Behavior." *Journal of Conflict Resolution* 66(4-5):595–622.

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.

Preston, Thomas. 2001. *The President and His Inner Circle: Leadership Style and the Advisory Process in Foreign Policy Making.* New York: Columbia University Press.

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.

Renshon, Jonathan and Stanley A Renshon. 2008. "The theory and practice of foreign policy decision making." *Political Psychology* 29(4):509–536.

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. 2015. "War and the Inner Circle: Democratic Elites and the Politics of Using Force." *Security Studies* 24(3):466–501.

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. 2022. "Informing the Leader: Bureaucracies and International Crises." *American Political Science Review* 116(4):1460–1476.

Schultz, Kenneth A. 2005. "The Politics of Risking Peace: Do Hawks or Doves Deliver the Olive Branch?" *International Organization* 59(1):1–38.

Snyder, Richard C., H.W. Bruck and Burton Sapin. 1962. Decision-Making as an Approach to the Study of International Politics. In *Foreign Policy Decision-Making*, ed. Richard C. Snyder, H.W. Bruck and Burton Sapin. Glencoe, IL: The Free Press.

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.

Taber, Charles S. and Milton Lodge. 2006. "Motivated Skepticism in the Evaluation of Political Beliefs." *American Journal of Political Science* 50(3):755–69.

Tomz, Michael, Jessica Weeks and Keren Yarhi-Milo. 2020. "Public Opinion and Decisions about Military Force in Democracies." *International Organization* 74(1):119–143.

Walker, Stephen G., Mark Schafer and Michael D. Young. 2003. *The Psychological Assessment of Political Leaders: With Profiles of Saddam Hussein and Bill Clinton*. Ann Arbor: University of Michigan Press.

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

July 11, 2023

Contents

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 missingness reflects the extended timeline for declassification as opposed to substantive considerations that might sway the relative influence of presidents versus advisers. As we note in the paper, work by Preston (2001) suggests that because Reagan and Carter were relatively inexperienced in foreign policy, they tended to delegate more to their advisers, such that we should expect their underrepresentation here should actually make it harder to find evidence in favor of our adviser model.

1.3 Meeting Statistics

Table A1 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	3,114	77	66	11
Eisenhower	337	359	22,172	315	261	54
Kennedy	66	89	12,072	90	75	15
Johnson	72	75	1,463	26	21	5
Nixon	60	90	4,463	29	15	14
Ford	34	42	4,708	9	4	5
Carter	18	41	1,372	18	12	6
Reagan	79	153	2,961	63	45	18
Total	791	977	52,325	627	499	128

Table A1: Coverage of formal NSC records for each administration.

Table A2 displays summary statistics on all informal meetings included in our analysis.

Admin.	**Records**	**Speech Acts**	**Decisions**	**Conf. Adv.**	**Coop. Adv.**
Truman	90	514	25	22	3
Eisenhower	552	5,708	102	68	34
Kennedy	184	2,924	34	25	9
Johnson	437	11,141	65	44	21
Nixon	393	27,497	46	28	18
Ford	204	3,572	47	16	31
Carter	30	770	4	0	4
Reagan	4	53	0	0	0
Total	1,894	52,179	323	203	120

Table A2: Coverage of informal meeting records for each administration.

2 Meeting Record Segmentation

A significant step in our analysis requires the conversion of raw meeting records into "speech acts," which are the uninterrupted words spoken by a single individual during a meeting.

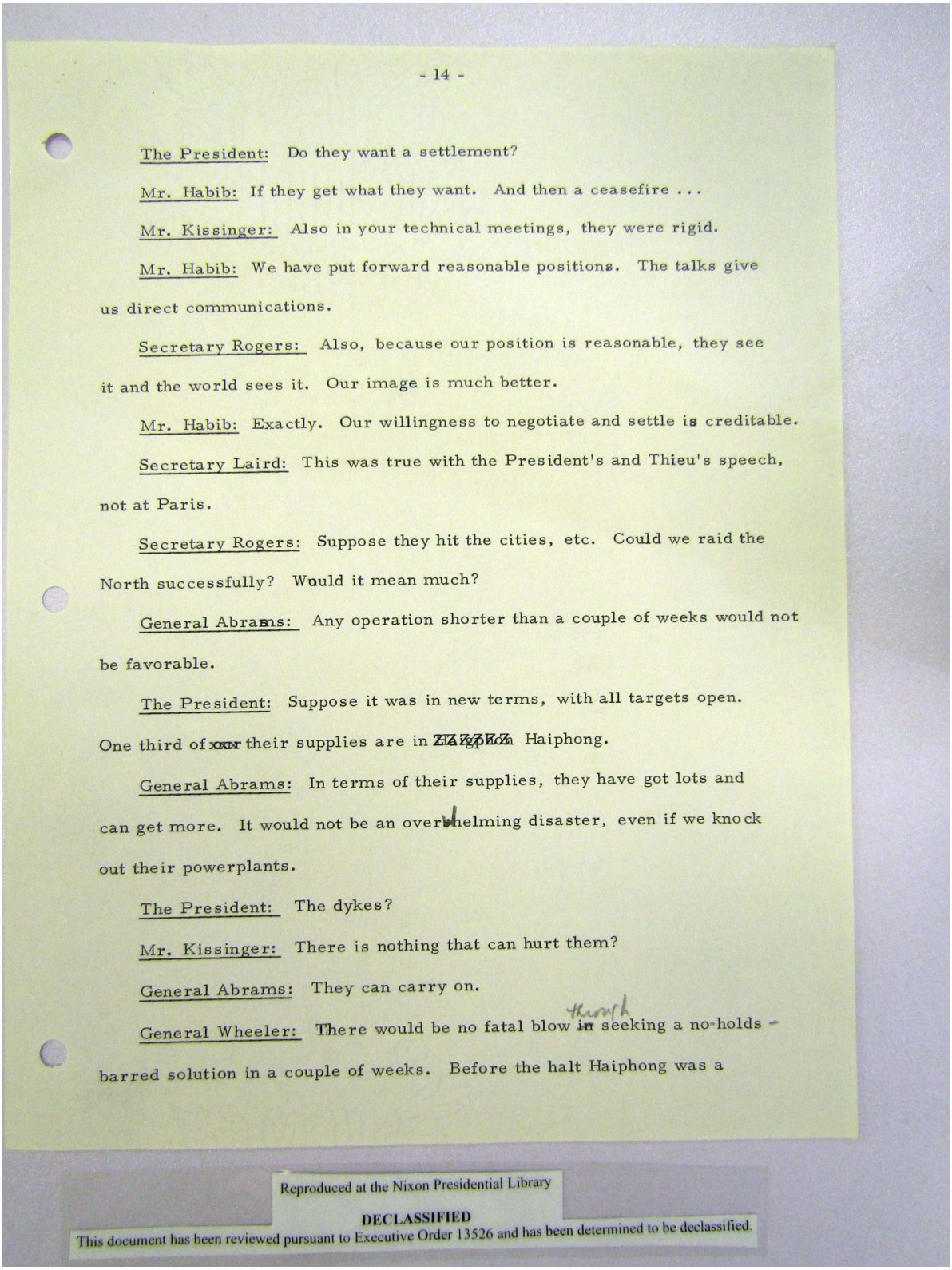
For meeting records that are in the form of transcripts, this task is relatively straightforward. Figure A1 displays one page from a meeting transcript during the Nixon administration. The document is formatted clearly so that each speech act and associated actor is clearly identified. We use our actor data to identify the full name of the individual based on titles or last names. The raw material on this page is converted into speech act data as reflected in Table A3.

The task is more complex for meeting records that are in the form of minutes. Figure A2 represents a page from an NSC meeting during the Eisenhower administration. As is clear from the image, paragraphs do not directly correspond to speech acts. We split these documents into speech acts by first identifying the use of key titles and last names. The light red rectangles in Figure A2 identify the relevant terms in the page. The text is then split at the sentence level according to these terms. We then use our actor data to identify the full name of each individual associated with a sentence or cluster of sentences, using either their last name or the title they held at the time.

Two adjustments are worth noting. First, we do not use information from introductory clauses to determine splits in the meeting records. For instance, the first sentence of the first full paragraph includes an introductory clause that mentions Dillon Anderson's comments (which were the subject of the previous paragraph of the meeting record). But it is clear from the full sentence that the speech act is attributed to President Eisenhower. In the infrequent cases where introductory clauses are used, they are discarded when identifying relevant record splits. Second, note that some splits using this procedure will divide statements made by the same person. The third full paragraph in Figure A2 begins with "The Secretary of State," and it then mentions "Secretary Dulles" in the fifth sentence. Secretary John Foster Dulles is the Secretary of State. As such, once all individuals are identified by name using our actor data, we reintegrate any sequential sentences that were originally divided but refer to the same person.

This process converts Figure A2 into speech act data reflected in Table A4.

Figure A1: Page of an NSC meeting record from September 12, 1969.



The President: Do they want a settlement?

Mr. Habib: If they get what they want. And then a ceasefire ...

Mr. Kissinger: Also in your technical meetings, they were rigid.

Mr. Habib: We have put forward reasonable positions. The talks give us direct communications.

Secretary Rogers: Also, because our position is reasonable, they see it and the world sees it. Our image is much better.

Mr. Habib: Exactly. Our willingness to negotiate and settle is creditable.

Secretary Laird: This was true with the President's and Thieu's speech, not at Paris.

Secretary Rogers: Suppose they hit the cities, etc. Could we raid the North successfully? Would it mean much?

General Abrams: Any operation shorter than a couple of weeks would not be favorable.

The President: Suppose it was in new terms, with all targets open. One third of their supplies are in Haiphong.

General Abrams: In terms of their supplies, they have got lots and can get more. It would not be an overwhelming disaster, even if we knock out their powerplants.

The President: The dykes?

Mr. Kissinger: There is nothing that can hurt them?

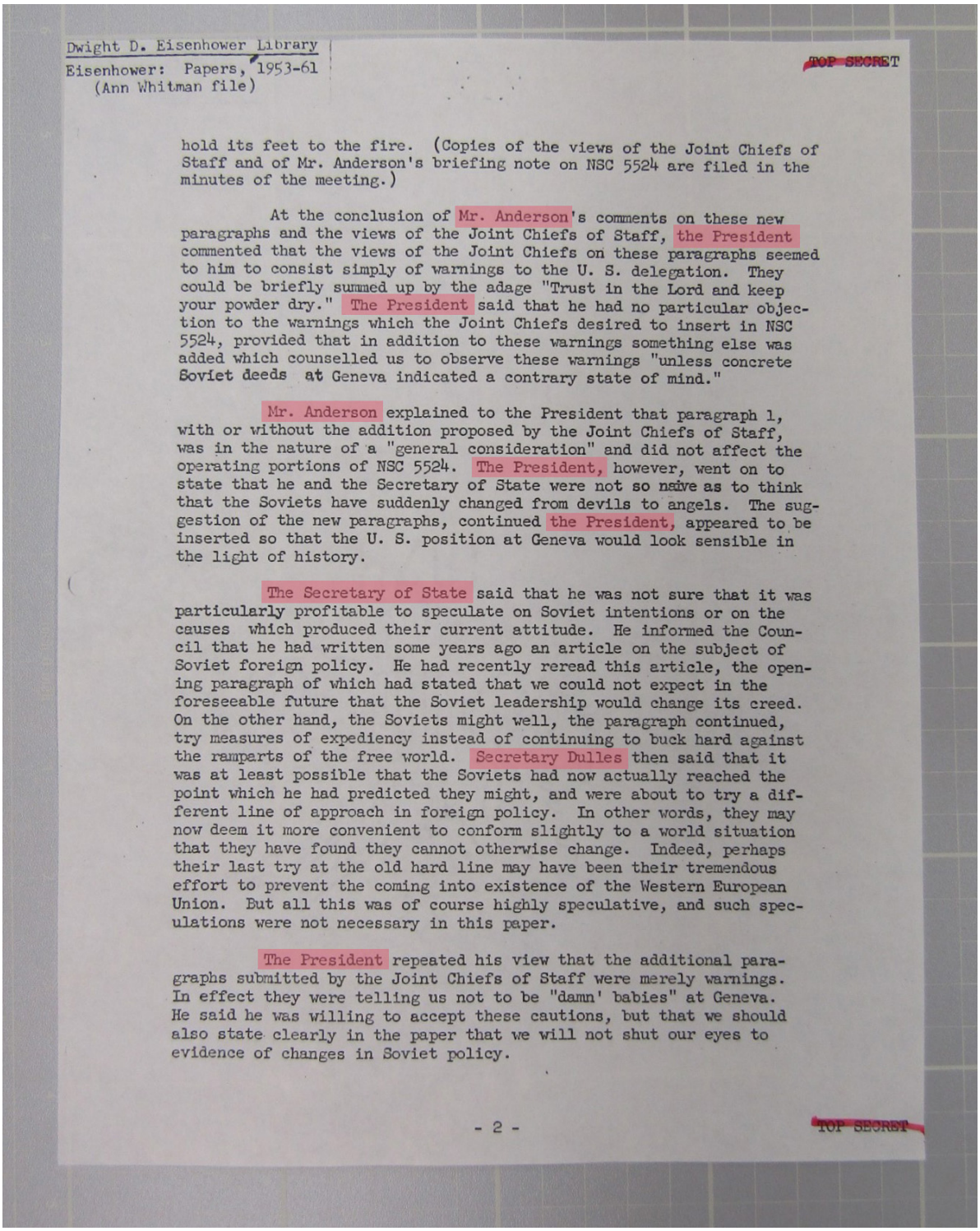
General Abrams: They can carry on.

General Wheeler: There would be no fatal blow through seeking a no-holds-barred solution in a couple of weeks. Before the halt Haiphong was a

Table A3: Speech act data from Figure A1.

Speech Act	Name
The President: Do they want a settlement?	Richard Nixon
Mr. Habib: If they get what they want. And then a ceasefire...	Philip Habib
Mr. Kissinger: Also in your technical meetings they were rigid.	Henry Kissinger
Mr. Habib: We have put forward reasonable positions. The talks give us direct communications.	Philip Habib
Secretary Rogers: Also because our position is reasonable they see it and the world sees it. Our image is much better.	William Rogers
Mr. Habib: Exactly. Our willingness to negotiate and settle is creditable.	Philip Habib
Secretary Laird: This was true with the President's and Thieu's speech not at Paris.	Melvin Laird
Secretary Rogers: Suppose they hit the cities etc. Could we raid the North successfully? Would it mean much?	William Rogers
General Abrams: Any operation shorter than a couple of weeks would not be favorable.	Creighton Abrams
The President: Suppose it was in new terms with all targets open. One third of nor their supplies are in Haiphong.	Richard Nixon
General Abrams: In terms of their supplies they have got lots and can get more. It would not be an overwhelming disaster even if we knock out their powerplants.	Creighton Abrams
The President: The dykes?	Richard Nixon
Mr. Kissinger: There is nothing that can hurt them?	Henry Kissinger
General Abrams: They can carry on.	Creighton Abrams
General Wheeler: There would be no fatal blow through seeking a no-holds-barred solution in a couple of weeks. Before the halt Haiphong was	Earle Wheeler

Figure A2: Page of an NSC meeting record from July 5, 1955. Terms indicating titles and last names are highlighted in light red.



Dwight D. Eisenhower Library
Eisenhower: Papers, 1953-61
(Ann Whitman file)

TOP SECRET

hold its feet to the fire. (Copies of the views of the Joint Chiefs of Staff and of Mr. Anderson's briefing note on NSC 5524 are filed in the minutes of the meeting.)

At the conclusion of Mr. Anderson's comments on these new paragraphs and the views of the Joint Chiefs of Staff, the President commented that the views of the Joint Chiefs on these paragraphs seemed to him to consist simply of warnings to the U. S. delegation. They could be briefly summed up by the adage "Trust in the Lord and keep your powder dry." The President said that he had no particular objection to the warnings which the Joint Chiefs desired to insert in NSC 5524, provided that in addition to these warnings something else was added which counselled us to observe these warnings "unless concrete Soviet deeds at Geneva indicated a contrary state of mind."

Mr. Anderson explained to the President that paragraph 1, with or without the addition proposed by the Joint Chiefs of Staff, was in the nature of a "general consideration" and did not affect the operating portions of NSC 5524. The President, however, went on to state that he and the Secretary of State were not so naïve as to think that the Soviets have suddenly changed from devils to angels. The suggestion of the new paragraphs, continued the President, appeared to be inserted so that the U. S. position at Geneva would look sensible in the light of history.

The Secretary of State said that he was not sure that it was particularly profitable to speculate on Soviet intentions or on the causes which produced their current attitude. He informed the Council that he had written some years ago an article on the subject of Soviet foreign policy. He had recently reread this article, the opening paragraph of which had stated that we could not expect in the foreseeable future that the Soviet leadership would change its creed. On the other hand, the Soviets might well, the paragraph continued, try measures of expediency instead of continuing to buck hard against the ramparts of the free world. Secretary Dulles then said that it was at least possible that the Soviets had now actually reached the point which he had predicted they might, and were about to try a different line of approach in foreign policy. In other words, they may now deem it more convenient to conform slightly to a world situation that they have found they cannot otherwise change. Indeed, perhaps their last try at the old hard line may have been their tremendous effort to prevent the coming into existence of the Western European Union. But all this was of course highly speculative, and such speculations were not necessary in this paper.

The President repeated his view that the additional paragraphs submitted by the Joint Chiefs of Staff were merely warnings. In effect they were telling us not to be "damn' babies" at Geneva. He said he was willing to accept these cautions, but that we should also state clearly in the paper that we will not shut our eyes to evidence of changes in Soviet policy.

- 2 -

TOP SECRET

Table A4: Speech act data from Figure A2.

Speech Act	Name
(At the conclusion of Mr. Anderson's comments on these new paragraphs and the views of the Joint Chiefs of Staff) the President commented that the views of the Joint Chiefs on these paragraphs seemed to him to consist simply of warnings to the U. S. delegation. They could be briefly summed up by the adage Trust in the Lord and keep your powder dry. The President said that he had no particular objection to the warnings which the Joint Chiefs desired to insert in NSC 5524 provided that in addition to these warnings something else was added which counselled us to observe these warnings unless concrete Soviet deeds at Geneva indicated a contrary state of mind.	Dwight Eisenhower
Mr. Anderson explained to the President that paragraph 1 with or without the addition proposed by the Joint Chiefs of Staff was in the nature of a general consideration and did not affect the operating portions of NSC 5524.	Dillon Anderson
The President however went on to state that he and the Secretary of State were not so naive as to think that the Soviets have suddenly changed from devils to angels. The suggestion of the new paragraphs continued the President appeared to be inserted so that the U. S. position at Geneva would look sensible in the light of history.	Dwight Eisenhower
The Secretary of State said that he was not sure that it was particularly profitable to speculate on Soviet intentions or on the causes which produced their current attitude. He informed the Council that he had written some years ago an article on the subject of Soviet foreign policy. He had recently reread this article the opening paragraph of which had stated that we could not expect in the foreseeable future that the Soviet leadership would change its creed. On the other hand the Soviets might well the paragraph continued try measures of expediency instead of continuing to buck hard against the ramparts of the free world. Secretary Dulles then said that it was at least possible that the Soviets had now actually reached the point which he had predicted they might and were about to try a different line of approach in foreign policy. In other words they may now deem it more convenient to conform slightly to a world situation that they have found they cannot otherwise change. Indeed perhaps their last try at the old hard line may have been their tremendous effort to prevent the coming into existence of the Western European Union. But all this was of course highly speculative and such speculations were not necessary in this paper.	John Dulles
The President repeated his view that the additional paragraphs submitted by the Joint Chiefs of Staff were merely warnings. In effect they were telling us not to be damn babies at Geneva. He said he was willing to accept these cautions but that we should also state clearly in the paper that we will not shut our eyes to evidence of changes in Soviet policy.	Dwight Eisenhower

3 Methodology for Measuring Decision Outcomes

3.1 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 an 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).

3.2 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 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 a 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 a 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-

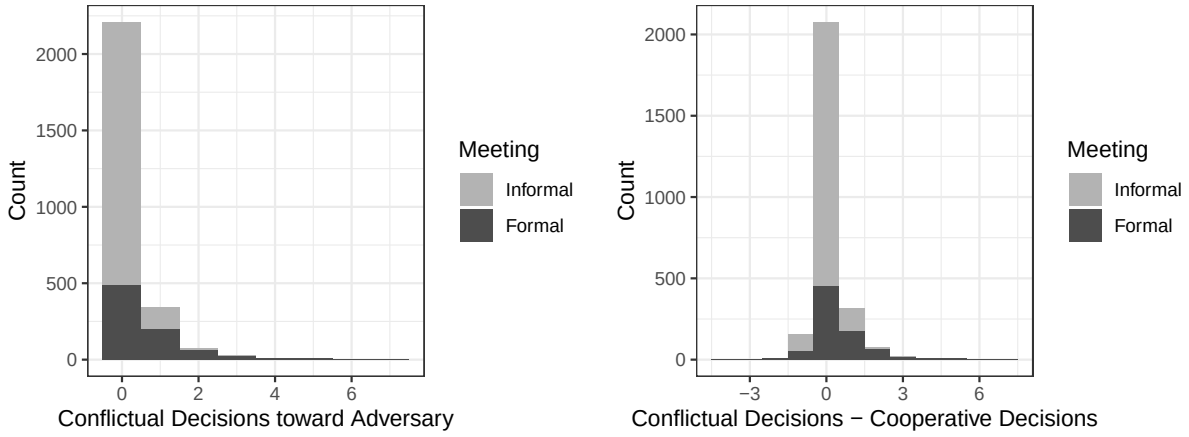
erative distinction. Moreover, the large presence of military planning decisions in our data does not have a natural analogue in existing event coding scales which further undercuts the validity of trying to apply those scales to our data.

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

3.3 Distribution of the Outcome Variable

Figure A3 illustrates the distribution of our two outcome variables of interest in the meeting-level analysis.

Figure A3: Distributions of Foreign Policy Decisions



4 Methodology for Measuring Decision-Maker Hawkishness

4.1 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

4.2 Foreign Policy Leadership Project Survey

The following attributes of individuals were extracted from 2,119 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.

