How the Scope of a Demand Conveys Resolve

Robert F. Trager†

Word Count: 9,029

How does the scope of costless threats convey information about resolve to adversaries? Analysis of a model similar to Fearon (1995) demonstrates that higher demands increase perceptions of a state’s resolve to fight for more favorable outcomes when either (a) settling on a negotiated solution is the outcome of a bargaining process in which both sides take part, rather than a take-it-or-leave-it offer from one side or (b) the issue space exhibits indivisibilities. Interestingly, compromise offers will be made even though they increase an adversary’s perception that the compromising state would be willing to make an even greater concession. In contrast to many other signaling mechanisms described in the literature, signaling of this sort does not depend on risking war and often reduces the probability of conflict.

† Assistant Professor, University of California, Los Angeles, rtrager@ucla.edu. Please do not cite without permission. I am grateful to Robert Jervis, Dov Levin, Chad Nelson, Barry O’Neill, Bob Pape and seminar participants at UCLA and the University of Chicago for trenchant comments and advice. Any missteps are mine.
In 1859, with Piedmont inspiring insurrections in Austria’s Italian provinces, the Austrians demanded the demobilization of the Piedmontese. A compromise proposal was floated according to which the Austrians would pull their troops back from the boarder in return for Piedmontese disarmament. The Austrians rejected the compromise, insisting on disarmament without an Austrian withdrawal. With only these two options on the table, the Austrian decision not to compromise conveyed information about the conditions under which Austria would go to war. Observing the Austrian stance communicated through private diplomatic channels, the British ambassador to Austria, Lord Loftus, concluded: “I have not the smallest hope that the Austrian Government will agree to any such [compromise].”\(^1\) Loftus may have drawn this inference because he believed that Austria, having made the threat, would not have wanted to be caught in a bluff, but this could be said of every threat and statesmen sometimes believe that threats lack credibility.\(^2\) Although a range of factors certainly affected Loftus’ conclusion, he likely made the following simple inference: in demanding more, Austria had given up the opportunity to achieve a compromise solution that Austria knew Piedmont was much more likely to have conceded without fighting; therefore, Austria is resolved to fight for the more substantial demand. Through this mechanism, the scope of state demands commonly conveys information about resolve to adversaries in

\(^1\)British Parliamentary Papers, 1859, v. XXXII, 213.

\(^2\)For a discussion of the reputational mechanism for private diplomacy, see Sartori (2005). Kurizaki (2007) shows that private threat-making is consistent with the threatened party assigning a non-zero probability to the threat’s credibility. For analysis of other mechanisms private diplomacy, see for instance Jervis (1970), Fearon (1995), and Trager (2010).
Despite literature in international relations that argues the contrary, such simple inferences are quite rational in diplomatic relations. In this article, I analyze a model similar to Fearon (1995) in order to demonstrate that higher demands can increase perceptions of a state’s resolve to fight for more favorable outcomes when either (a) settling on a negotiated solution is the outcome of a bargaining process in which both sides take part, rather than a take-it-or-leave-it offer from one side or (b) the issue space exhibits indivisibilities. Either circumstance allows the sender of a signal to capture some of the gains from bargaining to avoid war. This results in an essentially equivalent signaling mechanism in both cases. Incentives for less resolved types to represent themselves as resolved are not sufficient to eliminate informative costless signaling in international political contexts. Interestingly, compromise offers will be made even though they increase an adversary’s perception that the compromising state would be willing to make an even greater concession.

The reason states can draw inferences from cheap talk diplomacy in such contexts is simply that when states demand a lot when they make threats, they sometimes lower their chances of getting somewhat less without having to fight for it. Thus, when states do make large demands, they run a risk, and since states would not be willing to run this risk unless a large concession (rather than an intermediate compromise) were sufficiently important, these

---

3Note that, as in the Fearon model, the options available to the sides must be more than simply backing down or not backing down. If the target of a threat has only this binary choice, there can be no question of the scope of the demand. The result is that signaling will not be possible in contexts like the one analyzed here.
threats convey information. The mechanism by which resolve is signaled to an adversary is therefore different from the mechanism in most other models in the literature. In other models, it is often a state's willingness to risk war that conveys resolve. Here, the signaler decides whether to initiate a conflict and the least resolved signaler may be unwilling to imitate the behavior of more resolved signalers. The reason is that the less resolved signalers are unwilling to risk getting less from the negotiation even though there is no possibility of conflict.

Further, when one side believes the other is sufficiently unlikely to make a full concession, then demanding a full concession allows the other side to be certain that the demanding state will fight unless such a concession is offered. This contrasts with most cheap talk models in the literature in which demands never or rarely convey certain information about a state's willingness to fight.

The model also allows us to understand when offers of compromise will be made even though they signal a measure of weakness. This occurs only when an adversary is believed

---

4. In a general class of crisis bargaining games with one-sided incomplete information, Banks (1990) demonstrates that more resolved states will run an equal or greater risk of conflict compared to less resolved states. As Fearon (1995) illustrates, however, more resolved states will not always be able to incur such increased risks to signal resolve. Further, as Fey and Ramsay (2011) demonstrate, Banks’ result does not extend to the two-sided incomplete information context. Slantchev (2010a, 136-141) shows that, in many models in the literature, resolve is conveyed by incurring a greater risk of conflict.

5. Slantchev (2010b) and Trager (2010) show why states would sometimes feign weakness in order to catch an adversary unprepared, but not why states might allow an adversary to infer that the signaling state is weaker than it is even when the adversary cannot make substantial preparations for conflict.
highly unlikely to accede to maximalist demands. In such cases, states that are unwilling to fight if they are offered no concession and those states that would choose not to fight if they are offered just a slight concession both send the same signal while only states that are willing to fight unless they receive a large concession send that signal. Thus, the model explains why states are sometimes willing to make offers that signal a form of weakness, namely, that the state is more likely to be unwilling to fight for even a small concession.

