

The Credibility of Public and Private Signals: A Document-Based Approach

AZUSA KATAGIRI *Nanyang Technological University*

ERIC MIN *University of California, Los Angeles*

Crisis bargaining literature has predominantly used formal and qualitative methods to debate the relative efficacy of actions, public words, and private words. These approaches have overlooked the reality that policymakers are bombarded with information and struggle to adduce actual signals from endless noise. Material actions are therefore more effective than any diplomatic communication in shaping elites' perceptions. Moreover, while ostensibly "costless," private messages provide a more precise communication channel than public and "costly" pronouncements. Over 18,000 declassified documents from the Berlin Crisis of 1958–63 reflecting private statements, public statements, and White House evaluations of Soviet resolve are digitized and processed using statistical learning techniques to assess these claims. The results indicate that material actions have greater influence on the White House than either public or private statements; that public statements are noisier than private statements; and that private statements have a larger effect on evaluations of resolve than public statements.

Scholarship on international relations has long discussed the differing effects, if any, between public and private forms of diplomacy. Following the work of Schelling (1966), academics have broadly accepted that the credibility of threats made during crises is tied to their costliness. Despite the memorable nature of public gestures such as troop mobilizations, blockades, and the like, they are relatively infrequent. Much of the everyday administration of interstate diplomacy never rises to such dramatic heights, but instead remains in the realm of verbal exchanges.

Audience cost theory (Fearon 1994; Schultz 2001; Smith 1998; Tarar and Leventoglu 2012) has extended the logic of costly signals to this more peaceful arena, suggesting that public statements carry more weight by means of tying hands through highly visible commitments (Schelling 1960; Snyder 1972). Diplomatic communication that occurs out of view is comparatively cheap, as actors suffer fewer sanctions for renegeing on their claims. Another contingent of work has challenged

the general premise, utility, and evidence of audience cost theory (Downes and Sechser 2012; Snyder and Borghard 2011; Trachtenberg 2012). Despite active debate on the relative efficacy of public and private diplomatic signaling, few empirical studies beyond formal models and surveys have firmly put this comparison to a test, nor have they ascertained whether private channels can be just as costly as public ones (Kurizaki 2007).

Costs are a dimension to shaping a signal's credibility, but they are certainly not the only one. Many studies on crisis bargaining assume that all sides are focused on a specific signal in question and have abstracted away the actual informational environment in which signals are sent—one that is defined by tremendous volumes of information and purposes, all of which makes it difficult to separate actual signals from the commotion of everyday governance. Scholars have often overlooked how levels of *noise* vary across different forms of diplomatic signaling. This is a critical and consequential distinction. In the frenzied setting of real-life diplomacy, private statements are a more direct and concise manner in which states can communicate. The ability to ostensibly tie one's hands through a public statement is undermined by the higher degree of noisiness that public messages must overcome to be perceived as credible, or perceived at all.

In this article, we use the Berlin Crisis of 1958–63 as a testing ground for these claims regarding the complexity of the actual diplomatic environment. Full sets of declassified documents provide an exhaustive view of the United States government's evaluations of threat during this period, which was arguably one of the most dangerous times for American—if not global—security. We digitize, process, and analyze over 18,000 documents from the Department of State, Foreign Broadcast Information Service, and White House to create dynamic data on private and public diplomatic signals sent from the Soviet Union and East Germany to the United States, as well as a measure of American evaluations of Soviet resolve with respect to Berlin. These data allow for a uniquely detailed investigation of conflict dynamics.

Azusa Katagiri, Assistant Professor of Public Policy and Global Affairs, School of Social Sciences, Nanyang Technological University, azusak@ntu.edu.sg.

Eric Min, Assistant Professor of Political Science, 3254 Bunche Hall, University of California, eric.min@ucla.edu.

Names are in alphabetical order. We thank Ken Schultz, Jim Fearon, Scott Sagan, Justin Grimmer, Janet Box-Steffensmeier, Jeffrey Friedman, and Robert Gulotty; participants at the 2015 Annual Meeting of the Society for Political Methodology, 2015 Annual Meeting of the American Political Science Association, 2017 Annual Convention of the International Studies Association, and 2017 Annual Meeting of the Midwest Political Science Association; and Sabine Carey and our four anonymous reviewers for their thoughtful comments and encouragement. Azusa Katagiri acknowledges the support of the Suntory Foundation. Eric Min expresses gratitude for the support of the National Science Foundation's Graduate Research Fellowship #DGE-114747. Lastly, we both thank the Stanford Institute for Research in the Social Sciences, as well as Mai Hasegawa for valuable research assistance. Replication files are available at the American Political Science Review Dataverse: <https://doi.org/10.7910/DVN/1ZOEEX>.

Received: June 9, 2017; revised: April 25, 2018; accepted: August 13, 2018. First published online: November 13, 2018.

The new resources collectively challenge some primary facets of contemporary crisis bargaining theory. Three findings are especially salient. First, the intense scholarly debate about the relative effects of public and private diplomatic signals may be overstated. In the shadow of costly material actions, neither public nor private statements have a substantively noteworthy effect on shaping American perceptions of threat. Second, private diplomatic statements are more focused than public statements. Third, *private statements are more effective than public statements in affecting the perceived credibility of threats.*

Our study makes two primary and coequal contributions. First, we propose an alternative set of theoretical mechanisms for why private diplomatic signals are nontrivial and perhaps a more useful source of information. Second, we present the first systematic and quantitative analysis of the competing effects of different signals during crisis diplomacy. Our process establishes a framework that can be applied to other crises, opening the door to more empirically driven scholarship on the machinations of interstate interactions.

THE TRANSMISSION OF SIGNALS

Signals are statements or actions that convey information with the intent to influence a receiver's image of the sender (Jervis 1976). During a crisis, adversaries send each other signals about their intentions with hopes of reaching a more favorable bargain for themselves.

Signals do not always succeed in transmitting an intended message. The sender and the receiver each face unique obstacles in realizing this goal. Because senders have incentives to misrepresent their resolve and engage in cheap talk, they face a challenge of making their signals or threats credible to a receiver. This leads to the well-established rationalist claim that signals more credibly reflect private information when they are costly (Schelling 1966). Such costs can emerge in two ways: either through tying hands or sinking costs. The former reshapes costs for future decisions, while the latter immediately burns material resources (Fearon 1997).

This contrast typically breaks down into a distinction of words versus deeds. Words can be separated even further to public and private statements. This is especially salient for audience cost theory, which suggests that public declarations carry more credibility because they tie hands and implicate potential future costs for backing down (Fearon 1994; Schultz 2001; Tomz 2007).¹

This logic of costly signaling would predict that private statements are generally the least credible form of signal because they involve neither form of cost (Fearon 1995; Smith 1998; Tingley and Walter 2011). Indeed, costless communication should be least effective when

the sender and receiver have divergent interests (Kydd 2003; Morrow 1994). The relative credibility of public statements or threats (hand-tying) and public actions (cost-sinking) is not clear in periods preceding crisis initiation. During a crisis, however, reshaping future costs through hand-tying is more effective in augmenting the ex ante probability of conflict, and therefore more credible and likely to stem further escalation compared to material costs that are sunk immediately (Fearon 1994, 1997; Fuhrmann and Sechser 2014).² As such, public statements should wield greater credibility than public actions in the midst of a crisis. We thus get the following ranking of signal credibility:

Private statements < Actions ≤ Public statements

Some recent literature argues that private statements can also be costly because policymakers value maintaining a reputation for honesty and having the ability to avoid public escalation (Guisinger and Smith 2002; Kurizaki 2007; Ramsay 2011; Sartori 2002, 2005; Yarhi-Milo 2013).³ These claims complicate the logic and implications of costly signaling theory, and they make predictions unclear or more dependent on the specific crisis's context.

Moreover, a signal only impacts behavior or beliefs when a receiver processes it. While rational approaches tend to assume that signals are common knowledge and properly understood (Kertzer 2016; Kurizaki 2016), we know that receivers do not always interpret a sender's signal in the intended manner (Lebow 1981; Levy 1983; Quek 2016; Snyder and Diesing 1977). Explanations for signal misperception generally stem from psychological and cognitive approaches that focus on the individual. Biases and limitations of the human mind result in bounded rationality and misinterpretation of observed information (Simon 1947). In many cases, this misinterpretation is nonrandom and hews to the predispositions of the particular receiver, fueling belief perseverance and confirmation bias (Duelfer and Dyson 2011; Jervis 1976; Mercer 1996).

A Noisy Intersection

The previous overview indicates that costly signaling and misperception research emphasize distinct insights. Costly signaling theories point out that senders have multiple channels of communication with a receiver, and that these channels likely vary in their impacts. Meanwhile, studies of misperception note that receivers do not always properly interpret these signals. These two perspectives focus on different actors and forms of uncertainty and are not mutually exclusive (Kurizaki

¹ The distinction between words and deeds also has precedent in the literature that challenges audience cost theory. Levy et al. (2015) and Snyder and Borghard (2011), for example, debate the merits of audience costs by discussing the consistency between a leader's public statements and subsequent actions taken.

² We acknowledge that this discussion of public words and actions is stylized and based on ideal types. Actions likely have some residual hand-tying effects. Nevertheless, statements are unlikely to involve material costs. Our view of actions as mainly cost-sinking is not new (see Fearon (1997) and Slantchev (2011)). This perspective provides an imperfect but tractable way to operationally distinguish hand-tying and cost-sinking.

³ Seminal economic models also show how cheap talk can be informative (Crawford and Sobel 1982; Farrell and Gibbons 1989).

2016). However, little has been done to bridge the gap between them. We particularly lack an understanding of how the communication channel used to send a signal can impact its (mis)perception, and thus, its credibility.

A signal's effect not only rests with its costliness or its relation to an individual's predispositions but also on how well its key message stands out from the background. A signal must overcome *noise*—that is, fluctuations of false or irrelevant information that hinder perception of an intended signal.

We are not the first to suggest that noise impedes signaling. A host of studies, mostly using formal models, have analyzed how noise can impact strategic interactions in bargaining (Fey and Ramsay 2007; Handel 1977; Johns 2006; Meirowitz and Sartori 2008; Slantchev 2006). That said, these studies only model one channel of information, and thus, one configuration of noise.