4.2.1 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)

[1] We 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)

4.3 Predictive Model Descriptive Statistics

We generate predicted measures of actor hawkishness using a boosted linear regression model. The prediction process begins by first randomly splitting the FPLP data into two samples: one representing 70% of the data, and the other being the remaining 30%. Broadly speaking, the 70% sample is used to train the model, and the other 30%, called the test set, is used to measure the model's predictive performance on data that was not used in any way to create the original model itself.

The boosted linear regression model features a hyperparameter regarding the number of boosting iterations that should be applied to a model (*mstop*). We test four different potential values of *mstop*: 150, 250, 350, and 450. In order to determine which value of *mstop* is "best," we rely on five-fold cross-validation. We take our 70% sample of the FPLP data and split it once again into five equally apportioned subsamples. Four of the subsamples are deemed the training set, while the remaining fifth subsample is considered the validation set. The linear model is fit to the training set using the predetermined values of *mstop* iterations of boosting, and then these models are used to generate predictions for the validation set. The predictions are compared to the known actual outcome values of interest in the validation set to produce a measure of out-of-sample performance. This process is done five times so that each of the subsamples is used as a validation set, and the average performance across all five folds is the model's overall performance for the given dataset and *mstop* iterations. Once this process is repeated for all values of *mstop*, the model with the best overall performance is chosen as the model used to predict hawkishness for our own actor dataset. This optimal model is applied to the original test set to produce a final set of statistics regarding the model's out-of-sample performance.

Below, we report some descriptive statistics on the final boosted linear model used to predict actor hawkishness.

4.3.1 Estimated Effects of Predictors

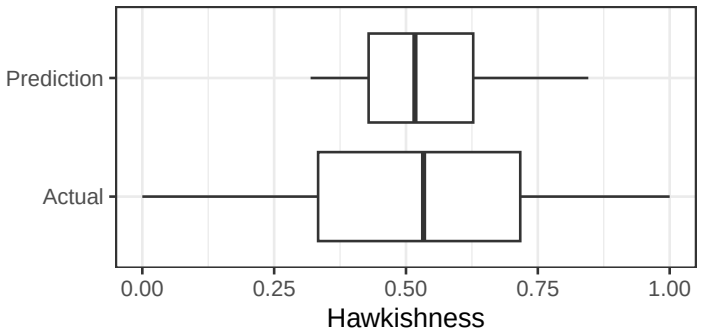
Table A5 reports the distribution of estimated coefficients across the 1,000 bootstrapped models of hawkishness. Recall that these coefficients reflect the estimated impact of each of these predictors on the hawkishness of a FPLP survey respondent. The table shows that several forms of higher

education, as well as being a Democrat (in 1976) are consistently associated with lower levels of hawkishness. Meanwhile, being a military officer or Republican in (in 1976) are tied to being more hawkish.

Table A5: Distribution of estimated coefficients across 1,000 bootstrapped models of hawkishness.

Predictor	**2.5%ile**	**Mean**	**97.5%ile**
Born 1910-1919	-0.048	-0.021	-0.002
Born 1920-1929	-0.033	-0.009	0.012
Born 1930-1939	-0.031	-0.008	0.007
Born on/after 1940	0.005	0.024	0.047
Male	-0.026	0.003	0.030
College Graduate	-0.040	-0.012	0.021
Law Degree	-0.057	-0.024	0.008
MA	-0.097	-0.062	-0.025
MBA	-0.069	-0.014	0.043
MD	-0.117	-0.054	-0.004
PhD	-0.124	-0.093	-0.058
WWII	-0.019	0.002	0.022
Korean War	-0.004	0.021	0.049
Vietnam War	0.002	0.030	0.070
Foreign Service Officer	-0.081	-0.007	0.051
Military Officer	0.110	0.153	0.191
Military Service	-0.018	0.005	0.024
Democratic	-0.116	-0.093	-0.072
Republican	0.070	0.091	0.111

Figure A4: Actual versus predicted values of hawkishness for FPLP respondents.



4.3.2 Prediction Error

To what extent do the bootstrapped models of hawkishness accurately predict an individual's hawkishness? We assess this by analyzing the models' out-of-sample performance. Each supervised model is trained on a random bootstrapped sample of 70% of the FPLP dataset; out-of-sample performance is then measured by applying this model to the held-out 30% of the survey data, for which actual levels of hawkishness (measured using the 15 questions listed below) are known. Any and all predictions of hawkishness for each FPLP survey respondent across the 1,000 bootstrapped iterations are averaged and then compared to the actual value of hawkishness.

Figure A4 illustrates the distribution of predicted hawkishness for FPLP survey respondents compared to the distribution of their actual hawkishness according to their responses. It is clear that the supervised model does not generate predictions that are as extreme as reality. Indeed, the predictions appear to be a compressed version of actual values.

Figure A5 provides more context by plotting each respondent's hawkishness measure to their prediction error. The model makes larger errors when individuals have more extreme measures of dovishness or hawkishness. Ardent hawks are underestimated, while ardent doves are overestimated. The Pearson correlation between respondents' predicted and actual hawkishness is 0.53. This value is quite strong given the task at hand and indicates that, despite the compression of the hawkishness scale, predicted values generally align in their ordering compared to actual ones.

Figure A5: Actual hawkishness versus prediction error for FPLP respondents.

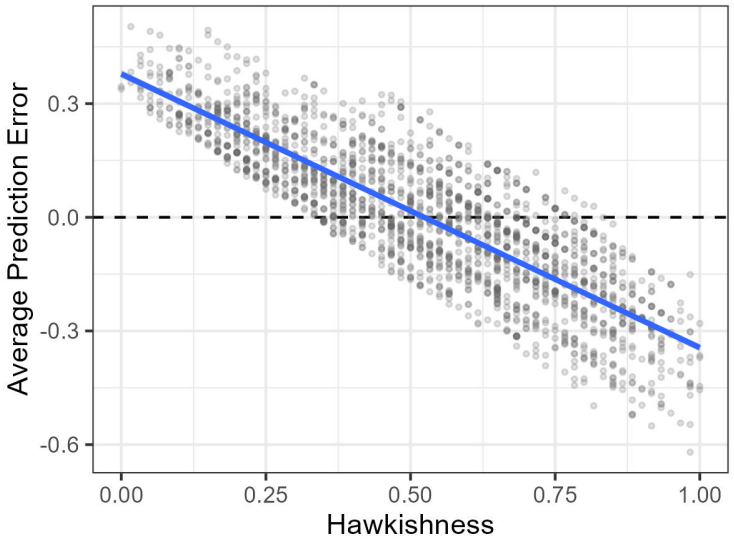


Table A6: Out-of-sample performance metrics for two linear models.

Model	RMSE	R^2	MAE	MAESD
Boosted linear	0.1993	0.2806	0.1629	0.1149
OLS	0.1994	0.2800	0.1627	0.1153

4.4 Comparison to OLS Model

We impute measures of hawkishness using a boosted linear regression model. It is necessary to use a linear model, as this is required to make our adjustments to the estimated coefficients for hawkishness according to party (using data from Jeong 2018). However, one may wonder how our boosted linear regression model compares to a more straightforward ordinary least squares (OLS) model. Table A6 reports metrics regarding out-of-sample performance for the two linear models. The models are extremely similar in performance. However, we opt to use the boosted linear model, as it features barely lower RMSE, slightly higher R^2, and mildly smaller standard errors for mean absolute error.

Appendix §5.4 shows that our main findings are unaffected by whether we use hawkishness measures based on the boosted linear regression model or the OLS model.

4.5 Comparison of Decision-Makers and FPLP Respondents

Table A7 provides a comparison of the mean values of all variables used to predict individual-level hawkishness. Note that all variables listed in the table are binary. In contrast to the average FPLP survey respondent, our set of decision-makers tend to be older, more likely to have gone to law school, less likely to have earned a PhD, more likely to be a foreign service officer, less likely to be a military officer, and somewhat less likely to be a Democrat. Imbalance between these two datasets is not material to whether the predictive model is effectively fitting the training data. As such, most of these differences should have no bearing on the effectiveness of our predictive model or the nature of the actual predictions it generates for our own actor dataset.

One exception to this rule is if our training data does not offer enough information to estimate a predictor's effect – and that predictor is relevant in the dataset for which we want to generate predictions. The FPLP data feature very few foreign service officers (FSOs) and individuals born before 1910, while our actor dataset has a substantial number of people who fit those two categories. The scant number of FSOs and those born prior to 1910 may impact our ability to find a consistent or systematic estimate of either trait on an individual's hawkishness, since there are limited data to train a model on these features. The lack of statistical significance for FSOs in Table A5 may perhaps reflect this.[2] This may mean that the effect of being an FSO or being born before 1910 is not fully accounted for in our predicted hawkishness measures. If we assume that FSOs would tend to be less hawkish (much as we suspect and find that military officers are systematically more hawkish), then this means that our hawkishness variable could be slightly biased upward for actors working in the State Department.[3] We have no strong ex ante expectations about whether being born before 1910 should be associated with being hawkish or dovish, so we are less certain about what impact this may have on our predictions for actors born in this earlier period. However, the relatively mild or non-existent impacts associated with birth decade in Table A5 suggest that the lack of data on pre-1910 individuals would unduly impact our measures.

[2] "Born before 1910" was used as a baseline category for age and is therefore not reported on Table A5.

[3] Because our model does not have much data about FSOs, this raises the possibility that the estimated effect of -0.007 is not only imprecise but also too small in magnitude relative to a hypothetical scenario where more FSOs had responded to the FPLP.

Table A7: Comparison of mean values across our NSC decision-maker dataset and the FPLP survey.

Predictor	**Actor Mean**	**FPLP Mean**
Born before 1910	0.424	0.028
Born 1910-1919	0.242	0.249
Born 1920-1929	0.192	0.115
Born 1930-1939	0.077	0.204
Born on/after 1940	0.064	0.403
Male	0.990	0.903
College Graduate	0.249	0.365
Law Degree	0.284	0.084
MA	0.109	0.162
MBA	0.004	0.022
MD	0.001	0.026
PhD	0.130	0.261
WWII	0.325	0.331
Korean War	0.047	0.085
Vietnam War	0.032	0.236
Foreign Service Officer	0.188	0.008
Military Officer	0.106	0.209
Military Service	0.583	0.671
Democratic	0.237	0.370
Republican	0.288	0.275

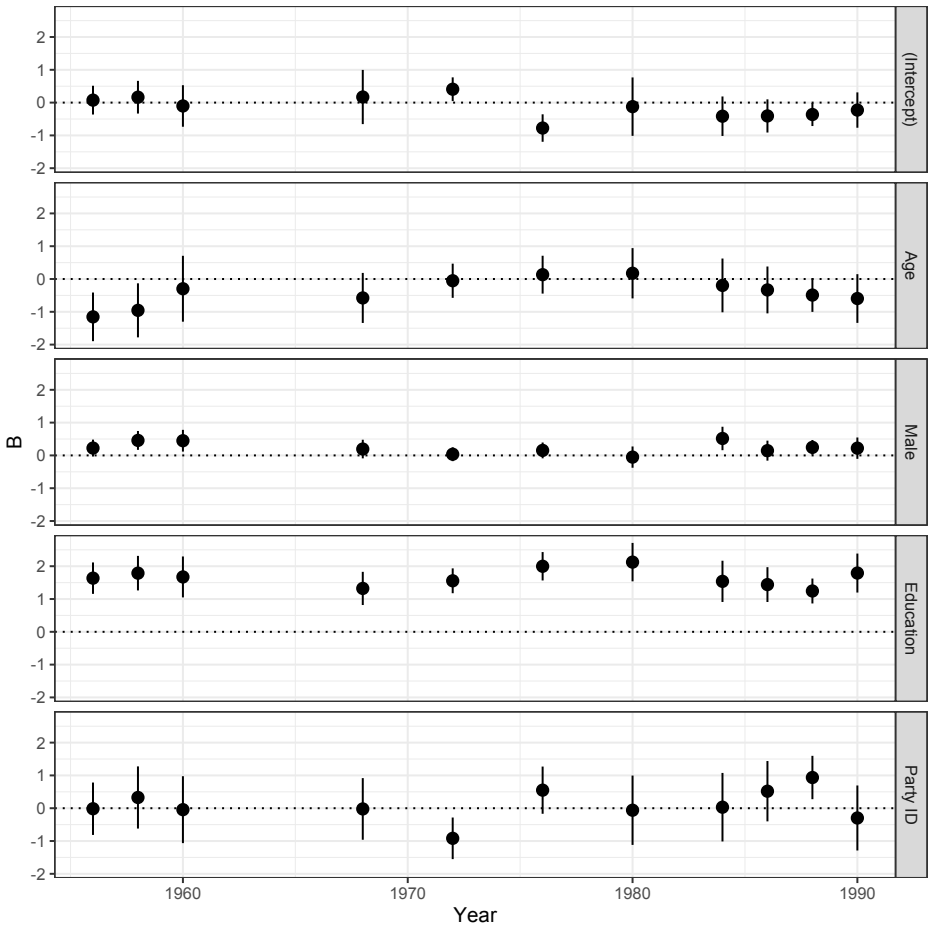
4.6 Hawkishness and Other Biographical Features

In our predictive model of hawkishness, we introduce an adjustment for party affiliation, as existing scholarship notes that the relationship between party and hawkishness changed over the course of the Cold War. This may raise the question of whether other adjustments might be necessary if there were similar reasons why the relationship between hawkishness and the other biographical characteristics—such as age, gender, education, or organizational affiliation—might changed over the course of time. While we have strong theoretical reasons to expect the association between partisanship and foreign policy preferences changes markedly over the Cold War (even going so far as to reverse in sign), we do not have similarly strong theoretical expectations about time trends in the effects of these other biographical characteristics.

In the case of age and gender, we note that much of the existing literature on underpinning *why* older, male individuals skew more hawkish emphasize time-invariant factors such as neurobiological processes (e.g., McDermott et al., 2009; McDermott, 2014). In the case of gender, the point is also somewhat moot given the fact that the advisers in our dataset were almost exclusively male for the entire period. A more plausible scenario might be that the relationship between education and hawkishness changed over time, but we are again unaware of any existing study that has directly demonstrated this empirically.

We can also offer some tentative evidence that shows that the effects of age, gender, and education on foreign policy predispositions do not vary dramatically in our time span. The American National Election Studies, which began fielding on the mass public early in the Cold War, do not have consistent measures of hawkishness but do include a measure of another foreign policy disposition—isolationism/internationalism (e.g. Klingberg, 1952; Braumoeller, 2010; Kertzer, 2013)—which was first included in 1956, and fielded periodically throughout the duration of the Cold War. We estimate a separate logistic regression for each year, regressing respondents' expressed internationalism on a series of demographic characteristics, including respondents' age, education, and gender — the coefficient estimates and 95% confidence intervals for which are presented in Figure A6. As Figure A6 shows, at least with regard to this foreign policy orientation, there is little evidence that these demographic characteristics have effects that substantially differ in magnitude over time. While caution should be taken in extrapolating from one foreign pol-

Figure A6: Coefficient estimates for demographic correlates of isolationism/internationalism fairly stable over time



Note: Coefficient estimates and 95% CIs from logistic regressions estimating the association between isolationism/internationalism and a set of demographic covariates in American National Election Survey data from 1956-1990.

icy orientation to another, the temporal dynamics shown here are consistent with the theoretical intuitions expressed above.

A final possibility is that organizational affiliation, particularly military status, might be associated with more hawkish attitudes during some periods, but not in others. Two points are worth noting, however. First, the bulk of the literature on civil-military relations tends to emphasize that hawkish dispositions tend to cluster in military organizations across a wide variety of contexts (Posen, 1984; Snyder, 1989). Surveys of military organizations from the 1970s (Etheredge, 1978), 1990s (Feaver and Gelpi, 2011), and 2000s (Jost, Meshkin and Schub, 2022) tend to yield similar findings. Second, the existing literature emphasizes that there are important cohort effects to serving in particular military conflicts (e.g., World War II as opposed to the Vietnam War). One of the benefits of the FPLP, however, is that it includes these characteristics—such that we are accounting for the cohort effects of different generations of military officers within our sample.

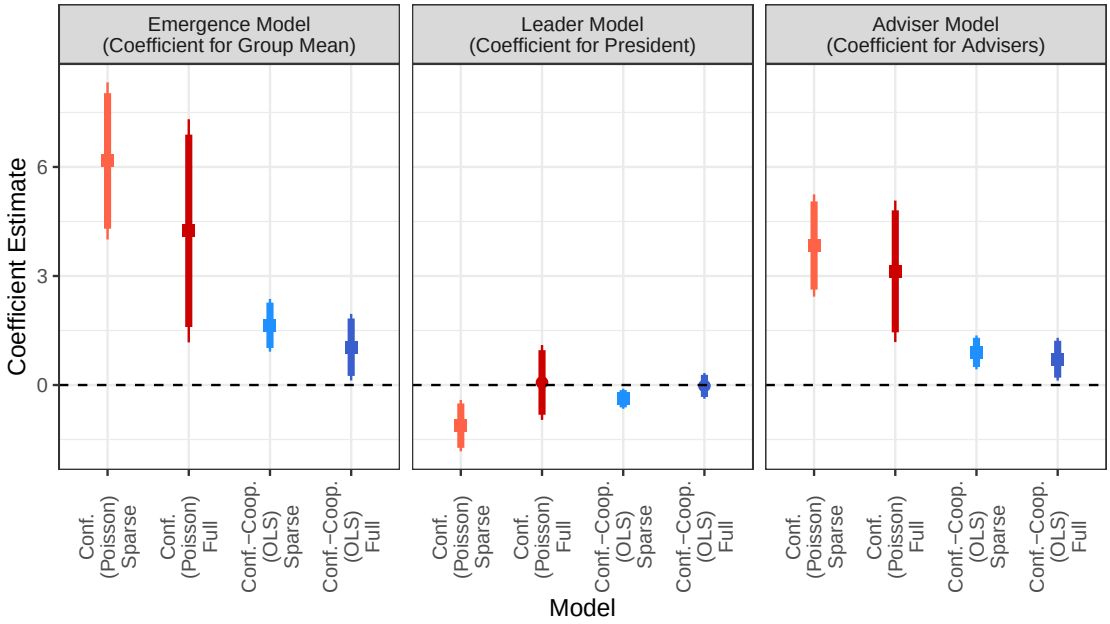
5 Full Results and Robustness Checks

In this section, we provide full results from our primary analysis, and then we conduct a series of robustness checks that uphold and further contextualize the validity of our main findings.

5.1 Full Results

Table 5 in the main text reports the results of fully-specified models. Tables A8 and A9 below supply results for these models, as well as bivariate specifications. Figure A7 illustrates the magnitude and statistical significance of the central hawkishness measures. All results in the remainder of this document will report both the sparse and full versions of models.

Figure A7: Summary of Three Models of Trait Aggregation



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. Square points indicate 95% statistical significance.