**Inferences Based on the Scope of Threats**

Sometimes, when a state makes several demands on another state, these demands are separate one from another such that they can be thought of as referring to separate issue dimensions. In other cases, however, issues are linked in such a way that acquiescence in one demand may imply acceptance of another. If a state demands the cession of a territory on its border, it may be impractical for the threatened state to offer to cede half the territory that does not border on the threatening state but keep half the territory that does border on the threatening state. The amount of territory demanded is a question of degree. The threatened state will then want to know whether the threatening state would settle for a compromise: perhaps the half of the territory nearest to the threatening state would be a sufficient concession to avoid war.

A striking example comes from the negotiations prior to the first Gulf War. In February of 1991, the U.S. promised to begin a ground offensive unless Iraq withdrew from Kuwait.

---

<sup>6</sup>Trager (2011).
City in two days and from Kuwait in seven. The U.S. demand consisted of two essential parts: the requirement to leave Kuwait and the specific timetable for withdrawal. Iraq had already offered to leave Kuwait in 21 days and Kuwait City in 4. The importance of the second U.S. demand for a shorter timetable was that Iraq would not be able to unwind its positions and leave with its equipment in the shorter time-frame. By this point in the conflict, a key U.S. goal was to degrade the ability of the Iraqi army to threaten its neighbors. The Iraqi regime accepted the first U.S. demand, but not the accelerated timetable. The U.S.-Iraqi negotiations concerned questions of degree: Iraq could not agree to the timetable, but refuse to leave Kuwait. Iraq was convinced by this stage that the U.S. would fight to restore Kuwaiti independence, but Iraq did not know whether the U.S. would go to war only to degrade the capability of the Iraqi army.

Previous literature has argued that nothing can be learned from statements like the U.S. ultimatum in this case, unless particular conditions apply. Fearon (1995) presents a model of costless threats in which no information can be conveyed. Fearon (1994) argues that threats acquire influence when they are made publicly, before audiences that would punish leaders for abrogating a commitment. Sartori (2005), Jervis (1970), Schelling (1966) and Schelling (1980) argue that the U.S. statement could have conveyed information if it engaged the U.S. bargaining reputation such that failing to follow through on a commitment would impair U.S.

---

7 Even after the war began, when the Iraqi regime announced it was abandoning Kuwait, President Bush said U.S. forces would continue to attack Iraqi soldiers who did not lay down their arms. See Pape (1996, 216-219).
credibility in the future. Unless the U.S. were particularly concerned with its credibility vis-
a-vis the Iraqi regime, however, this too would seem to require a public threat. Kurizaki (2007) shows that private threats will not be entirely devoid of credibility when a public threat would create audience costs in the threatened state, making it harder for that state to comply with the demand. Jervis (1970), Nicolson (1963) and others discuss additional mechanisms of information transmission.8

I will show below that costless statements in such contexts can convey information even when previously described mechanisms are unavailable. Signaling is likely to be particularly effective when the signaling state is believed to be highly resolved to fight at least for a partial concession from the second state, and the second state is believed sufficiently likely to be resolved to fight rather than make a full concession to the signaling state. The claim is not, however, that in such cases states are likely to reach a compromise, which would be unsurprising. It is rather the quite different claim that in this context, attempts at communication will change the threatened states beliefs about the threatening states intentions. In fact, in

8For other mechanisms of diplomatic communication, in addition to those cited earlier, see, for example, Kydd (1997); Powell (1990); Ramsay (2004); Morrow (1989); Nye (2004); Watson (1982); Leventoglu and Tarar (2005). Constructivist scholars are interested in intersubjective understandings and their role in influencing state behavior, but have focused on norms, which tend to be seen as changing only over the long term. Wendt (1999, 339) for instance, notes that a normative structure, “by definition, is self fulfilling. Its no wonder why they last. Few constructivists have therefore focused on how perceptions of threat are constructed in particular encounters. See, however, the articles on epistemic communities in Hass (1992), as well as Adler and Barnett (1998); Katzenstein (1996); Davis Cross (2007) and Der Derian (1987).
such cases, when the threatening state demands a full concession, the threatened state will know for sure that the threat is credible even though it did not believe the threatening state would fight for sure for a full concession before the threat was made.

To understand the intuition for the signaling dynamics in the models presented below, consider what the Iraqi government could have learned from the scope of the U.S. demand. First, note that Iraq was itself very unlikely to comply with a demand that involved such a significant degrading of its military capability. Second, with the U.S. insisting on the accelerated time table, the Iraqi government would never choose to unilaterally remove its troops from Kuwait at the slower timetable because Iraqi forces leaving their chosen and prepared positions would have been the more vulnerable to U.S. attack. Thus, by insisting on the greater Iraqi concession of the accelerated time table, the U.S. ensured that no partial Iraqi concession would be forthcoming. Suppose, by contrast, that the U.S. government insisted only on the slower Iraqi withdrawal. This, the Iraqis were clearly much more likely to do. From the point of view of the U.S. government, it only made sense to insist on the accelerated time table if the accelerated time table were of sufficient importance to U.S. policy-makers. Put in this way, it is obvious that Iraqi decision-makers could conclude from U.S. statements, which were not wholly public at the time, that U.S. resolve to attrit the Iraqi army was high. In diplomatic crises, similar dynamics to these recur. Of course, governments are always confronted with a range of signals and indices; this is but one.9

9For a discussion of signals, which senders can manipulate, and indices, which senders cannot, see Jervis (1970).
A Bargaining Model with a Discrete Set of Compromise Outcomes

In order to most closely relate the discussion to previous literature on these topics, the first model I shall describe is identical to the well known model of cheap talk communication in Fearon (1995), except in these three respects: (1) states have a discrete set of compromise solutions available, (2) both sides are uncertain about the others’ resolve and (3) states are uncertain about each others utilities over compromise outcomes rather than over each others’ costs of war. These modifications in the model often better fit the facts of international politics. Negotiations often center on a few discrete options, and this is sometimes because only a discrete set of options are practical without linking discussions to other issue areas. Nevertheless, I shall consider a divisible issue space in the succeeding section. Second, no statesman could claim to have certain knowledge of how adversaries weigh proposed compromise solutions against one another and this is precisely what diplomats often strive to communicate.