We contend that the three aforementioned channels—actions, public statements, and private statements—vary in the degree of noise surrounding their signals, which influences policymakers' ability to process them. Such distinctions rely on an understanding of the real-world practice and nature of diplomacy. Despite the risk of making our view of crisis bargaining less parsimonious, the implications we derive challenge conventional wisdom about the credibility of signals.

DIPLOMATIC SIGNALS IN PRACTICE

Signals can only affect beliefs of a potential receiver when they are noticed and then interpreted.⁴ The quantity and quality of signals affect each of these steps.

Quantity

Most rationalist studies of crisis bargaining presume that a signal arises in a vacuum so that both the sender and receiver recognize and evaluate it (Kurizaki 2016).⁵ This seems innocuous, but all signals are part of a larger flurry of bureaucratic, political, and administrative activity. In such a frenetic and multidimensional environment, policymakers constantly struggle to keep up with the deluge of incoming raw information—not only for one single issue, but many at once.⁶ Kissinger (1979) makes this point:

High office teaches decision-making, not substance. Cabinet members are soon overwhelmed by the insistent demands of running their departments....The novice Secretary of State thus finds on his desk not policy analyses or

options but stacks of dispatches which he is asked to initial and to do so urgently, if you please. He can scarcely know enough about all the subjects to which they refer, or perhaps about any of them, to form an opinion. (30–31)

Intelligence communities are acutely concerned with the ability to filter signals from noise, which is commonly referred to as “the Roberta Wohlstetter problem” (Dahl 2013; Wohlstetter 1962).⁷ Larger amounts of data can lead to information overload (Simon 1947), which engenders selective attention and incomplete updating (Jones and Baumgartner 2005; Yarhi-Milo 2014), all of which causes poorer decision-making (Holsti, Brody, and North 1964; Schroder, Driver, and Streufert 1967).

The problem of processing too much information becomes exacerbated as more intermediaries become involved in compiling and filtering signals to send up the chain of command (Finel and Lord 1999). Each overwhelmed individual is less likely to correctly find, interpret, and report signals to their superiors, especially during times of crisis (Snyder and Diesing 1977). Even at the highest levels of decision-making, actors that ostensibly get the most filtered versions of information—including the National Security Adviser and the Secretary of State—are hard-pressed to tread water (Brzezinski 1983).

In these cases of informational overload, much more visible and vivid signals will be likelier to break through the noise (Vertzberger 1990; Yarhi-Milo 2014). We would thus expect that policymakers take greater notice of material actions, which are less frequent and more conspicuous, rather than any statements. This motivates the first hypothesis:

Hypothesis 1. *Public and private diplomatic statements both have smaller effects on evaluations of resolve than material actions.*

Quality

Hypothesis 1 predicts that public and private statements are less influential on elites' beliefs than actions, but this does not suggest that both forms of diplomatic communication are equally (in)effective. Audience cost theory avers that public statements are better suited to tie leaders' hands and thus generate credible commitments (Fearon 1994; Schelling 1966; Schultz 2001; Tomz 2007). However, in a more realistic diplomatic environment, public signals have two characteristics that undercut this assertion.

First, and related to the previous discussion, public pronouncements are relatively high in volume. The Central Intelligence Agency's Foreign Broadcast Information Service (FBIS), to which we return later, was established to collect and translate all publicly available information such as official statements, press releases, and radio broadcasts emanating from adversarial nations (Roop 1969). For the issue of Berlin alone, the FBIS records over 10,700 entries from the Soviet Union between 1958 and 1963.

⁴ Wickens (1992) makes a similar point regarding human information processing.

⁵ Note that the standard crisis bargaining model may include uncertainty over preferences, but a State *A*'s challenge of the status quo is unambiguously understood as such and forces a State *B* to respond (Lewis and Schultz 2003).

⁶ Of 172 National Security Council meetings between 1958 and Kennedy's assassination in November 1963, only 13 (about 7%) explicitly discuss Germany or Berlin. Of these 13 meetings, only two focus solely on the topic.

⁷ The 9/11 Commission Report also heavily references Wohlstetter's original work on the Pearl Harbor attack.

Second, public statements must inevitably deal with multiple audiences. A sender may intend to relay a specific message to a specific receiver, but nothing can stop other domestic or international actors from noticing, processing, and acting upon this signal (Lake 2010/11; Vertzberger 1990). For example, Nixon's "silent majority" speech on November 3, 1969 was an appeal to the American public to support peace with honor in Vietnam. This did not stop the North Vietnamese from hearing the speech and accusing the United States of using "perfidious tricks" to purposefully undermine secret peace talks taking place at that time. In public statements of their own, the North Vietnamese called Nixon "stupid and naive" and said his speech was "a pack of lies to justify the war of aggression against the Vietnamese people."⁸

Despite the inability to limit who sees a public statement, a government will still endeavor to communicate with particular audiences. A state's desire to communicate with each of these different constituencies will likely produce an overall set of messages that appear unfocused and perhaps contradictory (Jönsson 1996; Mitchell 2000). Even when a public declaration appears strident, the fact that it comes from such a mixed and multipurpose pool can make decision-makers discount its relevance. A wide spectrum of statements also raises the likelihood of misperception since individuals can more easily engage in their own forms of confirmation bias.

In that context, private signals prove useful because they allow states to send more direct, restrained, and selective messages of intent to a receiver. Diplomatic communities emphasize the importance of the *tête-à-tête*—"head-to-head" conversations—as a manner to communicate without distractions or political theater (Perlmutter 1975; Russell 2000). This exact line of reasoning motivated the now-famous exchange of letters between President Kennedy and Chairman Khrushchev in late 1961, as both leaders attempted to address immense anxieties and escalating rhetoric regarding Berlin. Responding to Khrushchev's initial message from September 29, Kennedy wrote back the following on October 16:

I am gratified by your letter and your decision to suggest this additional means of communication. Certainly you are correct in emphasizing that this correspondence must be kept wholly private, not to be hinted at in public statements, much less disclosed to the press....I think it is very important that these letters provide us with an opportunity for a personal, informal but meaningful exchange of views. There are sufficient channels now existing between our two governments for the more formal and official communications and public statements of position....Neither of us will be induced by a letter to desert or subvert his own cause. So these letters can be free from the polemics of the "cold war" debate. That debate will, of course, proceed, but you and I can write messages which will be directed only to each other.

⁸ See "Hanoi Charges Nixon of Duplicity on Secret Talks" in the *New York Times*, November 7, 1969.

A more detailed historical case may lend more evidence. On October 20, 1973, Kissinger arrived in Moscow and agreed to establish a ceasefire to stop the Yom Kippur War, which had erupted two weeks earlier. After only one or two days, Israel violated the ceasefire and proceeded to encircle the Egyptian Army—an event that both the Americans and Soviets had sought to avoid. Kissinger feared that the Soviets would see this as a purposeful act of deception (Blechman and Hart 1988). Indeed, the Soviets alerted their forces, and Brezhnev publicly stated that the USSR would unilaterally impose a ceasefire if the United States chose not to join the effort. Around that time, American intelligence indicated that a Soviet ship carrying radioactive material was near American vessels in the Mediterranean.

At 11:41 P.M. on October 24, Nixon's administration publicly put all military commands on high alert, or DEFCON 3, to show American resolve to the Soviets (Sagan 1985). By October 25, the global DEFCON 3 was widely reported in the media (Kissinger 1982). Brezhnev and the Politburo were perplexed by this seemingly unprovoked action. Brezhnev had no intent of engaging in conflict and believed that the United States would read through the lines of his fierce public rhetoric. Moreover, many of the puzzled Soviet leaders concluded that Nixon's alert was designed for domestic political purposes to distract from the unraveling Watergate ordeal, while others underestimated the severity of a DEFCON 3 alert (Lebow and Stein 1994). Had it not been for Brezhnev's personal opposition toward military escalation, the Politburo was likely to have mobilized more forces in response to the nuclear alert.

Both the United States and Soviet Union misread the intent of their adversary's public signals. It was only through private communications, free of posturing, and second-guessing, that the standoff was peacefully resolved.

Private statements may therefore be relatively more focused and informed than public statements. This argument rests on an observable mechanism: The need to cater to multiple audiences will cause elites to produce an enormous number of public signals and interpretations that, when considered together, generate a noisier overall impression than that produced through private channels alone.

This leads to two related hypotheses.

Hypothesis 2. *Public statements feature greater variance across signals than private statements.*

Hypothesis 3. *Private statements are relatively more effective than public statements in shaping evaluations of resolve.*

The theoretical foundation of Hypothesis 2 is, to our knowledge, unique in contemporary theories of crisis bargaining. Although some scholars have pointed out that diplomatic signals are harder to interpret than many rationalist theories presume (Barston 1988; Lebow 2001), no relative distinctions of this sort have been made between public and private channels. Most arguments on public and private diplomacy tend to be based on costs. Little to no emphasis is placed on

practical and inadvertent considerations of noise in diplomatic signaling.

In summary, our discussion produces a ranking of signal credibility that stands in direct contrast with costly signaling and audience cost theories:

Public statements < Private statements < Actions

Our argument provides another practical justification for why private diplomacy exists and remains valuable. Perhaps due to the intuitive strategic logic of public hand-tying, even scholars promoting the importance of private diplomacy only claim that private communications can be *just* as effective as public ones *under certain circumstances*. Emphasizing the noisiness of signals leads to a stronger claim.

DATA

We evaluate our hypotheses using four sets of data that directly correspond to material (cost-sinking) actions, private statements, public statements, and elite assessments of these signals. The Berlin Crisis of 1958 to 1963 supplies these resources.

A document-based, within-case research design provides the most direct test of our claims.⁹ Declassified documents supply raw material to capture the concepts of interest to our framework at a level of temporal and conceptual precision that would not be feasible in a larger-*N* design. The application of supervised learning methods can also reveal insights that may be missed, either by bias or oversight, through a purely qualitative approach.

The Berlin Crisis

The Berlin Crisis engendered grave concerns about the outbreak of nuclear conflict and was arguably one of the most serious periods of sustained tension in recent history.