Table A8: Effect of Mean Participant Hawkishness and President's Hawkishness on Foreign Policy Decisions

	Emergence Model				*Leader Model*			
	Conflictual Decisions		Conflictual − Cooperative		Conflictual Decisions		Conflictual − Cooperative	
	Poisson		*OLS*		*Poisson*		*OLS*	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Mean Hawkishness	6.166***	4.243***	1.642***	1.040**				
	(1.103)	(1.566)	(0.369)	(0.467)				
President's Hawkishness					−1.118***	0.071	−0.377***	−0.024
					(0.361)	(0.526)	(0.141)	(0.180)
No. of Attendees		0.069***		0.038***		0.075***		0.042***
		(0.017)		(0.007)		(0.016)		(0.006)
Proportion - Defense		2.081***		0.297**		1.328***		0.198*
		(0.364)		(0.125)		(0.327)		(0.106)
Proportion - Intelligence		1.580***		0.260		1.310**		0.202
		(0.604)		(0.167)		(0.575)		(0.148)
Proportion - Military		0.936*		0.192		1.333***		0.327**
		(0.561)		(0.185)		(0.452)		(0.137)
Proportion - State		0.242		−0.016		0.047		0.031
		(0.359)		(0.090)		(0.350)		(0.078)
Diplomatic Experience		0.123***		0.004		0.085*		0.003
		(0.047)		(0.016)		(0.045)		(0.013)
Intelligence Experience		−0.122**		−0.044**		−0.089*		−0.048***
		(0.053)		(0.018)		(0.046)		(0.016)
Military Experience		0.041		−0.050*		0.133		−0.010
		(0.092)		(0.026)		(0.082)		(0.021)
5-Year MID Challenges		−0.267		−0.110		0.083		−0.047
		(0.232)		(0.078)		(0.148)		(0.041)
US CINC		6.427***		2.831***		2.759***		1.026***
		(2.356)		(0.919)		(1.069)		(0.318)
Formal	1.473***	0.622***	0.357***	0.139***	1.782***	0.783***	0.429***	0.205***
	(0.091)	(0.136)	(0.034)	(0.051)	(0.083)	(0.129)	(0.031)	(0.044)
Constant	−5.557***	−7.219***	−0.827***	−1.320***	−1.695***	−4.314***	0.227***	−0.298
	(0.646)	(1.270)	(0.218)	(0.420)	(0.186)	(0.669)	(0.070)	(0.193)
Administration FEs	✓	✓	✓	✓				
Agenda Items	✓	✓	✓	✓	✓	✓	✓	✓
Observations	2,685	2,685	2,685	2,685	2,685	2,685	2,685	2,685

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

Table A9: Effect of Adviser Hawkishness on Foreign Policy Decisions

	Adviser Model				*Advisers + Admin. FEs*			
	Conflictual Decisions		Conflictual − Cooperative		Conflictual Decisions		Conflictual − Cooperative	
	Poisson		*OLS*		*Poisson*		*OLS*	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Advisers' Hawkishness (Acts)	3.839***	3.128***	0.898***	0.710**	4.401***	2.893**	0.948***	0.708**
	(0.718)	(0.993)	(0.238)	(0.300)	(0.841)	(1.198)	(0.277)	(0.345)
President's Hawkishness	−1.721***	−0.835	−0.545***	−0.138				
	(0.371)	(0.600)	(0.148)	(0.224)				
No. of Attendees		0.090***		0.042***		0.078***		0.041***
		(0.017)		(0.007)		(0.017)		(0.007)
Proportion - Defense		1.592***		0.182		2.052***		0.293**
		(0.340)		(0.121)		(0.367)		(0.127)
Proportion - Intelligence		1.597***		0.247		1.636***		0.268
		(0.587)		(0.168)		(0.606)		(0.170)
Proportion - Military		0.932**		0.193		1.173**		0.237
		(0.465)		(0.162)		(0.526)		(0.180)
Proportion - State		0.278		−0.011		0.233		−0.013
		(0.360)		(0.090)		(0.364)		(0.093)
Diplomatic Experience		0.071		−0.009		0.119**		0.003
		(0.045)		(0.015)		(0.047)		(0.016)
Intelligence Experience		−0.112**		−0.039**		−0.120**		−0.043**
		(0.047)		(0.018)		(0.053)		(0.019)
Military Experience		0.056		−0.045*		−0.001		−0.062**
		(0.085)		(0.025)		(0.092)		(0.027)
5-Year MID Challenges		0.036		0.001		−0.260		−0.116
		(0.150)		(0.046)		(0.233)		(0.080)
US CINC		2.072*		1.337***		6.769***		2.909***
		(1.099)		(0.366)		(2.360)		(0.933)
Formal	1.660***	0.763***	0.403***	0.148***	1.553***	0.643***	0.385***	0.145***
	(0.085)	(0.130)	(0.032)	(0.050)	(0.089)	(0.137)	(0.034)	(0.052)
Constant	−3.356***	−5.092***	−0.144	−0.664***	−4.570***	−6.474***	−0.420**	−1.109***
	(0.368)	(0.727)	(0.123)	(0.229)	(0.506)	(1.153)	(0.165)	(0.384)
Administration FEs					✓	✓	✓	✓
Agenda Items	✓	✓	✓	✓	✓	✓	✓	✓
Observations	2,650	2,650	2,650	2,650	2,650	2,650	2,650	2,650

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

5.2 Disaggregating Results by Meeting Type

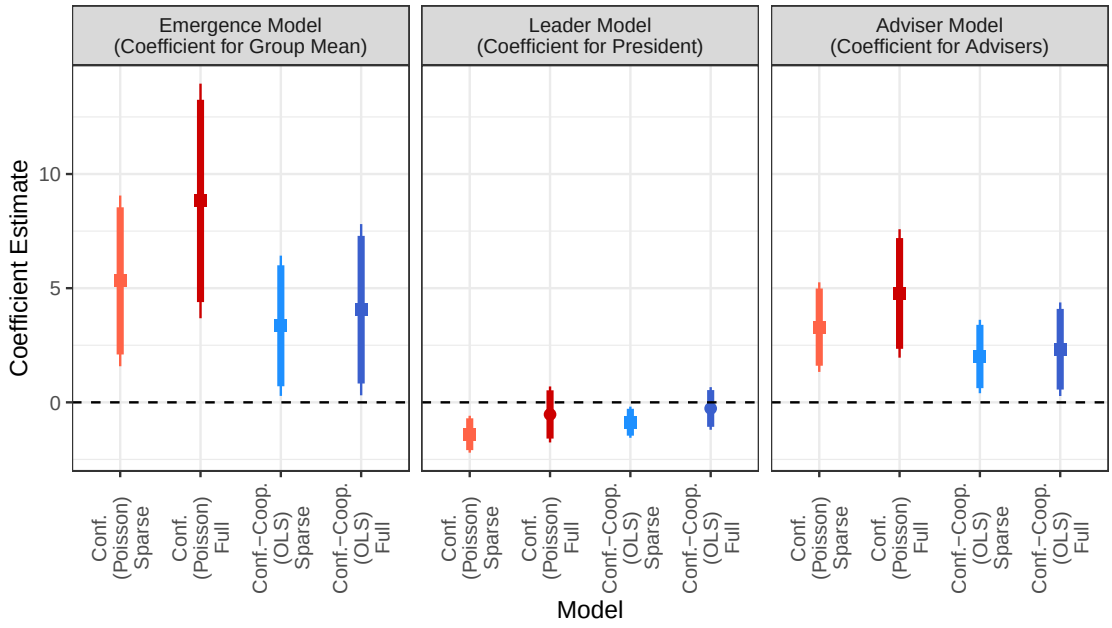
As noted, our main analysis includes both formal NSC meetings as well as informal gatherings, as national security decisions were made in both contexts. Yet we may also believe that differences exist in the dynamics of these separate settings. We can explore potential differences by replicating the previous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

Tables A10 and A11 present full statistical results from regressions that only use data from informal meetings. Figure A8 displays the corresponding coefficient plots. Tables A12 and A13 present full statistical results from regressions that only use data from informal meetings. Figure A9 displays the corresponding coefficient plots.

The results make two distinctions clear. First, the influence of advisers is stronger and more statistically significant in formal meetings compared to informal gatherings. Second, we find no consistent evidence of a statistically significant relationship between the hawkishness of leaders and conflictual decisions in informal meetings. This contrasts with the negative coefficients for several models associated with formal meetings.

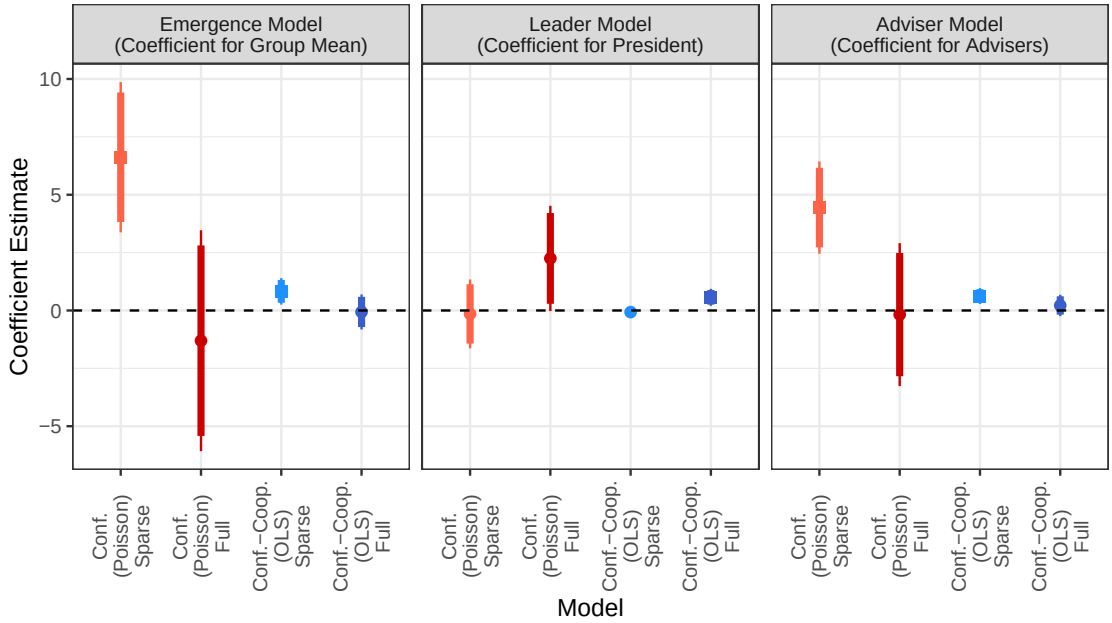
These findings enhance our understanding of foreign policy decision-making in critical ways. First, we find that advisers matter more in foreign policy decision-making than work that focuses only on leaders at the expense of advisers might suggest. We also find very little evidence for—and considerable evidence against—the emergence model, 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 A8: 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. Square points indicate 95% statistical significance.

Figure A9: 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. Square points indicate 95% statistical significance.

Table A10: Effect of Mean Participant Hawkishness and President's Hawkishness on Foreign Policy Decisions, Using Formal Meetings

	Emergence Model				*Leader Model*			
	Conflictual Decisions		Conflictual − Cooperative		Conflictual Decisions		Conflictual − Cooperative	
	Poisson		*OLS*		*Poisson*		*OLS*	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Mean Hawkishness	5.319***	8.821***	3.354**	4.058**				
	(1.907)	(2.621)	(1.567)	(1.913)				
President's Hawkishness					−1.396***	−0.530	−0.871**	−0.267
					(0.412)	(0.625)	(0.348)	(0.476)
No. of Attendees		0.074***		0.056***		0.087***		0.066***
		(0.022)		(0.018)		(0.020)		(0.018)
Proportion - Defense		1.612***		0.902**		0.399		0.314
		(0.542)		(0.416)		(0.480)		(0.387)
Proportion - Intelligence		0.544		0.583		0.212		0.570
		(1.019)		(0.665)		(0.966)		(0.645)
Proportion - Military		−0.455		0.168		−0.041		0.371
		(0.962)		(0.750)		(0.812)		(0.664)
Proportion - State		−0.418		0.243		−0.758		0.007
		(0.660)		(0.455)		(0.641)		(0.444)
Diplomatic Experience		0.098		0.014		0.061		−0.001
		(0.062)		(0.043)		(0.059)		(0.042)
Intelligence Experience		−0.070		−0.005		−0.094*		−0.031
		(0.075)		(0.051)		(0.056)		(0.042)
Military Experience		−0.119		−0.114		0.061		−0.047
		(0.144)		(0.104)		(0.130)		(0.094)
5-Year MID Challenges		−0.474		−0.263		0.060		0.034
		(0.296)		(0.214)		(0.207)		(0.141)
US CINC		8.172***		6.175***		3.035**		2.360**
		(2.890)		(2.216)		(1.391)		(1.015)
Constant	−3.686***	−8.477***	−1.467	−3.849***	0.220	−2.433**	0.902***	−0.631
	(1.107)	(1.930)	(0.896)	(1.419)	(0.203)	(1.011)	(0.177)	(0.701)
Administration FEs	✓	✓	✓	✓				
Agenda Items	✓	✓	✓	✓	✓	✓	✓	✓
Observations	791	791	791	791	791	791	791	791

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

Table A11: Effect of Adviser Hawkishness on Foreign Policy Decisions, Using Formal Meetings

	Adviser Model				*Advisers + Admin. FEs*			
	Conflictual Decisions		Conflictual − Cooperative		Conflictual Decisions		Conflictual − Cooperative	
	Poisson		*OLS*		*Poisson*		*OLS*	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Advisers' Hawkishness (Acts)	3.295***	4.769***	2.008**	2.327**	5.698***	6.597***	3.371***	3.408**
	(1.001)	(1.435)	(0.820)	(1.045)	(1.473)	(1.859)	(1.226)	(1.397)
President's Hawkishness	−1.914***	−2.044***	−1.231***	−1.123*				
	(0.434)	(0.783)	(0.373)	(0.604)				
No. of Attendees		0.106***		0.074***		0.083***		0.061***
		(0.021)		(0.018)		(0.022)		(0.018)
Proportion - Defense		0.754		0.476		1.584***		0.865**
		(0.498)		(0.401)		(0.547)		(0.421)
Proportion - Intelligence		0.621		0.715		0.502		0.586
		(0.977)		(0.655)		(1.039)		(0.674)
Proportion - Military		−0.237		0.270		0.289		0.469
		(0.814)		(0.665)		(0.889)		(0.699)
Proportion - State		−0.288		0.166		−0.339		0.206
		(0.653)		(0.468)		(0.676)		(0.475)
Diplomatic Experience		0.048		−0.011		0.082		0.004
		(0.059)		(0.043)		(0.063)		(0.044)
Intelligence Experience		−0.150**		−0.055		−0.084		−0.012
		(0.060)		(0.043)		(0.075)		(0.051)
Military Experience		−0.072		−0.115		−0.190		−0.164
		(0.136)		(0.101)		(0.146)		(0.107)
5-Year MID Challenges		−0.069		−0.049		−0.412		−0.267
		(0.213)		(0.145)		(0.298)		(0.215)
US CINC		1.891		1.729		8.739***		6.402***
		(1.463)		(1.052)		(2.901)		(2.221)
Constant	−1.308**	−3.369***	0.001	−0.919	−3.906***	−7.258***	−1.470**	−3.355***
	(0.510)	(1.094)	(0.415)	(0.750)	(0.863)	(1.679)	(0.702)	(1.239)
Administration FEs					✓	✓	✓	✓
Agenda Items	✓	✓	✓	✓	✓	✓	✓	✓
Observations	787	787	787	787	787	787	787	787

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

Table A12: Effect of Mean Participant Hawkishness and President's Hawkishness on Foreign Policy Decisions, Using Informal Meetings

	Emergence Model				*Leader Model*			
	Conflictual Decisions		Conflictual − Cooperative		Conflictual Decisions		Conflictual − Cooperative	
	Poisson		*OLS*		*Poisson*		*OLS*	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Mean Hawkishness	6.620***	−1.305	0.826***	−0.064				
	(1.655)	(2.435)	(0.291)	(0.386)				
President's Hawkishness					−0.149	2.248*	−0.070	0.568***
					(0.759)	(1.160)	(0.118)	(0.187)
No. of Attendees		0.028		0.011*		0.001		0.007
		(0.039)		(0.007)		(0.036)		(0.006)
Proportion - Defense		2.649***		0.271***		2.590***		0.260***
		(0.529)		(0.097)		(0.510)		(0.095)
Proportion - Intelligence		2.739***		0.299**		2.766***		0.254**
		(0.814)		(0.125)		(0.779)		(0.123)
Proportion - Military		2.141***		0.328**		1.515**		0.276**
		(0.774)		(0.138)		(0.591)		(0.111)
Proportion - State		0.646		−0.009		0.570		−0.045
		(0.516)		(0.067)		(0.487)		(0.064)
Diplomatic Experience		0.073		−0.003		0.089		0.004
		(0.083)		(0.013)		(0.078)		(0.012)
Intelligence Experience		−0.071		−0.009		−0.104		−0.008
		(0.097)		(0.017)		(0.092)		(0.017)
Military Experience		0.231		0.009		0.366***		0.017
		(0.147)		(0.021)		(0.132)		(0.018)
5-Year MID Challenges		0.305		0.075		−0.127		0.065
		(0.456)		(0.067)		(0.261)		(0.041)
US CINC		−5.545		−1.212		0.892		0.733**
		(5.296)		(0.854)		(1.918)		(0.319)
Constant	−5.506***	−1.609	−0.294	0.314	−2.161***	−5.362***	0.078	−0.688***
	(1.056)	(2.454)	(0.185)	(0.392)	(0.374)	(1.605)	(0.059)	(0.236)
Administration FEs	✓	✓	✓	✓				
Agenda Items	✓	✓	✓	✓	✓	✓	✓	✓
Observations	1,894	1,894	1,894	1,894	1,894	1,894	1,894	1,894

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

Table A13: Effect of Adviser Hawkishness on Foreign Policy Decisions, Using Informal Meetings

	Adviser Model				*Advisers + Admin. FEs*			
	Conflictual Decisions		Conflictual − Cooperative		Conflictual Decisions		Conflictual − Cooperative	
	Poisson		*OLS*		*Poisson*		*OLS*	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Advisers' Hawkishness (Acts)	4.442***	−0.179	0.624***	0.217	3.523***	−2.239	0.525***	−0.001
	(1.018)	(1.575)	(0.175)	(0.235)	(1.094)	(1.742)	(0.196)	(0.256)
President's Hawkishness	−0.830	2.186*	−0.180	0.505**				
	(0.756)	(1.208)	(0.124)	(0.201)				
No. of Attendees		−0.00001		0.009		0.022		0.012*
		(0.039)		(0.007)		(0.039)		(0.007)
Proportion - Defense		2.536***		0.268***		2.544***		0.269***
		(0.522)		(0.097)		(0.541)		(0.099)
Proportion - Intelligence		2.772***		0.271**		2.682***		0.307**
		(0.788)		(0.126)		(0.821)		(0.127)
Proportion - Military		1.513**		0.231*		2.383***		0.314**
		(0.649)		(0.121)		(0.748)		(0.134)
Proportion - State		0.523		−0.033		0.530		−0.008
		(0.509)		(0.067)		(0.529)		(0.070)
Diplomatic Experience		0.092		0.002		0.086		−0.004
		(0.080)		(0.013)		(0.083)		(0.013)
Intelligence Experience		−0.107		−0.008		−0.073		−0.009
		(0.092)		(0.017)		(0.097)		(0.017)
Military Experience		0.371***		0.011		0.250*		0.009
		(0.139)		(0.019)		(0.148)		(0.022)
5-Year MID Challenges		−0.154		0.068		0.336		0.076
		(0.265)		(0.042)		(0.457)		(0.068)
US CINC		1.062		0.702**		−5.585		−1.251
		(2.007)		(0.339)		(5.305)		(0.876)
Constant	−4.098***	−5.217***	−0.183*	−0.750***	−3.442***	−1.154	−0.081	0.297
	(0.592)	(1.699)	(0.095)	(0.248)	(0.698)	(2.260)	(0.126)	(0.359)
Administration FEs					✓	✓	✓	✓
Agenda Items	✓	✓	✓	✓	✓	✓	✓	✓
Observations	1,863	1,863	1,863	1,863	1,863	1,863	1,863	1,863

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

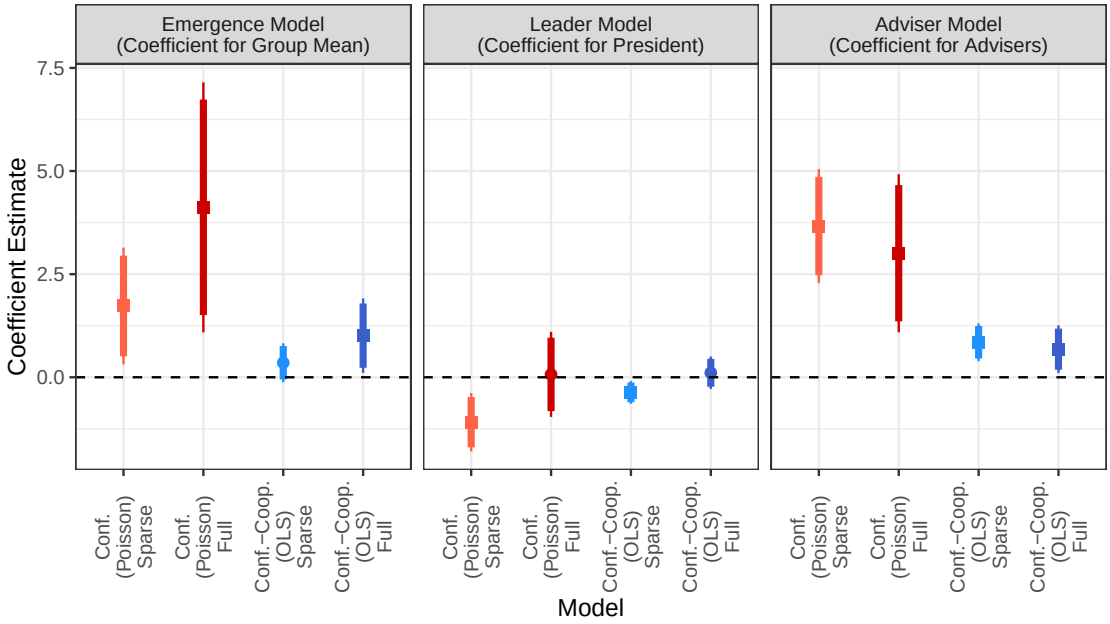
5.3 Propagating Uncertainty of Hawkishness Measures

The hawkishness measures used in our core analysis are constructed through a bootstrap process, described in subsection 2.2 of the main text. The bootstrapping process producees 1,000 predictions of each actor's hawkishness based on different sets of training data. Our paper calculates the mean value of hawkishness for each observation across these 1,000 observations, and this mean is our key variable. While this approach is relatively efficient and does account for potential variation across all bootstrap iterations, it does not fully propagate the uncertainty inherent to our estimates through our primary analysis. To what extent could the wider array of predicted measures of hawkishness from our bootstrapping process potentially impact our findings?

To address this question, we go back to our 1,000 bootstrap iterations. For each iteration, we take the predicted measures of hawkishness for each of our actor-years and construct meeting-level data necessary for our main analysis. We then run our series of Poisson and OLS models on this meeting-level data, producing a coefficient estimate and associated standard error for each variable. This process is performed 1,000 times (once for each dataset produced from each bootstrap iteration).

Tables A14 and A15 report the average coefficient estimates and standard errors across all 1,000 iterations of each model; Figure A10 illustrates the coefficients for relevant measures of hawkishness. The vast majority of our primary findings, and particularly those related to the adviser model, remain unchanged even when performing a more computationally-intensive form of analysis that explicitly propagates the uncertainty of our hawkishness measures.

Figure A10: Summary of Three Models of Trait Aggregation, Propagating Uncertainty from Bootstrapping



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. Square points indicate 95% statistical significance.

