

Flipping Channels: Diplomatic Signals in Front of Multiple Audiences

Rough working paper prepared for ISA 2022; please do not distribute or cite

Azusa Katagiri and Eric Min*
Osaka University and University of California, Los Angeles

March 23, 2022

Abstract

How do states choose to use public or private signals to communicate their intentions? Formal literature, particularly involving notions of costly signaling, has focused on differences in these two channels primarily in terms of costs to renegeing on any statements made. Moreover, a great deal of work has only considered signaling interactions between two actors in a vacuum. We contend that the presence of third parties that are involved in disputes with a specific sender will affect the sender's decision to direct more of their signals of resolve through public or private channels. To explore this claim, we conduct a computational analysis of 19,000 diplomatic statements from the Berlin Crisis of 1958 to 1963, which allows unprecedented access to public and private signals exchanged between the United States and Soviet Union. Through the use of structural topic models and text similarity measures, we find that the Soviets send more signals, which are both more aggressive and focused in their content, via private channels to the US when embroiled in direct disputes with the US. However, these signals shift over to public channels when Soviets are involved in multiple disputes with states around the world. Our contribution is two-fold. Substantively, we outline new mechanisms to understand how different forms of diplomatic communication differ beyond their costliness. Methodologically, we provide a unique example of how statistical learning methods can be applied to study the strategic logic of diplomacy.

*Names are in alphabetical order. We thank attendees of the UCLA Computational International Relations Workshop (February 2020) for their helpful comments. 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 Mathew Yeo Jie Sheng for excellent research assistance.

Over the last several decades, literature in international relations (IR) has drawn distinctions between signals that are relayed in public from those that are exchanged behind closed doors. The predominant view, captured by theories of costly signaling and audience costs, suggest that signals are most credible when they are costly. These costs are most effectively accrued by launching threats in public (Fearon 1994; Schelling 1966; Schultz 2001; Tomz 2007). In contrast, statements made in private are thus cheap talk with little consequence if the sender pulls back. More contemporary works have countered this assertion, explaining how private threats can also implicate reputational costs and have credibility because they could lead to embarrassing revelations or even conflict (Guisinger and Smith 2002; Kurizaki 2007; Ramsay 2011; Sartori 2002, 2005; Yarhi-Milo 2013b). This line of work has been productive, but it harbors an implicit assumption that public and private diplomacy mainly differ in a single dimension: their costliness. While evidence is clear that costs matter, this may be an overly reductive view of the strategies that motivate various forms of diplomatic signaling, as well as the varied roles that each type of communication can play.

We argue that public and private signals also differ in that the former can be used to signal to multiple audiences at once. This feature of public diplomatic communication comes with its own costs and benefits. The potential cost is that third-party actors may interpret a signal meant for another target as being relevant to them, which could lead to policy changes that the sender of the original signal would find counterproductive and perhaps even dangerous. This concern is especially salient when an actor is in a dispute with one specific adversary and does not want to affect relations with others. Yet at the same time, a potential benefit is that *because* a public signal can be viewed and interpreted by multiple audiences, making focused public statements of resolve can relay relevant information to several targets at once. We argue that these trade-offs influence an actor's decision to lean more heavily into public signals or into private signals. When dealing with just one specific adversary, states will prefer to use private diplomatic channels to convey their signals of resolve and overall intentions, as this option minimizes the chances of misinterpretation by extraneous audiences. In contrast, when a state is confronting multiple adversaries at once, the risks to using public diplomatic signals will fall while the appeal to doing so will grow.

We test these claims using 19,000 diplomatic statements from the Berlin Crisis of 1958 to 1963, which document how the two key actors of the Cold War—the United States and the Soviet Union—sent signals to the international community and each other. Harnessing a combination

of structural topic models (STMs) and document similarity measures, we find support for these claims.

This work, even in its less developed form, makes several contributions. First, we outline ways in which public and private diplomatic signaling can differ beyond the ostensible costs associated with renegeing on threats made in either channel. Second, we emphasize the importance of third-party actors in understanding conflict diplomacy strategy. Third, we propose an argument that scrutinizes the balance that states may take in terms of public and private signals, rather than treating the decision as entirely binary.¹ Fourth, our paper also makes an empirical contribution by demonstrating the value of applying text-as-data methods to large sets of archival documents. Performing this interdisciplinary work on declassified documents, which ties together quantitative and historical approaches, helps us glean valuable insight on how real policymakers thought and communicated during unique moments of international security affairs. Finally, our analysis disaggregates the tone of diplomatic signals from their content. Even though our results are largely similar regardless of whether we analyze tone or topic, we believe that this is an important distinction that can be more effectively leveraged in subsequent research.

Dealing with Multiple Audiences

Diplomatic communication is the foundation of international affairs. A critical challenge that a sender has when communicating with an intended target is making the target believe that the sender's words are credible. This is particularly salient in bargaining interactions where an object or issue is at dispute. Each side may have incentives to dissemble and exaggerate its capabilities or commitment with hopes of producing a more favorable outcome for itself. Given that an actor has enormous liberty when choosing what to say, how does the actor make its diplomatic messages and signals credible to the receiver?

The predominant and orthodox view suggests that diplomatic signals are most effectively made credible when they are costly Schelling (1966), and signals are most costly when they are launched in public. By making threats in public, an actor ties their hands and creates costs if they do not follow through with their words (Fearon 1994; Tarar and Leventoğlu 2012).² Meanwhile, words

¹In that respect, the current title of this paper draft may be a bit misleading.

²For a deeper discussion and extension of costs associated with signaling, see Quek (2021).

exchanged behind closed doors are considered to be cheap because they can easily be taken back without as many repercussions.

More recent literature has challenged several aspects of this classic argument, suggesting that private signals can be effective vehicles of credible signals. Scholars have noted that private messages still implicate costs because they can be revealed after the fact (Carson and Yarhi-Milo 2017; Kurizaki 2007) and because reputations for honesty are extremely valuable (Sartori 2002, 2005; Trager 2017).³ Diplomatic interactions that take place in person, which allow for the interpretation of emotions and other cues, can also be deemed credible even if they involve only two individuals (Holmes 2013; Wong 2016). Moreover, statements made in private may be more precise and easy to interpret than the overwhelming and relatively noisy flow of signals characteristic of real-world diplomacy in the public realm (Katagiri and Min 2019).⁴

Most of these works, much like the classic view of costly signaling theory, analyze an interaction which takes place between only two actors—the sender and the receiver—who are presumed to understand that any signals exchanged are exclusively meant for one another. To be sure, costs to renegeing from public threats may indeed stem from international audiences that become emboldened to challenge actors with a bad reputation for resolve (Weisiger and Yarhi-Milo 2015), but far more emphasis is typically placed on domestic political ramifications (Debs and Weiss 2016; Fearon 1994; Weeks 2008).

The dynamics and strategies of signaling may become more complex when we acknowledge another major distinction between private and public signals: the latter are exchanged in front of numerous additional audiences that are not necessarily the intended target of a specific message. In the realm of social psychology, scholars have noted that individuals often present themselves in different ways depending on the audience with which they are interacting. The ability to present oneself becomes constrained and more complicated when that individual finds themselves in front of multiple audiences, each of which is accustomed to seeing the individual in a different light. This is the “multiple audience problem” (Fleming et al. 1990; Fleming and Darley 1991; Van Boven et al. 2000).