As in Fearon (1995), the game described here has two players, a “Signaler” and a “Tar-

\footnote{O’Neill (2001) argues against thinking about the “issue space” negotiated over as a space in which one could define a sensible measure of distance between the possible outcomes. I take the standard approach here for simplicity and, again, to most closely relate the results to previous scholarship.}

\footnote{With additional restrictions on the type utility functions, this model is equivalent to one in which uncertainty is modeled as being over the costs of war. The choice to model uncertainty as being over player utilities for compromise outcomes was made largely for clarity of exposition.}

8
get” indexed by \( i \in I \equiv \{s, t\} \), and four stages. In the first stage, Nature draws utility functions \( u^z_i(x) \) for each player \( i \) over outcomes in the bargaining space \( X \equiv [0, 1] \) (with generic element \( x \)) according to the independent, commonly known, discrete distribution functions \( h_i(u^z_i) \). For each player, there are three possible utility functions, so \( h_i(u^z_i) \) has support \( \{u^1_i(x), u^2_i(x), u^3_i(x)\} \). The players have directly opposed preferences over the set of compromise outcomes, so \( u^z_s(x) \) is strictly increasing in \( x \) for all \( z_s \), while \( u^z_t(x) \) is strictly decreasing in \( x \) for all \( z_t \). Player utility functions are private information of each player.

In the second stage, the Signaler has the opportunity to send a message \( m_y \) to the Target from a large but finite set of messages indexed by \( y \). Then, the Target chooses one of three settlement outcomes \( \{x_1, x_2, x_3\} \), where \( 0 < x_1 < x_2 < x_3 < 1 \).\(^{12}\) In the final stage, the Signaler chooses \( r \in R \equiv \{0, 1\} \), where \( r = 1 \) indicates a decision to initiate a war and \( r = 0 \) indicates peace, and then the game ends. If war occurs, the Signaler wins with probability \( p \) and the Target wins with probability \( 1 - p \), and the victorious player attains its most preferred outcome in \( X \), 1 for the Signaler and 0 for the Target. Players have commonly known costs of fighting \( c_i \). \( \mu(u^z_s | m_y) \) represents the Target’s updated beliefs about the Signaler’s type \( u^z_s \) given signal \( m_y \) in a particular perfect Bayesian equilibrium.

In order to most closely relate this model to other models in the literature, I shall assume that uncertainty about preferences relates only to player preferences over the compromise

\(^{12}\)The assumption that the Target cannot choose the extremes of the bargaining space, which would guarantee war for sure by the assumptions made below, simplifies the cases to consider in the analysis. This assumption does not have substantively important implications, however, and in particular, allowing the Target to choose an extreme outcome does not eliminate costless signaling equilibria in the model.
outcomes rather than to the extreme outcomes. Thus, we set \( u^z_s(0) = u^{zt}_t(1) = 0 \) and \( u^z_s(1) = u^{zt}_t(0) = 1 \) for all \( z_s \) and \( z_t \). This implies that expected utilities for war are \( w_s = p - c_s \) for the Signaler and \( w_t = 1 - p - c_t \) for the Target.

I also make several assumptions about player preferences. First, players prefer war to their least preferred outcome in \( X \) (formally, \( w_i > 0 \ \forall i \)) and prefer at least one of the three possible negotiated outcomes in the interior of \( X \) to war (formally, \( u^z_s(x_3) > w_s \ \forall z_s \) and \( u^{zt}_t(x_1) > w_t \ \forall z_t \)). Second, each player is uncertain whether the other will or will not fight if offered each of the compromise positions. Formally, for the Signaler, this implies \( u^1_s(x_1) > w_s, u^2_s(x_2) > w_s > u^2_s(x_1) \) and \( w_s > u^3_s(x_2) \). For the Target, this implies \( u^1_t(x_3) > w_t, u^2_t(x_2) > w_t > u^2_t(x_1) \) and \( w_t > u^3_t(x_2) \). Figure 1 is an example of Signaler-type utility functions that satisfy these assumptions. Note that this sort of uncertainty implies that there may be no negotiated solution that both sides prefer to war.

The first proposition gives sufficient conditions for an equilibrium to exist in which the Target’s beliefs are affected by the Signaler’s cheap talk message. So long as both sides are sufficiently unlikely to be the least resolved type (the type unwilling to fight even if it is offered its least preferred of the three settlement outcomes), an equilibrium exists in which the least resolved and middle resolved Signaler’s claim they will fight unless the Target offers at least \( x_2 \) and only the most resolved Signaler’s claim to be willing to fight unless they are offered the maximal concession \( x_3 \). Thus, when the Signaler says it will fight unless it is
offered a maximal concession, the Target knows for sure that this is true. When the Signaler admits that it will settle for $x_2$, the Target revises upwards its belief that, in fact, the Signaler would not go to war if it were offered only $x_1$.

**Proposition 1:** If $h_i(u^1_i)$ is sufficiently low for all $i$, a perfect Bayesian equilibrium exists in which the signals $m_2$ and $m_3$ are sent with positive probability and,

1. $\mu(u^3_s \mid m_3) = 1 \& \mu(u^3_s \mid m_y) = 0 \forall y \neq 3$
2. $\mu(u^2_s \mid m_2) = \frac{h_s(u^2_s)}{h_s(u^1_s) + h_s(u^2_s)} < 1 \& \mu(u^1_s \mid m_2) = \frac{h_s(u^1_s)}{h_s(u^1_s) + h_s(u^2_s)} > h_s(u^1_s)$
3. $\mu(u^1_s \mid m_y) = 1 \forall y \neq m_2, m_3$

When the Signaler is less convinced that the Target is not the least resolved type, an equilibrium with informative signals but different properties from those described in Proposition 1 can also exist.\(^\text{13}\) Here, the least resolved and most resolved Signaler’s send the same signal, while Signaler’s that would fight unless they are offered at least $x_2$ send a unique signal.