Table A14: Effect of Mean Participant Hawkishness and President's Hawkishness on Foreign Policy Decisions, Propagating Uncertainty from Bootstrapping

	Emergence Model				*Leader Model*			
	Conflictual Decisions		Conflictual − Cooperative		Conflictual Decisions		Conflictual − Cooperative	
	Poisson		*OLS*		*Poisson*		*OLS*	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Mean Hawkishness	1.728**	4.120***	0.352	1.009**				
	(0.722)	(1.546)	(0.240)	(0.461)				
President's Hawkishness					−1.090***	0.069	−0.366***	0.109
					(0.361)	(0.526)	(0.141)	(0.200)
No. of Attendees		0.069***		0.038***		0.075***		0.038***
		(0.017)		(0.007)		(0.016)		(0.007)
Proportion - Defense		2.076***		0.297**		1.329***		0.137
		(0.364)		(0.125)		(0.327)		(0.118)
Proportion - Intelligence		1.577***		0.260		1.310**		0.196
		(0.603)		(0.167)		(0.575)		(0.165)
Proportion - Military		0.959*		0.198		1.333***		0.310**
		(0.560)		(0.185)		(0.452)		(0.153)
Proportion - State		0.236		−0.017		0.047		−0.047
		(0.361)		(0.091)		(0.350)		(0.086)
Diplomatic Experience		0.124***		0.004		0.085*		−0.004
		(0.047)		(0.016)		(0.045)		(0.015)
Intelligence Experience		−0.121**		−0.044**		−0.089*		−0.037**
		(0.053)		(0.018)		(0.046)		(0.018)
Military Experience		0.041		−0.050*		0.133		−0.024
		(0.092)		(0.026)		(0.082)		(0.023)
5-Year MID Challenges		−0.265		−0.109		0.081		0.005
		(0.232)		(0.078)		(0.148)		(0.046)
US CINC		6.439***		2.831***		2.755**		1.481***
		(2.356)		(0.919)		(1.072)		(0.355)
Formal	1.713***	0.622***	0.412***	0.139***	1.780***	0.783***	0.429***	0.144***
	(0.087)	(0.136)	(0.032)	(0.051)	(0.083)	(0.129)	(0.031)	(0.049)
Constant	−3.104***	−7.156***	−0.132	−1.303***	−1.709***	−4.309***	0.221***	−0.513**
	(0.373)	(1.264)	(0.121)	(0.418)	(0.186)	(0.669)	(0.070)	(0.213)
Administration FEs	✓	✓	✓	✓				
Agenda Items	✓	✓	✓	✓	✓	✓	✓	✓
Observations	2,685	2,685	2,685	2,685	2,685	2,685	2,685	2,685

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

Table A15: Effect of Adviser Hawkishness on Foreign Policy Decisions, Propagating Uncertainty from Bootstrapping

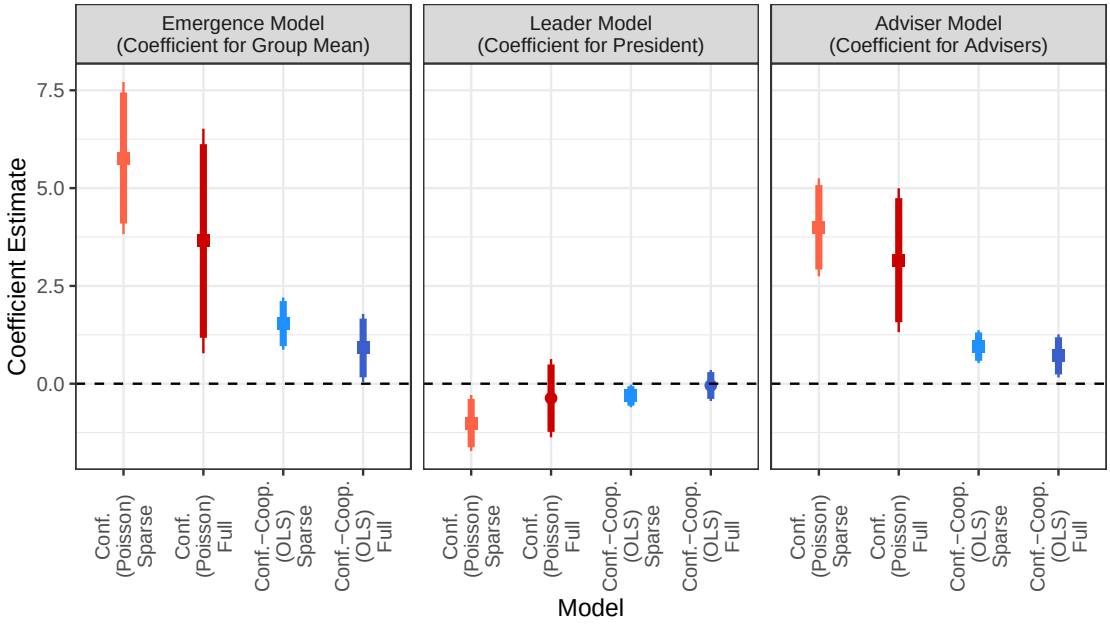
	Adviser Model				*Advisers + Admin. FEs*			
	Conflictual Decisions		Conflictual − Cooperative		Conflictual Decisions		Conflictual − Cooperative	
	Poisson		*OLS*		*Poisson*		*OLS*	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Advisers' Hawkishness (Acts)	3.667***	3.007***	0.849***	0.680**	4.196***	2.766**	0.892***	0.675**
	(0.706)	(0.977)	(0.233)	(0.295)	(0.829)	(1.174)	(0.271)	(0.337)
President's Hawkishness	−1.674***	−0.792	−0.526***	−0.130				
	(0.372)	(0.599)	(0.148)	(0.223)				
No. of Attendees		0.089***		0.042***		0.078***		0.040***
		(0.017)		(0.007)		(0.017)		(0.007)
Proportion - Defense		1.584***		0.181		2.045***		0.291**
		(0.340)		(0.121)		(0.367)		(0.127)
Proportion - Intelligence		1.588***		0.245		1.631***		0.266
		(0.587)		(0.168)		(0.606)		(0.170)
Proportion - Military		0.949**		0.198		1.199**		0.245
		(0.466)		(0.162)		(0.525)		(0.180)
Proportion - State		0.268		−0.013		0.225		−0.014
		(0.361)		(0.091)		(0.365)		(0.094)
Diplomatic Experience		0.072		−0.008		0.120**		0.003
		(0.045)		(0.015)		(0.047)		(0.016)
Intelligence Experience		−0.111**		−0.039**		−0.120**		−0.043**
		(0.047)		(0.018)		(0.053)		(0.019)
Military Experience		0.060		−0.044*		−0.000		−0.062**
		(0.085)		(0.024)		(0.092)		(0.027)
5-Year MID Challenges		0.037		0.0004		−0.258		−0.116
		(0.149)		(0.046)		(0.233)		(0.080)
US CINC		2.101*		1.345***		6.772***		2.909***
		(1.101)		(0.368)		(2.361)		(0.933)
Formal	1.664***	0.762***	0.404***	0.148***	1.555***	0.642***	0.385***	0.145***
	(0.085)	(0.130)	(0.032)	(0.050)	(0.089)	(0.137)	(0.034)	(0.052)
Constant	−3.291***	−5.065***	−0.128	−0.656***	−4.451***	−6.408***	−0.387**	−1.091***
	(0.363)	(0.726)	(0.121)	(0.229)	(0.499)	(1.146)	(0.162)	(0.382)
Administration FEs					✓	✓	✓	✓
Agenda Items	✓	✓	✓	✓	✓	✓	✓	✓
Observations	2,650	2,650	2,650	2,650	2,650	2,650	2,650	2,650

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

5.4 Using an OLS Predictive Model

In Appendix §4.4, we addressed why we chose to use a boosted linear regression model to produce hawkishness measures instead of an ordinary least squares (OLS) model. Tables A16 and A17, summarized in Figure A11, replicate our analysis using the hawkishness measures generated with an OLS model. The results are very similar to what we report in our main findings.

Figure A11: Summary of Three Models of Trait Aggregation, Using OLS 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. Square points indicate 95% statistical significance.

Table A16: Effect of Mean Participant Hawkishness and President's Hawkishness on Foreign Policy Decisions, Using OLS Supervised Learning Model

	Emergence Model				*Leader Model*			
	Conflictual Decisions		Conflictual − Cooperative		Conflictual Decisions		Conflictual − Cooperative	
	Poisson		*OLS*		*Poisson*		*OLS*	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Mean Hawkishness	5.769***	3.648**	1.537***	0.914**				
	(0.991)	(1.464)	(0.341)	(0.444)				
President's Hawkishness					−1.005***	−0.369	−0.312**	−0.046
					(0.365)	(0.510)	(0.147)	(0.203)
No. of Attendees		0.068***		0.038***		0.074***		0.038***
		(0.017)		(0.007)		(0.016)		(0.007)
Proportion - Defense		2.062***		0.296**		1.363***		0.137
		(0.363)		(0.125)		(0.328)		(0.118)
Proportion - Intelligence		1.524**		0.244		1.322**		0.190
		(0.603)		(0.167)		(0.574)		(0.165)
Proportion - Military		0.924		0.188		1.334***		0.316**
		(0.578)		(0.190)		(0.453)		(0.153)
Proportion - State		0.212		−0.018		0.028		−0.055
		(0.359)		(0.090)		(0.350)		(0.086)
Diplomatic Experience		0.123***		0.004		0.091**		−0.003
		(0.047)		(0.016)		(0.045)		(0.015)
Intelligence Experience		−0.116**		−0.042**		−0.100**		−0.038**
		(0.053)		(0.018)		(0.046)		(0.018)
Military Experience		0.042		−0.050*		0.127		−0.028
		(0.092)		(0.027)		(0.082)		(0.023)
5-Year MID Challenges		−0.271		−0.108		0.015		−0.014
		(0.232)		(0.078)		(0.149)		(0.047)
US CINC		6.226***		2.779***		2.557**		1.421***
		(2.366)		(0.920)		(1.023)		(0.339)
Formal	1.472***	0.628***	0.357***	0.140***	1.789***	0.810***	0.431***	0.155***
	(0.090)	(0.136)	(0.034)	(0.051)	(0.083)	(0.128)	(0.031)	(0.050)
Constant	−5.517***	−6.922***	−0.818***	−1.266***	−1.742***	−3.897***	0.198***	−0.372*
	(0.616)	(1.239)	(0.213)	(0.417)	(0.191)	(0.637)	(0.074)	(0.205)
Administration FEs	✓	✓	✓	✓				
Agenda Items	✓	✓	✓	✓	✓	✓	✓	✓
Observations	2,685	2,685	2,685	2,685	2,685	2,685	2,685	2,685

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

Table A17: Effect of Adviser Hawkishness on Foreign Policy Decisions, Using OLS Supervised Learning Model

	Adviser Model				*Advisers + Admin. FEs*			
	Conflictual Decisions		Conflictual − Cooperative		Conflictual Decisions		Conflictual − Cooperative	
	Poisson		*OLS*		*Poisson*		*OLS*	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Advisers' Hawkishness (Acts)	3.999***	3.157***	0.949***	0.712**	4.023***	2.620**	0.868***	0.609*
	(0.638)	(0.938)	(0.214)	(0.281)	(0.755)	(1.121)	(0.252)	(0.320)
President's Hawkishness	−1.510***	−1.212**	−0.458***	−0.282				
	(0.365)	(0.570)	(0.152)	(0.222)				
No. of Attendees		0.090***		0.042***		0.078***		0.040***
		(0.016)		(0.007)		(0.017)		(0.007)
Proportion - Defense		1.641***		0.184		2.045***		0.289**
		(0.341)		(0.121)		(0.367)		(0.127)
Proportion - Intelligence		1.557***		0.222		1.587***		0.250
		(0.583)		(0.167)		(0.604)		(0.169)
Proportion - Military		0.807*		0.170		1.130**		0.236
		(0.474)		(0.164)		(0.539)		(0.184)
Proportion - State		0.262		−0.020		0.225		−0.015
		(0.359)		(0.090)		(0.364)		(0.093)
Diplomatic Experience		0.075*		−0.008		0.119**		0.003
		(0.045)		(0.015)		(0.048)		(0.016)
Intelligence Experience		−0.121**		−0.039**		−0.118**		−0.041**
		(0.047)		(0.018)		(0.053)		(0.019)
Military Experience		0.051		−0.050**		0.002		−0.062**
		(0.085)		(0.024)		(0.092)		(0.027)
5-Year MID Challenges		−0.052		−0.023		−0.259		−0.114
		(0.150)		(0.047)		(0.233)		(0.080)
US CINC		1.546		1.227***		6.606***		2.856***
		(1.069)		(0.355)		(2.365)		(0.933)
Formal	1.651***	0.797***	0.399***	0.161***	1.560***	0.646***	0.385***	0.146***
	(0.085)	(0.130)	(0.032)	(0.050)	(0.089)	(0.137)	(0.034)	(0.052)
Constant	−3.590***	−4.645***	−0.221*	−0.514**	−4.464***	−6.347***	−0.396**	−1.056***
	(0.359)	(0.686)	(0.122)	(0.218)	(0.478)	(1.134)	(0.158)	(0.378)
Administration FEs					✓	✓	✓	✓
Agenda Items	✓	✓	✓	✓	✓	✓	✓	✓
Observations	2,650	2,650	2,650	2,650	2,650	2,650	2,650	2,650

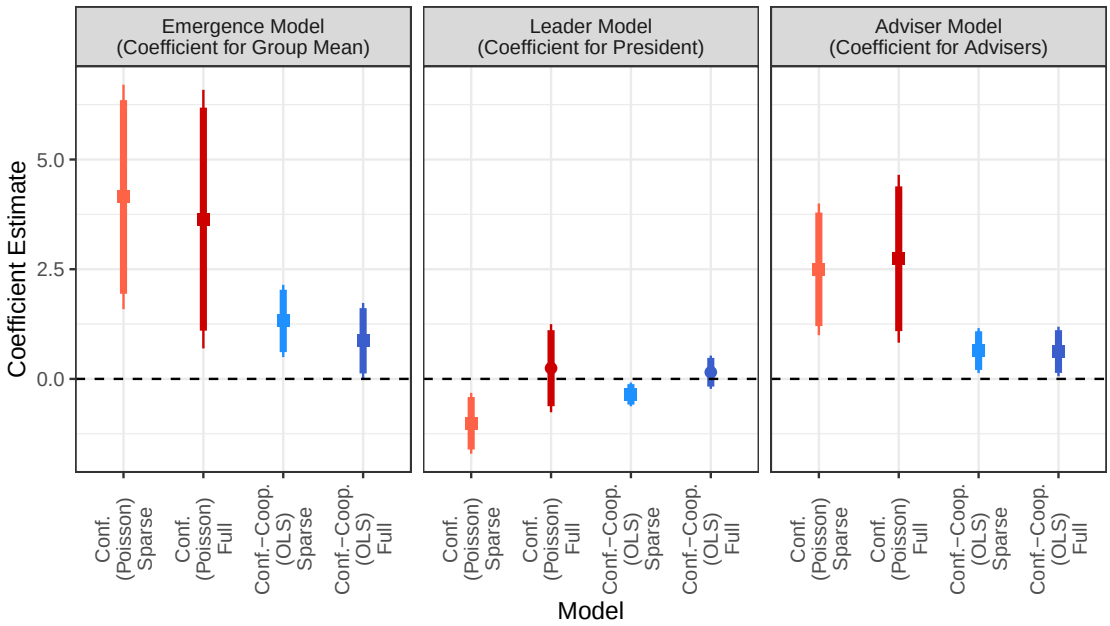
Note: $^{*}p < 0.1$; $^{**}p < 0.05$; $^{***}p < 0.01$

5.5 Removing Bureaucratic Affiliations

As an additional robustness check, 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 A18 and A19, as well as Figure A12, indicate that our main results are generally upheld with this revised measure, though some coefficients for advisers lose statistical significance in the fully specified models. The correlation between our original hawkishness measure and that produced without incorporating information on bureaucratic affiliation is 0.954. This reinforces that adviser preferences come from more than simply their institutional role.

Figure A12: 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. Square points indicate 95% statistical significance.

Table A18: Effect of Mean Participant Hawkishness and President's Hawkishness on Foreign Policy Decisions, Removing Bureaucratic Affiliation from Supervised Learning Model

	Emergence Model				*Leader Model*			
	Conflictual Decisions		Conflictual − Cooperative		Conflictual Decisions		Conflictual − Cooperative	
	Poisson		*OLS*		*Poisson*		*OLS*	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Mean Hawkishness	4.147***	3.641**	1.321***	0.869**				
	(1.305)	(1.504)	(0.420)	(0.440)				
President's Hawkishness					−1.011***	0.244	−0.353**	0.152
					(0.354)	(0.512)	(0.138)	(0.193)
No. of Attendees		0.071***		0.039***		0.075***		0.038***
		(0.017)		(0.007)		(0.016)		(0.007)
Proportion - Defense		2.047***		0.292**		1.320***		0.138
		(0.363)		(0.125)		(0.327)		(0.118)
Proportion - Intelligence		1.556***		0.254		1.305**		0.198
		(0.602)		(0.167)		(0.575)		(0.165)
Proportion - Military		1.647***		0.368**		1.332***		0.309**
		(0.476)		(0.160)		(0.451)		(0.152)
Proportion - State		0.215		−0.020		0.054		−0.044
		(0.359)		(0.090)		(0.350)		(0.086)
Diplomatic Experience		0.127***		0.005		0.083*		−0.004
		(0.047)		(0.016)		(0.045)		(0.015)
Intelligence Experience		−0.122**		−0.044**		−0.084*		−0.036**
		(0.053)		(0.018)		(0.046)		(0.018)
Military Experience		0.022		−0.055**		0.136*		−0.023
		(0.091)		(0.026)		(0.082)		(0.023)
5-Year MID Challenges		−0.255		−0.108		0.108		0.011
		(0.232)		(0.078)		(0.147)		(0.045)
US CINC		6.603***		2.870***		2.892***		1.510***
		(2.353)		(0.919)		(1.071)		(0.354)
Formal	1.526***	0.630***	0.372***	0.140***	1.783***	0.772***	0.430***	0.141***
	(0.091)	(0.136)	(0.034)	(0.051)	(0.083)	(0.129)	(0.031)	(0.049)
Constant	−4.422***	−6.927***	−0.649***	−1.233***	−1.735***	−4.489***	0.219***	−0.556***
	(0.754)	(1.256)	(0.248)	(0.414)	(0.186)	(0.666)	(0.071)	(0.211)
Administration FEs	✓	✓	✓	✓				
Agenda Items	✓	✓	✓	✓	✓	✓	✓	✓
Observations	2,685	2,685	2,685	2,685	2,685	2,685	2,685	2,685

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

Table A19: Effect of Adviser Hawkishness on Foreign Policy Decisions, Removing Bureaucratic Affiliation from Supervised Learning Model

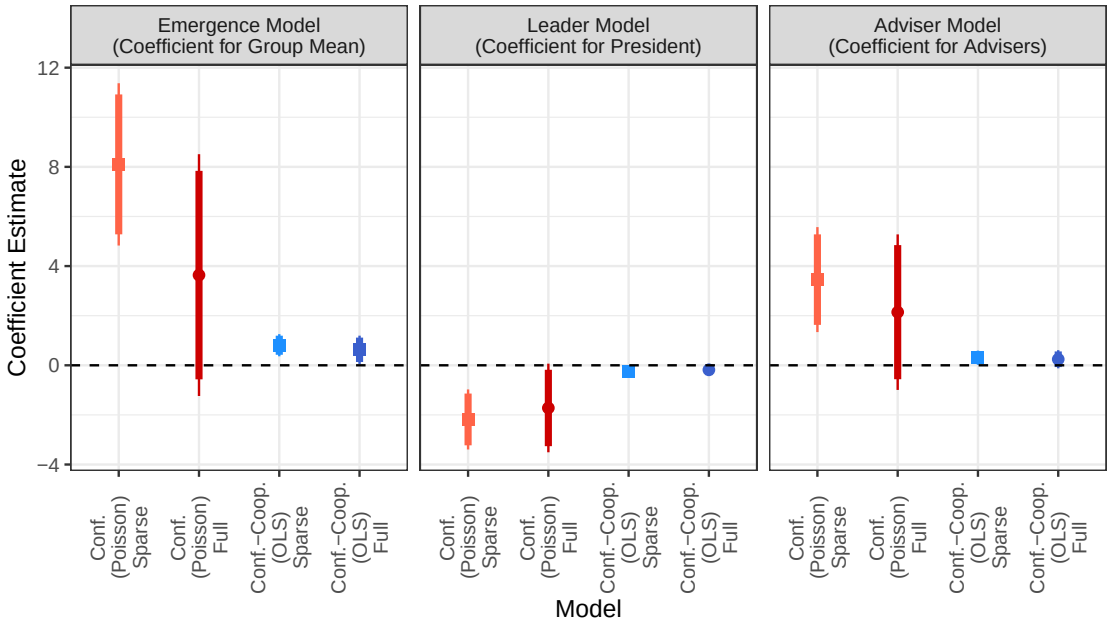
	Adviser Model				*Advisers + Admin. FEs*			
	Conflictual Decisions		Conflictual − Cooperative		Conflictual Decisions		Conflictual − Cooperative	
	Poisson		*OLS*		*Poisson*		*OLS*	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Advisers' Hawkishness (Acts)	2.497***	2.739***	0.646**	0.624**	3.223***	2.487**	0.750**	0.645**
	(0.765)	(0.976)	(0.261)	(0.288)	(1.020)	(1.191)	(0.315)	(0.326)
President's Hawkishness	−1.544***	−0.558	−0.505***	−0.077				
	(0.385)	(0.587)	(0.150)	(0.217)				
No. of Attendees		0.088***		0.042***		0.076***		0.041***
		(0.017)		(0.007)		(0.017)		(0.007)
Proportion - Defense		1.534***		0.169		2.006***		0.284**
		(0.339)		(0.120)		(0.365)		(0.126)
Proportion - Intelligence		1.553***		0.239		1.599***		0.260
		(0.585)		(0.167)		(0.603)		(0.170)
Proportion - Military		1.528***		0.346**		1.745***		0.398**
		(0.457)		(0.155)		(0.473)		(0.161)
Proportion - State		0.230		−0.016		0.180		−0.019
		(0.358)		(0.090)		(0.361)		(0.093)
Diplomatic Experience		0.075*		−0.009		0.125***		0.004
		(0.045)		(0.015)		(0.047)		(0.016)
Intelligence Experience		−0.103**		−0.038**		−0.116**		−0.042**
		(0.047)		(0.018)		(0.053)		(0.019)
Military Experience		0.054		−0.045*		−0.012		−0.066**
		(0.087)		(0.025)		(0.093)		(0.027)
5-Year MID Challenges		0.067		0.005		−0.250		−0.114
		(0.148)		(0.046)		(0.233)		(0.080)
US CINC		2.536**		1.418***		6.883***		2.941***
		(1.084)		(0.364)		(2.357)		(0.933)
Formal	1.698***	0.736***	0.413***	0.143***	1.551***	0.630***	0.388***	0.143***
	(0.086)	(0.130)	(0.032)	(0.050)	(0.089)	(0.137)	(0.034)	(0.052)
Constant	−2.728***	−5.180***	−0.030	−0.679***	−3.882***	−6.249***	−0.308*	−1.078***
	(0.362)	(0.723)	(0.126)	(0.224)	(0.595)	(1.147)	(0.186)	(0.380)
Administration FEs					✓	✓	✓	✓
Agenda Items	✓	✓	✓	✓	✓	✓	✓	✓
Observations	2,650	2,650	2,650	2,650	2,650	2,650	2,650	2,650

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

5.6 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 702 conflictual decisions and 248 cooperative decisions. Once we remove any decisions that were coded as involving the Soviet Union, we are left with 272 conflictual decisions and 84 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 A20 and A21, as well as Figure A13, show that our Poisson models maintain the same findings remain even when only analyzing decisions targeted at all other countries. Most of the OLS models do not produce significant findings for hawkishness, suggesting that US-Soviet relations may have had more of an impact on willingness to cooperate.