³States may also prefer to send more of their costly signals outside of the public spotlight to avoid backlash over undesirable policies; see McManus and Yarhi-Milo (2017).

⁴Other scholars have more directly challenged assumptions and implications of audience cost theory. See, for example, Downes and Sechser (2012); Kertzer and Brutger (2016); Levy et al. (2015); Trachtenberg (2012).

Social psychologists often consider this dilemma in the sense of roles that people adopt, such as being a boss in front of one's employees while being a child in front of one's parents. Yet the underlying dilemma is not limited to individual people. States in the global system must also wrestle with the question of how to present themselves to others, knowing full well that signals they send could be seen by actors who are not the intended targets of a message.

A sender may undermine their own goal with respect to a particular receiver if their self-representation with that one receiver is at odds with how they want to represent themselves and their own intentions in front of another actor.⁵ Third parties that are not meant to be part of an interaction may still see a particular signal and erroneously interpret it as being directed at them. This can affect the third party's assessment of the sender's reputation for either aggression or restraint, which could trigger policies against the sender's interests in either the original interaction or in subsequent interactions with the third party (Carson and Yarhi-Milo 2017; Lake 2010/11; Vertzberger 1990; Yarhi-Milo 2014).

Anticipating the potential for these mix-ups, senders of diplomatic signals may have incentives to make their public statements more vague (Jönsson and Hall 2002). This can subsequently dilute the impact that public statements have on the intended (or, for that matter, any other) receiver's assessment of threat. Fearon (1997) touches upon this possibility, suggesting that leaders may opt for "partial signals of commitment" (84) when trying to strike a balance of signals in front of more than one audience.⁶ During possible or active crises, actors seeking to revise the status quo in their relations with one actor may seek to avoid creating unintended opposition in third parties that could increase the potential costs of realizing one's policy aims. In recent work, Wolford (2020) formalizes this argument and suggests that the most effective way to maintain a reputation for restraint with third-party audiences may be to send ambiguous signals in public or to communicate the most serious signals away from public scrutiny.

This line of reasoning suggests that states have incentives to exercise caution, even when they are attempting to gain an upper hand in a bargaining interaction with another state. When an actor is in a dispute with one specific adversary, they should tend to communicate more of their

⁵In recent work, Pu (2019) explores how China has navigated this multiple audience problem—simultaneously attempting to promote its status as a world power while simultaneously claiming to be a developing nation.

⁶This is also consistent with Snyder and Borghard (2011), who observe that states frequently hedge their threats to avoid being trapped by (and to avoid suffering audience costs if they pull back from) their own words.

hostility, indications of their intentions, and their policy priorities using private channels. Doing so minimizes the likelihood of creating unnecessary and unwarranted push-back from extraneous parties.⁷

However, the potential liabilities to signaling in front of multiple audiences when dealing with a single adversary become mitigated when the actor is actually attempting to confront and influence multiple audiences. The fact that signals sent regarding one specific issue area could flow over into others can become a benefit in the correct circumstances. Public signals provide a more efficient way to express resolve and profess information when there are numerous hot spots that demand attention. The actor in question would have greater incentive to relay the same signal of resolve to multiple audiences, as doing so may reinforce the actor's position and help all of these audiences update their beliefs in a manner the sending actor desires.⁸

We therefore propose two observable implications that would align with our argument. The first is that *actors should rely more heavily on private diplomatic signals when they are involved in disputes with a specific adversary*. The second is that *actors should rely more heavily on public diplomatic signals when they are involved in more disputes implicating more simultaneous adversaries*.

In the next section, we describe the data that can allow us to assess these predictions.

Data

We assess the dynamics of signals using a digitized set of 19,000 archival documents covering the Berlin Crisis of 1958 to 1963. Despite its name, the Berlin Crisis is better understood as several flash points distributed over a relatively quieter six-year period.⁹ Due to the sufficient passage of time, diplomatic cables from this five-year period are now almost fully declassified. They provide

⁷The argument we make here is consistent with previous game-theoretical findings, which suggest that the existence of another audience can make an actor more restrained and honest in how they communicate with their intended target (Farrell and Gibbons 1989). See also Goltsman and Pavlov (2011).

⁸There are some echoes here of Walter (2006), who contends that governments are less prone to make concessions to separatist groups when they are aware of additional separatists who may infer weakness if the government makes any concessions. In a similar vein, Clare and Danilovic (2010) suggest that states facing multiple rivals may choose to cultivate a reputation for resolve by initiating and escalating disputes.

⁹The most threatening period during this time, and what usually comes to mind when people consider the crisis, is the time between June 4, 1961 and November 9, 1961. This window captures the unsuccessful Vienna Summit where Khrushchev issued Kennedy an ultimatum to withdraw from West Berlin, the construction of the Berlin Wall, and the Checkpoint Charlie incident where opposing tanks faced off at the border.

an unprecedented and comprehensive view of diplomacy during one of the most serious moments of tension in modern American history.¹⁰ The documents, which represent diplomatic signals, vary on two different dimensions: whether they are from the Soviet Union to the United States or vice versa, and whether they are public or private.

Multiple collections of archival data capture public signals. Missives from the Soviet Union are from the Foreign Broadcast Information Service (FBIS), which tracks public statements by the Soviet Union. The FBIS began as an open-source intelligence project originally led by the Central Intelligence Agency, and its main goal was to translate and record foreign countries' public statements made via radio and press releases. Policymakers relied heavily on this information stream to infer its adversaries' intentions (Leetaru 2010). These records fully capture *public diplomatic statements* by the Soviet Union. The original English translations of all FBIS entries, both in photographed and digital text formats, are housed online on the NewsBank service. All documents collected involve the topic of Germany and/or Berlin.

Public statements from the United States come from two sources. The first is public statements made by Presidents Eisenhower and Kennedy. These come from the Public Papers of the Presidents of the United States (PPP) collections, which are available online through the University of Michigan Digital Library.¹¹ The second are press conferences (PC) led by the United States Department of State, which come from the National Archives II at College Park, Maryland.¹²

Private statements between the Soviet Union and the United States are preserved in records from the Department of State (DOS), which has a collection of declassified telegrams also maintained at the National Archives II in College Park, Maryland.¹³ These documents are not available online. As such, we manually photographed the entirety of Berlin-related collections from the Department of State during this time period. Each photographed page was then converted into computer-readable text using optical character recognition (OCR) software. We review these documents to exclude unclassified records and then disaggregate the remaining data into incoming telegrams (Soviet Union to the United States; DOS_{in}) and outgoing telegrams (United States to the Soviet Union; DOS_{out}). Incoming messages typically involve the United States Embassies in

¹⁰See Katagiri and Min (2019) for a brief summary of the five-year period.

¹¹The access to the database is at <https://quod.lib.mich.edu/p/ppotpus/>.

¹²Daily News Conferences, Office of Press Relations, Department of State, Vol. 17-31 (1958-1963).

¹³Record Group 59: Department of State Records, Central Files.

Direction	Data Source	Type	# Docs	# Segments
USSR to US	DOS _{in}	Private	4,012	7,777
	FBIS	Public	10,714	13,576
US to USSR	DOS _{out}	Private	903	1,259
	PPP	Public	2,641	10,088
	PC	Public	1,006	22,148

Table 1: Data sources.

Bonn and Moscow, as well as the U.S. Mission Berlin (which was the State Department’s substitute for having an embassy in West Berlin). Each document summarizes meetings with foreign government officials and other important messages to be exchanged between the recipient state’s government and the American government. The DOS collections therefore represent *private diplomatic statements* between the Soviet Union to the United States.