As World War II ended, the four victorious powers—the United States, the United Kingdom, France, and the Soviet Union—divided the German capital into four sectors. The Soviet sector stood alone as East Berlin, while the remaining three were jointly considered West Berlin. The rest of Germany was also split on similar terms. Berlin (and thus, the Allies' West Berlin) lay deeply embedded in East Germany, more than 100 miles behind Soviet lines. The Western powers originally established a presence in Berlin because they assumed all of Germany would be overseen collectively. However, mounting Cold War tensions increased the salience of the boundary between West and East Germany and made Western presence in Berlin increasingly troublesome and geographically symbolic (Trachtenberg 1999).

For many years after, but especially between 1958 and 1963, the right of Western access to West Berlin was

treated as a fundamental testing ground of resolve. In a letter to Secretary of State John Foster Dulles, American ambassador to West Germany James B. Conant dubbed Berlin a “superdomino” for which American weakness would reverberate across the entirety of Germany and Europe (DiGangi et al. 1992, 376–81).

The Western allies began efforts to reform West German currency in 1948. At that time, the Soviet Union initiated a blockade that closed ground routes in and out of West Berlin, forcing the three allies to deliver supplies using the Berlin airlift. In 1949, the Western allies helped found the Federal Republic of Germany, or FRG (which technically excluded West Berlin, even though these sectors publicly aligned themselves with the West), and the Soviets helped establish the German Democratic Republic, or GDR. The GDR declared its capital to be East Berlin. The Western powers refused to recognize what they saw as an intentionally provocative choice. The FRG established its capital in Bonn.

Restricted access to West Berlin underscored the escalated tension from 1958 to 1963.¹⁰ Leading up to 1958, Soviet Premier Khrushchev had grown weary of diplomatic tap-dancing regarding Berlin, which appeared to be going nowhere (Kempe 2011). On November 10, 1958, Khrushchev made his first ultimatum, publicly granting Eisenhower and the Western allies six months to withdraw from and demilitarize their share of Berlin (Williamson 2012). If this did not occur, the Soviets would turn all lines of communication and control over to East Germany, threatening all Western access to West Berlin.¹¹ As early as March 1959, Acting Secretary of State Christian Herter (working in the stead of an ailing John Foster Dulles) indicated that the United States would have to issue the “ultimate threat” of nuclear war to defend its interests (Burr 1994).

By May of 1959, the Western allies had remained resolute, and Khrushchev had withdrawn his ultimatum. A meeting of the four foreign ministers in July failed to make significant progress on the Berlin question but ended on a note of mutual desire for a peaceful resolution to be further discussed at a Paris summit planned for May 14, 1960. However, due to the U-2 Incident on May 1, in which an American reconnaissance plane was shot down over Soviet territory, this meeting was dead on arrival (Barker 1963).

The Kennedy administration took office in January 1961 and initially adopted a reactive stance on Berlin. A meeting in Vienna between Khrushchev and Kennedy on June 4, 1961 started on friendly terms but quickly devolved when Khrushchev, according to Kennedy

⁹ Trager (2017) uses a similar strategy to analyze diplomatic communications among the European great powers between 1900 and 1914.

¹⁰ Many references focus on late 1961 as the Berlin Crisis. We use the more inclusive view where Khrushchev's ultimatum is the key trigger event. Nevertheless, as we show later, our main findings are equally valid for this shorter period.

¹¹ Zubok (1993) finds that this declaration was “ninety percent improvisation” (12), apparently uttered with hopes to resolve the German issue peacefully. However, on November 27, the Soviet Union sent the Western allies a diplomatic note formally restating these terms (Newman 2007).

himself, “went berserk” (Smyser 2009, 65). This resulted in a second Soviet ultimatum: If the Western allies did not immediately agree to a peace treaty proposing reunification on Communist terms, the Soviet Union would sign a separate peace treaty with East Germany, cutting off access to West Berlin. Berlin contingency planners began to seriously discuss the possibility of nuclear conflict (Williamson 2012, 215). On July 25, 1961, President Kennedy even made a nationally televised address in which he explained Soviet attempts to cut off access to West Berlin, bringing up the imminent threat of nuclear war and discussing measures to make sure that all American citizens had access to fallout shelters.¹²

On August 12, 1961, East German head of state Walter Ulbricht authorized an order to close the Berlin border with a wall, stemming extensive emigration of East Germans (Harrison 2011). A barbed wire fence was stretched across the border and later replaced with the notorious concrete barrier.

In the following months, American forces experienced harassment at checkpoints between East and West Berlin. Slowly escalating activities, often resembling a game of chicken, boiled over at Checkpoint Charlie on October 22, 1961. Three days of posturing peaked when Soviet and American tanks sat pointed at one another, 100 yards apart, before both sides backed down. This standoff was one of the most dangerous moments of the Cold War, at least in Europe (Tauschweizer 2006).

From late 1961 through 1962, the United States government engaged in sporadic negotiations with both its Western allies and the Soviets. In January and March 1962, Secretary of State Dean Rusk met with Soviet Foreign Minister Andrei Gromyko in a series of fruitless talks, and attempts to negotiate a solution ended by the summer. The Soviet Union’s retreat from the Cuban Missile Crisis months later irrevocably dampened Khrushchev’s clout and diluted Soviet diplomatic leverage regarding Berlin. This loss of face helped remove remaining obstacles to the Limited Nuclear Test-Ban Treaty, a years-long effort that had been held up by both parties linking concessions to the Berlin question. The treaty was signed on August 5, 1963 and went into effect on October 10, 1963. Although the wall would not fall for another 36 years, the most heightened period of tension regarding Berlin had passed.

This overview indicates two reasons why the Berlin Crisis is ideal for studying crisis diplomacy. First, this five-year period is one of substantial historical import, punctuated by multiple moments that could have sparked major hostilities involving nuclear weapons. While the term “Berlin Crisis” suggests a single phase of hostility, the time span is better characterized by several distinct flash points, as well as sustained periods of relative calm, that each provide evidence on the different effects of public and private diplomacy during the everyday administration of policymaking. Moreover,

even though American policymakers during this time may have perceived a higher than average level of resolve, actual perceptions of threat fluctuated over time. Second, studying the Cold War and a location that served as a geographical and symbolic focal point for this period provides a best-case scenario for finding larger effects of diplomatic statements—particularly public ones—in shaping perceptions.

Beyond its substantive import, the Berlin Crisis is well-documented in archival collections and is almost fully declassified.¹³ This allows us to obtain a comprehensive set of documents that reflect the uncensored and instantaneous messages and thoughts of the policymaking elite, without retrospective, censored, or historical biases. To our knowledge, this sort of sweeping document-based approach does not exist in studies of international security.

Data Sources

Our analysis requires measures of four concepts: material actions, private statements, public statements, and White House assessments of these signals. The last three rely on archival documents. We take each of these in turn.

Actions

We account for material actions from the Berlin Crisis using headlines and abstracts from the *New York Times* (NYT).¹⁴ Between January 1, 1958 and December 31, 1963, the *New York Times* has 14,178 articles regarding Berlin. Of these, 1,601 articles use one of several terms that reflect material action and conflict. These include “suspend,” “seize,” “ambush,” “raid,” and the like.¹⁵ We manually code whether each of these 1,601 articles reported on costly military action reflecting hostility. This results in 91 events representing five types of activity: the construction of the wall (1); nuclear or missile tests (7); shooting down of a plane (1); a blockade (1); and detainment or halting of military convoys and transports (81). The first plot of Figure 2 illustrates the frequency of these material actions at the weekly level.

Private Statements

Private diplomatic statements are captured through declassified telegrams from the U.S. Department of

¹² This address occurs more than a year before the Cuban Missile Crisis and one month before construction begins on the Berlin Wall.

¹³ Documents such as those used in this article are mostly declassified up until 1978. Due to the dilatory nature of declassification, many documents on Berlin were only released in the last two decades. A substantial collection of highly sensitive materials were released in 2011—the fiftieth anniversary of the Berlin Wall.

¹⁴ This set of data does not rely on declassified documents.

¹⁵ These terms come from the Conflict and Peace Data Bank, or COPDAB (Azar 1982). While COPDAB presents ready-made data for superpower activity, we create new event data for two reasons. First, COPDAB was released decades ago and provides no information on sources. Second, COPDAB may be incomplete. Appendix G supplies information on these issues but also shows that most of our findings are intact using COPDAB.

State (DOS), obtained at the National Archives II in College Park, Maryland.¹⁶ These collections predominantly involve incoming messages from the United States Embassies in Bonn and Moscow, as well as the U.S. Mission Berlin.¹⁷ The cables include summaries of private conversations, meetings with foreign government officials, and noteworthy information that the outposts send to the capital for discussion behind closed doors. They are not public reports.

We take several measures to ensure that we only track relevant private statements made by Soviets to the United States. First, we only use documents classified at the “Confidential” level and above.¹⁸ 0.7% were unclassified and thus removed. Second, we exclude all outgoing cables. Third, we only utilize memoranda of conversation that involve meetings with Soviet officials. Contemporaneous records of intra-governmental meetings or deliberations, which do not reflect private diplomatic statements from the Soviet Union to the United States, are omitted.

Public Statements

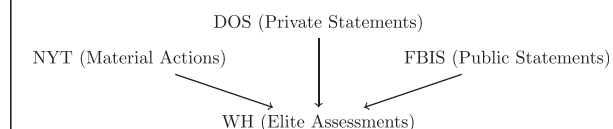
Records of public statements during the Berlin Crisis come from the Foreign Broadcast Information Service (FBIS). The FBIS was an open-source intelligence system originally housed in the Central Intelligence Agency that recorded and translated foreign countries’ official public statements made through radio and press agency releases. Policymakers sought to use this massive information stream to track how adversaries discussed recent events, as well as how the tone or content of their statements reflected their intentions (Leetaru 2010). During the Cold War, the FBIS placed tremendous focus on tracking public Soviet statements. Images of the original English translations are available online via NewsBank.

Elite Assessments of Resolve

American elites’ evaluations of the Soviet Union’s resolve come from declassified internal White House (WH) documents, collected from both the Dwight D. Eisenhower Presidential Library in Abilene, Kansas and the John F. Kennedy Presidential Library in Boston, Massachusetts. We gather all archived collections categorized under “Berlin.” Figure 1 shows how these sets of data map to our theoretical framework.

Both DOS and WH documents are individually photographed at these archives. All images are then processed using optical character recognition (OCR) software, which converts each image into computer-

FIGURE 1. Dynamics and Sources



readable digital text.¹⁹ Hand-written documents, few in number, are omitted from this process.²⁰ Online FBIS records include OCR-processed text in their metadata, which are extracted.