Figure A13: 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. Square points indicate 95% statistical significance.

Table A20: Effect of Mean Participant Hawkishness and President's Hawkishness on Foreign Policy Decisions, Removing Decisions Involving the USSR

	Emergence Model				*Leader Model*			
	Conflictual Decisions		Conflictual − Cooperative		Conflictual Decisions		Conflictual − Cooperative	
	Poisson		*OLS*		*Poisson*		*OLS*	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Mean Hawkishness	8.100***	3.637	0.809***	0.624**				
	(1.669)	(2.487)	(0.228)	(0.291)				
President's Hawkishness					−2.183***	−1.721*	−0.260***	−0.183
					(0.617)	(0.911)	(0.086)	(0.125)
No. of Attendees		0.015		0.004		0.021		0.006
		(0.030)		(0.004)		(0.028)		(0.004)
Proportion - Defense		2.090***		0.151*		1.326**		0.077
		(0.584)		(0.078)		(0.523)		(0.073)
Proportion - Intelligence		2.678***		0.166		2.135***		0.136
		(0.836)		(0.104)		(0.807)		(0.102)
Proportion - Military		1.096		0.025		1.439**		0.132
		(0.867)		(0.115)		(0.688)		(0.095)
Proportion - State		0.425		−0.010		0.265		−0.022
		(0.544)		(0.056)		(0.524)		(0.054)
Diplomatic Experience		0.174**		0.009		0.116		0.004
		(0.078)		(0.010)		(0.072)		(0.009)
Intelligence Experience		−0.075		−0.0004		−0.005		0.006
		(0.076)		(0.011)		(0.070)		(0.011)
Military Experience		0.197		0.007		0.271**		0.010
		(0.142)		(0.016)		(0.129)		(0.014)
5-Year MID Challenges		−1.341***		−0.129***		−0.472**		−0.039
		(0.404)		(0.049)		(0.237)		(0.029)
US CINC		4.505		1.395**		−0.147		0.206
		(4.044)		(0.572)		(1.709)		(0.220)
Formal	0.960***	0.219	0.101***	0.031	1.153***	0.382**	0.120***	0.039
	(0.137)	(0.206)	(0.021)	(0.032)	(0.122)	(0.189)	(0.019)	(0.031)
Constant	−7.603***	−6.866***	−0.463***	−0.707***	−1.735***	−3.253***	0.162***	0.043
	(1.000)	(2.038)	(0.134)	(0.262)	(0.303)	(1.077)	(0.043)	(0.133)
Administration FEs	✓	✓	✓	✓				
Agenda Items	✓	✓	✓	✓	✓	✓	✓	✓
Observations	2,650	2,650	2,650	2,650	2,650	2,650	2,650	2,650

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

Table A21: Effect of Adviser Hawkishness on Foreign Policy Decisions, Removing Decisions Involving the USSR

	Emergence Model				*Leader Model*			
	Conflictual Decisions		Conflictual − Cooperative		Conflictual Decisions		Conflictual − Cooperative	
	Poisson		*OLS*		*Poisson*		*OLS*	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Advisers' Hawkishness (Acts)	3.455***	2.142	0.305**	0.243	5.123***	1.623	0.482***	0.303
	(1.080)	(1.600)	(0.146)	(0.187)	(1.209)	(1.870)	(0.171)	(0.214)
President's Hawkishness	−2.688***	−2.381**	−0.320***	−0.277**				
	(0.625)	(1.022)	(0.091)	(0.139)				
No. of Attendees		0.033		0.007		0.020		0.005
		(0.029)		(0.004)		(0.031)		(0.005)
Proportion - Defense		1.522***		0.093		2.055***		0.146*
		(0.546)		(0.075)		(0.593)		(0.079)
Proportion - Intelligence		2.314***		0.155		2.776***		0.170
		(0.822)		(0.104)		(0.838)		(0.106)
Proportion - Military		1.117		0.092		1.459*		0.083
		(0.725)		(0.101)		(0.825)		(0.112)
Proportion - State		0.442		−0.007		0.407		−0.010
		(0.544)		(0.056)		(0.552)		(0.058)
Diplomatic Experience		0.105		0.002		0.182**		0.010
		(0.073)		(0.009)		(0.078)		(0.010)
Intelligence Experience		−0.016		0.005		−0.082		−0.001
		(0.071)		(0.011)		(0.076)		(0.012)
Military Experience		0.220		0.003		0.180		0.001
		(0.136)		(0.015)		(0.145)		(0.017)
5-Year MID Challenges		−0.489**		−0.042		−1.325***		−0.131***
		(0.239)		(0.029)		(0.407)		(0.050)
US CINC		−0.793		0.151		4.886		1.435**
		(1.771)		(0.228)		(4.050)		(0.580)
Formal	1.047***	0.375**	0.112***	0.044	1.089***	0.236	0.116***	0.037
	(0.126)	(0.191)	(0.020)	(0.031)	(0.133)	(0.207)	(0.021)	(0.032)
Constant	−3.247***	−3.759***	0.037	−0.001	−5.906***	−5.881***	−0.274***	−0.519**
	(0.571)	(1.165)	(0.075)	(0.142)	(0.746)	(1.884)	(0.102)	(0.239)
Administration FEs					✓	✓	✓	✓
Agenda Items	✓	✓	✓	✓	✓	✓	✓	✓
Observations	2,650	2,650	2,650	2,650	2,650	2,650	2,650	2,650

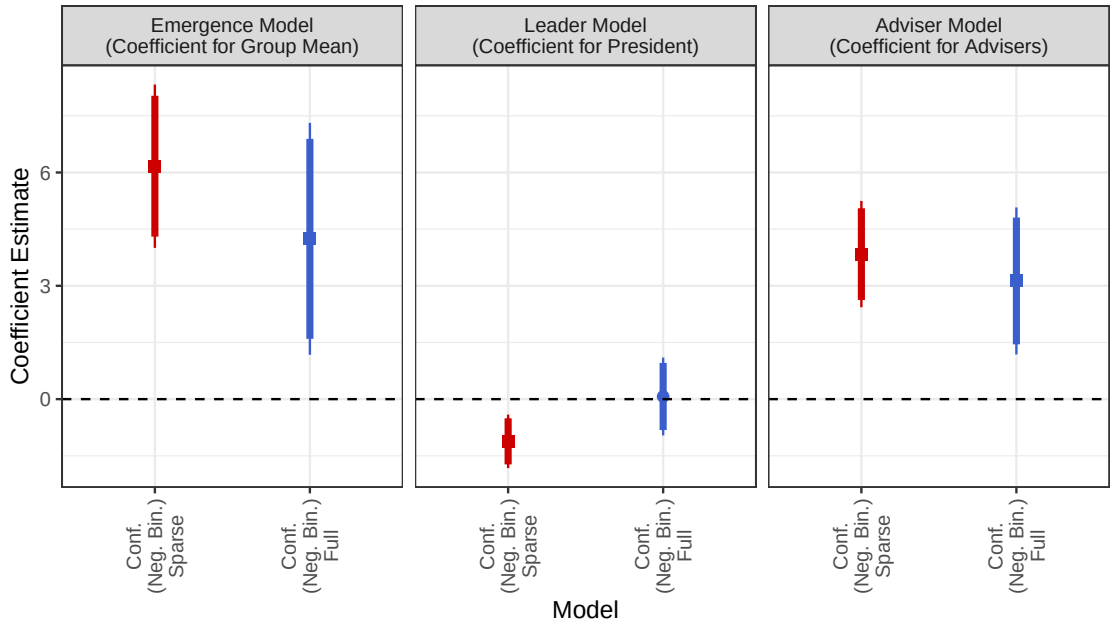
Note: $^{*}p < 0.1$; $^{**}p < 0.05$; $^{***}p < 0.01$

5.7 Negative Binomial

The analysis in the main text focuses both on the number of conflictual decisions made in a meeting, as well as the ratio between the number of conflictual and cooperative decisions made in a meeting. For the former, the analysis in the main text 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 0.261, while the variance is 0.490. However, dispersion tests suggest that a negative binomial model may still be appropriate to account for overdispersion.

Tables A22 and A23 therefore replicate the Poisson models in the main text but instead rely on negative binomial models. Figure A14 illustrates the corresponding coefficient plots. All results are consistent with the findings from simpler Poisson models.

Figure A14: 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. Square points indicate 95% statistical significance.

Table A22: Effect of Mean Participant Hawkishness and President's Hawkishness on Foreign Policy Decisions, Using Negative Binomial Models

	Emergence Model		*Leader Model*	
	Conflictual Decisions			
	(1)	(2)	(3)	(4)
Mean Hawkishness	6.703***	4.181**		
	(1.281)	(1.736)		
President's Hawkishness			−1.008**	0.537
			(0.452)	(0.616)
No. of Attendees		0.072***		0.074***
		(0.020)		(0.020)
Proportion - Defense		2.253***		1.527***
		(0.408)		(0.376)
Proportion - Intelligence		1.576**		1.286**
		(0.669)		(0.646)
Proportion - Military		1.054*		1.356***
		(0.625)		(0.512)
Proportion - State		0.363		0.148
		(0.392)		(0.386)
Diplomatic Experience		0.114**		0.077
		(0.053)		(0.051)
Intelligence Experience		−0.120**		−0.089*
		(0.060)		(0.054)
Military Experience		0.040		0.153*
		(0.102)		(0.092)
5-Year MID Challenges		−0.192		0.123
		(0.262)		(0.167)
US CINC		6.067**		2.557**
		(2.709)		(1.220)
President's Hawkishness			−1.008**	0.537
			(0.452)	(0.616)
Formal	1.444***	0.600***	1.778***	0.749***
	(0.104)	(0.153)	(0.096)	(0.148)
Constant	−5.815***	−7.237***	−1.747***	−4.715***
	(0.752)	(1.433)	(0.229)	(0.762)
Administration FEs	✓	✓		
Agenda Items	✓	✓	✓	✓
Observations	2,685	2,685	2,685	2,685

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

Table A23: Effect of Adviser Hawkishness on Foreign Policy Decisions, Using Negative Binomial Models

	Adviser Model		*Advisers + Admin. FEs*	
	Conflictual Decisions			
	(1)	(2)	(3)	(4)
Advisers' Hawkishness (Acts)	4.002***	3.070***	4.525***	2.383*
	(0.846)	(1.116)	(0.971)	(1.333)
President's Hawkishness	−1.656***	−0.419		
	(0.463)	(0.698)		
No. of Attendees		0.089***		0.081***
		(0.020)		(0.021)
Proportion - Defense		1.797***		2.194***
		(0.389)		(0.410)
Proportion - Intelligence		1.573**		1.621**
		(0.658)		(0.670)
Proportion - Military		0.950*		1.366**
		(0.529)		(0.593)
Proportion - State		0.363		0.324
		(0.397)		(0.397)
Diplomatic Experience		0.063		0.116**
		(0.052)		(0.053)
Intelligence Experience		−0.112**		−0.120**
		(0.055)		(0.060)
Military Experience		0.077		−0.002
		(0.096)		(0.102)
5-Year MID Challenges		0.074		−0.188
		(0.168)		(0.264)
US CINC		1.831		6.430**
		(1.248)		(2.713)
Formal	1.655***	0.725***	1.525***	0.613***
	(0.098)	(0.149)	(0.102)	(0.154)
Constant	−3.471***	−5.423***	−4.571***	−6.231***
	(0.437)	(0.823)	(0.583)	(1.299)
Administration FEs			✓	✓
Agenda Items	✓	✓	✓	✓
Observations	2,650	2,650	2,650	2,650

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

5.8 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.[4] 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 A24 (which is a duplicate of Table 7 in the main text but is provided here for convenience) displays the results of two Poisson models that regress new MIDs (hostility level 3 and above) featuring 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.

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

Table A24: Effect of NSC Principals' Hawkishness on MIDs, Using Monthly Data

	Dependent variable:	
	Onset of MIDs involving US	
	(1)	(2)
Advisers' hawkishness	6.816**	11.198**
	(2.734)	(4.629)
President's hawkishness	−3.510**	−3.795
	(1.593)	(2.832)
War ongoing		−0.377**
		(0.153)
Deaths per capita in last war (logged)		0.163
		(0.111)
Months since last war (logged)		0.294
		(0.198)
Victory in last war		−1.070
		(0.760)
MID challenges to US in last 5 years		−0.035
		(0.049)
Average MID outcome in last 5 years		−55.273
		(33.947)
Economic recession		−0.194
		(0.216)
Unified government		0.404
		(0.258)
US material capabilities		−5.518
		(6.850)
President's tenure (logged months)		−0.016
		(0.090)
Constant	−2.989***	−3.126
	(0.836)	(2.942)
Observations	501	501

Note: $^{*}p < 0.1$; $^{**}p < 0.05$; $^{***}p < 0.01$. Outcome variable is the number of US MIDs beginning in a given month. Advisers' hawkishness reflects average hawkishness score of the most senior advisers in the administration in a given month.

5.8.1 Propagating Uncertainty of Hawkishness Measures in Time-Unit Analysis

In Appendix § 5.3, we explain how our main analysis does not fully propagate the uncertainty inherent to our predicted measures of hawkishness. It may be worth checking whether the uncertainty in our hawkishness measures would affect the strength of our time-unit analysis.

We once again go back to our 1,000 bootstrap iterations. For each iteration, we take the predicted measures of hawkishness for each of our actor-years and construct monthly-level data. We then run our regression models on this monthly data, producing a coefficient estimate and associated standard error for each variable. This process is performed 1,000 times (once for each monthly level dataset produced from each bootstrap iteration).

Table A25 reports the average estimated coefficient and standard error for each variable in each model across all 1,000 iterations. Results are highly similar to those in Table A24.

5.9 Statutory Members Only

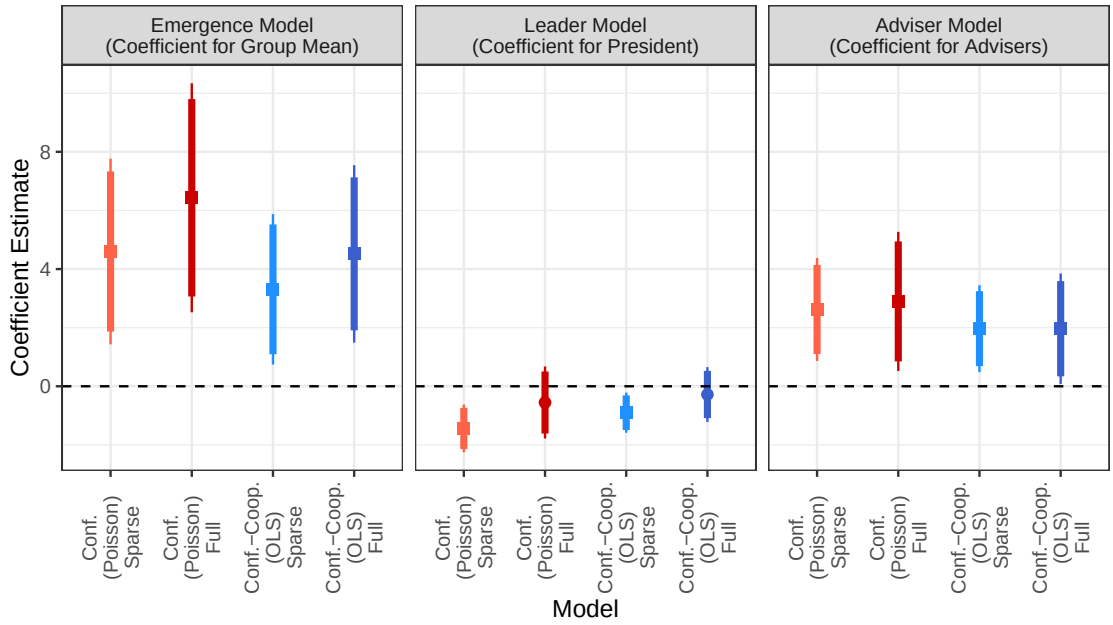
One may be concerned that the president shapes the outcomes of policy discussions by choosing which specific advisers attend a meeting. To address this potential selection issue, we replicate our meeting-level analysis after recalculating our measures of adviser hawkishness, proportion of attendees according to bureaucratic affiliation, and levels of diplomatic/intelligence/military experience using only information on statutory members of the NSC, who are ostensibly required to be at all meetings. These include the president, the vice president, the Secretary of State, the Secretary of Defense, Director of Central Intelligence, Chairman of the Joint Chiefs of Staff, and the National Security Advisor. We limit this analysis to only formal NSC meetings, where the statutory nature of these advisers is relevant. Tables A26 and A27, which are visually summarized by Figure A15, show that our results remain robust.

Table A25: Effect of NSC Principals' Hawkishness on MIDs, Using Monthly Data, Propagating Uncertainty from Bootstrapping

	Dependent variable:	
	Onset of MIDs involving US	
	(1)	(2)
Advisers' hawkishness	6.465**	10.703**
	(2.702)	(4.519)
President's hawkishness	−3.315**	−3.497
	(1.587)	(2.772)
War ongoing		−0.380**
		(0.154)
Deaths per capita in last war (logged)		0.159
		(0.111)
Months since last war (logged)		0.295
		(0.198)
Victory in last war		−1.073
		(0.763)
MID challenges to US in last 5 years		−0.034
		(0.049)
Average MID outcome in last 5 years		−56.079*
		(33.991)
Economic recession		−0.190
		(0.216)
Unified government		0.392
		(0.257)
US material capabilities		−5.216
		(6.827)
President's tenure (logged months)		−0.017
		(0.090)
Constant	−2.897***	−3.105
	(0.824)	(2.953)
Observations	501	501

Note: $^{*}p < 0.1$; $^{**}p < 0.05$; $^{***}p < 0.01$. Outcome variable is the number of US MIDs beginning in a given month. Advisers' hawkishness reflects average hawkishness score of the most senior advisers in the administration in a given month.

Figure A15: Summary of Three Models of Trait Aggregation, Using Only Statutory NSC Members in 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. Square points indicate 95% statistical significance.

Table A26: Effect of Mean Participant Hawkishness and President's Hawkishness on Foreign Policy Decisions, Using Only Statutory NSC Members in Formal Meetings

	Emergence Model				*Leader Model*			
	Conflictual Decisions		Conflictual − Cooperative		Conflictual Decisions		Conflictual − Cooperative	
	Poisson		*OLS*		*Poisson*		*OLS*	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Mean Hawkishness	4.599***	6.432***	3.308**	4.518***				
	(1.617)	(1.993)	(1.310)	(1.545)				
President's Hawkishness					−1.440***	−0.554	−0.901**	−0.282
					(0.415)	(0.626)	(0.350)	(0.476)
No. of Attendees		0.077***		0.059***		0.088***		0.066***
		(0.021)		(0.018)		(0.020)		(0.018)
Proportion - Defense		1.411***		0.850**		0.290		0.244
		(0.543)		(0.414)		(0.489)		(0.392)
Proportion - Intelligence		0.835		0.830		0.208		0.568
		(1.017)		(0.668)		(0.967)		(0.645)
Proportion - Military		−0.215		−0.029		−0.335		0.181
		(0.942)		(0.732)		(0.833)		(0.678)
Proportion - State		−0.376		0.279		−0.741		0.017
		(0.672)		(0.456)		(0.642)		(0.446)
Diplomatic Experience		0.089		0.008		0.061		−0.001
		(0.062)		(0.043)		(0.059)		(0.043)
Intelligence Experience		−0.102		−0.024		−0.093*		−0.030
		(0.075)		(0.051)		(0.056)		(0.042)
Military Experience		−0.151		−0.131		0.070		−0.039
		(0.146)		(0.106)		(0.131)		(0.096)
5-Year MID Challenges		−0.458		−0.304		0.067		0.035
		(0.301)		(0.215)		(0.208)		(0.141)
US CINC		8.629***		6.294***		3.078**		2.375**
		(2.912)		(2.216)		(1.394)		(1.016)
Constant	−3.508***	−7.454***	−1.610**	−4.269***	0.238	−2.456**	0.914***	−0.645
	(1.014)	(1.748)	(0.812)	(1.301)	(0.204)	(1.017)	(0.178)	(0.705)
Administration FEs	✓	✓	✓	✓				
Agenda Items	✓	✓	✓	✓	✓	✓	✓	✓
Observations	788	788	788	788	788	788	788	788

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

Table A27: Effect of Adviser Hawkishness on Foreign Policy Decisions, Using Only Statutory NSC Members in Formal Meetings

	Adviser Model				*Advisers + Admin. FEs*			
	Conflictual Decisions		Conflictual − Cooperative		Conflictual Decisions		Conflictual − Cooperative	
	Poisson		*OLS*		*Poisson*		*OLS*	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Advisers' Hawkishness (Acts)	2.619***	2.896**	1.966***	1.964**	4.242***	4.096***	2.937***	3.076***
	(0.898)	(1.211)	(0.757)	(0.961)	(1.198)	(1.429)	(1.022)	(1.143)
President's Hawkishness	−2.294***	−2.047**	−1.597***	−1.381**				
	(0.496)	(0.862)	(0.430)	(0.703)				
No. of Attendees		0.100***		0.073***		0.080***		0.059***
		(0.020)		(0.018)		(0.022)		(0.018)
Proportion - Defense		0.452		0.336		1.257**		0.732*
		(0.502)		(0.407)		(0.548)		(0.424)
Proportion - Intelligence		0.695		0.883		0.794		0.811
		(0.989)		(0.676)		(1.028)		(0.689)
Proportion - Military		−0.420		0.092		0.428		0.339
		(0.837)		(0.684)		(0.896)		(0.709)
Proportion - State		−0.435		0.115		−0.365		0.193
		(0.659)		(0.476)		(0.686)		(0.483)
Diplomatic Experience		0.045		−0.005		0.084		0.009
		(0.060)		(0.044)		(0.063)		(0.045)
Intelligence Experience		−0.135**		−0.059		−0.105		−0.026
		(0.059)		(0.044)		(0.076)		(0.052)
Military Experience		−0.014		−0.093		−0.192		−0.158
		(0.138)		(0.105)		(0.151)		(0.112)
5-Year MID Challenges		−0.039		−0.057		−0.430		−0.304
		(0.215)		(0.148)		(0.303)		(0.218)
US CINC		1.887		1.406		8.440***		6.014***
		(1.489)		(1.114)		(2.938)		(2.244)
Constant	−0.767*	−2.578**	0.194	−0.598	−3.295***	−5.840***	−1.388**	−3.167***
	(0.410)	(1.065)	(0.340)	(0.750)	(0.757)	(1.583)	(0.637)	(1.181)
Administration FEs					✓	✓	✓	✓
Agenda Items	✓	✓	✓	✓	✓	✓	✓	✓
Observations	775	775	775	775	775	775	775	775

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

5.10 Crisis Period Analysis

The analysis in the main text examines the effects of advisers on foreign policy decisions by coding all of the substantive decisions made in these meetings along a continuum from conflictual to cooperative. This allows us to avoid the aggregation bias that ensues when we study foreign policy decision-making by only focusing on major uses of force. At the same time, however, one might wonder whether this more comprehensive analysis also exacerbates the importance of advisers, who might be less influential in high stakes crisis situations. We consider this unlikely — if anything, leaders should presumably have more reason to turn to advisers in high-stakes crises, not less — but we test this empirically by turning to the International Crisis Behavior dataset to identify crisis periods for the United States. Of the 2,685 total meetings in our dataset, 895 (33%) occur during crisis periods. In Table A28, we replicate the adviser models (Models 5 through 8) in the main text but include an interaction term between advisers' hawkishness and whether a meeting takes place during a crisis period. The interaction term is statistically significant and negative in Models 1 and 3, potentially suggesting that advisers are not as influential in generating conflictual foreign policy decisions during times of crisis. However, the magnitude of this interaction effect continues to be outweighed by the first-order effect of advisers. The interaction bears no significance in the three other models. Overall, these findings suggest weak evidence that advisers matter significantly less during crises.