\(^{13}\)Note that the equilibrium described in Proposition 1 will not be eliminated by standard perfect Bayesian equilibrium refinements. The intuitive criterion and divinity refinements do not apply to cheap talk games and the equilibrium is neologism-proof. To see that the Proposition 1 equilibrium is neologism-proof, note that according to the players’ strategies described in the proof in the appendix, (1) $u^3_s$ sends a unique signal and prefers to have its type revealed and thus would not want to deviate to any neologism, (2) no type other than $u^1_s$ would prefer to deviate to the same neologism as $u^1_s$ and thus only neologisms in which $u^1_s$ sends a unique signal could be self-signaling but $u^1_s$ has no incentive to deviate to such a signal, and (3) in response to type $u^2_s$’s equilibrium signal, the Target behaves as it would if it were sure that the Sender’s type were $u^2_s$, which means this type of Signaler has no incentive to deviate to a neologism either alone or with $u^1_s$. For a clear discussion of this equilibrium refinement, see Farrell (1993).
signal. Thus, in this case, admitting a willingness to compromise on $x_2$ does not increase the Target’s belief that the Signaler would accept $x_1$ over war as well. Rather, the Target knows for sure following such a signal that an offer of $x_2$ will be accepted and avoid war while an offer of only $x_1$ will not. Proposition 2 gives sufficient conditions for an equilibrium of this type.

**Proposition 2:** If $h_s(u_s^1)$ is sufficiently low and $h_t(u_t^1) \in \left( \frac{h_s(u_s^2)[u_s^1(x_2) - u_s^1(x_1)]}{u_s^1(x_3) - u_s^1(x_2)}, \frac{h_s(u_s^2)[u_s^1(x_2) - w_s^1]}{u_s^2(x_3) - u_s^2(x_2)} \right)$, a perfect Bayesian equilibrium exists in which the signals $m_2$ and $m_3$ are sent with positive probability and,

1. $\mu(u_s^3 | m_3) = \frac{h_s(u_s^3)}{h_s(u_s^2) + h_s(u_s^3)} \& \mu(u_s^3 | m_y) = 0 \forall y \neq 3$

2. $\mu(u_s^2 | m_2) = 1 \& \mu(u_s^1 | m_2) = 0$

3. $\mu(u_s^1 | m_y) = 1 \forall y \neq m_2, m_3$

Thus, when the Signaler believes the Target is somewhat more likely to accept the Signaler’s most preferred negotiated solution, the signaling dynamics may change to those described in Proposition 2. This will not necessarily be the case, however, because the parameters of the model may be such that no range of the sort specified in the proposition for $h_t(u_t^1)$ exists. In fact, it is reasonable to doubt whether equilibria of this sort closely track many situations in international politics. The reason is that in these equilibria, the least resolved Signaler’s must prefer to gamble that they will get their most preferred outcome (by sending the same signal as the most resolved types) rather than achieve $x_2$ for sure, while Signalers that are somewhat more resolved would prefer $x_2$ for sure rather than take a
similar gamble. Why should states that are willing to fight for a better deal be less willing to take such risks? There is no reason to expect this to be the case in general. Nevertheless, Corollary 1 suggests a way to understand when we might expect to observe equilibria of this sort: when \( u_1^s \) and \( u_2^s \) are equivalent at \( x_2 \) and \( x_3 \) and the Signaler is not too likely to be the least resolved type, then an equilibrium of the type described in Proposition 2 exists when the likelihood that the Target is the least resolved type is in a middle range.

**Corollary 1:** If \( u_1^s(x_2) = u_2^s(x_2) \) and \( u_1^s(x_3) = u_2^s(x_3) \) and \( h_s(u_1^s) \) is sufficiently low, then a range of values of \( h_t(u_1^t) \) exists such that an equilibrium of the type described in Proposition 2 exists.

Proposition 3 demonstrates that whenever the Target is believed sufficiently unlikely to make the maximalist concession \( x_3 \), any separating or semi-separating equilibrium will have the properties described in Proposition 1. This means that the model gives us a strong empirical expectation: when there is reason to believe that the states involved are each unlikely to accept their least preferred among the three settlement outcomes, signaling will have the properties described in Proposition 1.

**Proposition 3:** For \( h_t(u_1^t) \) sufficiently low, any pure strategy equilibrium in which Signalers do not pool on a single message has the properties described in Proposition 1.

In cheap talk models of this sort, so long as the probability that the Signaler is the least resolved type is sufficiently low, the possibility of communication never increases the probability of war and sometimes, communication makes war less likely. This result is proved as Proposition 4. Note that this is in contrast to models of signaling based on reputation.
and other models in the literature.\textsuperscript{14}

**Proposition 4:** For $h_s(u_s)$ sufficiently low, no pure strategy PBE of the game with the cheap talk stage has a higher probability of war than the same game without communication.

### A Bargaining Model with a Divisible Issue Space

The effect of the partial issue indivisibility in the previous model is to ensure that the Target does not capture all the gains from agreeing on a negotiated solution that both prefer to war. Many bargaining contexts will result in outcomes in which the gains from agreement are shared between the parties, however. As long as the gains from agreement are shared among the players and uncertainty is the sort characterized above, the results related to costless signaling described above will hold. The dynamics will be essentially the same as those described in the partially divisible good case.

To see this, consider a model in which the issue space can be infinitely divided and which is identical to the one described above, except in these respects. Following the Signaler’s message, the players simultaneously announce a compromise position $a_i \in X$. If the players announce different positions, the status quo position $x_1$ remains, where $x_1$ satisfies the conditions on $x_1$ from the previous model. If the players announce the same position, this agreement becomes the new status quo. Following the announcements, paralleling the previous model, the Signaler decides whether or not to go to war.

The simple protocol of announcing a point of agreement is a sensible representation of

\textsuperscript{14}Sartori (2005), Slantchev (2010a, 136-141).
bargaining if both sides must cooperate to implement or work out a compromise agreement. The result, as the formal analysis in the appendix demonstrates, is that equilibria exist in which both sides share in a potential bargaining surplus. And the result of that is that costless signaling is possible. Many other bargaining protocols besides simply agreeing on an announcement would yield similar results. For instance, Rubinstein models and Nash Bargaining also produce outcomes where both sides share in the bargaining surplus.\footnote{For an overview of such models, see Ausubel, Cramton and Deneckere (2002).}

Thus, in this modified model, although the issue space is infinitely divisible, costless signaling is possible. As in the case of discrete options, the players will have ideas about what particular bargaining outcomes are likely under different courses of action. Whether these outcomes are exogenously given or arrived at through bargaining does not greatly affect the signaling dynamics as long as both sides are expected to share in the bargaining surplus. As before, as long as the Target is not overly likely to be the least resolved type, informative signaling of the type described in Proposition 1 is possible. This result is proved as Proposition 5.