These data represent the most comprehensive coverage of the Berlin Crisis thus far. Although we cannot guarantee that all sensitive documents are in our data, several observations suggest that this is not a serious issue. First, the collections include red inserts which indicate when entire documents are or were classified. All of these documents are now accessible. Second, less than 1% of declassified documents feature any form of sanitizing at the paragraph or sentence levels. Third, the documents that are currently available reveal secrets that would have been incredibly sensitive at the time. This includes memoranda discussing detailed logistics for an American first strike against the Soviet Union, aimed at initiating a general war.²¹ It is hard to envision documents that could contain more delicate information.

Measuring Signals and Perceptions

These raw diplomatic documents involving public statements, private statements, and elite assessments must be translated into quantitative data that gauge the resolve of the Soviet Union and East Germany/Berlin. We generate measures of observed Soviet resolve through their public and private diplomatic statements, captured, respectively, in FBIS and DOS records. For these two collections, we seek an indication of whether each memorandum or publicly released message reflects Soviet resolve—that is, willingness to use force to achieve its aims in Berlin. To see whether these signals had any effect on evaluations of resolve, we would then turn to White House documents to see whether policymakers appear to express concerns regarding Soviet resolve.

Given the number of documents involved, coding by hand would be difficult but not impossible. This study nevertheless uses supervised statistical learning models to automate much of this process. We do so for three reasons. First, manually classifying thousands of documents

¹⁶ Appendix A has a full listing of sources.

¹⁷ Mission Berlin was the State Department’s substitute for an embassy in West Berlin.

¹⁸ “Confidential” is the lowest classification level for government information and documents. This involves information for which “the unauthorized disclosure of which reasonably could be expected to cause damage to the national security that the original classification authority is able to identify or describe” (United States Department of State 2005, 2).

¹⁹ OCR accuracy rates at the character level had a mean of 98.0% and median of 99.6% per page. At the word level, accuracy rates had a mean of 95.8% and median of 98.3%.

²⁰ This is theoretically motivated. Even during the 1950s, official documents meant for circulation were always typed.

²¹ Memorandum from Carl Kaysen to General Maxwell Taylor, Military Representative to the President. “Strategic Air Planning and Berlin.” September 5, 1961.

increases the likelihood that coding becomes inconsistent and colored by a researcher's personal expectations. Second, automated coding may help us find patterns that counter conventional wisdom, none of which would be discovered if conventional wisdom dictated the coding process. Third, we seek to demonstrate a procedure that can adapt to even larger sets of documents. Unlike statistical learning methods, manual classification is not scalable.

Supervised learning involves manually coding a random sample of the data and then training a statistical model to identify what features of each observation best explain this manual classification. The model is then cross-validated to assess how capably it makes predictions on additional manually coded data that was not used to train it, which ensures that the model is not overfitting the training data. Conditional on performing sufficiently well, this model is used to predict the variable of interest for the remaining observations for which the value of the variable is unknown.

The technical details of this process are explained elsewhere.²² Here, it suffices to say that the three pools of raw text are converted into quantitative data. Cables and releases are first split into 300-word segments to ensure that long documents are analyzed more properly. The text in each segment undergoes standard pre-processing. This includes the removal of stop words such as “a” and “the,” as well as the stemming of words into tokens (e.g., converting “talks,” “talked,” and “talking” into “talk”). The number of remaining tokens in each segment is counted and recorded. This produces a document-term matrix for each collection of cables, where each row represents a 300-word segment, and each column contains how many times a token is used. These token counts are the main variables used to train the models and to generate predicted values of Soviet resolve for each segment.

Coding Criteria

Each subset of training documents was classified for Soviet resolve, or more specifically, willingness to use force with respect to Berlin.²³

Such statements typically fall into one of two categories: threats to close off the Western allies' access to West Berlin and threats to engage in armed conflict. These threats can be both explicit and implicit. An explicit threat proposes shutting down access or initiating hostilities in concrete terms. Khrushchev's ultimatums are explicit threats, but such declarations are relatively infrequent. More common is an implicit threat, which vaguely intimates the possibility of either event, and does so using hostile language. For instance,

a DOS document from February 9, 1962 conveys a Soviet message about West Berlin: “No matter how much the United States and its allies refer to their privileges of occupation, there can be no two opinions about this. This land and air space belong to GDR.” An FBIS entry from August 2, 1961 warns: “Those who raise their arms against us will be destroyed on their own territory.” We base these classifications on the specific context of the Berlin Crisis, which permits more precise measurement.

For a signal to be credible, it must be perceived in the sense that the receiver notices and processes it.²⁴ Two White House documents illustrate how elites discern and consider Soviet resolve that is conveyed through private and public statements. We use examples mentioning direct and specific signals to best illustrate our process, but White House documents that infer Soviet resolve more generally also qualify.

Private Statement (August 24, 1961):

[In a private conversation,] Khrushchev said that the West was now threatening to cut off trade and even to go to war if the Soviet government signed a peace treaty. If other countries strengthened their military forces in Germany, the Soviet Union would do the same and could always have forces in a position to protect the territory of its ally the GDR. But in any case, modern wars would be fought with nuclear weapons. The Soviet Union and also the USA would no doubt lose tens of millions but the Soviet Union would certainly go to war if the Western Allies tried to force their way through to West Berlin after the Soviet Union had signed a peace treaty with the GDR. It would however be ridiculous for two hundred million people to die over two million Berliners.

Public Statement (November 1, 1961):

In [Khrushchev's] July 8 speech, he attributed motives of military pressure against the Communist Bloc. In response among other threats, he spoke of a 100 megaton super-H bomb which he said had been devised. From other reports as well, we learn that Khrushchev was especially stung by this speech. On August 7, Khrushchev made a speech in which he stressed the horrendous consequences of a nuclear war, a speech in this respect unusual for delivery to a domestic Soviet audience.

Table 1 summarizes all three sets of archival data. The hand-coded segments are processed using a predictive model that finds relationships between the tokens used in a segment and the segment's classification. Of many possible models, the balanced random forest model exhibited the best overall performance compared to several other alternatives.²⁵

²² See Friedman, Tibshirani, and Hastie (2009).

²³ The authors and a third party produced two sets of codings. In terms of intercoder reliability, accuracies were 0.89 (DOS), 0.91 (FBIS), and 0.85 (WH); F_1 scores were 0.93 (DOS), 0.97 (FBIS), and 0.91 (WH). Discrepancies were resolved via further investigation and discussion, and these revised codings were used in the trained model.

²⁴ We leave aside the question of whether the receiver explicitly believes the signal. Our emphasis is on a more minimal standard of whether the receiver acknowledges a signal of resolve in the first place. Moreover, subsequent discussion of a signal likely indicates that at least some policymakers take the message seriously.

²⁵ Breiman (2001) and Chen, Liaw, and Breiman (2004) provide details on this method. Technical results from our analysis are in Appendix C.

TABLE 1. Totals for Archival Data

Collection	Documents	Segments	Hand-coded
DOS	4,012	7,777	804
FBIS	10,714	13,576	503
WH	3,725	15,499	602

Predicted Data

The balanced random forest model is applied to the three full document-term matrices to create a predicted probability of each segment exhibiting signals or perceptions of Soviet resolve. Depending on the hypothesis tested, we either utilize this raw predicted probability or dichotomize it using a cutpoint of 0.5.²⁶ Through this process, we generate three sets of time series data that reflect expressions of Soviet resolve in private Soviet diplomatic statements (DOS), expressions of Soviet resolve in public Soviet diplomatic statements (FBIS), and the White House elites' perceptions of Soviet resolve involving the Berlin Crisis (WH).

To ascertain the face validity of the new data, we identify eight key moments of tension or political importance during the Berlin Crisis. These include the following:

- (a) 11/10/58: Khrushchev's first ultimatum
- (b) 6/4/61: Vienna summit; Khrushchev's second ultimatum
- (c) 8/13/61: Start of construction of the Berlin Wall
- (d) 10/22/61: Checkpoint Charlie standoff
- (e) 8/17/62: Killing of Peter Fechter at the Berlin Wall²⁷
- (f) 9/25/62: Berlin Air Corridor incident
- (g) 10/10/63: 15-hour detainment of U.S. Army convoys
- (h) 11/4/63: Autobahn Tailgate Crisis

Figure 2 displays the NYT action data, as well as the three sets of predicted data. The eight aforementioned events are marked. Both the collected and generated data aptly identify these qualitatively important moments, which increases our general confidence in the data's overall utility.

There may be some concern that using the entirety of the DOS and FBIS documents would bias results in our favor. It is unlikely that WH elites placed equal weight on every single statement made in private or particularly in public, so using all documents would artificially inflate noise. As such, we create filtered DOS and FBIS measures that only count documents with statements

attributed to prominent Soviet entities. For both DOS and FBIS, we track only *high-ranking* individuals, which include Politburo members and other key Soviet officials. For FBIS, we create an additional intermediate measure based on *governmental* sources. This includes all high-ranking individuals, as well as official Soviet media outlets and spokespeople.²⁸ Figure 2 shows the high-ranking measures in light gray. Our analysis will generally focus on high-ranking DOS and FBIS measures, as they most closely reflect a plausible set of verbal signals that would be noticed by and raise concern among American policymakers.

ANALYSIS

We now proceed to evaluate evidence related to the proposed hypotheses. We first address Hypothesis 2 regarding the noisiness of public and private diplomatic statements. We then turn to Hypotheses 1 and 3, which speak to the absolute and relative impacts of different signaling channels on elites' evaluations of Soviet resolve.

Variation Across Statements

Figure 3 displays overall distributions of predicted codings of Soviet resolve in public and private statements from high-ranking officials.²⁹ The two are clearly not alike: The distribution of FBIS documents is especially asymmetric, and a greater proportion of private signals tend to be tied to higher levels of perceived hostility.³⁰

The across-signal variances for public and private statements from high-level officials are 0.032 and 0.027, respectively. To ensure that this difference is significant, we apply a Fligner–Killeen test for the homogeneity of variances across samples, which is robust to deviations from non-normality. The test strongly rejects the null of homogeneous variances ($\chi^2 = 41.896, p \ll 0.001$).³¹ This result is consistent with Hypothesis 2. Since words stated in public address multiple audiences in different ways, public statements tend to provide a less focused overall signal than private ones.