Table A28: Effect of Adviser Hawkishness on Foreign Policy Decisions, Including Interaction Term for Crisis Periods

	Adviser Model		*Advisers + Admin. FEs*	
	Conf	Conf. − Coop.	Conf.	Conf. − Coop.
	Poisson	*OLS*	*Poisson*	*OLS*
	(1)	(2)	(3)	(4)
Advisers' Hawkishness (Acts)	4.860***	0.716**	3.970***	0.654*
	(1.182)	(0.336)	(1.377)	(0.384)
Crisis	2.429***	0.104	1.922**	0.076
	(0.777)	(0.236)	(0.789)	(0.239)
Advisers × Crisis	−3.926***	0.015	−2.905**	0.089
	(1.446)	(0.457)	(1.475)	(0.464)
President's Hawkishness	−0.810	−0.083		
	(0.609)	(0.224)		
No. of Attendees	0.085***	0.042***	0.073***	0.040***
	(0.017)	(0.007)	(0.017)	(0.007)
Proportion - Defense	1.474***	0.178	1.990***	0.301**
	(0.343)	(0.121)	(0.368)	(0.127)
Proportion - Intelligence	1.670***	0.261	1.702***	0.275
	(0.594)	(0.168)	(0.613)	(0.170)
Proportion - Military	0.827*	0.180	1.177**	0.243
	(0.465)	(0.162)	(0.527)	(0.180)
Proportion - State	0.336	0.002	0.307	−0.003
	(0.360)	(0.090)	(0.363)	(0.093)
Diplomatic Experience	0.059	−0.012	0.113**	0.002
	(0.046)	(0.015)	(0.048)	(0.016)
Intelligence Experience	−0.104**	−0.038**	−0.110**	−0.042**
	(0.048)	(0.018)	(0.054)	(0.019)
Military Experience	0.069	−0.045*	−0.004	−0.066**
	(0.086)	(0.024)	(0.093)	(0.027)
5-Year MID Challenges	0.012	0.015	−0.240	−0.080
	(0.154)	(0.047)	(0.235)	(0.080)
US CINC	1.856*	1.197***	6.838***	3.122***
	(1.124)	(0.369)	(2.386)	(0.932)
Formal	0.779***	0.152***	0.648***	0.147***
	(0.131)	(0.050)	(0.138)	(0.052)
Constant	−6.063***	−0.720***	−7.212***	−1.232***
	(0.811)	(0.243)	(1.207)	(0.395)
Administration FEs			✓	✓
Agenda Items	✓	✓	✓	✓
Observations	2,650	2,650	2,650	2,650

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

5.11 Adviser Experience and Dispositional Distance Analysis

The theoretical model advanced in the paper is that advisers' predispositions shape leader decisions by affecting the information, analysis and policy recommendations advisers offer leaders during deliberations. The ill-structured nature of foreign policy problems means leaders turn to advisers for counsel, and the kind of counsel advisers supply depends on advisers' dispositions, such that leaders who surround themselves with a team of hawks will end up making different decisions than those who have a more diverse advisory team.

In this respect, our argument intersects with recent work by Saunders (2017, 2018), which also studies the interplay between leaders and advisers. Saunders's interests are somewhat different than ours, focusing on the *quality* of foreign policy decision rather than merely their substance, the aggregation of biases rather than the aggregation of traits,[5] emphasizing a causal mechanism of advisory influence rooted in domestic politics rather than deliberation, and reputations rather than dispositions. Nonetheless, two interesting points of dialog emerge. One concerns the potential role of experience. Saunders shows that relatively inexperienced leaders are less equipped to monitor their advisers, resulting in suboptimal policy choices. An implication in our theoretical framework would be that when the balance of experience favors leaders rather than advisers, advisers' predispositions should exert less of an impact on the foreign policy decisions reached, since more experienced leaders should be more likely to get their way.

We test this proposition by calculating the average number of years of foreign policy experience that advisers in each meeting had prior to the meeting, in either military, diplomatic, or intelligence roles. We then compare this average to the amount of experience possessed by the president. We then can calculate (Leader Experience - Average Adviser Experience) and call this the *leader-adviser experience gap*.[6]

Table A29 replicates the adviser model results from the main text, but includes an interaction term between advisers' hawkishness and the adviser-leader experience gap. The individual terms

[5] On the distinction between the aggregation of cognitive biases and trait aggregation, see Kertzer et al. (2022). Trait aggregation concerns the mapping between the distribution of traits among group members and the group's behavior (e.g. if a group has two extroverted members and one introverted member, how extraverted is the group's behavior?), whereas the aggregation of biases concerns whether cognitive biases that appear at the individual level also hold at the collective level.

[6] Meetings without the President in attendance are excluded from this analysis.

are scaled in order to reduce multicollinearity between them and their interaction term.[7] It is worth noting that the president's hawkishness measure introduces some multicollinearity issues in Models 1 and 2 regardless of scaling, so the standard errors are somewhat inflated. As such, Models 3 and 4 revisit the adviser model after removing measures of the president's hawkishness. Results suggest that the experience gap does not significantly moderate the effects of adviser traits: hawkish advisers are not significantly less able to push foreign policy decisions in a more conflictual direction when serving a more experienced leader than a less experienced one.

[7]The correlation between the experience gap and the interaction is 0.994. This correlation drops to 0.106 when using scaled variables.

Table A29: Effect of Adviser Hawkishness on Foreign Policy Decisions, Including Interaction Term for Leader-Adviser Experience Gap

	Adviser Model		*Adviser Model − Admin. FEs*		*Advisers + Admin. FEs*	
	Conf	Conf. − Coop.	Conf.	Conf. − Coop.	Conf	Conf. − Coop.
	Poisson	*OLS*	*Poisson*	*OLS*	*Poisson*	*OLS*
	(1)	(2)	(3)	(4)	(5)	(6)
Advisers' Hawkishness (Acts)	0.082	0.031	0.162***	0.044**	0.190**	0.044*
	(0.074)	(0.021)	(0.063)	(0.019)	(0.084)	(0.023)
Experience Gap	0.258***	0.047*	0.140**	0.028	0.436*	0.013
	(0.090)	(0.025)	(0.066)	(0.021)	(0.248)	(0.059)
Advisers × Exp. Gap	0.054	0.018	0.013	0.012	0.027	0.015
	(0.063)	(0.018)	(0.059)	(0.017)	(0.063)	(0.018)
President's Hawkishness	1.947**	0.383				
	(0.980)	(0.292)				
No. of Attendees	0.079***	0.041***	0.087***	0.043***	0.058***	0.041***
	(0.018)	(0.007)	(0.017)	(0.007)	(0.021)	(0.008)
Proportion - Defense	2.380***	0.333**	2.431***	0.331**	2.671***	0.402***
	(0.386)	(0.134)	(0.386)	(0.134)	(0.413)	(0.139)
Proportion - Intelligence	1.745***	0.242	1.736***	0.236	1.675**	0.260
	(0.640)	(0.176)	(0.641)	(0.176)	(0.664)	(0.178)
Proportion - Military	1.566**	0.309	1.025*	0.211	1.404**	0.182
	(0.611)	(0.199)	(0.544)	(0.185)	(0.681)	(0.210)
Proportion - State	0.311	−0.010	0.349	−0.010	0.293	−0.006
	(0.384)	(0.094)	(0.384)	(0.094)	(0.391)	(0.097)
Diplomatic Experience	0.102**	−0.006	0.087*	−0.009	0.129**	−0.003
	(0.051)	(0.016)	(0.050)	(0.016)	(0.055)	(0.018)
Intelligence Experience	−0.135***	−0.039**	−0.148***	−0.041**	−0.114**	−0.044**
	(0.050)	(0.019)	(0.050)	(0.019)	(0.056)	(0.019)
Military Experience	0.121	−0.040	0.085	−0.047*	0.172	−0.048
	(0.096)	(0.027)	(0.093)	(0.026)	(0.119)	(0.034)
5-Year MID Challenges	−0.029	−0.001	−0.141	−0.027	−0.486*	−0.136
	(0.153)	(0.048)	(0.143)	(0.044)	(0.253)	(0.084)
US CINC	1.664	1.399***	1.160	1.298***	5.929**	2.787***
	(1.138)	(0.396)	(1.118)	(0.388)	(2.537)	(0.966)
Formal	0.674***	0.133**	0.746***	0.148***	0.556***	0.127**
	(0.137)	(0.052)	(0.132)	(0.051)	(0.143)	(0.054)
Constant	−4.974***	−0.607**	−3.571***	−0.320**	−4.559***	−0.673*
	(0.857)	(0.254)	(0.483)	(0.128)	(1.042)	(0.356)
Administration FEs					✓	✓
Agenda Items	✓	✓	✓	✓	✓	✓
Observations	2,543	2,543	2,543	2,543	2,543	2,543

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

Another point of dialog concerning the contingent nature of advisory influence on leaders concerns the dispositional distance between leaders and advisers. Here, the existing literature suggests competing predictions. Kertzer, Rathbun and Rathbun (2020) proposes a theory of motivated reasoning in foreign policy, in which actors update the most in response to information they already agree with. If advisers are providing recommendations or perspectives consistent with their dispositions, this suggests leaders should respond the most to advisers whose dispositions match their own: hawks should pay the most attention to hawkish advisers, and doves should be the most sensitive to dovish advisers. This would imply a significant positive interaction between leader and adviser hawkishness.[8] In contrast, Saunders (2018) offers an important a theory of domestic politics of foreign policy in which leaders with reputations for dovishness have political incentives to care the most about the recommendations from their most hawkish advisers (and leaders with hawkish reputations are the most sensitive to their most dovish advisers). If we presume reputations for hawkishness or dovishness are at least partially rooted in actual hawkish or dovish policy preferences themselves, this implies a significant negative interaction between leader and adviser hawkishness.[9]

Table A30 replicates the analysis from the main text, but including an interaction term between leader and adviser hawkishness. The hawkishness variables are scaled in order to reduce multicollinearity between these individual measures and the interaction term.[10] We fail to find evidence that the effect of adviser traits on foreign policy behavior varies as a function of the leader's own level of hawkishness. We see this as further evidence in support of our theoretical model, which suggests that in deliberative contexts, leaders' need for counsel means that hawkish and dovish advisers can influence hawkish and dovish leaders alike.

[8] Of course, it is possible that advisers go against type and provide recommendations contrary to their dispositions (e.g. Mattes and Weeks, 2019; Kreps, Saunders and Schultz, 2018); in section 3.2 of the main paper, we show that even though advisers have the ability to go against type, on average, hawkish advisers are still more likely to make hawkish policy recommendations than doves are (which presumably is what allows hawks to maintain hawkish reputations in the first place!). See Kertzer and Brooks, 2021).

[9] We also test the effect of dispositional gaps using a model that interacts the president's political party and adviser hawkishness, in case leader reputation stems from party affiliation rather than policy preference. The results of this alternative specification remain similar.

[10] The correlation between the president's hawkishness and the interaction is 0.898. This correlation drops to 0.310 when using scaled variables.

Table A30: Effect of Interaction between Adviser Hawkishness and Leader Hawkishness on Foreign Policy Decisions

	Dependent variable:			
	Conflict		Conflict − Cooperation	
	Poisson		*OLS*	
	(1)	(2)	(3)	(4)
Advisers' Hawkishness (Acts)	0.406***	0.192***	0.105***	0.046**
	(0.038)	(0.063)	(0.015)	(0.019)
President's Hawkishness	−0.215***	−0.120*	−0.058***	−0.010
	(0.047)	(0.069)	(0.016)	(0.023)
Adv. Hawk. × Pres. Hawk.	0.050	0.050	0.002	−0.010
	(0.038)	(0.047)	(0.015)	(0.015)
No. of Attendees		0.091***		0.042***
		(0.017)		(0.007)
Proportion - Defense		1.624***		0.184
		(0.342)		(0.121)
Proportion - Intelligence		1.572***		0.251
		(0.586)		(0.168)
Proportion - Military		1.024**		0.176
		(0.473)		(0.164)
Proportion - State		0.269		−0.011
		(0.360)		(0.090)
Diplomatic Experience		0.070		−0.008
		(0.045)		(0.015)
Intelligence Experience		−0.110**		−0.040**
		(0.047)		(0.018)
Military Experience		0.047		−0.045*
		(0.085)		(0.025)
5-Year MID Challenges		0.012		0.005
		(0.152)		(0.047)
US CINC		2.028*		1.359***
		(1.097)		(0.367)
Formal		0.762***		0.152***
		(0.130)		(0.051)
Constant	−1.427***	−3.833***	0.171***	−0.375***
	(0.043)	(0.518)	(0.015)	(0.143)
Administration FEs				
Agenda Items	✓	✓	✓	✓
Observations	2,650	2,650	2,650	2,650

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

6 Modeling Adviser Appointment

In order for adviser predispositions to affect foreign policy decisions, advisers cannot merely be clones of their leaders. 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 would simply represent an extension of the leader's disposition.

One of the premises of the adviser model advanced in the paper is that even though advisers aren't selected randomly, hawkishness is one of many different considerations leaders weigh when selecting advisers, alongside considerations like experience, qualifications, personal connections, perceived loyalty, and so on (Jost, 2023). Since these traits are bundled together, appointment decisions for one position often depend on appointment choices for other positions (e.g. George W. Bush appointed Donald Rumsfeld to check Colin Powell), and different presidents vary in what they look for in their advisers, we should not expect that all advisers will automatically be dispositional mimeographs of the leaders that appoint them, and there should be variation in the predispositions inside each administration. We demonstrate this point empirically in the paper by using our hawkishness scores and biographical data to estimate a set of mixed-effect models with administration random effects, which find that the intraclass correlation in hawkishness ranges between 0.043 and 0.179 – that is, there is approximately 5.6 to 23 times greater variation in hawkishness *within* individual administrations than there is between them. We would expect a far lower figure if hawkish leaders were only hiring hawks, and dovish leaders only hiring doves.

We also build on this point by utilizing a dataset we constructed for another part of this project, which compares the hawkishness scores of NSC principals with the hawkishness scores of other individuals who had been considered for the same position, but ultimately not selected. This dataset is the subject of a separate paper that examines the politics of presidential appointments to national security positions in more detail, so we only briefly discuss it here, but Table A31 presents the results of a set of logistic regressions modeling adviser appointment to 13 distinct positions filled between the Eisenhower and Reagan administrations (for Secretaries of State, Defense, Ambassador to the United Nations, etc.), where where we calculate the distance between each candidate's predicted hawkishness and the predicted hawkishness of the respective president. Models 1 and 2 feature administration fixed effects, while Models 3 and 4 employ administration random effects.

Table A31: Logistic Regressions Regarding Distance from President and Eventual Appointment

	Dependent variable:			
	Appointed			
	(1)	(2)	(3)	(4)
Distance from President	−3.974	−0.450	−1.671	−1.671
	(4.907)	(5.794)	(4.272)	(4.272)
Co-partisanship		1.242		
		(0.881)		
Previously Worked		0.348		
		(0.872)		
Worked on Campaign		−2.034*		
		(1.119)		
Constant	−0.314	−0.654	−1.314**	−1.314**
	(1.054)	(1.309)	(0.510)	(0.510)
Fixed Effects (Admin.)	✓	✓		
Random Effects (Admin.)			✓	✓
Observations	70	70	70	70

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

We find that once we control for other appointment considerations (such as whether the candidate is of the same partisan affiliation as the president, prior experience, whether they previously worked for the president), hawkishness distance from the president is not a statistically significant predictor of whether an adviser is appointed.

7 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.

One point to note from the outset is that these results are not an artifact of high correlation

Table A32: Variance inflation factors for key covariates.

	Emergence Model		*Leader Model*		*Adviser Model*		*Advisers + Admin. FEs*	
	Conf	Conf. − Coop.	Conf.	Conf. − Coop.	Conf	Conf. − Coop.	Conf.	Conf. − Coop.
	Poisson	*OLS*	*Poisson*	*OLS*	*Poisson*	*OLS*	*Poisson*	*OLS*
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Pres. Hawkishness	1.148	2.077	1.155	1.951				
Adv. Hawkishness	1.099	2.828	1.083	2.643	1.312	1.776	1.258	1.608
Admin. FEs					1.046	1.518	1.042	1.510

between leader and adviser hawkishness, which could hypothetically produce unstable and highly inflated results. The correlation between leader hawkishness and advisers' overall hawkishness (weighted by speech acts) at the meeting level is 0.277.

Moreover, an analysis of variance inflation factors (VIFs) across all models that include both leader and adviser hawkishness do not provide any evidence of concerning multicollinearity involving the hawkishness measures. Table A32 reports VIFs across all eight models in Table 5 in the main text. (Note that in Models 5 through 8, we report $\text{GVIF}^{1/(2*Df)}$ to account for administration fixed effects, which are a categorical variable.) All of these statistics are well within a reasonable range that does not introduce concerns about multicollinearity.

In the discussion below, we 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 — such that out findings are actually consistent with theirs. We then present a number of potential explanations for these results.

7.1 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.[11] We took the responses of all 14 experts and calculated an

[11] For 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

average score for each president, to produce an alternative measure of presidential hawkishness we can use to replicate the results in the main text.[12]

Table A33 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 and negative coefficients for president's hawkishness or coefficients without any statistical significance. Figure A16 visually reinforces this point.

on Cold War-era US foreign policy, were assured anonymity and confidentiality of their responses.

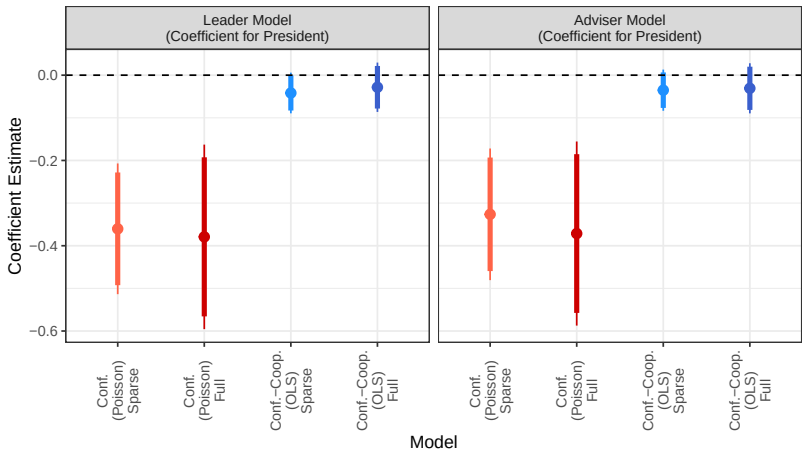
[12]One expert did not provide a score for President Ford. Ford's average score is therefore based on 13 responses. Note that we do not have expert codings of Carter or Reagan, which results in the omission of 131 meetings, or 5% of all meetings in the main paper.