All documents underwent standard text preprocessing and were broken down into 300-word segments to help capture heterogeneity of topics and content in longer statements. Table 1 summarizes these sets of data.

We are far more confident of the completeness of our Soviet data than our US data. As such, our analysis primarily focuses on explaining Soviet signaling behavior, and we use data on US signaling behavior as control variables. The US data may not be complete enough to stand alone in its own analysis, but we believe it is informative enough to at least provide valuable background on the interaction between the two superpowers.

The Berlin Crisis represents a rather difficult case for our argument. This five-year period constitutes an ideal case where both elites and the general public were highly cognizant of tensions regarding Berlin. Conditions were ripe for the Soviets to make aggressive public statements that would easily gain attention not only in the United States but around the world. The Soviets certainly did have prominent moments of public threat-making; consider Khrushchev’s first ultimatum against the allied powers in November 1958, which marked the start of the entire crisis. Those moments notwithstanding, any evidence that the Soviets stepped back from public signals and instead preferred to communicate with the United States privately would suggest that stand in tension would the idea that public threats are the only useful coin of the diplomatic realm or that third-party audiences are irrelevant to understanding signaling strategies.

Measuring Tone and Topics

To assess our predicted observable implications, we must convert these raw documents into measures of diplomatic behavior. Our goal here is to create multiple measures of behavior that each serve as plausible and distinct reflections of the extent to which actors relied on public or private channels to communicate their intentions.

We currently have measures for three underlying dimensions: quantity, tone, and topic(s). We gauge the characteristics of each dimension separately for public and private signals, which produces daily-level measures for each stream of communication. By comparing these daily-level measures between public and private signals, we can determine the extent to which actors relied relatively more heavily on one channel versus the other.

It is very straightforward to measure the quantity of public and private signals, so we do not discuss that in much detail here. Subfigure 5a illustrates the general trends for Soviet signaling over the entirety of our time frame. However, measuring the tone and topics in each document requires additional effort and discussion. We explain the process to quantify each of these two latter dimensions below.

Tone

To create a measure of tone, we apply the Linguistic Inquiry and Word Count, or LIWC (Tausczik and Pennebaker 2010).¹⁴ LIWC is well-established tool that identifies the use of words in numerous psychologically relevant categories. LIWC includes dozens of separate dictionaries that reflect different psychological, social, and emotional states. For our purposes, we use LIWC to calculate what percentage of words are included in its dictionary of terms associated with the emotion of anger.

LIWC appears to produce measures of anger that seem both useful and plausible. While future drafts of this paper will supply additional evidence to this effect, for now we note that the following FBIS entry from December 18, 1962 is one of the highest scoring signals in terms of anger:

EXPLOSIVE ATTACK. The new bombing attack by West Berlin provocateurs against the GDR state border is the result of the agitation speeches by the frontline city politi-

¹⁴Our analysis relied on the 2015 version of the LIWC dictionary.

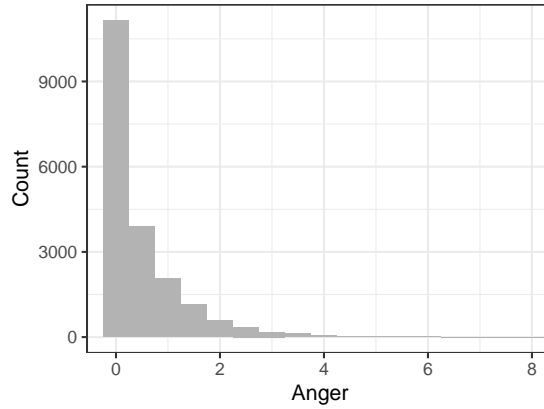


Figure 1: Prevalence of words indicating anger in diplomatic documents.

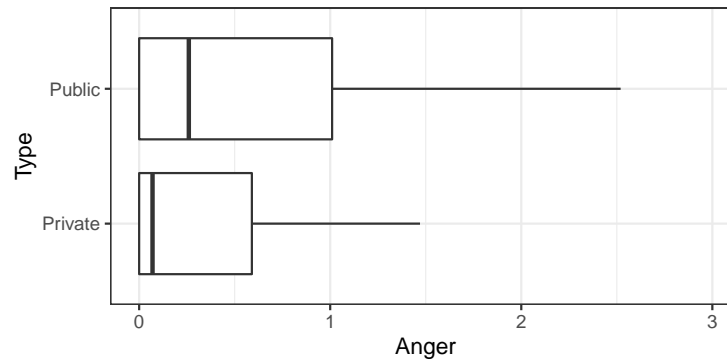


Figure 2: Prevalence of words indicating anger in diplomatic documents, disaggregated by type. Outliers are removed.

cians in which people again and again are asked to attack the wall The incited rowdies prepare the bombing attacks with the consent of the West Berlin police and Senate authorities Foreign journalists in West Berlin asked whether Brandt wants to cause additional troubles for the West Berlin people because the East might feel induced to take countermeasures in view of such provocations East Berlin.

Figure 1 displays the distribution of anger in the diplomatic documents. In the majority of signals, fewer than 1% of words are considered to express anger. However, the observed percentages can go as high as 16%. Figure 2 disaggregates the anger measure by whether the signal is public or private. The gap between the two is distinct; public messages are disproportionately more reliant on words of anger than messages exchanged in private.

The wide gap in anger between public and private diplomatic messages is striking in its own right. The fact that public signals evince such higher levels of general anger, but that the Berlin

Crisis did not result in repeated catastrophes, suggests that public signals are often not properly acknowledged or are too imprecise to cause alarm (Katagiri and Min 2019). At the same time, it may be a reflection of the fact that diplomats realize that some aspects of public communication are products of “theatricality” (Jönsson 1996, 27).¹⁵ Diplomacy in the public eye may be subject to performative incentives, and importantly, international diplomatic actors engage in tacit cooperation to keep this order and shared set of behavioral expectations afloat (Sending et al. 2015).¹⁶ This is certainly not to argue that all public diplomatic threats are easily dismissed. But it does suggest another important manner in which public and private channels qualitatively differ.¹⁷

Topics and Topic Similarity

To assess the content of individual diplomatic signals, we must systematically determine the subject matter of each statement segment. We use a structural topic model (STM) to perform this task (Roberts et al. 2017). STMs are generative and largely unsupervised models of word counts that attempt to identify topics (mixtures of words) across a set of text corpora, as well as the distribution of these topics within each piece of text.

There is no hard and fast rule for the appropriate number of N topics that should be in a STM. For our purposes, a model with 40 topics strikes a useful balance of exclusivity (where words are more unique to their topics) and semantic coherence (where the most probable words in a topic co-occur). Beyond these metrics, the topics that emerge from this model are also largely sensible to us, both in general and in the context of the Berlin Crisis. Table 2 some of the most relevant topics with the three most indicative words according to the FREX measure, which identifies words that are both most frequent and exclusive to a particular topic.

How much do documents vary in their composition of topics? Figure 3 speaks to this question, splitting the data apart by whether signals are public or private. The diagram firmly shows that

¹⁵See also Ringmar (2012); Shimazu (2014). An authoritative text on diplomacy, *Satow’s Diplomatic Practice*, almost appears to argue that public communication does not satisfy a classical or traditional notion of diplomacy: “[T]he problem is that if actual negotiation is carried out entirely in the public eye ... it quickly turns into a travesty of efficient procedure and runs the risk of betraying any constructive purpose for which it may have been conceived” (Roberts 2009, 14-15).