Diplomatic Statements' Small Effects

Due to the high and arguably overwhelming volume of words that define real-world politics, Hypothesis 1 postulated that neither public nor private statements should have a large impact on policymakers' evaluations of resolve when compared to material actions. Hypothesis 3 went one step further: If Hypothesis 2 were true and

²⁶ Appendix B provides example segments with predicted probabilities. Appendix C illustrates the distribution of predicted values for all segments.

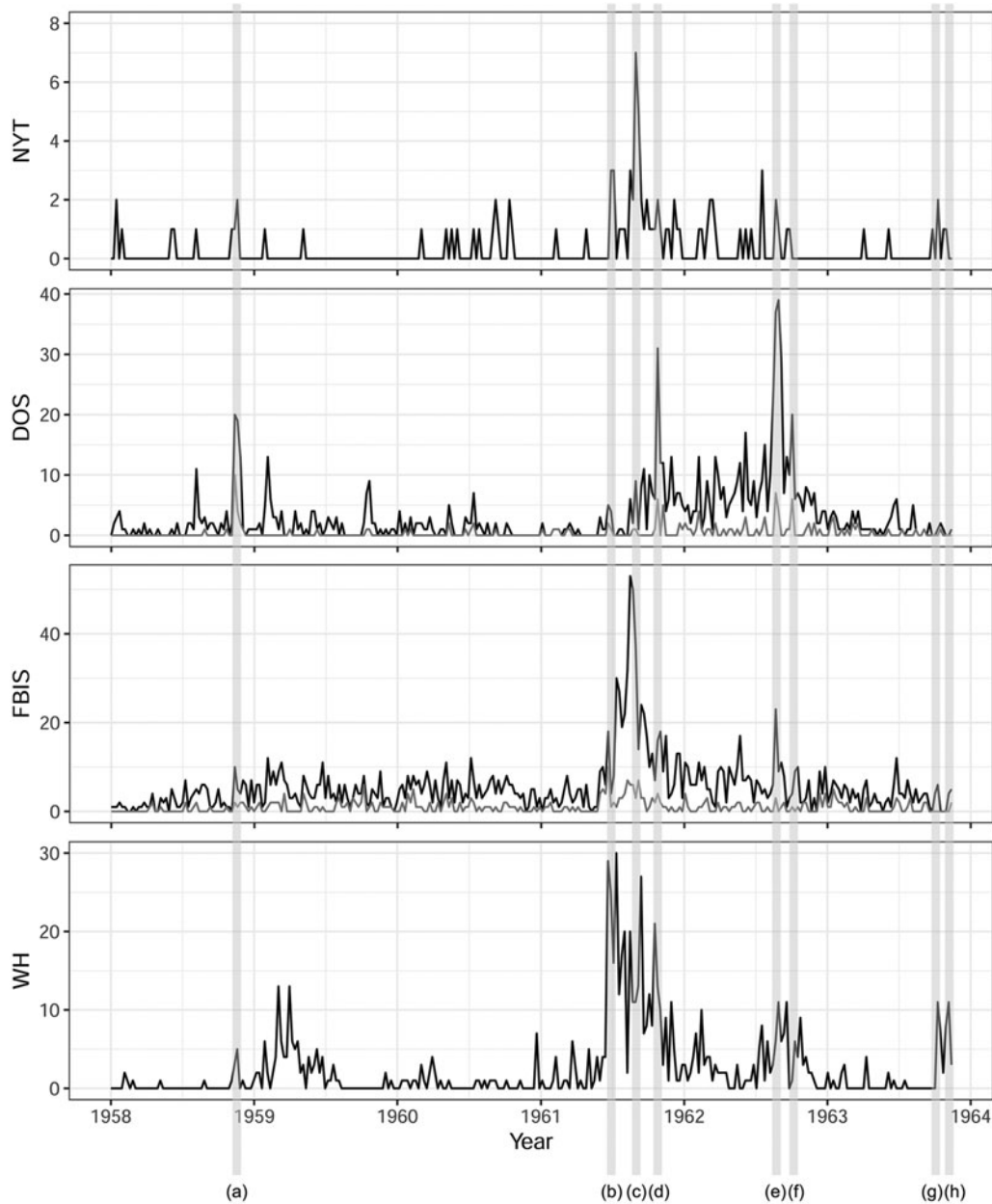
²⁷ Fechter was an East German who was shot while attempting to traverse the Berlin Wall. His death was one of the first involving the barrier. Many historical accounts of Berlin Crisis do not mention this incident, even though the death caused immense anxiety in West Berlin (as evidenced by the spike in the DOS measure in Figure 2). This emphasizes the risks involved with retrospectively identifying "important" events.

²⁸ Appendix D has a full list of individuals, as well as counts of relevant documents.

²⁹ Whenever we present analyses only based on measures using only high-ranking officials, Appendix E provides replications based on all available documents.

³⁰ A two-sample Kolmogorov–Smirnov test soundly rejects the null that the distributions are the same: $D = 0.334, p \ll 0.001$. When using all documents, $D = 0.241, p \ll 0.001$.

³¹ Using all statements, variances are 0.038 (FBIS) and 0.032 (DOS). These data also reject the null: $\chi^2 = 142.480, p \ll 0.001$.

FIGURE 2. Weekly Number of Documents Expressing Soviet Resolve (DOS, FBIS, WH), as Well as Material Actions (NYT)

Notes: For FBIS and DOS, light gray lines indicate measures using only documents from high-level Soviet officials. Vertical lines mark qualitatively noteworthy moments in the Berlin Crisis; see main text.

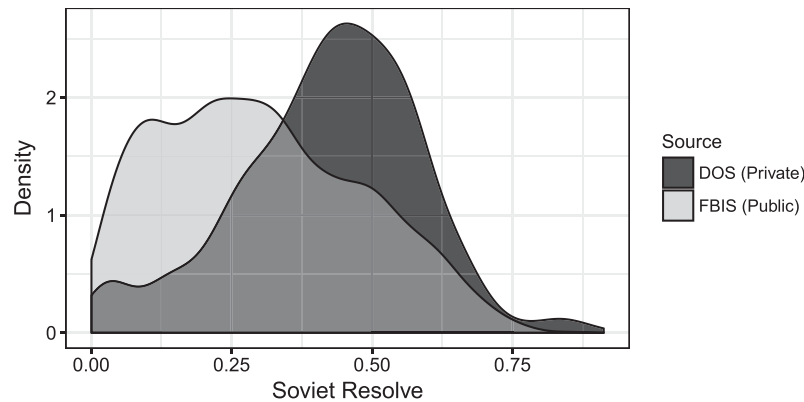
public verbal signals were noisier than private missives, as appears to be the case, then private signals should have relatively stronger effects in shaping perceptions.

We first test these notions using negative binomial regressions.³² The primary outcome variable is the number of White House memoranda that are determined to convey perceptions of Soviet resolve. The unit

³² The White House measure's mean is 2.479 and variance is 21.897, ruling out a Poisson model.

of analysis is the week. This level is chosen to allow for dynamic activity while also overcoming some estimation challenges that arise from the high frequency of zeroes in daily-level data.

The data are stationary, which allows us to use the variables in their original forms. Stationarity also permits our analysis to assume that temporal co-occurrence or proximity of signals indicates the processing of diplomatic statements, actions by elites, or both. That is, if a spike in private statements of resolve via DOS co-occurs

FIGURE 3. Predicted Probabilities of Soviet Resolve in Public and Private Statements

Notes: Using only documents from high-ranking Soviet officials. For a figure using all documents, see Appendix E.

TABLE 2. Results From Negative Binomial Regressions

	Dependent variable: White House					
	(1)	(2)	(3)	(4)	(5)	(6)
Private (DOS)						
<i>All cables</i>	0.033** (0.015)	0.032** (0.015)	0.031** (0.015)			
<i>High-rank</i>				0.158** (0.067)	0.152** (0.067)	0.131** (0.067)
Public (FBIS)						
<i>All statements</i>	0.001 (0.014)			-0.001 (0.014)		
<i>Governmental</i>		0.017 (0.026)			0.012 (0.026)	
<i>High-rank</i>			0.079 (0.052)			0.065 (0.052)
<i>Actions (NYT)</i>	0.219** (0.106)	0.212** (0.108)	0.231** (0.101)	0.225** (0.106)	0.218** (0.108)	0.236** (0.101)
Constant	-1.387*** (0.303)	-1.396*** (0.302)	-1.421*** (0.301)	-1.443*** (0.312)	-1.445*** (0.310)	-1.445*** (0.308)
Lagged DVs	3	3	3	3	3	3
Year FEs	Yes	Yes	Yes	Yes	Yes	Yes
Observations	304	304	304	304	304	304
Log-likelihood	-511.396 0.952*** (0.150)	-511.166 0.956*** (0.151)	-509.703 0.986*** (0.158)	-511.050 0.944*** (0.147)	-510.939 0.948*** (0.148)	-510.020 0.969*** (0.153)
AIC	1046.791	1046.332	1043.406	1046.099	1045.878	1044.041

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

with an uptick in the WH measure, we presume that those private statements helped engender the increase.³³

Table 2 displays initial results. All models include three weeks of lags in the dependent variable, as this minimizes the Akaike information criterion (AIC). Model 1 uses the entirety of DOS, FBIS, and NYT data.

³³ In Appendix F, we address concerns about spurious findings by testing alternative specifications of the explanatory variables and performing placebo tests.

We see that material actions have a strongly positive and statistically significant effect on WH perceptions of Soviet resolve. Private statements also wield a positive effect, while public statements have essentially no effect. Actions have a magnitude that is more than six times greater than a private signal. Put together, these findings corroborate Hypothesis 1 and also speak to Hypothesis 3.