**Proposition 5**: Proposition 1 holds in the modified game.

**Discussion**

In these models, communication is possible between adversaries because some Signaler types admit that they’re not the most resolved of the possible Signaler types. These types are willing to reveal this information, in spite of the fact that they also have incentive to mis-
represent themselves as more resolved than they are in order to achieve a more favorable bargain, because demanding more sometimes also entails a risk they are unwilling to run. The risk is that an increased demand will be resisted where a more moderate demand would not be, with the result that the parties will not agree on a negotiated solution and a Signaler who had misrepresented its preferences may even be forced to fight rather than accept the unilateral action by the Target or the no agreement outcome. Thus, this form of signaling can only occur in contexts in which the scope of demands is an issue, that is, when more than two non-conflict outcomes are possible. If the only choice of the Signaler is to threaten or not threaten and the only choice of the Target is to back down or not, in the signaling context analyzed here, no communication of resolve will be possible because demanding more carries no risk of getting even less at the bargaining table and therefore the incentives to misrepresent imply that no semi-separating equilibria exist.

Unlike in many other models of diplomatic signaling, therefore, it is not through risking conflict that information is conveyed. In the Proposition 1 equilibrium, when the most resolved Signaler-types send a signal that other types are unwilling to send, the probability that war occurs decreases. This is because only the Signaler has the option to initiate conflict and when resolved Signalers send such a signal, they are more likely to get a better offer and therefore less likely to initiate a conflict. Further, for most parameterizations of the model that seem reasonable, the existence of signaling mechanisms of the type described here either does not increase the likelihood of conflict or actually causes the probability of

---

16See Slantchev (2010a, 136-141) for a discussion of other models in the literature on this point.
war to decline.

An example of the effect of communication on the probability of conflict is shown in Figure 2. On the left of the figure, communication strictly decreases the probability of war. As it becomes more likely that the Signaler is the most resolved type, the benefit of communication increases up to a point and then the probability of conflict in the equilibrium without communication falls discontinuously to be equal to the likelihood of conflict in the cheap talk equilibrium. The reason for this is that, as the Target becomes more convinced that the Signaler is the most resolved type, the Target becomes less willing to risk conflict even without a credible signal of resolve. On the right of the figure, as the Target becomes still more convinced that the Signaler is the most resolved type, the Target becomes less convinced that the Signaler’s resolve is in a middle range. Therefore, the Target becomes less likely to be willing to offer a middle range compromise in response to a less than maximalist demand by the Signaler. The result is that the Signaler becomes unwilling to make less than maximalist demands, the communication equilibrium breaks down and the probability of conflict increases dramatically.\textsuperscript{17}

Figure 2 about here.

\textsuperscript{17}Parameter values for this simulation are $p = .5$, $c_i = .15 \forall i$, $x_1 = .25$, $x_2 = .5$, $x_3 = .75$, $u_1^s(x) = x^\frac{5}{3}$, $u_1^t(x) = (1 - x)^\frac{2}{3}$, $u_2^s(x) = x^{\frac{3}{2}}$, $u_2^t(x) = (1 - x)^{\frac{1}{2}}$, $u_3^s(x) = x^2$, $u_3^t(x) = (1 - x)^2$, $h_s(u_1^s) = .3$, $h_s(u_2^s) = 1 - h_s(u_2^s) - h_s(u_3^s)$, $h_t(u_1^s) = .2$, $h_t(u_1^t) = .4$, $h_t(u_3^t) = .4$. 

17
the fact that in the Fearon model, the Target captures all of the benefit from avoiding war. The Target’s disproportionate share of the benefits of peace results from (1) the take-it-or-leave-it aspect of the model and (2) the infinitely divisible issue space. This is because, in that model, when the Target of a threat knows the Signaler’s type, its take-it-or-leave-it equilibrium offer leaves the Signaler indifferent between choosing war or peace. As a result, the Signaler has no incentive to reveal information: although peace may result, the nature of the peace is such that the Signaler finds war an equally compelling alternative. If either of these two assumptions are relaxed, the Target need not capture all of the gains from a negotiated solution, the Signaler can therefore have an incentive to reveal its type, and the scope of a costless demand can convey information to the Target.

While the take-it-or-leave-it model with an infinitely divisible issue space is interesting to study, it is likely that in most bargaining contexts, the players expect the benefits of peace to be spread more evenly between the players. When states are highly resolved and willing to go to war over a particular set of issues, they still often have a strong preference for getting their way through the threat rather than the costly use of force. For this reason, resolved states are thought to have an incentive to reveal their types. In the Fearon model, by contrast, not only unresolved, but also resolved states have no incentive to reveal their types. If they do, they end up with their war payoff and if they are not able to, they still get their war payoff. Thus, what prevents costless signals from conveying information is not

---

18The discussion here does not directly bear on, and does not constitute a critique of, the central theses of Fearon (1995) related to the nature of rationalist explanations for war.
merely that unresolved types have an incentive to misrepresent themselves as resolved (that is true in the model analyzed above), but also that both resolved and unresolved types have no incentive to make their signals credible.

In many real world crises, a model in which even successful signalers receive no benefit from success does not appear to correspond actors’ understandings of the situation. In the Cuban Missile Crisis, for instance, although the Kennedy administration was willing to initiate conflict with the Soviet Union through an air strike on Cuba, members of the administration were very glad to have forced the removal of the missiles without having to do so. That U.S. policy-makers were so relieved at the resolution of the crisis indicates that the Soviets were unable to capture all of the gains from peace in this case, as Targets are in the majority of instances of successful coercion.