¹⁶McConnell (2018) argues that “unofficial” representatives from marginalized communities and weak states understand and endeavor to replicate these norms of diplomatic decorum.

¹⁷One other potential explanation for this gap, according to our own argument, might be that the data capture a time when the Soviets faced a very high number of active adversaries relative to the number of disputes they had with the United States. However, the data do not bear out this claim. The majority of disputes the Soviets faced in this period involved the United States.

	Topic	FREX ₁	FREX ₂	FREX ₃
1	War and peace	peac	imperialist	coexist
7	Convoys	vehicl	convoy	checkpoint
9	Flights	flight	aircraft	corridor
13	Atomic weapons	nuclear	atom	test
18	Financial assistance	fund	million	dollar
21	Military/Defense	defens	arm	militari
27	Berlin’s status	western	power	occup
30	Negotiation	propos	negoti	accept
32	Trade	trade	market	tariff
35	Summits	confer	summit	geneva

Table 2: Topics with most indicative tokens. The labels in the “Topic” column are determined by the authors and are not inherent to the model’s output itself.

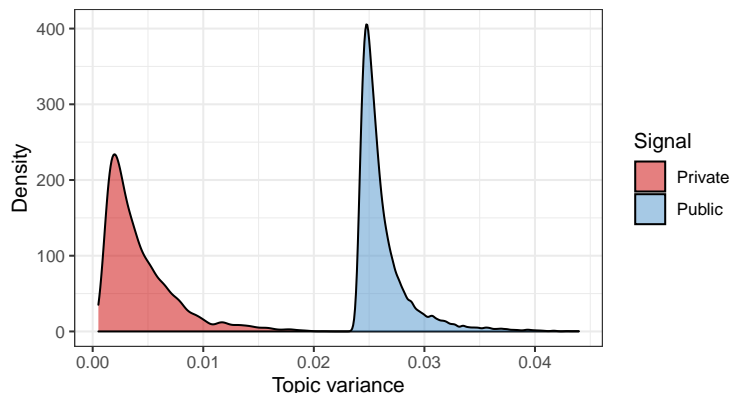


Figure 3: Variance of topic prevalence in private and public documents.

public statements feature far more variation in their content. A simple t -test finds this difference to be highly statistically significant ($p \ll 0.01$). It is unsurprising that public interactions could touch upon a much wider array of issues. However, this obvious finding is also a testament to how public and private channels frequently serve different purposes.

Figure 4 shows the relationship between angry words and topics in the documents. The distributions in the plots suggest that the anger measure based on LIWC is plausible; topics involving the militaristic affairs, atomic weapons, corridor issues, espionage, and war/peace are associated with higher proportions of words expressing anger. Note that the anger measure is not identical to the prevalence of these more “hawkish” topics. The correlation between anger and these listed topics is, at maximum, 0.36. Meanwhile, more innocuous subjects like trade, law, and financial aid feature quite low levels of anger. The intuitive relationships between the topics and associated tones

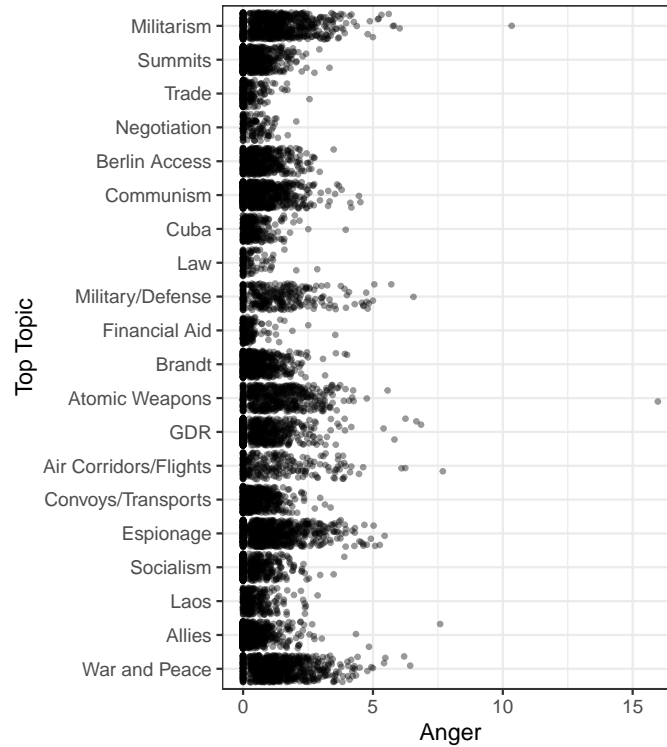


Figure 4: Prevalence of words expressing anger in documents, conditional on the most common topic in the document.

simultaneously lend confidence to both measures while emphasizing the value of disaggregating them.

With the STM output in hand, we can create measures of document similarity in terms of the topics discussed in them. Recall that the STM involves 40 topics ($N = 40$). Each document processed through this model is associated with a vector of length 40, where each entry reflects the prevalence of a particular topic. Since topic prevalence is expressed as a probability, each 40-dimensional vector adds up to 1.

In order to find the similarity of topic compositions between two documents, we can calculate the pair’s cosine similarity.¹⁸ A cosine similarity of 0 indicates that the two vectors are completely orthogonal and do not match in any way. As a cosine similarity approaches 1, the pair of documents have topic prevalences that are more closely aligned.

¹⁸Formally, for two non-zero vectors d_1 and d_2 , cosine similarity is defined as

$$\cos(\theta) = \frac{d_1 \cdot d_2}{\|d_1\| \cdot \|d_2\|}$$

For each day in our data, we gathered all public (or private) documents made from that day as well as the previous week. We then calculate the number of total documents, the documents' average level of expressed anger, and the average cosine similarity between all pairs of documents.

Figure 5 encapsulates all the aforementioned measures, illustrating each as a function of time. We see that public statements are generally higher in volume as well as their levels of expressed anger through the Berlin Crisis. Straightforward t -tests indicate that these differences are statistically significant for both measures ($p \ll 0.001$). Subfigure 5c shows that private signals are generally more similar to each other than is the case for public signals. This general gap is also highly statistically significant ($p \ll 0.001$). Given that private signals generally feature less variance (see Figure 3), this is a natural result.

These distinctions strengthen and contextualize the finding that private signals of resolve cause more alarm than those in public (Katagiri and Min 2019). Public signals are less concerning to elites *even though* their overall tone tends to be hostile and their frequency is considerably higher. To some extent, this may be because public signals are also more varied in the subjects they cover. But this could also reflect an understanding that much of what is stated in public is performative and not to be taken as seriously.