Models 2 through 5 perform the analysis with different combinations of filtered data that only look at governmental or high-ranking sources. Model 6 uses

TABLE 3. Results From PAR(7) Models

	<i>Dependent variable: White House</i>			
	All DOS/FBIS		High-ranking DOS/FBIS	
	(1)	(2)	(3)	(4)
Private (DOS)	0.043*** (0.011)	0.033*** (0.013)	0.108*** (0.028)	0.078** (0.038)
Public (FBIS)	−0.025 (0.024)	−0.010 (0.027)	−0.003 (0.083)	0.044 (0.068)
Actions (NYT)	0.289** (0.110)	0.241* (0.131)	0.221*** (0.065)	0.190** (0.082)
Election period		0.034 (0.368)		0.145 (0.276)
US-USSR MIDs		0.152 (0.095)		0.166* (0.096)
Kennedy		−0.003 (0.210)		0.036 (0.196)
ρ_1	0.142*** (0.034)	0.138*** (0.034)	0.141*** (0.035)	0.138*** (0.034)
ρ_2	0.135*** (0.034)	0.130*** (0.034)	0.131*** (0.034)	0.127*** (0.033)
ρ_3	0.124*** (0.033)	0.121*** (0.033)	0.121*** (0.033)	0.118*** (0.032)
ρ_4	0.053* (0.032)	0.053* (0.031)	0.055* (0.031)	0.054* (0.031)
ρ_5	0.053* (0.032)	0.055* (0.031)	0.054* (0.031)	0.056* (0.030)
ρ_6	0.077** (0.031)	0.078** (0.030)	0.079** (0.031)	0.079** (0.030)
ρ_7	0.031 (0.031)	0.034 (0.031)	0.034 (0.031)	0.037 (0.030)
Constant	0.217 (0.190)	−0.078 (0.328)	0.293 (0.183)	−0.115 (0.346)
Observations	300	300	300	300
Log-likelihood	−506.350	−505.402	−508.443	−506.566
AIC	1032.061	1036.803	1036.886	1039.131

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

statements from high-ranking officials for both DOS and FBIS. The key results not only remain, but the substantive significance of private signals grows. A material action has only double the impact of a private threat involving a high-ranking official.

The PAR(p) Model

Brandt and Williams (2001) argue that most traditional approaches for studying count data, including negative binomial regressions, do not properly account for systematic dynamics, leading to model misspecification and potential inefficiencies. They instead propose a Poisson autoregressive model, or PAR(p), to properly analyze dynamic count data. Our main results utilize this technique.

In service of a full analysis, we control for a series of potential confounders. First, we capture federal *elections* with a binary variable that takes the value 1 in the two months preceding each presidential and midterm election in November of 1958, 1960, and 1962. Second, we include a running count of *militarized interstate disputes* between the United States and the Soviet

Union that do not involve Berlin (Palmer et al. 2015). Third, a dummy variable for the *Kennedy administration* begins on January 20, 1961.³⁴

Table 3 shows the results using four PAR(7) models.³⁵ Models 1 and 2 use all DOS and FBIS data. Both strongly support Hypotheses 1 and 3. Material actions have larger impacts on elite reactions—and thus have greater credibility—than verbal statements, and private statements are relatively more credible than public statements. Models 3 and 4 are limited to high-ranking statements and exhibit the same findings. As was the case with negative binomial models, private statements from high-ranking officials are higher in credibility than public statements, and their effect is slightly less than half that of material actions. Controls in Models 2 and 4 do not impact the results.³⁶

³⁴ Appendix E contains results from negative binomial regressions that include these control variables. The results do not change.

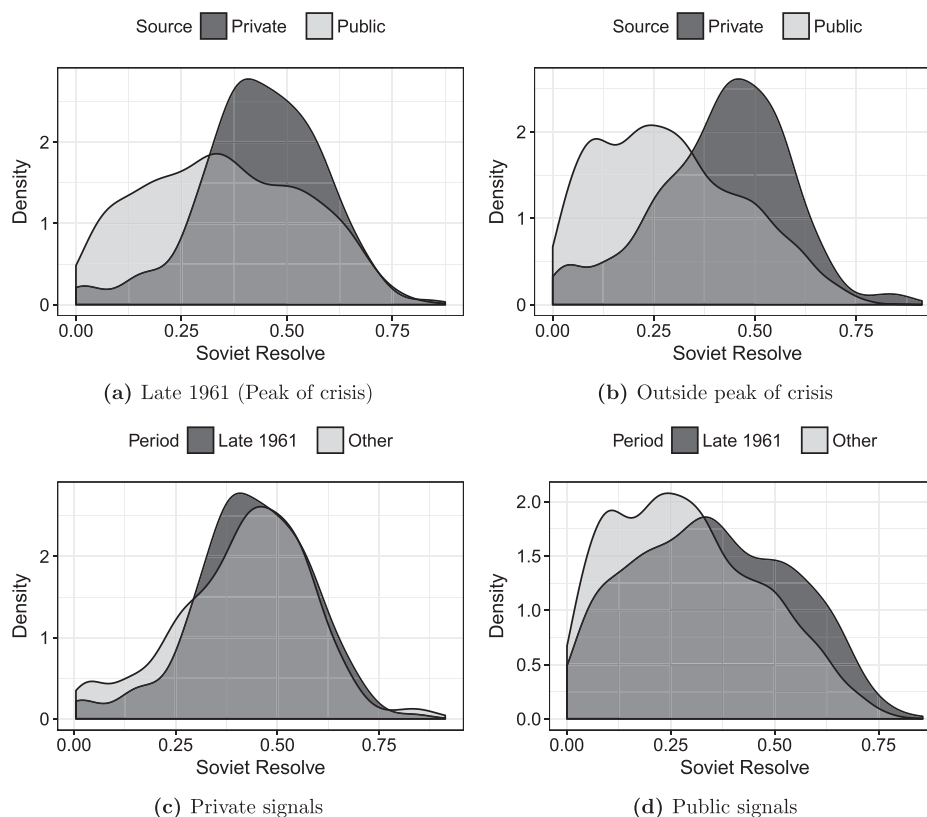
³⁵ Using seven autoregressive terms ($p = 7$) minimizes AIC.

³⁶ Appendix F performs additional checks using moving averages, one-week lags, and combined FBIS/NYT measures. Results are generally unchanged.

TABLE 4. Short-Term and Long-Term Effects on the WH Variable

	Model	All, No controls (1)	All, Controls (2)	High, No controls (3)	High, Controls (4)
Short-term	DOS	0.028 (1.112%)	0.019 (0.756%)	0.077 (3.117%)	0.052 (2.116%)
	FBIS	-0.016 (-0.655%)	-0.006 (-0.231%)	-0.002 (-0.077%)	0.029 (1.184%)
	NYT	0.188 (7.568%)	0.137 (5.507%)	0.159 (6.414%)	0.128 (5.174%)
Long-term	DOS	0.072 (2.923%)	0.048 (1.928%)	0.200 (8.068%)	0.134 (5.415%)
	FBIS	-0.042 (-1.705%)	-0.015 (-0.588%)	-0.005 (-0.200%)	0.075 (3.031%)
	NYT	0.488 (19.703%)	0.348 (14.053%)	0.412 (16.602%)	0.328 (13.244%)

Note: Percentage changes from the overall mean are in parentheses.

FIGURE 4. Distributions of Soviet Resolve by Diplomatic Channel and Time Period

Notes: Using only documents from high-ranking Soviet officials. For a figure using all documents, see Appendix E.

Table 4 provides more intuitive versions of the PAR (7) models' results by determining the short-term and long-term effects of a one-unit increase of a signal on the White House's evaluations of Soviet resolve. In other words, we determine the instantaneous and long-run impacts of one additional DOS (private) or FBIS (public) statement expressing resolve on the number of White House documents expressing perceived hostility.

The effects of FBIS signals are so imprecise and small that the estimate is often negative. However, private signals generate changes ranging from a 0.756% to 2.923% increase from the mean number of White

House documents, which is 2.479. The range increases from 2.116% to 8.068% when using only private statements from high-ranking officials. In comparison, costly and hostile actions have a markedly larger effect on White House documents, deviating from the mean by anywhere between 5.174% and 19.703%. In an "ideal" setting, one's prior belief may be that a single threatening statement from private channels should result in approximately one subsequent White House document conveying this message. The results presented here indicate an inefficient process borne out of a more realistic understanding of the vast and

TABLE 5. Mean Levels of Soviet Resolve by Diplomatic Channel and Time Period

	Noncrisis	Crisis	t-test	KS test
DOS	0.418	0.437	$p = 0.110$	$p = 0.083$
FBIS	0.291	0.350	$p \ll 0.001$	$p \ll 0.001$
t-test	$p \ll 0.001$	$p \ll 0.001$		
KS test	$p \ll 0.001$	$p \ll 0.001$		

Notes: Using only documents from high-ranking Soviet officials. Results of t-tests and Kolmogorov-Smirnov two-sample tests are presented for each row and column. For results using all documents, see Appendix E.

cacophonous informational environment in which diplomacy occurs.

Overall, this analysis supports Hypothesis 3, which predicts that private statements have a larger impact on elites' evaluations than public statements. Hypothesis 1, which postulates that both forms of diplomatic signal should have smaller effects than cost-sinking material actions, also finds support. These results provide the first systematic evidence of a hierarchy in signals with respect to shaping evaluations of an adversary's intentions: material actions on top, private diplomatic messages in the middle, and public diplomatic messages on the bottom. Such a ranking belies the implications of the canonical costly signaling perspective.

The Core of the Crisis

Crisis scholars may object to the notion of the Berlin Crisis lasting five years. A crisis, particularly as envisioned in game-theoretic terms, is a more distinct event defined by an initial challenge and then a sequence of interactions to address it (Snyder and Diesing 1977). The informational environment in immensely concentrated moments of tension may theoretically be starker and set a better stage for public hand-tying. As such, the analysis thus far may not be an honest reflection of dynamics during a "crisis" as understood by formal models.

We address this concern by examining a subset of the five-year range most strongly associated with the Berlin Crisis: June 4, 1961 to November 9, 1961. This five-month period is book-ended by the calamitous Vienna summit, where Khrushchev made his second ultimatum, and a Soviet proposal for a compromise solution on Berlin (Leng 2000).³⁷ Between these two events, the Berlin Wall emerged, and the two sides' tanks faced off at Checkpoint Charlie. We utilize data from between June 1 and November 15, 1961 to make our assessments.

Figure 4 illustrates overall distributions of Soviet resolve across time periods and diplomatic channels.³⁸

³⁷ See "Soviet Modifies Berlin Proposal; U.S. Unimpressed" on the front page of the *New York Times* on November 10, 1961. While the plan received a tepid response, it signified a pacific shift in the Soviet approach to the Berlin problem.