One reason the gains from cooperation are shared in real world cases is that no continuously divisible issue space exists in many instances or, for complex reasons outside of the questions considered here, in practice, actors do not consider the issue space divisible. Fearon (1995, 389-90) and Powell (2006) note that mechanisms, such as randomization devices or alternating possession of the good, exist to make the issue space divisible, and that therefore indivisibility is generally not by itself a non-rational explanation for war. They also emphasize, however, that for complex reasons, states often act as if issues are indivisible. This may result from the construction of the conflict,19 from inherent features of the issues involved,20

19 Goddard (2006)

20 Hassner (2003) Even territory is often less easily divisible than it might appear. Often, given military
from the logic of commitment problems,\textsuperscript{21} from the positions of other powers that limit the options of the two states,\textsuperscript{22} or because the set of potential issues of contention in particular cases is not large enough.\textsuperscript{23} Yet another reason issues may be indivisible in practice has to do with bargaining reputation. As Schelling has pointed out, the U.S. would have a hard time offering to give up California and then establishing a credible commitment to give up no additional territory. It is the unavailability of other salient lines that leads to a sort of indivisibility of the initial grouping of territory.\textsuperscript{24} Still another reason for indivisibility relates to what moral codes allow leaders to demand and offer each other. Hitler could demand the Sudetenland, and even that the West stand aside while he occupy Czechoslovakia in the technologies available at the time of a crisis, some groupings of territory are considered much more defensible, and thus much more valuable than others. A striking example is Hitler’s calculation in 1941 that since war with the U.S. was likely, control of the entire European landmass was essential to preserve German territorial gains. This made Hitler less willing to compromise on lesser territorial gains and resulted in the decision to attack Russia. On this points, see, for instance, Kershaw (2007, 54-90). Fearon (1995, 389-90) discusses the rise of nationalism in making territory less easily divisible.

\textsuperscript{21}Powell (2006).

\textsuperscript{22}Fearon (1995, 389-90) mentions the practice of compensating a state for territorial acquisition of a rival state with territory somewhere else as a means of making the issue space convex. In many cases, however, this proved impossible because of the attitude of 3rd powers. France’s call for compensation in Belgium for Russian gains in the Ottoman territories prior to the Crimean War, for instance, was objectionable to Britain. See Puryear (1931, 272).

\textsuperscript{23}The implications of “issue linkages” are discussed in Fearon (1995, 389-90) and AUTHOR.

\textsuperscript{24}Schelling (1966, Chapter 2).
name of protecting Germans in the Sudetenland, but he could not demand “three quarters of Czechoslovakia”. At the Munich conference and in diplomacy that preceded it, neither side considered such demands and offers.

Besides issue indivisibility, another reason Targets often are not expected to garner all gains from cooperation is that bargaining outcomes that are preferable to both sides often require actions taken by both sides. Each side may be willing to withdraw its troops from a border, but only if the other side also does so. Saddam Hussein may have been willing to withdraw his forces from Kuwait on a particular timetable, but only in return for a public guarantee of their safety from the United States. If neither side can achieve the preferred compromise outcomes unilaterally, we should expect that the sides negotiate and - if they reach agreement - that they share the bargaining surplus between them.

A second driver of the signaling dynamics of the model is the assumption that both sides are uncertain whether the other would or would not be willing to fight rather than accept some compromise solution. As a result, when the Target believes the Signaler is a type that would be unwilling to make a compromise that the Target would accept, the Target no longer has any interest in making concessions to the Signaler. This gives less resolved Signalers a disincentive to misrepresent themselves as more resolved than they are.

The assumption that both sides are uncertain whether the other would accept a compromise appears uncontroversial in that it faithfully represents the subjective states of international actors in crises. However, the implication of the assumption is that it is possible that no compromise exists that both sides prefer to conflict. Put in this way, the assumption
appears more controversial, but we can think of this assumption as merely a simplification of a more complex strategic process that is not modeled explicitly. To see this, note that virtually all crisis bargaining models share the following property: for some combinations of player types, war occurs with certainty. Whether this results from an explicitly modeled strategic context or directly from the preferences of actors is not important for the present inquiry. Thus, the assumption that player preferences can produce war could be justified by embedding many prominent crisis bargaining models from the literature in the game presented here. This would substantially complicated the analysis, of course, as well as obscure the dynamics of the current framework, which is why the simpler strategy was adopted above.

It is thus in a restricted set of cases that the logic of Fearon’s take-it-or-leave-it model might be expected to operate: when the Target of a threat can unilaterally choose from a set of options that approximate a divisible issue space and when the signaler has few options short of conflict or acquiescence. In such cases, costless signals will not convey information. One might suspect, however, that such cases are relatively rare in international politics either because, in practice, inherent indivisibilities exist or because compromise outcomes preferable to conflict require the give and take agreement of both sides. Of course, signals may nevertheless convey little to adversaries for other reasons. A claim to be willing to fight for a large concession, for instance, will likely signal little to the adversary when all sides believe that the state making the claim believes that the adversary is likely to be willing

---

make the large concession. Such threats may be effective even when they are not credible in themselves.\footnote{See Fearon (2002) for a discussion of this point.}

The analysis here also demonstrates why states would make compromises even though doing so will often result in the perception that they would accept outcomes involving even greater concessions - thereby encouraging an adversary to demand these additional concessions. In the informative equilibria characterized in Propositions 1 and 5, when the Target sees the Signaler send the signal that only issue one is important, the Target knows the Signaler is less likely to go war if the Target makes no concession at all to the Signaler. Nevertheless, the Signaler sometimes still prefers to send the signal that it would settle for the compromise position. The reason is that by demanding too much, that which the Target is relatively unlikely to give up, the Signaler risks getting nothing at all. By demanding only the compromise, the Signaler increases its chances of getting something. Proposition 3 demonstrates that this dynamic will be associated with any offer of compromise when the players are each believed, prior to the diplomatic signal, to be sufficiently unlikely to be willing to settle for their least preferred compromise outcome.

**Conclusion**

The model demonstrates that leaders can learn a great deal about the intentions of other leaders from the scope of the demands other leaders make. Scope of demand conveys information because of the risks involved in larger demands. The demanding state may end
up worse off than if it had asked only for a compromise outcome that is more likely to be forthcoming. This conclusion is particularly certain when compromise outcomes require the give and take agreement of both sides or the issue space consists of discrete alternatives and when adversaries each believe the other relatively unlikely to accept the negotiated outcome that each most prefers. In such cases, compromises will be offered even though these compromises increase the perception that the compromiser would be willing to settle for even less than the compromise offered. This appears to be a commonly used and intuitive signaling mechanism in international politics.