Table A33: Effect of President's Hawkishness on Foreign Policy Decisions, Using Expert Codings

	Leader Model				*Adviser Model*			
	Conflictual Decisions		Conflictual − Cooperative		Conflictual Decisions		Conflictual − Cooperative	
	Poisson		*OLS*		*Poisson*		*OLS*	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
President's Hawkishness	−0.360***	−0.379***	−0.042*	−0.028	−0.326***	−0.371***	−0.035	−0.031
	(0.078)	(0.110)	(0.024)	(0.030)	(0.079)	(0.110)	(0.025)	(0.030)
Advisers' Hawkishness (Acts)					3.389***	1.916*	0.710***	0.570*
					(0.776)	(1.061)	(0.242)	(0.306)
No. of Attendees		0.082***		0.038***		0.089***		0.041***
		(0.017)		(0.007)		(0.017)		(0.007)
Proportion - Defense		1.476***		0.125		1.596***		0.162
		(0.335)		(0.117)		(0.345)		(0.121)
Proportion - Intelligence		1.299**		0.192		1.402**		0.235
		(0.589)		(0.167)		(0.600)		(0.170)
Proportion - Military		1.681***		0.294*		1.332***		0.186
		(0.465)		(0.152)		(0.497)		(0.164)
Proportion - State		−0.050		−0.096		0.077		−0.055
		(0.363)		(0.088)		(0.375)		(0.094)
Diplomatic Experience		0.129***		0.001		0.109**		−0.006
		(0.048)		(0.015)		(0.050)		(0.016)
Intelligence Experience		−0.066		−0.033*		−0.059		−0.032*
		(0.049)		(0.018)		(0.049)		(0.018)
Military Experience		−0.012		−0.033		−0.031		−0.045*
		(0.087)		(0.024)		(0.088)		(0.025)
5-Year MID Challenges		−0.050		−0.039		0.012		−0.012
		(0.129)		(0.040)		(0.136)		(0.044)
US CINC		5.259***		1.675***		4.898***		1.623***
		(1.172)		(0.377)		(1.191)		(0.388)
Formal	1.699***	0.601***	0.432***	0.131**	1.606***	0.568***	0.414***	0.127**
	(0.086)	(0.135)	(0.032)	(0.051)	(0.087)	(0.135)	(0.033)	(0.052)
Constant	−1.313***	−3.235***	0.154**	−0.293**	−3.123***	−4.248***	−0.221	−0.596***
	(0.203)	(0.563)	(0.065)	(0.144)	(0.466)	(0.803)	(0.143)	(0.220)
Administration FEs								
Agenda Items	✓	✓	✓	✓	✓	✓	✓	✓
Observations	2,554	2,554	2,554	2,554	2,521	2,521	2,521	2,521

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

Figure A16: 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. Square points indicate 95% statistical significance.

7.2 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, 2021), 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 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.[13] 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 point out that patterns that hold in the aggregate in cross-national models may not hold in specific, theoretically important, 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.[14] We discuss each in turn.

7.2.1 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

[13] Yarhi-Milo's results are reported in Table 4.5 of the original book (pp. 89-90).

[14] We 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).

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 A34 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 A34: Replication of Models in Table 1 of Horowitz and Stam (2014), Only US Presidents

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

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

7.2.2 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.[15]

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 7.2.1 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 A35 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 A36 reports an analysis of severe MID initiation. The coefficients of all four measures of latent hawkishness are negative and not statistically significant. Third, Table A37 shows that there are negative and statistically significant relationships between each latent hawkishness measure and war initiation.

Finally, Table A38 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

[15]Each 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.

7.3 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, 2023), 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.

7.4 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 7.1, we replicate our results using the expert survey, the results of which (in Appendix 7.1) similarly show a weak or negative effect for presidential hawkishness. As noted in

Table A35: Extension of Carter and Smith (2020): MID Initiation, Only US Presidents

	Dependent variable:				
	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	−0.356	−0.021	0.441	0.448
	(0.569)	(0.731)	(0.672)	(0.363)	(0.363)
Leader age	0.014	0.004	0.007	0.033	0.037
	(0.030)	(0.029)	(0.032)	(0.029)	(0.030)
Material capabilities	−10.212***	−9.218***	−9.710***	−8.134**	−8.160**
	(3.443)	(3.219)	(3.215)	(3.191)	(3.180)
Tau B with system leader	0.791**	0.707*	0.762**	0.822**	0.806**
	(0.366)	(0.367)	(0.361)	(0.365)	(0.366)
Time in office	0.128	0.133	0.142	0.164	0.151
	(0.210)	(0.208)	(0.209)	(0.214)	(0.214)
Five-year MID challenge lag	−0.103	−0.128	−0.104	−0.061	−0.078
	(0.429)	(0.430)	(0.429)	(0.436)	(0.436)
Constant	1.061	1.701	1.446	0.700	0.472
	(1.490)	(1.647)	(1.761)	(1.502)	(1.521)
Observations	178	178	178	178	178
Peace Year Splines	✓	✓	✓	✓	✓
Only US Presidents	✓	✓	✓	✓	✓

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

Table A36: Extension of Carter and Smith (2020): Severe MID Initiation, Only US Presidents

	Dependent variable:				
	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	0.976	1.073	0.740	0.748
	(0.787)	(1.001)	(0.970)	(0.567)	(0.567)
Leader age	−0.024	−0.016	−0.006	−0.013	−0.008
	(0.044)	(0.043)	(0.048)	(0.042)	(0.044)
Material capabilities	1.629	1.128	0.974	3.435	3.411
	(4.443)	(4.256)	(4.252)	(4.555)	(4.531)
Tau B with system leader	0.044	0.095	0.071	0.096	0.078
	(0.515)	(0.517)	(0.508)	(0.522)	(0.519)
Time in office	0.012	0.012	−0.020	0.050	0.035
	(0.307)	(0.299)	(0.305)	(0.306)	(0.306)
Five-year MID challenge lag	1.983*	1.989*	1.959*	2.046*	2.026*
	(1.094)	(1.091)	(1.090)	(1.104)	(1.102)
Constant	−2.995	−3.569	−4.062	−3.276	−3.488
	(2.251)	(2.482)	(2.759)	(2.234)	(2.278)
Observations	178	178	178	178	178
Peace Year Splines	*No*	*No*	*No*	*No*	*No*
Only US Presidents	✓	✓	✓	✓	✓

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

Table A37: Extension of Carter and Smith (2020): Interstate War, Only US Presidents

	Dependent variable:				
	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***	267.652***	27.308**	4.272***	4.713***
	(1.081)	(89.740)	(10.648)	(1.592)	(1.656)
Leader age	0.033	−0.087	0.307***	0.188***	0.256***
	(0.044)	(0.068)	(0.109)	(0.058)	(0.065)
Material capabilities	−13.930	−42.495***	−6.438	−13.965	−13.934
	(11.909)	(15.271)	(13.621)	(14.728)	(15.902)
Tau B with system leader	0.523	1.452	−0.647	2.267	2.167*
	(1.226)	(0.928)	(1.178)	(1.390)	(1.264)
Time in office	0.296	1.996***	0.355	0.689	0.534
	(0.559)	(0.566)	(0.701)	(0.568)	(0.605)
Five-year MID challenge lag	−16.581***	−16.917***	−17.538***	−17.557***	−17.438***
	(0.843)	(0.617)	(0.865)	(0.764)	(0.685)
Constant	−3.474	−134.039***	−34.495***	−6.525***	−10.307***
	(2.603)	(44.404)	(11.699)	(2.442)	(2.850)
Observations	126	126	126	126	126
Peace Year Splines	*No*	*No*	*No*	*No*	*No*
Clustered SEs (leaders)	✓	✓	✓	✓	✓
Only US Presidents	✓	✓	✓	✓	✓

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

Table A38: 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	−0.20

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

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.

As discussed in the main text, however, we took a number of steps to ensure that that our measures of leader hawkishness were generally consistent with expert consensus (when one existed) and that our results were not an artifact of one particular measurement strategy or leader.

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.

7.5 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 §5.8 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.

7.6 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 5.2 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.

8 Probing the Deliberation Mechanism

8.1 Seeking Counsel and Expressing Dissent

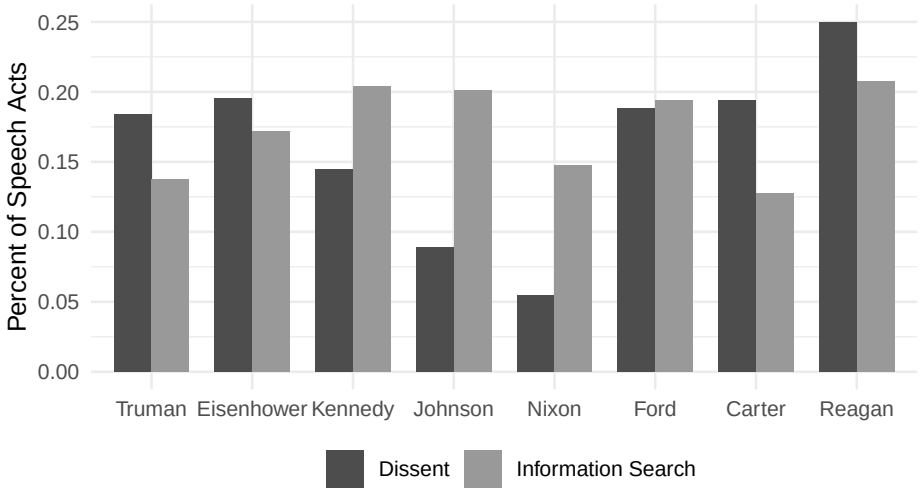
Drawing upon coding methodologies from studying deliberation elsewhere in political science (e.g., Parthasarathy, Rao and Palaniswamy, 2019), we developed a coding scheme to identify speech acts that exhibited *seeking counsel*, defined as a textual indication that the speaker requested input from another 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.[16]

Figure A17 plots deliberation patterns by administration. While seeking counsel and expressing 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 (McDermott, 2007, Chapter 6) restricted policy debates between advisers.

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

Figure A17: Deliberation Patterns across Administrations



8.2 Dictionary Approach to Speaker Topics

Table A39 displays the full set of terms used to count the prevalence of key topics of interest in our study. We generated an initial list of terms based on our substantive familiarity with how speakers addressed each topic within the corpus. Beyond capturing the topic of interest, we selected terms that could best differentiate between the topics. For instance, we selected active terms for *military violence* (e.g., attack, invade) while opting for era-relevant weapons for *military balance* (e.g., carrier, missile). We refined term lists through an iterative process that examined high scoring texts and dropping or adding terms to improve topic coherence. As is evident, the words are all stemmed. Counts of these stemmed terms are performed on stemmed versions of the original speech act texts. These counts are then divided by the total number of words in a text in order to produce a measure of the *proportion* of words associated with a topic.

In order to ensure that the analysis is not unduly affected by extremely short speech acts (where a single word from a topic could dramatically change the proportion), we take two additional measures. First, we aggregate speech acts to the meeting-adviser level, where all speech acts made by a single person in an individual meeting are combined into a single text. This converts our speech act data, which has 104,504 speech acts, into a dataset of 14,814 clusters of speech acts made by advisers in individual meetings. Second, we filter this dataset down to only observations with at least 50 words. This shrinks the dataset to 11,609 observations. Note, however, that these 11,609 observations account for 100,089 speech acts, which represent almost 96% of the original

Table A39: Key Terms for Dictionary Approach to Identifying Key Speaker Topics

Topic	Terms
Military violence	attack, bomb, fight, invad, strik, deploy, offens, retali, retaliatori, counterforc, escal
International threats	aggress, aggressor, anxieti, belliger, compet, enemi, hostil, risk, threat, war
Military balance	atom, bomber, capabl, carrier, cruis, fighter, icbm, mirv, missil, nuclear, satellit, silo, submarin, tank
Diplomacy	diplomaci, diplomat, forum, meet, negoti, peac, summit, talk, treati
Adversary interests	castro, china, communism, communist, khrushchev, mao, moscow, prc, soviet, stalin, ussr, vietnam

Table A40: Descriptive Statistics for Topic Propensities

Topic	Min.	Q1	Med.	Mean	Q3	Max.
Military violence	0.000	0.000	0.000	0.002	0.003	0.071
International threat	0.000	0.000	0.000	0.003	0.005	0.082
Military balance	0.000	0.000	0.000	0.003	0.002	0.094
Diplomacy	0.000	0.000	0.002	0.004	0.007	0.071
Adversary interests	0.000	0.000	0.002	0.006	0.009	0.091

data.

Table A40 reports the distribution of the five topic proportions. It is not surprising that these values are quite low and close to zero.

Table A41 displays results for regressions that use speech acts aggregated at the adviser-meeting level 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 acts. These models include two important control variables: whether the speech acts were made in a formal meeting, and whether the speech acts were recorded in the form of a transcript. We may expect topics raised to differ slightly depending on whether discussions take place in an official NSC meeting or a less formal environment. Further, we may believe that speech acts which are recorded in the form of transcripts are distinct from speech acts in the form of meeting minutes. The former is more spontaneous and could be made of shorter utterances, while the latter is a summary written after the fact and could thus be longer. Even with the inclusion of these controls, we see quite consistent evidence that advisers provide information that leans into topics which are consistent with their predispositions.

Table A41: Speaker Hawkishness and Speech Act Content

	Dependent variable:				
	Violence	Threat	Balance	Diplomacy	Adversary
	(1)	(2)	(3)	(4)	(5)
Speaker Hawkishness	0.0024***	0.0030***	0.0037***	−0.0034***	−0.0047***
	(0.0006)	(0.0006)	(0.0008)	(0.0007)	(0.0010)
Formal	0.0009***	0.0008***	0.0027***	−0.0015***	0.0032***
	(0.0001)	(0.0001)	(0.0002)	(0.0001)	(0.0002)
Transcript	0.0014***	−0.0002	0.0001	0.00004	−0.0010***
	(0.0002)	(0.0002)	(0.0002)	(0.0002)	(0.0003)
Eisenhower	0.0004**	0.0007***	0.0031***	−0.0003	0.0022***
	(0.0002)	(0.0002)	(0.0003)	(0.0002)	(0.0003)
Kennedy	0.0012***	−0.0002	0.0042***	0.0009***	0.0033***
	(0.0003)	(0.0003)	(0.0003)	(0.0003)	(0.0004)
Johnson	0.0041***	0.0013***	0.0022***	0.0008**	0.0055***
	(0.0003)	(0.0003)	(0.0003)	(0.0003)	(0.0004)
Nixon	0.0006**	0.0002	0.0035***	0.0015***	0.0041***
	(0.0003)	(0.0003)	(0.0004)	(0.0003)	(0.0005)
Ford	−0.0005*	−0.0001	0.0055***	0.0006*	0.0030***
	(0.0003)	(0.0003)	(0.0004)	(0.0004)	(0.0005)
Carter	−0.0003	0.0007	0.0020***	0.0030***	0.0030***
	(0.0004)	(0.0004)	(0.0006)	(0.0005)	(0.0007)
Reagan	−0.0012***	0.0001	0.0012***	0.0030***	0.0020***
	(0.0003)	(0.0003)	(0.0004)	(0.0004)	(0.0005)
Constant	−0.0007*	0.0008*	−0.0035***	0.0065***	0.0041***
	(0.0004)	(0.0004)	(0.0005)	(0.0005)	(0.0007)
Observations	11,307	11,307	11,307	11,307	11,307

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

8.3 President's Hawkishness and Topic Prevalence

One may be concerned that our findings, which indicate that hawkish advisers talk more about issues of military violence and international threats (and also that dovish advisers talk more about diplomacy and adversary interests), are driven by a president's strategic decisions about who attends the meetings. Perhaps a more hawkish (dovish) president simply invites more hawkish (dovish) individuals to the meetings, which results in more speech acts that are related to hawkish (dovish) topics.

To examine this possibility, we analyze the relationship between the president's hawkishness and the prevalence of topics discussed in the meetings that they oversee. We take our adviser-meeting data and calculate the average prevalence of each topic for each meeting. These calculations only use data on advisers; all observations reflecting presidents are omitted. These data are then analyzed using a series of regressions similar to those used in the main text. Because the president's hawkishness is the independent variable, administration fixed effects are taken out of the models.

Table A42 reports the results of these regressions. We see that more hawkish presidents tend to oversee meetings that have less discussion involving military violence, international threats, or adversary interests. Equivalently, more dovish presidents are more likely to have meetings that involve these three topics. The findings for military violence and international threat go against what we would expect to see if presidents simply chose advisers that were likely to discuss topics that hew toward the president's interests. The result for adversary interests does align with a potential story where hawkish (dovish) presidents hear less (more) from their advisers about a topic that is associated with being dovish. Nonetheless, the overall findings in Table A42 do not indicate a systematic attempt by leaders to only hear about topics that align with their underlying degree of hawkishness.

Table A42: President's Hawkishness and Speech Act Content

	Dependent variable:				
	Violence	Threat	Balance	Diplomacy	Adversary
	(1)	(2)	(3)	(4)	(5)
President's Hawkishness	-0.0067^{***}	-0.0030^{***}	-0.0020^{*}	0.0015	-0.0053^{***}
	(0.0009)	(0.0008)	(0.0011)	(0.0013)	(0.0015)
Formal	0.0009^{***}	0.0010^{***}	0.0024^{***}	-0.0016^{***}	0.0028^{***}
	(0.0002)	(0.0002)	(0.0002)	(0.0003)	(0.0003)
Transcript	0.0019^{***}	-0.00001	0.0003	0.0004	0.0005^{*}
	(0.0002)	(0.0002)	(0.0002)	(0.0003)	(0.0003)
Constant	0.0045^{***}	0.0039^{***}	0.0024^{***}	0.0046^{***}	0.0072^{***}
	(0.0004)	(0.0004)	(0.0005)	(0.0006)	(0.0007)
Observations	2,578	2,578	2,578	2,578	2,578

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

8.4 Propagating Uncertainty of Hawkishness Measures in Topic Analysis

In Appendix § 5.3, we explain how our main analysis does not fully propagate the uncertainty inherent to our predicted measures of hawkishness. Here, we investigate whether the uncertainty of our hawkishness measures has any bearing on our findings regarding the substance of advisers' counsel.

We replicate the analysis shown in Appendix § 8.2 (and visually summarized by Figure 6 in the main text). However, we now run this analysis on each of the 1,000 bootstrapped iterations of our hawkishness data, and we collect the coefficient estimates and standard errors from each of the models that is run.

Table A43 reports the average estimated coefficient and standard error for each variable in each model across all 1,000 iterations. Results are highly similar to those in Table A41 in Appendix 8.2 and Figure 6 in the main text.

Table A43: Hawkishness and Speech Act Content, Propagating Uncertainty from Bootstrapping

	Dependent variable:				
	Violence	Threat	Balance	Diplomacy	Adversary
	(1)	(2)	(3)	(4)	(5)
Speaker Hawkishness	0.002^{***}	0.003^{***}	0.004^{***}	-0.003^{***}	-0.004^{***}
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
Formal	0.001^{***}	0.001^{***}	0.003^{***}	-0.001^{***}	0.003^{***}
	(0.0001)	(0.0001)	(0.0002)	(0.0001)	(0.0002)
Transcript	0.001^{***}	-0.000	0.0001	0.0000	-0.001^{***}
	(0.0002)	(0.0002)	(0.0002)	(0.0002)	(0.0003)
Eisenhower	0.0004^{**}	0.001^{***}	0.003^{***}	-0.000	0.002^{***}
	(0.0002)	(0.0002)	(0.0003)	(0.0002)	(0.0003)
Kennedy	0.001^{***}	-0.000	0.004^{***}	0.001^{***}	0.003^{***}
	(0.0003)	(0.0003)	(0.0003)	(0.0003)	(0.0004)
Johnson	0.004^{***}	0.001^{***}	0.002^{***}	0.001^{***}	0.006^{***}
	(0.0003)	(0.0003)	(0.0003)	(0.0003)	(0.0004)
Nixon	0.001^{**}	0.0002	0.003^{***}	0.002^{***}	0.004^{***}
	(0.0003)	(0.0003)	(0.0004)	(0.0003)	(0.0005)
Ford	-0.001^{*}	-0.000	0.005^{***}	0.001^{*}	0.003^{***}
	(0.0003)	(0.0003)	(0.0004)	(0.0004)	(0.001)
Carter	-0.000	0.001	0.002^{***}	0.003^{***}	0.003^{***}
	(0.0004)	(0.0004)	(0.001)	(0.001)	(0.001)
Reagan	-0.001^{***}	0.0001	0.001^{***}	0.003^{***}	0.002^{***}
	(0.0003)	(0.0003)	(0.0004)	(0.0004)	(0.001)
Constant	-0.001	0.001^{**}	-0.003^{***}	0.006^{***}	0.004^{***}
	(0.0004)	(0.0004)	(0.001)	(0.0005)	(0.001)
Observations	11,307	11,307	11,307	11,307	11,307

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

8.5 Topic Model Approach to Speaker Topics

While the main text uses a straightforward dictionary-based approach to test whether hawkish and dovish advisers exhibit different speech patterns on these topics, an alternative approach would be to utilize topic modeling. To that end, we employ a semi-supervised text analysis method called the keyword-assisted topic model (keyATM) (Eshima, Imai and Sasaki, 2023). 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. The fitted model then assigns topic propensities to each unit of text (in our case, all speech acts at the adviser-meeting level).

8.5.1 Keywords

Table A44 enumerates all keywords provided to search for our five primary topics.[17]

Table A44: Supplied Keywords for KeyATM Topics

Topic	Terms
Military violence	attack, bomb, fight, invad, strik, deploy, offens, retali, retaliatori, counterforc, escal
International threats	aggress, aggressor, anxieti, belliger, compet, enemi, hostil, risk, threat, war
Military balance	atom, bomber, capabl, carrier, cruis, fighter, icbm, mirv, missil, nuclear, satellit, silo, submarin, tank
Diplomacy	diplomaci, diplomat, forum, meet, negoti, peac, summit, talk, treati
Adversary interests	castro, china, communism, communist, khrushchev, mao, moscow, prc, soviet, stalin, ussr, vietnam

Our keyATM model sought to identify these five topics but also allowed for 20 additional topics that the model would independently and inductively identify. Tables A45 and A46 list the most frequent terms that were associated with the 35 topics in the model. Words with "[X]" beside them represent terms that were included in that category in Table A44, while words with "[Letter]" represent terms that were associated with one of the five pre-defined topics but also showed up as common terms for other topics.

[17] This table of keywords is identical to the dictionary of terms we used in our main analysis; see Table A39.