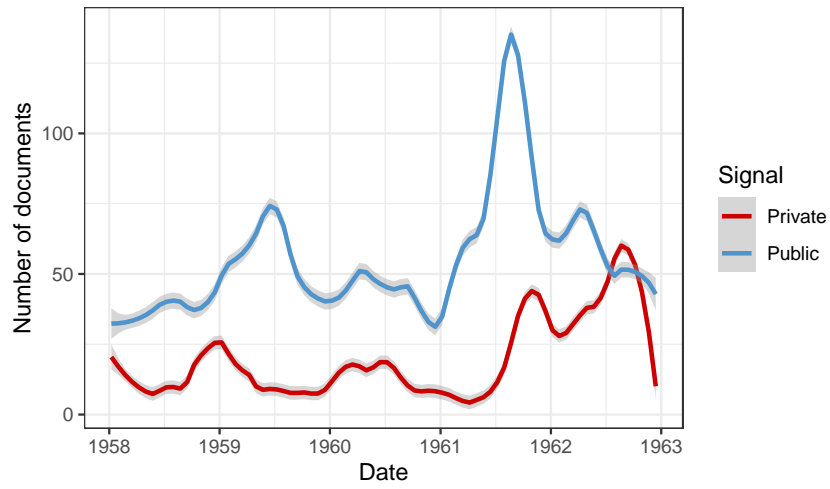
Analysis

Our analysis investigates the extent to which Soviets relied on public versus private signals to discuss Berlin, depending on the nature of confrontations it faced with the United States or with other parties.

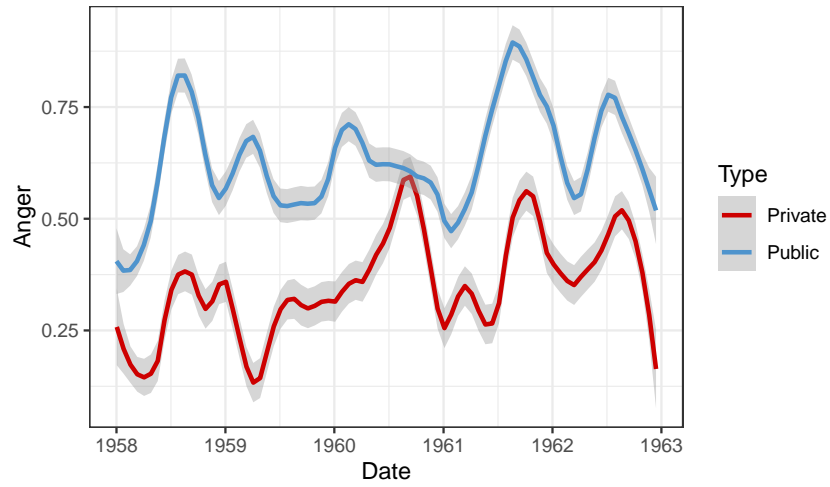
To measure reliance on public or private signals, we construct three dependent variables which are based on the measures illustrated in Figure 5. For each dimension of signaling on some day t , we subtract the value of the private dimension from the value of the public dimension—that is, $X_{\text{Public},t} - X_{\text{Private},t}$, where X is the number of documents, expressions of anger, or similarity of topics. Note that this means lower or more negative values are associated with greater reliance on private channels, while higher or more positive values are associated with greater reliance on public channels. We use a series of OLS regressions to analyze these measures.¹⁹

¹⁹Even though the number of documents may be a count variable, which would appear to necessitate a Poisson model, recall that the outcome variable is the comparison of document volume between public and private channels.

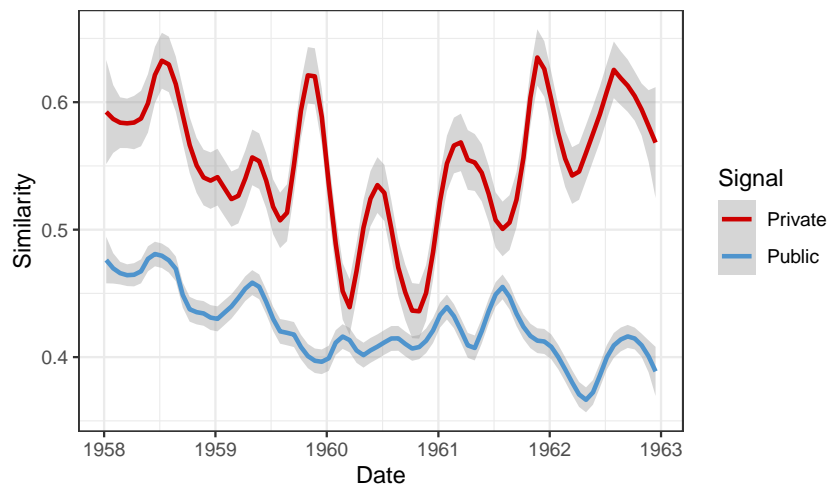
Figure 5: Public and private Soviet signaling, using loess curves.



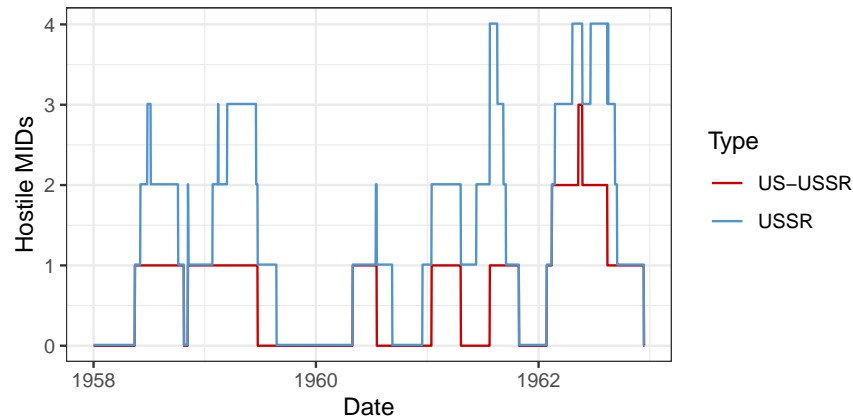
(a) Number of daily documents.



(b) Prevalence of words expressing anger.



(c) Similarity of topics.

Figure 6: MIDs involving the Soviet Union.

Two explanatory variables capture contemporaneous confrontations. Both are derived from the Militarized Interstate Dispute dataset. To look at specific disagreements between the Soviet Union and the United States, we track the number of hostile militarized interstate disputes (MIDs) between the Soviet Union and the United States at the daily level. To look at general crises that are may have preoccupied the Soviet Union, we track the number of hostile militarized interstate disputes involving the Soviet Union, regardless of whether the United States was also involved. Inclusion of US-USSR MIDs in the global count is important, as our theoretical argument would suggest that any disputes embroiling the Soviets should motivate them to send out public signals for other audiences. We limit our analysis to MIDs which are recorded as having a hostility level of three or higher. Figure 6 illustrates these two measures.

It is worth noting that some of the MIDs that do not involve the United States still implicate the United Kingdom and France, both of which were part of the Western alliance administering West Berlin. Between 1958 and 1963, the Soviets were parties to 26 hostile MIDs that involved a total of 45 adversarial states. Thirteen of these adversarial roles (29%) were the United States, four (9%) were the United Kingdom, and two (4%) were France.

We include a series of control variables to account for potential confounders. The first set of variables recognize the fact that Soviet signals do not exist in a vacuum; they are part of a strategic interaction with other states, with the United States being the most relevant party. Diplomatic signals regarding Berlin may also be shaped by material actions taking place in Berlin. Such activities are not only noteworthy in sinking costs, but also in being more vivid signals that are

likely to attract more attention (Yarhi-Milo 2013a). We account for material actions taken by both sides, which are based on *New York Times* articles and were compiled by Katagiri and Min (2019). We use this data to include variables for the one-week period *before Soviet actions*, *after Soviet actions*, and *after US actions*. We include a variable for the time period before Soviet actions to account for the possibility that diplomatic signals could be manipulated in anticipation of material actions. An analogous variable is not added for the United States under the assumption that the Soviets could not fully anticipate American actions.