Appendix

Proof of Proposition 1: I will show that for \( h_i(u^1_i) \) sufficiently low for all \( i \), the following strategies and beliefs constitute a perfect Bayesian equilibrium. The Signaler’s strategy is:

for \( u^1_i \), send \( m_2 \) and choose \( r = 0 \); for \( u^2_3 \), send \( m_2 \) and choose \( r = 1 \) following \( x_1 \) and \( r = 0 \) otherwise; for \( u^3_3 \), send \( m_3 \) and choose \( r = 1 \) following \( x_1 \) and \( x_2 \) and \( r = 0 \) otherwise.

The Target’s strategy is: for \( u^1_1 \), choose \( x_1 \) following \( m_y \) \( \forall y \neq 2, 3 \), \( x_2 \) following \( m_2 \), and \( x_3 \) following \( m_3 \); for \( u^3_2 \), choose \( x_1 \) following \( m_y \) \( \forall y \neq 2, 3 \), \( x_2 \) following \( m_2 \), and \( x_1 \) following \( m_3 \); for \( u^3_3 \), choose \( x_1 \) following \( m_y \) \( \forall y \). The Target’s updated beliefs are:

\[
\mu(u^1_s \mid m_y) = 1, \\
\mu(u^2_s \mid m_y) = 0 \; \forall y \neq 2, 3; \\
\mu(u^1_s \mid m_2) = \frac{h_s(u^1_s)}{h_s(u^1_s)+h_s(u^2_s)}; \\
\mu(u^2_s \mid m_2) = \frac{h_s(u^2_s)}{h_s(u^1_s)+h_s(u^2_s)}, \\
\mu(u^3_s \mid m_2) = 0; \\
\mu(u^1_s \mid m_3) = \mu(u^2_s \mid m_3) = 0, \text{ and } \mu(u^3_s \mid m_3) = 1.
\]

We need not specify the Signaler’s updated beliefs except to say that they must accord with Bayes’ rule.

Note that beliefs following signals \( m_2 \) and \( m_3 \) follow directly from Bayes’ rule and the
Signaler’s strategy, and that the Target’s beliefs at other information sets are unconstrained in a PBE. Also note that these beliefs imply the properties described in the proposition. Optimality of the Signaler’s strategy at the nodes following the Target’s action follows from a direct comparison of the payoffs.

To see that the Signaler’s action at the signaling stage is optimal, first note that no Signaler-type can do better by deviating to a signal other than $m_2$ or $m_3$ because this guarantees the action $x_1$ on the part of the Target. Second, note that $u^3_s$ Signaler-types certainly cannot do better by deviating. Given the Target’s strategy and beliefs and the other components of the Signaler’s strategy, this leaves two conditions that must be satisfied for the Signaler’s strategy to be optimal:

$$Eu_s(m_2 | u^1_s) = [h_t(u^1_t) + h_t(u^2_t)]u^1_s(x_2) + [1 - h_t(u^1_t) - h_t(u^2_t)]u^1_s(x_1) \geq$$

$$Eu_s(m_3 | u^1_s) = h_t(u^1_t)u^1_s(x_3) + [1 - h_t(u^1_t)]u^1_s(x_1)$$

$$Eu_s(m_2 | u^2_s) = [h_t(u^1_t) + h_t(u^2_t)]u^2_s(x_2) + [1 - h_t(u^1_t) - h_t(u^2_t)]w_s \geq$$

$$Eu_s(m_3 | u^2_s) = h_t(u^1_t)u^2_s(x_3) + [1 - h_t(u^1_t)]w_s$$

Both conditions hold for sufficiently low $h_t(u^1_t)$.

To see that the Target’s strategy is optimal, note that $x_1$ is clearly an optimal choice for $u^3_t$ following any signal and that $x_1$ is optimal for any Target type, given the Target’s beliefs, following any signal other than $m_2$ or $m_3$. For Target type $u^1_t$, following $m_2$, $x_2$ gives utility $u^1_t(x_2)$ with certainty, so this Target type never chooses $x_3$ following $m_2$. Thus, the Target’s
strategy is optimal in the information set following $m_2$ if the following condition holds:

$$Eu_t(x_2 \mid u_t^1, m_2) = u_t^1(x_2) \geq Eu_t(x_1 \mid u_t^1, m_2) = \frac{h_s(u_s^1)}{h_s(u_s^1) + h_s(u_s^2)} u_t^1(x_1) + \frac{h_s(u_s^2)}{h_s(u_s^1) + h_s(u_s^2)} w_t$$

(3)

Target type $u_t^1$’s strategy is optimal following $m_3$ because, given the Target’s beliefs, any other choice results in war, which is a worse outcome for this type. For Target type $u_t^2$, its strategy following $m_3$ is optimal because a different choice produces either $u_t^2(x_3) < w_t$ or $w_t$, which is the same as the utility that results from its equilibrium strategy. For this Target type, following $m_2$, $x_3$ is clearly not preferred to $x_2$ since the latter is certain to be accepted by the Signaler. Thus, the strategy ascribed to the Target is optimal for $u_t^2$ if the following condition holds:

$$Eu_t(x_2 \mid u_t^2, m_2) = u_t^2(x_2) \geq Eu_t(x_1 \mid u_t^2, m_2) = \frac{h_s(u_s^1)}{h_s(u_s^1) + h_s(u_s^2)} u_t^2(x_1) + \frac{h_s(u_s^2)}{h_s(u_s^1) + h_s(u_s^2)} w_t$$

(4)

Conditions (3) and (4) are both satisfied for sufficiently low $h_s(u_s^1)$. ■

The proof of Proposition 2 follows the form of the proof of Proposition 1 closely and is therefore omitted.

**Lemma 1:** No fully separating equilibria exist in the model with discrete compromise options and only two semi-separating equilibria are possible: where $u_s^1$ pools with $u_s^2$ but not $u_s^3$, and where $u_s^1$ pools with $u_s^3$ but not $u_s^2$.