Table A45: Most Frequent Terms for All keyATM Topics (Part 1)

Topic	Most Frequent Terms
Military violence (V)	general, secretari, vietnam [A], rusk, north, south, mcnamara, vietnames, bomb [X], will, wheeler, forc, attack [X], hanoi, militari, air, troop, said, men, clifford
International threat (T)	said, state, forc, general, militari, secretari, admir, war [X], unit, presid, point, defens, chief, radford, use, plan, weapon, view, action, agre
Military balance (B)	missil [X], soviet [A], secretari, will, capabl [X], agreement, limit, system, option, can, ford, number, forc, weapon, test, us, program, mirv [X], nuclear [X], defens
Diplomacy (D)	will, kissing, presid, go, get, want, can, think, us, say, said, talk [X], now, make, one, thing, give, know, see, time
Adversary (A)	soviet [X], mr, state, communist [X], govern, point, dull, situat, unit, countri, new, might, continu, world, said, now, polici, one, also, problem
Other_1	mr, said, dillon, herter, secretari, dull, lao, govern, cuba, congo, khrushchev [5], situat, report, un, french, forc, souvanna, plan, castro [A], phoumi
Other_2	unclear, note, think, ship, well, one, missil [B], uh, cuba, time, go, site, john, today, report, can, novemb, two, yes, soviet [A]
Other_3	secretari, countri, million, econom, year, aid, anderson, develop, problem, program, world, assist, will, oil, price, increas, foreign, trade, market, food
Other_4	israel, isra, arab, egypt, east, middl, egyptian, nasser, said, jordan, soviet [A], settlement, un, oil, situat, syria, move, british, sadat, state
Other_5	haig, scowcroft, agnew, sir, agent, connal, yes, use, chemic, right, biolog, think, one, martin, gas, now, tough, incapacit, warfar, year
Other_6	roger, schlesing, can, will, jone, ship, island, think, american, get, oper, use, sea, time, go, colbi, marin, weinberg, take, hour
Other_7	secretari, will, vanc, africa, pakistan, issu, us, brzezinski, note, indian, soviet [A], dr, need, india, african, south, support, un, problem, ambassador
Other_8	cuba, note, soviet [A], missil [B], cuban, inspect, castro [A], presid, khrushchev [A], state, stevenson, mr, weapon, point, meet [D], bundi, mccloy, ship, russian, use
Other_9	secretari, said, mr, lovett, state, council, presid, depart, clement, marshal, nsc, paper, propos, felt, acheson, suggest, staff, general, forrest, problem
Other_10	turkey, turk, greek, turkish, cyprus, ball, diem, greec, govern, mat, said, situat, general, militari, mr, forc, u., coup, presus, will

Notes: [X] indicates a supplied keyword for that topic; [Letter] indicates a keyword for a predefined topic.

Table A46: Most Frequent Terms for All keyATM Topics (Part 2)

Topic	Most Frequent Terms
Other_11	secretari, dull, state, french, unit, humphrey, british, said, council, point, propos, indochina, cours, confer, repli, posit, accord, hoover, nation, agreement
Other_12	nato, german, said, secretari, europ, forc, meet [D], european, french, mr, germani, berlin, de, gaull, presid, discuss, british, alli, franc, western
Other_13	dr, flem, program, said, satellit [B], space, killian, studi, scientif, stockpil, earth, reloc, scienc, defens, kistiakowski, point, present, warn, council, report
Other_14	smith, general, intellig, said, attorney, shultz, agenc, depart, servic, recommend, cia, oper, inform, secur, port, respons, case, committe, report, command
Other_15	presid, said, ask, vice, thought, meet [D], whether, peopl, discuss, make, ambassador, get, take, congress, want, go, suggest, question, agre, say
Other_16	dull, iran, secretari, iraq, said, rumsfeld, nasser, shah, iranian, japanes, oil, british, japan, situat, lebanon, east, middl, arab, govern, macarthur
Other_17	think, just, go, well, now, got, get, right, know, thing, say, yes, presid, can, want, yeah, one, good, come, like
Other_18	program, said, year, militari, assist, defens, countri, budget, increas, mr, cost, billion, requir, strauss, secretari, fund, develop, million, expenditur, product
Other_19	senat, said, russel, presid, pres, ball, fulbright, committe, mansfield, rostow, ask, want, stenni, bill, jackson, get, dirksen, sec, statement, talk [D]
Other_20	mr, paragraph, gray, council, nsc, polici, cutler, propos, board, plan, suggest, paper, languag, meet [D], read, defens, state, brief, action, general
Other_21	nixon, well, yeah, go, right, know, unclear, think, thing, just, got, now, mean, say, get, see, oh, henri, want, can
Other_22	presid, mr, said, test, secretari, ask, propos, discuss, meet [D], agre, mccone, statement, ambassador, state, question, inspect, suggest, thought, matter, inform
Other_23	secretari, trade, state, said, control, soviet [A], depart, acheson, mr, item, list, polici, agre, oil, alli, propos, action, export, bloc, countri
Other_24	unclear, think, go, say, well, get, now, just, thing, can, want, got, right, yeah, one, put, see, take, letter, realli
Other_25	presid, said, ask, state, whether, get, inquir, agre, go, comment, thought, suggest, view, question, express, vice, might, seem, make, discuss
Other_26	governor, stassen, shelter, program, unit, said, adoula, peterson, state, defens, squadron, hoegh, attack [V], civil, peopl, tshomb, fallout, inspect, feder, disarma
Other_27	secretari, wilson, defens, said, program, budget, depart, billion, forc, year, mcelroy, fy, million, mr, figur, cost, armi, militari, cut, expenditur
Other_28	said, mr, secretari, mann, gate, canal, ambassador, presid, panama, johnson, treati [3], thought, agre, ask, royal, state, panamanian, linowitz, felt, canadian
Other_29	latin, america, laird, american, cuba, cuban, need, allen, countri, secretari, castro [A], brazil, action, support, govern, problem, republ, will, econom, nicaragua
Other_30	mr, council, cutler, report, plan, state, nsc, program, defens, recommend, board, depart, point, meet [D], secretari, brief, presid, secur, present, polici

Note: [Letter] indicates a keyword for a predefined topic.

8.5.2 Topic Validation

In order to ensure that our keyATM produced reasonable results, we perform a validation test. For each of our five primary topics of interest, we identified the 70 adviser-meeting texts which the keyATM determined had the highest propensities for the topic. This produced a pool of 350 total adviser-meeting texts. In order to help with the qualitative coding which was to come, we focus only on texts of at least 100 words in length. This reduces our validation data to 239 observations.

These 239 total texts were given to a research assistant, who was provided qualitative information about the five topics and the keyATM, and tasked with attempting to identify which of the five key topics each text primarily reflected.

Table A47 shows the correspondence between the research assistant's determined topics and the actual topics. Overall accuracy is 72% (171 out of 239 are coded correctly). Table A48 reports more detailed metrics for each topic. Performance is quite good overall, especially given in mind that the coding task involves five classes rather than just two. The "Threat" topic exhibits the least effective performance; texts that the keyATM classifies as strong examples of threat tended to be coded as instances of diplomacy or violence; quite a few texts which reflected adversary interests were coded by the research assistant as examples of threat. Given that threats are often discussed in the context of the other topics (e.g., an adviser discussing how to write a diplomatic statement responding to a perceived threat, or an adviser describing the threats posed by an adversary's military vessels), such miscodings involving threats are not surprising. The last two rows of Table A48 reports the overall performance metrics if any adviser-meeting texts deemed to reflect threat are removed. The second-to-last row reports statistics if all texts that the keyATM identified as exhibiting the threat topic are removed; the final row reports statistics if all texts that the research assistant identified as exhibiting the threat topic are removed. Across both variations, we see that overall predictive performance markedly improves, particularly in accuracy and the F1 score.

Overall, these results lend substantial credence to the keyATM's ability to identify our primary topics of interest.

Table A47: Comparison of hand-coded topics with topics extracted from keyATM.

		Actual				
		Violence	Threat	Balance	Diplomacy	Adversary
Predicted	Violence	29	6	0	3	0
	Threat	9	13	3	2	14
	Balance	4	3	44	0	0
	Diplomacy	0	9	9	53	3
	Adversary	1	1	0	1	32

Table A48: Performance metrics for hand-coded topics.

Topic	**Accuracy**	**Precision**	**Recall**	**Specificity**	**F1**	**Total**
Violence	0.904	0.763	0.674	0.954	0.716	43
Threat	0.803	0.317	0.406	0.865	0.356	32
Balance	0.921	0.863	0.786	0.962	0.822	56
Diplomacy	0.887	0.716	0.898	0.883	0.797	59
Adversary	0.916	0.914	0.653	0.984	0.762	49
Overall	0.715	0.746	0.715	0.933	0.722	239
Overall (no actual "Threat")	0.763	0.891	0.763	0.963	0.816	207
Overall (no guessed "Threat")	0.798	0.807	0.798	0.932	0.840	198

8.5.3 Counsel Congruence with Predispositions

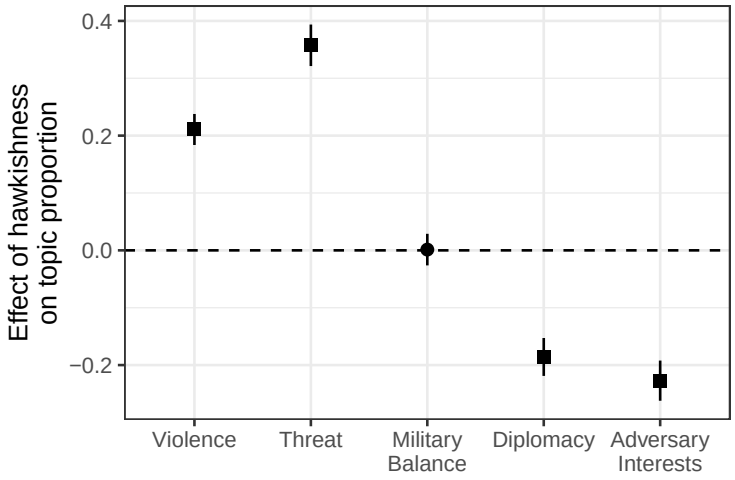
Table A49 displays results for regressions using topics as measured by our keyATM (instead of the dictionary-based approach). This replicates Table A41 in Appendix §8.2. Figure A18 below replicates Figure 6 in the main manuscript, except using the topic propensities drawn from our keyATM. We see that the vast majority of our results are unchanged. The one exception exists for military balance, which appears to not bear any strong relationship with advisers being hawkish or dovish.

Table A49: Hawkishness and Speech Act Content

	Violence	Threat	Balance	Diplomacy	Adversary
	(1)	(2)	(3)	(4)	(5)
Speaker Hawkishness	0.211***	0.357***	0.001	−0.186***	−0.227***
	(0.014)	(0.018)	(0.014)	(0.017)	(0.018)
Formal	−0.012***	0.035***	0.051***	−0.102***	0.011***
	(0.003)	(0.004)	(0.003)	(0.003)	(0.003)
Transcript	0.048***	−0.057***	−0.001	0.120***	−0.087***
	(0.004)	(0.005)	(0.004)	(0.004)	(0.005)
Eisenhower	0.002	−0.032***	0.028***	−0.022***	0.047***
	(0.005)	(0.006)	(0.005)	(0.006)	(0.006)
Kennedy	0.028***	−0.116***	0.049***	−0.012*	0.007
	(0.006)	(0.008)	(0.006)	(0.007)	(0.008)
Johnson	0.260***	−0.114***	0.039***	0.006	0.003
	(0.006)	(0.008)	(0.006)	(0.007)	(0.008)
Nixon	0.002	−0.163***	0.106***	0.189***	−0.0004
	(0.006)	(0.008)	(0.006)	(0.008)	(0.008)
Ford	−0.028***	−0.159***	0.170***	0.210***	0.001
	(0.007)	(0.010)	(0.007)	(0.009)	(0.009)
Carter	0.013	−0.133***	0.084***	0.088***	0.016
	(0.010)	(0.013)	(0.010)	(0.012)	(0.012)
Reagan	−0.040***	−0.204***	0.102***	0.133***	0.062***
	(0.007)	(0.009)	(0.007)	(0.008)	(0.009)
Constant	−0.094***	0.030**	−0.034***	0.251***	0.301***
	(0.009)	(0.012)	(0.009)	(0.011)	(0.012)
Observations	11,293	11,293	11,293	11,293	11,293

Notes: OLS regression with the adviser-meeting as the unit of analysis. Hawkishness scores are from the boosted linear approach and outcome scores are topic proportions from the keyATM analysis.

Figure A18: Effect of Speaker Hawkishness on Topic Proportions in Meeting-Adviser Speech Acts, Using KeyATM Model



Note: Plot shows marginal effect of moving from the least to most hawkish speaker within a fixed administration. Bands represent 95% confidence intervals. Square points indicate 95% statistical significance.

8.6 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 dissenting 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 address this concern in two ways. First, we use our dissent dataset, described in the main text. 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.

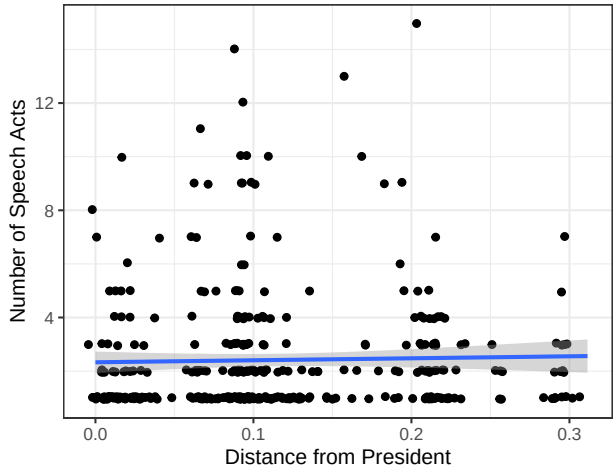


Figure A19: Relationship (jittered) between speech acts expressing dissent and the speaker's distance from the president in terms of hawkishness.

Figure A19 displays the distribution of these distances for speech acts that either do or not do express dissent across 270 randomly selected meetings. We do not find any evidence of a link between relative positions in hawkishness and willingness to express dissent during meetings. A simple bivariate Poisson regression indicates no statistically significant relationship ($p = 0.908$).

Second, we can also study self-censorship another way. The above analysis focuses on the probability of expressing dissent as a function of dispositional distance from the president. Yet self-censorship might also manifest itself in terms of advisers choosing not to speak at all. This is

difficult to capture with the full meeting dataset due to data restrictions: most meeting records only document attendees who participated in the discussion, rather than listing the full list of attendees. However, the Eisenhower administration does feature relatively comprehensive information on meeting attendance, irrespective of active participation.

For each NSC meeting in the Eisenhower administration, we identified every individual who attended. A total of 9,850 meeting-individual observations exist in these data. We were able to identify the name of the individual and/or their hawkishness in 83% of these observations. The remaining 17% are not identifiable, as the records only mention the attendee's position or department in a vague manner that disallows identification, or people who very infrequently attended meetings. (Examples include "Acting Secretary," "Staff Secretary," "Defense," "for the Chairman of the Joint Chiefs of Staff," and the like.) Our final dataset thus features 8,193 meeting-individual observations.

For each identified adviser, we calculate the difference between their predicted hawkishness measure and our hawkishness measure for President Eisenhower. The absolute value of this difference in hawkishness is what we call "distance from president." We can then analyze what effect this difference has, if any, on an adviser's willingness to speak in an NSC meeting.

Table A50 presents the results from two logit models, which model whether a meeting attendee spoke up at a meeting. We include a set of control variables in both models: whether the adviser is more hawkish than the president, and whether the adviser has a lower-level position (defined as positions underneath the cabinet level according to the United States Order of Precedence). In the second model, we also control for the logged number of meeting attendees. Across both models, we see that an important predictor of meeting participation is rank; individuals in lower-level positions are less likely to speak. Individuals are also less likely to speak as the number of meeting attendees increases. In both models, however, an adviser's distance from the president has a positive impact on whether they choose to speak in a meeting. Distance from the president in terms of hawkishness therefore does not appear to be responsible for outright self-censorship. Indeed, the effect goes *against* the notion that advisers remain silent when they hold policy stances that are distinct from the president.

Table A50: Propensity of Eisenhower Advisers to Speak in Meetings

	Speaks in Meeting	
	Logistic	
	(1)	(2)
Distance from President	4.499***	4.584***
	(1.506)	(1.521)
More Hawkish than President	0.291	0.263
	(1.049)	(1.059)
Lower-level Position	−1.174***	−1.138***
	(0.207)	(0.208)
Number of Attendees (logged)		−1.165***
		(0.132)
Constant	−0.004	3.888***
	(0.269)	(0.516)
Random Effects (Actor)	✓	✓
Observations	8,193	8,193

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

9 Agenda Items and Hawkishness

Table A51 displays the distribution of agenda items that were mentioned in at least one speech act in each meeting. Table A52 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 A51: Distribution of agenda items across all meetings.

Agenda Item	**No**	**Yes**
USSR	2,468	217
Asia	2,221	464
Middle East	2,321	364
Economy	2,459	226
Europe	2,307	378
International Institutions	2,648	37
Intelligence	2,611	74
Strategic Forces	2,496	189
Americas	2,431	254
Defense	2,426	259
Diplomacy	2,586	99
Organization	2,615	70
Vietnam	2,280	405
Policy	2,531	154
China	2,594	91
Africa	2,586	99
Latin America	2,669	16
Arms Control	2,555	130
North Africa	2,684	1
Other	2,653	32

Table A52: OLS regressions on the relationship between meeting topics and average hawkishness of meeting participants.

	Dependent variable:		
	Mean Hawkishness		
	All	Formal	Informal
	(1)	(2)	(3)
USSR	0.006**	−0.003	0.016
	(0.003)	(0.003)	(0.011)
Asia	0.010***	0.010***	0.015
	(0.002)	(0.002)	(0.011)
Middle East	−0.001	−0.003	0.005
	(0.002)	(0.002)	(0.011)
Economy	−0.001	−0.003	0.014
	(0.007)	(0.005)	(0.019)
Europe	0.003	0.001	0.010
	(0.002)	(0.002)	(0.011)
International Institutions	0.003	−0.008	0.012
	(0.006)	(0.006)	(0.013)
Intelligence	0.003	−0.002	0.018
	(0.004)	(0.004)	(0.013)
Strategic Forces	0.014***	0.001	0.024**
	(0.003)	(0.003)	(0.012)
Americas	0.003	0.0003	0.012
	(0.003)	(0.003)	(0.011)
Defense	0.016***	0.005**	0.025**
	(0.003)	(0.002)	(0.012)
Diplomacy	−0.001	−0.002	0.005
	(0.004)	(0.004)	(0.012)
Organization	0.012***	−0.013***	0.033***
	(0.004)	(0.004)	(0.012)
Vietnam	0.016***	0.006	0.024**
	(0.003)	(0.004)	(0.011)
Policy	0.027***	−0.001	0.047
	(0.004)	(0.003)	(0.039)
China	0.001	−0.011***	0.010
	(0.004)	(0.004)	(0.012)
Africa	0.001	−0.002	−0.00001
	(0.004)	(0.004)	(0.012)
Latin America	0.003	0.013	0.005
	(0.009)	(0.016)	(0.015)
Arms Control	0.005	−0.002	0.016
	(0.003)	(0.004)	(0.012)
North Africa	−0.024	−0.017	
	(0.036)	(0.024)	
Other	0.002	−0.005**	0.009
	(0.003)	(0.002)	(0.011)
Formal	0.013***		
	(0.002)		
Constant	0.488***	0.510***	0.537***
	(0.020)	(0.013)	(0.024)
Year FEs	✓	✓	✓
Observations	2,685	791	1,894

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. 2010. "The Myth of American Isolation." *Foreign Policy Analysis* 6:349–371.

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. 2023. "Keyword Assisted Topic Models." *American Journal of Political Science* Forthcoming.

Etheredge, Lloyd S. 1978. *A World of Men: The Private Sources of American Foreign Policy*. Cambridge: The MIT Press.

Feaver, Peter D. and Christopher Gelpi. 2011. *Choosing Your Battles: American Civil-Military Relations and the Use of Force*. Princeton: Princeton University Press.

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. 2021. "All the World's a Stage: US Presidential Narcissism and International Conflict." *International Studies Quarterly* 65(3):825–837.

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.

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. 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. 2023. *Bureaucracy and War: The Institutional Origins of Miscalculation in International Politics*. New York: Cambridge University Press.

Jost, Tyler, Kaine Meshkin and Robert Schub. 2022. "The Character and Origins of Military Attitudes on the Use of Force." *International Studies Quarterly* 66(2):sqac005.

Kertzer, Joshua D. 2013. "Making Sense of Isolationism: Foreign Policy Mood as a Multilevel Phenomenon." *Journal of Politics* 75(1):225–240.

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.

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.

Kertzer, Joshua D., Marcus Holmes, Brad L. LeVeck and Carly Wayne. 2022. "Hawkish Biases and Group Decision Making." *International Organization* 76(3):513–548.

Klingberg, Frank L. 1952. "The Historical Alternation of Moods in American Foreign Policy." *World Politics* 4(2):239–273.

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.

Krasner, Stephen D. 1972. "Are Bureaucracies Important? (Or Allison Wonderland)." *Foreign Policy* (7):159–179.

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.

McDermott, Rose. 2014. "The Biological Bases for Aggressiveness and Nonaggressiveness in Presidents." *Foreign Policy Analysis* 10(4):313–327.

McDermott, Rose, Dustin Tingley, Jonathan Cowden, Giovanni Frazzetto and Dominic D. P. Johnson. 2009. "Monoamine Oxidase A Gene (MAOA) Predicts Behavioral Aggression following Provocation." *Proceedings of the National Academy of Sciences* 106(7):2118–2123.

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.

Posen, Barry. 1984. *The Sources of Military Doctrine: France, Britain, and Germany between the World Wars.* Ithaca: Cornell University Press.

Potter, Philip B.K. 2007. "Does Experience Matter? American Presidential Experience, Age, and International Conflict." *Journal of Conflict Resolution* 51(3):351–378.

Preston, Thomas. 2001. *The President and His Inner Circle: Leadership Style and the Advisory Process in Foreign Policy Making.* New York: Columbia University Press.

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. 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.

Saunders, Elizabeth N. 2022. "Elites in the Making and Breaking of Foreign Policy." *Annual Review of Political Science* 25:219–240.

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.

Snyder, Jack. 1989. *The Ideology of the Offensive: Military Decision Making and the Disasters of 1914*. Cornell University Press.

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.