Moreover, some choices that the Soviets make regarding their diplomatic communications could be consequences of responding to American diplomatic activities. For each outcome variable, we therefore also include *lagged US variable*, which represents an analogous version of the variable using US documents as a control.

To account for the impact of material capabilities on Soviets' willingness to signal in particular ways, we also include measures of each state's material capabilities using the CINC score from the National Material Capabilities dataset.

Two variables reflect leadership in the United States. A dummy variable tracks whether the Soviets are dealing with the *Kennedy* administration, which came into office in January 1961. Another variable measures *presidential tenure* by tracking the (logged) number of days that either Eisenhower or Kennedy were in office. This helps to address the possibility that presidents and/or the Soviets gained experience dealing with an administration and learned to adjust their signaling strategies over time.

Finally, all models include a cubic spline for time.

Results

Table 3 summarizes the results from a series of models regarding Soviet signaling behavior during and regarding Berlin. Odd-numbered models are sparse analyses that only include the explanatory variables and the cubic spline for time; even-numbered models are full specifications with all relevant controls. Fully specified models for anger and topic similarity lose a significant number of observations because of the relative sparsity of our US diplomatic data. Nevertheless, results are unchanged by removing the lagged US variables.

Recall that the outcome variables are comparisons of document volume, anger, and topic similarity between public and private channels. Positive values indicate greater levels of these concepts in public signals, while negative values indicate greater levels in private signals.

A highly consistent story emerges across all models: Soviets rely more on public signals when they are juggling multiple crises around the world at once, and they rely more on private signals to the United States when an increasing number of crises involve the United States. One minor exception arises in Model 2, where the estimated coefficient for US-USSR MIDs is negative, which is consistent with our expectations, but is not statistically significant. The fact that higher numbers of global MIDs involving the USSR augment the Soviets' reliance on public messages is particularly noteworthy. Recall that our public signal data from the FBIS is explicitly limited to documents that mention Berlin. Even within these tight substantive confines, we see that other disputes outside of Berlin shape Soviet signaling behavior when discussing Berlin.

We have already seen that levels of anger are consistently higher in public than they are in private. However, the gap between these two streams of communication is more likely to close or even become reversed when the Soviets are embroiled in more disputes with the United States and in the aftermath of material actions perpetrated in Berlin by either side.

These effects are substantively significant. The average value for document balance is 23.28 (which means that, on average, there are 23.28 more public Soviet signals sent over the previous seven days compared to analogous private signals). As such, according to Model 2, each additional MID involving the Soviet Union is analogous to a 24% increase relative to the observed mean value. In the case of anger, the standard deviation across all observations is 0.304. Model 4 therefore suggests that each additional MID has an effect that is equivalent to increasing anger in public signals by 0.311 standard deviations; each additional MID involving the United States results in more anger in private signals by about 0.420 standard deviations. Similarly, for topic similarity, the standard deviation across the data is 0.131. Every MID is associated with an increase in topic similarity in public signals by a magnitude equivalent to 0.145 standard deviations; a MID involving the United States leads to more topic similarity in private documents by about 0.176 standard deviations.

Figure 7 makes the results slightly more tangible by producing predicted outcomes for several configurations of MIDs, based on the minimum and maximum number of MIDs observed in the

Table 3: Results of OLS regressions on Soviet signaling patterns.

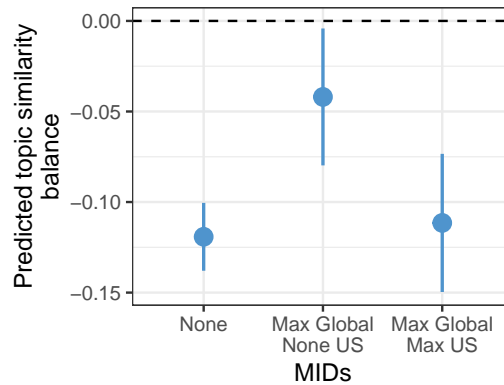
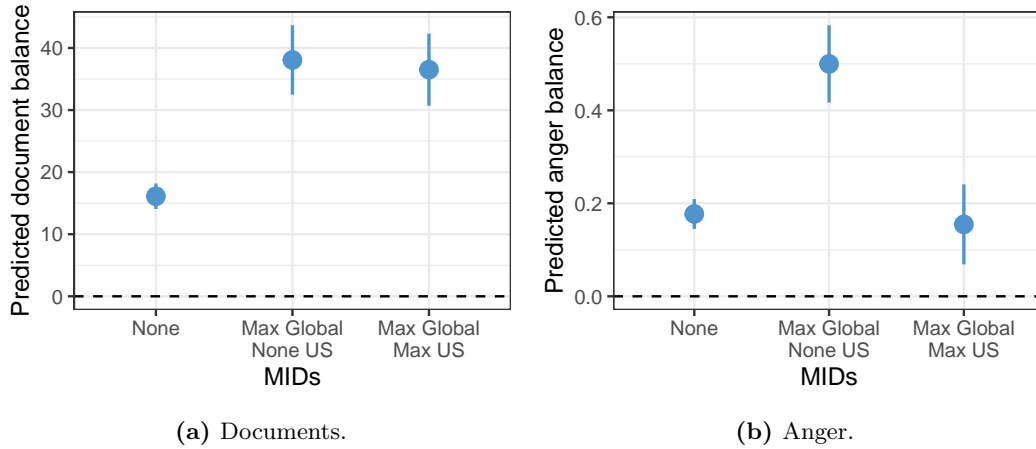
	<i>Dependent variable:</i>					
	Quantity		Anger (Tone)		Topic similarity	
	(1)	(2)	(3)	(4)	(5)	(6)
USSR MIDs	9.253*** (0.717)	5.488*** (0.697)	0.095*** (0.010)	0.081*** (0.010)	0.013*** (0.004)	0.019*** (0.004)
US-USSR MIDs	-7.470*** (1.420)	-0.520 (1.415)	-0.128*** (0.020)	-0.115*** (0.020)	-0.027*** (0.008)	-0.023*** (0.009)
Before USSR actions		7.921*** (1.271)		0.022 (0.018)		0.043*** (0.008)
After USSR actions		-2.232* (1.269)		-0.061*** (0.018)		-0.015* (0.008)
After US actions		-3.142 (1.955)		-0.036 (0.026)		-0.026** (0.011)
Lagged US documents		0.480*** (0.069)				
Lagged US anger				0.080*** (0.017)		
Lagged US topic similarity						0.002 (0.020)
USSR CINC		-12.808*** (2.409)		-0.026 (0.033)		-0.033** (0.014)
US CINC		6.813*** (2.120)		-0.205*** (0.033)		-0.126*** (0.018)
Kennedy		78.799*** (6.395)		1.010*** (0.107)		0.138* (0.074)
Presidential tenure		15.758*** (1.355)		0.177*** (0.025)		0.031 (0.020)
Constant	6.220*** (2.075)	-149.818*** (10.435)	0.298*** (0.028)	-0.767*** (0.190)	-0.131*** (0.012)	-0.184 (0.145)
Observations	1,803	1,802	1,796	1,313	1,758	978
Time cubic spline	✓	✓	✓	✓	✓	✓

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

data.²⁰ A potential fourth category, which involves the maximum number of US-USSR MIDs and no global count of MIDs, is not well-defined because the global count includes US-USSR disputes. The subfigures for anger and topic similarity show the cross-cutting effects of being involved in disputes around the world versus dealing with confrontations involving the United States.²¹ When dealing with multiple disputes around the world, public Soviet signals become significantly more likely to include expressions of anger compared to their private signals. However, as a large number

²⁰All other covariates are fixed at their median observed values. The highest number of any MIDs the Soviets dealt with at once, at least in our data, is 4. The maximum value of MIDs involving the US is 3.