³⁸ Again, note that this analysis uses DOS and FBIS measures based on high-ranking officials.

TABLE 6. Variances in Soviet Resolve Across Signals by Diplomatic Channel and Time Period

	Noncrisis	Crisis	FL test
DOS	0.027	0.021	$p = 0.064$
FBIS	0.031	0.035	$p = 0.004$
FL test	$p \ll 0.001$	$p \ll 0.001$	

Notes: Using only documents from high-ranking Soviet officials. Results of Fligner-Killeen tests are presented for each row and column. For results using all documents, see Appendix E.

Table 5 shows that levels of perceived hostility indeed become higher during the main crisis for both public and private statements, though statistical significance is quite tenuous for DOS. Private statements reflect higher perceived resolve than public statements throughout. Table 6 includes results for four Fligner-Killeen tests of homogeneity of variances and indicates that public statements have more noise than private statements, both in and out of the key crisis. Variation in public signals is also mildly higher during the crisis. In times of serious tension, senders may try to direct more threatening public statements toward their adversary but may not change (or may even soften) the nature of their public statements directed to other audiences, thus widening the spectrum of signals produced. Speculation notwithstanding, our statistical results cast further doubt on crisis bargaining scholarship's assumption that public signals should be especially direct and forceful during crisis to effectively relay credible threats. Hypothesis 2 remains valid regardless of whether we focus only on the peak crisis or not.

This 160-day period only produces 22 or 23 weekly observations, so we cannot perform a useful statistical analysis. But the findings provided here suggest that even though the intensity of diplomatic statements increases during extraordinary moments of disquiet, private communications continue to be a more precise source of information to understand intentions of an adversary. Importantly, the time between June and November 1961 represents a "perfect" scenario in which public declarations should, in theory, generate the most strategic and political traction.

CONCLUSION

More than fifty years ago, Thomas Schelling established a cogent framework for understanding conflict behavior. The notions of credible commitment and rationality have left an indelible mark on the study of crisis bargaining. In the last two decades, audience cost theory has become the primary prism through which scholars debate the effectiveness of public and private diplomatic signals in influencing perceptions. The predominant belief remains that public hand-tying is a costly act that conveys far greater commitment than a seemingly costless message relayed behind closed doors.

This logic is intuitive and compelling. However, without much systematic empirical data, conflict research

has increasingly relied on abstractions and has swept away a more thoughtful examination of the actual informational environment in which crisis diplomacy occurs. The tremendous volume of dissonant information that passes through a government, involving a constantly shifting array of priorities, pulls at some threads that hold together many contemporary views of crisis bargaining. Real-world diplomacy is hectic, so the transmission of information is noisy and misperception is inevitable. But critically, noise is not constant across diplomatic channels. Because public statements are more frequent and directed at multiple audiences, they are particularly noisy and ineffectual in changing elites' beliefs.

These are not superfluous details or theoretical window dressing. Many extant studies on crisis resolution may have overstated the absolute and relative effects of public and private diplomatic statements. By failing to consider the quantity and quality of information that policymakers confront, scholars have either argued that only public statements are effective, or that private statements can sometimes be effective. Two additional options are omitted: that neither signal has a substantial effect, and that private statements may be more effective than public ones.

We have used a novel combination of archival, statistical learning, and time series methods to evaluate these claims. The approach is new for crisis bargaining and permits one of the first quantitative analyses of the effect of public and private diplomacy on evaluations of resolve. Critically, we study the Berlin Crisis of 1958 to 1963—a period of immense danger involving a clear focal point and the possibility of nuclear war. The Berlin Crisis should be a textbook case of crisis bargaining that represents an ideal set of conditions for public and private communications to shape American elites' perceptions of the Soviet Union.³⁹ Even in this best-case scenario, we find that neither form of diplomatic signal has a large effect on the White House, that public statements are noisier than private statements, and that private signals are generally more efficacious in altering perceptions.

These findings do not mean that public signals are universally ineffective. There are clear instances where categorically hostile threats made in public should hold greater weight than an analogous statement made in secret. Nonetheless, such instances are rare and must be undeniably obvious to break through the noise and to be understood in the moment. Our contention, backed by new data, is that such crystal-clear scenarios are not the norm. Concerns about Soviet resolve did not only spike in response to smoking-gun events but constantly and uneasily fluctuated based on everyday diplomatic activity.

This study, the new text-based methods, and the general framework we have outlined set a foundation for further research on several important aspects of diplomacy. We mention two here.

The temporal sequencing of public statements, private statements, and material actions could help illuminate

how elites send, manipulate, and clarify information. For example, if a private signal is followed by a growth in public statements, this suggests a dynamic of escalation where private communication failed to stem an issue, pushing the disagreement into the public spotlight. Conversely, if public statements are followed by private statements, it may be that elites use private communications to highlight which public messages are important and directed at the adversary. If no temporal relationship exists, public and private messages may have fundamentally different information or goals. Even if these signals have no natural ordering, actors may use multiple channels simultaneously to ensure that their message is noticed and taken seriously. Any of these scenarios would be revealing and encourage scholars to more seriously consider the calculus of diplomatic communications (Trager 2017).

A key contention in this article is that elites are literally incapable of processing and acting upon every piece of information running across their desks. We have provided evidence that a consequence of this information overload—attenuated levels of signal processing and increased misperception—indeed exists, but we do not address when and why actors shift their focus to a specific topic of concern, or what consequences this has for beliefs and policymaking.⁴⁰ By collecting a wider array of archived government documents, future studies could use similar methods to examine the causes and effects of information processing and agenda-setting on foreign policy.

Tensions concerning North Korea, the South China Sea, Syria, and other locales continue to make crisis diplomacy relevant today. On the one hand, our findings are discouraging for the study of contemporary conflicts since we cannot fully observe private diplomatic statements exchanged in current affairs. On the other hand, the results also suggest that we should temper our reactions to various provocative statements that parties make in public. Regardless of their implications on the study of present-day issues or predominant theories of diplomacy, our theoretical understanding of information and our empirical contribution of document-based analysis establish a more rigorous and pragmatic approach to learning when, why, and how crises unfold.

SUPPLEMENTARY MATERIAL

To view supplementary material for this article, please visit <https://doi.org/10.1017/S0003055418000643>.

Replication materials can be found on Dataverse at: <https://doi.org/10.7910/DVN/1ZOEEX>.

REFERENCES

Azar, Edward E. 1982. *Codebook of the Conflict and Peace Data Bank (COPDAB)*. Center for International Development, University of Maryland, Ann Arbor, MI.

³⁹ Concerns about Cold War tensions escalating to nuclear hostilities motivated the work of Schelling and his contemporaries.

⁴⁰ Jones and Baumgartner (2005) and Yarhi-Milo (2014) speak to this subject.