²¹One may be concerned that the two MIDs measures are highly correlated. Indeed, the direct correlation between these two variables is 0.82. In the analysis to come, variance inflation factors suggest that these two variables do not introduce issues of collinearity.

Figure 7: Predicted outcomes at maximal values of MIDs.

(c) Topic similarity.

of these MIDs involve the United States, the Soviets gradually redirect more of their angry language away from public channels and toward private ones. The similarity of topics discussed in documents follows a similar pattern. The model suggests that the balance of behaviors between public and private channels—in terms of expressions of anger and similarity of topics—is roughly equivalent in a world where the Soviets face no disputes and a world where the Soviets face a large number of disputes, with the majority involving the United States. However, the occurrence of disputes, regardless of the adversary, leads to an increase in the Soviet Union’s volume of public signals compared to private ones.

Several controls reveal additional insights regarding the determinants of Soviet signaling behavior. Moments around material actions in Berlin appear to influence which channel Soviets use to communicate and how they choose to communicate. Models 2 and 6 suggest that Soviet elites were more prone to launch missives and to be more consistent in their choice of topics in public

communication in the week prior to the actions (such as halting military convoys) they enacted against the Western allies in Berlin. Model 4 then finds that expressions of anger become more prevalent in private Soviet communications after material actions by either side. If we assume material actions to be a sign of resolve, then this intimates a general story regarding how Soviets use different diplomatic channels sequentially: More similar public signals are followed by material actions, which are followed by more discontented private signals to the United States. The fact that private signals come after public messages and material actions suggests that the Soviets highlight and contextualize their public and somewhat more imprecise public gestures with additional signals sent directly and privately to the United States.

The Soviet Union appears to more readily discuss things in private as its material capabilities increase. Meanwhile, it sends more public signals but is more consistent in the topics it discusses in private when the United States increases its capabilities. Limitations exist to how deeply we can interpret these findings, as both states' capabilities consistently increased with each year.

The longer either President Eisenhower or Kennedy was in office, the more the Soviets leaned into angry and more frequent public signals. The Kennedy administration also experienced higher levels of public Soviet signaling than its predecessor.

“Conclusions” and Future Steps

It is too early for us to state firm conclusions in this rough draft. Nevertheless, the ideas and findings in this current paper highlight some important aspects of diplomatic signaling that are not fully appreciated by many works in IR. We find highly consistent patterns in Soviet diplomatic behavior during one of the most heightened periods of tension and existential risk in modern history. Perhaps going against standard expectations, the results portray a somewhat cautious party that had no interest in using its bullhorn when dealing only with the United States. The Soviets frequently hewed toward relaying their displeasure and exercising greater precision in the content of their messages when communicating in private with the White House. On one hand, the combination of hostility and clarity in private signals may look like a much more alarming form of signaling that would easily rattle policymakers (Katagiri and Min 2019). Yet on the other hand, being willing and able to transmit more precise signals that are less likely to trigger unintended consequences from

third-party audiences is a form of restraint compared to the more absolute but imprecise threats that could be levied in public.

While much of our analysis leverages the Berlin Crisis, we believe this research contributes to an important trend of studying the logic of “everyday” diplomacy that occurs outside the crisis environment (Trager 2017). There may certainly be moments when public threats become highly salient and commonly understood to be grave without any private communications backing them up. Belligerent public diplomacy may also be used to reshape public sentiment to be more predisposed to conflict. Nevertheless, focusing only on these major upheavals may provide a biased view of what diplomacy represents. In their seminal work on crisis bargaining, Snyder and Diesing (1977) explain their focus on crises because “[c]onflict is central to all politics, especially international politics, and crises are conflict episodes *par excellence*” (3). It is equally important to note that their goal was to have theory developed from this domain “extended outward, with appropriate qualifications, to the more peripheral aspects” (3). This paper suggests that some of the key assumptions and dimensions scholars have absorbed from crisis diplomacy may not be as relevant or complete in studying the vast majority of strategic but quotidian interactions between states.

This paper is a preliminary product. At most, our goal here was to demonstrate the relevance, feasibility, and promise of a research agenda that seriously addresses the strategy behind diplomacy and goes beyond the analysis of signals in isolation or in the context of only two actors. We envision some important additions and refinements in subsequent versions. Three are mentioned here.

Our argument and statistical analysis focused on the role that third-party states may have on an actor’s choice of signaling via public or private channels. We have left aside the issue of whether domestic audiences influence this calculus. After all, our core data focuses on the Soviet Union, which is an autocratic government throughout the conflict and giving us no leverage to investigate regime type. Existing theories would perhaps suggest that, relatively speaking, the Soviet Union should be less concerned about its domestic public. With one fewer audience of concern, the Soviets may be freer to engage in public diplomatic posturing. Some tentative models using our data suggest that the United States’ signaling behavior during the Berlin Crisis was similar to Soviet conduct, with one exception: American diplomatic signals communicated more of their anger through private channels regardless of whether active disputes involved the Soviet Union. We require additional information and checks to stand firmly behind this finding. However, if true, this distinction would

be consistent with the idea that elites in democratic countries had reasons to limit public verbal demonstrations of hostility. The idea that both states behave similarly when it comes to topic similarity but differ with respect to tone (in terms of anger) underscores the value of measuring these two attributes of diplomatic signaling separately.

Additionally, our study only analyzes diplomatic interactions between the United States and the Soviet Union. These two states are obviously important in the context of the Cold War, but it may be valuable to verify whether our predictions hold for other countries during this same time period. For example, if we found evidence that the Soviets increased the hostility and focus of their private messages to the British when embroiled in disputes with Great Britain, we would have even stronger and broader evidence about how states balance their use of public and private messages.

Finally, the results we found for variables capturing the moments before and after material actions generated interesting suggestive findings regarding the sequencing of diplomatic signals. More research directly analyzing this topic may prove fruitful. Previous scholarship has already identified that the deployment of specific types of signals may depend on what other signals have already been sent (Leng and Walker 1982; Leng and Wheeler 1979; Snyder and Diesing 1977), but few empirical works have fully illuminated how some kinds of signals may precede or contextualize others. The dynamic manner in which public and private signals are shown to be harnessed in this paper suggest that we have much more to learn and appreciate about the complexities of diplomacy.