- Barker, Elisabeth. 1963. "The Berlin Crisis 1958–1962." *International Affairs* 39 (1): 59–73.
- Barston, R. P. 1988. *Modern Diplomacy*. London: Longman.
- Blechman, Barry M., and Douglas M. Hart. 1988. "Nuclear Weapons and the 1973 Middle East Crisis." In *The Use of Force: Military Power and International Politics*, eds. Robert J. Art and Kenneth N. Waltz. Lanham, MD: Rowman & Littlefield, 307–27.
- Brandt, Patrick T., and John T. Williams. 2001. "A Linear Poisson Autoregressive Model: The Poisson AR(p) Model." *Political Analysis* 9 (2): 164–84.
- Breiman, Leo. 2001. "Random Forest." *Machine Learning* 45 (1): 5–32.
- Brzezinski, Zbigniew. 1983. *Power and Principle: Memoirs of the National Security Adviser, 1977–1981*. New York: Farrar, Straus, Giroux.
- Burr, William. 1994. "Avoiding the Slippery Slope: The Eisenhower Administration and the Berlin Crisis, November 1958–1959." *Diplomatic History* 18 (2): 177–205.
- Chen, Chao, Andy Liaw, and Leo Breiman. 2004. Using Random Forest to Learn Imbalanced Data. Technical Report, University of California at Berkeley. <http://statistics.berkeley.edu/sites/default/files/tech-reports/666.pdf>.
- Crawford, Vincent P., and Joel Sobel. 1982. "Strategic Information Transmission." *Econometrica* 50 (6): 1431–51.
- Dahl, Erik J. 2013. *Intelligence and Surprise Attack: Failure and Success from Pearl Harbor to 9/11 and Beyond*. Washington, DC: Georgetown University Press.
- DiGangi, Roberta L., Lorraine Lees, Aaron Miller, and Charles S. Sampson. 1992. *Foreign Relations of the United States, 1955–1957, Volume XXVI: Central and Southeastern Europe*. Washington, DC: United States Government Publishing Office.
- Downes, Alexander B., and Todd S. Sechser. 2012. "The Illusion of Democratic Credibility." *International Organization* 66 (3): 457–89.
- Duelfer, Charles A., and Stephen B. Dyson. 2011. "Chronic Misperception and International Conflict." *International Security* 36 (1): 73–100.
- Farrell, Joseph, and Robert Gibbons. 1989. "Cheap Talk Can Matter in Bargaining." *Journal of Economic Theory* 48 (1): 221–37.
- Fearon, James D. 1994. "Domestic Political Audience Costs and the Escalation of International Disputes." *American Political Science Review* 88 (3): 577–92.
- Fearon, James D. 1995. "Rationalist Explanations for War." *International Organization* 49 (3): 379–414.
- Fearon, James D. 1997. "Signaling Foreign Policy Interests: Tying Hands versus Sunk Costs." *Journal of Conflict Resolution* 41 (1): 68–90.
- Fey, Mark, and Kristopher W. Ramsay. 2007. "Mutual Optimism and War." *American Journal of Political Science* 51 (4): 738–54.
- Finel, Bernard I., and Kristin M. Lord. 1999. "The Surprising Logic of Transparency." *International Studies Quarterly* 43 (2): 315–39.
- Friedman, Jerome H., Robert Tibshirani, and Trevor Hastie. 2009. *The Elements of Statistical Learning: Data Mining, Inference, and Prediction*. New York: Springer.
- Fuhrmann, Matthew, and Todd S. Sechser. 2014. "Signaling Alliance Commitments: Hand-Tying and Sunk Costs in Extended Nuclear Deterrence." *American Journal of Political Science* 58 (4): 919–35.
- Guisinger, Alexandra, and Alastair Smith. 2002. "Honest Threats: The Interaction of Reputation and Political Institutions in International Crises." *Journal of Conflict Resolution* 46 (2): 175–200.
- Handel, Michael I. 1977. "The Yom Kippur War and the Inevitability of Surprise." *International Studies Quarterly* 21 (3): 461–502.
- Harrison, Hope M. 2011. *Driving the Soviets Up the Wall*. Princeton, NJ: Princeton University Press.
- Holsti, Ole R., Richard A. Brody, and Robert C. North. 1964. "Measuring Affect and Action in International Reaction Models: Empirical Materials from the 1962 Cuban Crisis." *Journal of Peace Research* 1 (3/4): 170–90.
- Jervis, Robert. 1976. *Perception and Misperception in International Politics*. Princeton, NJ: Princeton University Press.
- Johns, Leslie. 2006. "Knowing the Unknown: Executive Evaluation and International Crisis Outcomes." *Journal of Conflict Resolution* 50 (2): 228–52.
- Jones, Bryan D., and Frank R. Baumgartner. 2005. *The Politics of Attention: How Government Prioritizes Problems*. Chicago, IL: University of Chicago Press.
- Jönsson, Christer. 1996. "Diplomatic Signaling in the Television Age." *Press/Politics* 1 (3): 24–40.
- Kempe, Frederick. 2011. *Berlin 1961: Kennedy, Khrushchev, and the Most Dangerous Place on Earth*. New York: Penguin Group.
- Kertzer, Joshua D. 2016. *Resolve in International Politics*. Princeton: Princeton University Press.
- Kissinger, Henry A. 1979. *White House Years*. Boston: Little, Brown.
- Kissinger, Henry A. 1982. *Years of Upheaval*. Boston: Little, Brown.
- Kurizaki, Shuhei. 2007. "Efficient Secrecy: Public versus Private Threats in Crisis Diplomacy." *American Political Science Review* 101 (3): 543–58.
- Kurizaki, Shuhei. 2016. "Signaling and Perception in International Crises: Two Approaches." *Journal of Theoretical Politics* 28 (4): 625–54.
- Kydd, Andrew. 2003. "Which Side Are You on? Bias, Credibility, and Mediation." *American Journal of Political Science* 47 (4): 597–611.
- Lake, David A. 2010/11. "Two Cheers for Bargaining Theory: Assessing Rationalist Explanations of the Iraq War." *International Security* 35 (3): 7–52.
- Lebow, Richard N. 1981. *Between Peace and War*. Baltimore, MD: Johns Hopkins Press.
- Lebow, Richard N. 2001. "Deterrence and Reassurance: Lessons from the Cold War." *Global Dialogue* 3 (4): 119–32.
- Lebow, Richard N., and Janice G. Stein. 1994. *We All Lost the Cold War*. Princeton, NJ: Princeton University Press.
- Leetaru, Kalev. 2010. "The Scope of FBIS and BBC Open-Source Media Coverage, 1979–2008." *Studies in Intelligence* 54 (1): 17–37.
- Leng, Russell J. 2000. *Bargaining and Learning in Recurring Crises: The Soviet-American, Egyptian-Israeli and Indo-Pakistani Rivalries*. Ann Arbor, MI: University of Michigan Press.
- Levy, Jack S. 1983. "Misperception and the Causes of War: Theoretical Linkages and Analytical Problems." *World Politics* 36 (1): 76–99.
- Levy, Jack S., Michael K. McKoy, Paul Poast, and Geoffrey P. R. Wallace. 2015. "Backing Out or Backing in? Commitment and Consistency in Audience Costs Theory." *American Journal of Political Science* 59 (4): 988–1001.
- Lewis, Jeffrey A., and Kenneth A. Schultz. 2003. "Revealing Preferences: Empirical Estimation of a Crisis Bargaining Game with Incomplete Information." *Political Analysis* 11 (4): 345–67.
- Meirowitz, Adam, and Anne E. Sartori. 2008. "Strategic Uncertainty as a Cause of War." *Quarterly Journal of Political Science* 3 (4): 327–52.
- Mercer, Jonathan. 1996. *Reputation and International Politics*. Ithaca, NY: Cornell University Press.
- Mitchell, Christopher. 2000. *Gestures of Conciliation: Factors Contributing to Successful Olive Branches*. New York: St. Martin's Press.
- Morrow, James D. 1994. "Modeling the Forms of International Cooperation: Distribution versus Information." *International Organization* 48 (3): 387–423.
- Newman, Kitty. 2007. *Macmillan, Khrushchev and the Berlin Crisis, 1958–1960*. New York: Routledge.
- Palmer, Glenn, Vito D'Orazio, Michael Kenwick, and Matthew Lane. 2015. "The MID4 Dataset, 2002–2010: Procedures, Coding Rules and Description." *Conflict Management and Peace Science* 32 (2): 222–42.
- Perlmutter, Amos. 1975. "Crisis Management: Kissinger's Middle East Negotiations." *International Studies Quarterly* 19 (3): 316–43.
- Quek, Kai. 2016. "Are Costly Signals More Credible? Evidence of Sender-Receiver Gaps." *Journal of Politics* 78 (3): 925–40.
- Ramsay, Kristopher W. 2011. "Cheap Talk Diplomacy, Voluntary Negotiations, and Variable Bargaining Power." *International Studies Quarterly* 55 (4): 1003–23.
- Roop, Joseph E. 1969. *Foreign Broadcast Information Service, History, Part I: 1941–1947*. Washington, DC: Central Intelligence Agency.
- Russell, Richard. 2000. "American Diplomatic Realism: A Tradition Practised and Preached by George F. Kennan." *Diplomacy & Statecraft* 11 (3): 159–82.
- Sagan, Scott D. 1985. "Nuclear Alerts and Crisis Management." *International Security* 9 (4): 99–139.
- Sartori, Anne E. 2002. "The Might of the Pen: A Reputational Theory of Communication in International Disputes." *International Organization* 56 (1): 121–49.

- Sartori, Anne E. 2005. *Deterrence by Diplomacy*. Princeton, NJ: Princeton University Press.
- Schelling, Thomas. 1960. *The Strategy of Conflict*. Cambridge, MA: Harvard University Press.
- Schelling, Thomas C. 1966. *Arms and Influence*. New Haven, CT: Yale University Press.
- Schroder, Harold M., Michael J. Driver, and Siegfried Streufert. 1967. *Human Information Processing: Individuals and Groups Functioning in Complex Social Situations*. New York: Holt, Rinehart and Winston.
- Schultz, Kenneth A. 2001. *Democracy and Coercive Diplomacy*. Cambridge, UK: Cambridge University Press.
- Simon, Herbert A. 1947. *Administrative Behavior: A Study of Decision-Making Processes in Administrative Organization*. New York: Macmillan Company.
- Slantchev, Branislav L. 2006. "Politicians, the Media, and Domestic Audience Costs." *International Studies Quarterly* 50 (2): 445–77.
- Slantchev, Branislav L. 2011. *Military Threats: The Costs of Coercion and the Price of Peace*. Cambridge, UK: Cambridge University Press.
- Smith, Alastair. 1998. "International Crises and Domestic Politics." *American Political Science Review* 92 (3): 623–38.
- Smyser, William R. 2009. *Kennedy and the Berlin Wall*. New York: Rowman & Littlefield.
- Snyder, Glenn H. 1972. "Crisis Bargaining." In *International Crises: Insights from Behavioral Research*, ed. Charles F. Hermann. Free Press, 217–56.
- Snyder, Glenn H., and Paul Diesing. 1977. *Conflict Among Nations: Bargaining, Decision Making, and System Structure in International Crises*. Princeton, NJ: Princeton University Press.
- Snyder, Jack, and Erica D. Borghard. 2011. "The Cost of Empty Threats: A Penny, Not a Pound." *American Political Science Review* 105 (3): 437–56.
- Tarar, Ahmer, and Bahar Leventoglu. 2012. "Limited Audience Costs in International Disputes." *Journal of Conflict Resolution* 57 (6): 1065–89.
- Tingley, Dustin H., and Barbara F. Walter. 2011. "Can Cheap Talk Deter? An Experimental Analysis." *Journal of Conflict Resolution* 55 (6): 996–1020.
- Tomz, Michael R. 2007. "Domestic Audience Costs in International Relations: An Experimental Approach." *International Organization* 61 (4): 821–40.
- Trachtenberg, Marc. 1999. *A Constructed Peace: The Making of the European Settlement 1945–1963*. Princeton, NJ: Princeton University Press.
- Trachtenberg, Marc. 2012. "Audience Costs: An Historical Analysis." *Security Studies* 21 (1): 3–42.
- Trager, Robert F. 2017. *Diplomacy: Communication and the Origins of International Order*. Cambridge, UK: Cambridge University Press.
- Trauschweizer, Ingo W. 2006. "Tanks at Checkpoint Charlie: Lucius Clay and the Berlin Crisis, 1961–62." *Cold War History* 6 (2): 205–28.
- United States Department of State. 2005. "Department of State Classification Guide (DSCG 05-01)." *Declassified document*.
- Vertzberger, Yaacov Y. I. 1990. *The World in Their Minds: Information Processing, Cognition, and Perception in Foreign Policy Decisionmaking*. Stanford, CA: Stanford University Press.
- Wickens, Christopher D. 1992. *Engineering Psychology and Human Performance*. New York: HarpersCollins.
- Williamson, Richard D. 2012. *First Steps Toward Detente: American Diplomacy in the Berlin Crisis, 1958–1963*. Lanham, MD: Lexington.
- Wohlstetter, Roberta. 1962. *Pearl Harbor: Warning and Decision*. Stanford, CA: Stanford University Press.
- Yarhi-Milo, Keren. 2013. "Tying Hands Behind Closed Doors: The Logic and Practice of Secret Reassurance." *Security Studies* 22 (3): 405–35.
- Yarhi-Milo, Keren. 2014. *Knowing the Adversary: Leaders, Intelligence, and Assessment of Intentions in International Relations*. Princeton: Princeton University Press.
- Zubok, Vladislav M. 1993. "Khrushchev and the Berlin Crisis (1958–1962)." Working Paper No. 6, Woodrow Wilson International Center for Scholars.