References

- Carson, Austin and Keren Yarhi-Milo. 2017. "Covert Communication: The Intelligibility and Credibility of Signaling in Secret." *Security Studies* 26(1): 124–156.
- Clare, Joe and Vesna Danilovic. 2010. "Multiple Audiences and Reputation Building." *Journal of Conflict Resolution* 54(6): 860–882.
- Debs, Alexandre and Jessica Chen Weiss. 2016. "Circumstances, Domestic Audiences, and Reputational Incentives in International Crisis Bargaining." *Journal of Conflict Resolution* 60(3): 403–433.
- Downes, Alexander B. and Todd S. Sechser. 2012. "The Illusion of Democratic Credibility." *International Organization* 66(3): 457–489.
- Farrell, Joseph and Robert Gibbons. 1989. "Cheap Talk with Two Audiences." *The American Economic Review* 79(5): 1214–1223.
- Fearon, James D. 1994. "Domestic Political Audience Costs and the Escalation of International Disputes." *American Political Science Review* 88(3): 577–592.
- Fearon, James D. 1997. "Signaling Foreign Policy Interests: Tying Hands Versus Sunk Costs." *Journal of Conflict Resolution* 41(1): 68–90.
- Fleming, John H. and John M. Darley. 1991. "Mixed Messages: The Multiple Audience Problem and Strategic Communication." *Social Cognition* 9(1): 25–46.
- Fleming, John H., John M. Darley, James L. Hilton, and Brian A. Kojetin. 1990. "Multiple Audience Problem: A Strategic Communication Perspective on Social Perception." *Journal of Personality and Social Psychology* 58(4): 593–609.
- Goltsman, Maria and Gregory Pavlov. 2011. "How to talk to multiple audiences." *Games and Economic Behavior* 72(1): 100–122.
- 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.

- Holmes, Marcus. 2013. "The Force of Face-to-Face Diplomacy: Mirror Neurons and the Problem of Intentions." *International Organization* 67(4): 829–861.
- Jönsson, Christer. 1996. "Diplomatic Signaling in the Television Age." *Press/Politics* 1(3): 24–40.
- Jönsson, Christer and Martin Hall. 2002. "Communication: An Essential Aspect of Diplomacy." *International Studies Perspectives* 4(2): 195–210.
- 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.
- Kertzer, Joshua D. and Ryan Brutger. 2016. "Decomposing Audience Costs: Bringing the Audience Back into Audience Cost Theory." *American Journal of Political Science* 60(1): 234–249.
- Kurizaki, Shuhei. 2007. "Efficient Secrecy: Public versus Private Threats in Crisis Diplomacy." *American Political Science Review* 101(3): 543–558.
- Lake, David A. 2010/11. "Two Cheers for Bargaining Theory: Assessing Rationalist Explanations of the Iraq War." *International Security* 35(3): 7–52.
- 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. and Stephen G. Walker. 1982. "Comparing Two Studies of Crisis Bargaining." *Journal of Conflict Resolution* 26(4): 571–591.
- Leng, Russell J. and Hugh G. Wheeler. 1979. "Influence Strategies, Success, and War." *Journal of Conflict Resolution* 23(4): 655–684.
- 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.
- McConnell, Fiona. 2018. "Performing Diplomatic Decorum: Repertoires of "Appropriate" Behavior in the Margins of International Diplomacy." *International Political Sociology* 12(4): 362–381.

- McManus, Roseanne W. and Keren Yarhi-Milo. 2017. "The Logic of "Offstage" Signaling: Domestic Politics, Regime Type, and Major Power-Protégé Relations." *International Organization* 71(4): 701–733.
- Pu, Xiaoyu. 2019. *Rebranding China: Contested Status Signaling in the Changing Global Order*. Stanford, CA: Stanford University Press.
- Quek, Kai. 2021. "Four Costly Signaling Mechanisms." *American Political Science Review* 115(2): 537–549.
- Ramsay, Kristopher W. 2011. "Cheap Talk Diplomacy, Voluntary Negotiations, and Variable Bargaining Power." *International Studies Quarterly* 55(4): 1003–1023.
- Ringmar, Erik. 2012. "Performing International Systems: Two East-Asian Alternatives to the Westphalian Order." *International Organization* 66(1): 1–25.
- Roberts, Ivor. 2009. *Satow's Diplomatic Practice*. Oxford: Oxford University Press.
- Roberts, Margaret E., Brandon M. Stewart, and Dustin Tingley. 2017. "stm: R Package for Structural Topic Models." *Journal of Statistical Software* .
- Sartori, Anne E. 2002. "The Might of the Pen: A Reputational Theory of Communication in International Disputes." *International Organization* 56(1): 121–149.
- Sartori, Anne E. 2005. *Deterrence by Diplomacy*. Princeton: Princeton University Press.
- Schelling, Thomas C. 1966. *Arms and Influence*. New Haven, CT: Yale University Press.
- Schultz, Kenneth A. 2001. *Democracy and Coercive Diplomacy*. Cambridge: Cambridge University Press.
- Sending, Ole Jacob, Vincent Pouliot, and Iver B. Neumann. 2015. "Introduction." In *Diplomacy: The Making of World Politics*, eds. Ole Jacob Sending, Vincent Pouliot, and Iver B. Neumann. Cambridge: Cambridge University Press, pages 1–28.
- Shimazu, Naoko. 2014. "Diplomacy As Theatre: Staging the Bandung Conference of 1955." *Modern Asian Studies* 48(1): 225–252.

- Snyder, Glenn H. and Paul Diesing. 1977. *Conflict Among Nations: Bargaining, Decision Making, and System Structure in International Crises*. Princeton: 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): 437–456.
- Tarar, Ahmer and Bahar Leventoğlu. 2012. "Limited Audience Costs in International Disputes." *Journal of Conflict Resolution* 57(6): 1065–1089.
- Tausczik, Yla R. and James W. Pennebaker. 2010. "The Psychological Meaning of Words: LIWC and Computerized Text Analysis Methods." *Journal of Language and Social Psychology* 29(1): 24–54.
- Tomz, Michael R. 2007. "Domestic Audience Costs in International Relations: An Experimental Approach." *International Organization* 61(4): 821–840.
- 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: Cambridge University Press.
- Van Boven, Leaf, Justin Kruger, Kenneth Savitsky, and Thomas Gilovich. 2000. "When Social Worlds Collide: Overconfidence in the Multiple Audience Problem." *Personality and Social Psychology Bulletin* 26(5): 619–628.
- Vertzberger, Yaacov Y.I. 1990. *The World in Their Minds: Information Processing, Cognition, and Perception in Foreign Policy Decisionmaking*. Stanford, CA: Stanford University Press.
- Walter, Barbara F. 2006. "Building Reputation: Why Governments Fight Some Separatists but Not Others." *American Journal of Political Science* 50(2): 313–330.
- Weeks, Jessica L. 2008. "Autocratic Audience Costs: Regime Type and Signaling Resolve." *International Organization* 62(1): 35–64.
- Weisiger, Alex and Keren Yarhi-Milo. 2015. "Revisiting Reputation: How Past Actions Matter in International Politics." *International Organization* 69(2): 473–495.

- Wolford, Scott. 2020. "War and diplomacy on the world stage: Crisis bargaining before third parties." *Journal of Theoretical Politics* 32(2): 235–261.
- Wong, Seanon S. 2016. "Emotions and the Communication of Intentions in Face-to-Face Diplomacy." *European Journal of International Relations* 22(1): 144–167.
- Yarhi-Milo, Keren. 2013a. "In the Eye of the Beholder: How Leaders and Intelligence Communities Assess the Intentions of Adversaries." *International Security* 38(1): 7–51.
- Yarhi-Milo, Keren. 2013b. "Tying Hands Behind Closed Doors: The Logic and Practice of Secret Reassurance." *Security Studies* 22(3): 405–435.
- Yarhi-Milo, Keren. 2014. *Knowing the Adversary: Leaders, Intelligence, and Assessment of Intentions in International Relations*. Princeton: Princeton University Press.