**Proof:** Note that in any PBE, the Signaler type $u_s^1$ never sends a signal that neither of the other two types send. In such an equilibrium, whether the other two types pool or send unique signals, it will always be optimal, for any set of Target beliefs consistent with Bayes’
rule, for Target-type \( u_t^1 \) to offer more than \( x_1 \) in response to the signal or signals sent by Signaler types \( u_s^2 \) and \( u_s^3 \) and no Target-types can offer less than \( x_1 \). Since in equilibrium all Target types must respond with \( x_1 \) following a unique signal sent by \( u_s^1 \), Signaler-type \( u_s^1 \) would prefer to deviate to the message assigned to one of the other Signaler types, which means that no such equilibrium exists. Thus, no fully separating equilibria exist and only the two semi-separating equilibria described in the Lemma are possible.

**Proof of Proposition 3:** The proof of Proposition 1 demonstrates that equilibria of the first type described in Lemma 1 have the properties described in the proposition. Thus, by Lemma 1, we need only show that, for sufficiently low \( h_t(u_t^1) \), an equilibrium does not exist in which \( u_s^1 \) pools with \( u_s^3 \).

In any such equilibrium, since Bayes’ rule implies that the Target believes it is facing a \( u_s^2 \) type Signaler following that type’s unique signal \( m_2 \), \( u_t^1 \) and \( u_t^2 \) type Targets must choose \( x_2 \) following \( m_2 \) and \( u_t^3 \) must choose \( x_1 \) (since other offers would be accepted and result in lower utility for this type than conflict). This implies that

\[
Eu_s(m_2 \mid u_s^1) = (h_t(u_t^1) + h_t(u_t^2))u_t^1(x_2) + (1 - h_t(u_t^1) + h_t(u_t^2))u_t^1(x_1).
\]

Following the common signal sent by \( u_s^1 \) and \( u_s^3 \), \( u_t^3 \) types must choose \( x_1 \) because there is a positive probability of either of the other offers being accepted, which leads to lower expected utility than conflict for this Target type. Note that no Target type offers \( x_2 \) following \( m_3 \) because an offer of \( x_1 \) yields the same probability of war and is preferable if accepted. Thus, following \( m_3 \), \( u_t^2 \) must choose \( x_1 \) because a choice of \( x_3 \) will be accepted in any PBE, leading to a strictly lower payoff. This implies that

\[
Eu_s(m_3 \mid u_s^1) \leq h_t(u_t^1)u_s^1(x_3) + (1 - h_t(u_t^1))u_s^1(x_1),
\]

27
which implies that the equilibrium condition, \( Eu_s(m_3 \mid u^1_s) \geq Eu_s(m_2 \mid u^1_s) \), cannot hold for sufficiently low \( h_t(u^1_t) \).

**Proof of Proposition 4:** By Lemma 1, the only equilibria in which cheap talk affects the outcome are (1) those in which the \( u^1_s \) type sends the same signal as the \( u^2_s \) type, while the \( u^3_s \) type sends a different signal and (2) those in which the \( u^1_s \) type sends the same signal as the \( u^3_s \) type, while the \( u^2_s \) type sends a different signal.

Consider the first form of signaling equilibrium. In any such equilibrium, following the signal sent by the \( u^1_s \) and \( u^2_s \) types (\( m_2 \) without loss of generality), \( u^2_t \) Target type must choose \( x_2 \) because if such Target’s choose \( x_1 \), then \( u^2_s \) strictly prefer to deviate from their equilibrium strategy to the signal sent by \( u^3_s \) types (\( m_3 \)). Further, for \( h_s(u^1_s) \) sufficiently low, the Target prefers \( x_2 \) to \( x_1 \) following \( m_2 \). In any PBE, \( u^3_t \) must choose \( x_1 \) following \( m_2 \). Following \( m_3, u^2_t \) and \( u^3_t \) certainly cannot offer \( x_3 \) given beliefs consistent with Bayes’ rule. These considerations imply that the probability of war in such cases is \( h_s(u^3_s)(h_t(u^2_t) + h_t(u^3_t)) + h_s(u^2_s)h_t(u^3_t) \).

In a PBE of a world without communication, clearly no Target type can make an offer that if likes less than going to war because there is a positive probability that the offer will be accepted. Thus, the equilibrium probability of war is at least \( h_s(u^3_s)(h_t(u^2_t) + h_t(u^3_t)) + h_s(u^2_s)h_t(u^3_t) \), which is the probability of war in the PBE with communication for sufficiently low \( h_s(u^1_s) \).

Now consider the second form of signaling equilibrium. In any equilibrium of this type, the Target type \( u^3_t \) must choose \( x_1 \) following the signal sent by \( u^2_s \), and then \( u^2_s \) must
choose war in a PBE. Following the signal sent by \( u_t^1 \) and \( u_s^3 \), Target types \( u_t^2 \) and \( u_t^3 \) must choose \( x_1 \) (no other feasible \( x \) that is preferred by these Target types to war has a higher chance of being accepted by the Signaler). Since Signaler types \( u_s^1 \) never elect to fight, these considerations imply that the probability of war in this equilibrium is equal to \( h_t(u_s^2)h_t(u_t^3) + h_s(u_s^3)(h_t(u_t^2) + h_t(u_s^2)) \), which equals the minimum probability of war in the game without cheap talk communication.

**Proof of Proposition 5:** Take the Signaler’s strategy to be: for \( u_s^1 \), send \( m_2 \), announce \( a_s = \frac{u_s^2 - 1(w_s) + u_s^2 - 1(w_t)}{2} \equiv \chi_2 \) following \( m_2 \), \( a_s = \frac{u_s^3 - 1(w_s) + u_s^3 - 1(w_t)}{2} \equiv \chi_3 \) following \( m_3 \), \( a_s = u_s^1 - 1(w_s) \equiv \chi_1 \) following any signal other than \( m = 2, 3 \), where the -1 superscripts represent the inverse of the functions, and choose \( r = 0 \); for \( u_s^2 \), send \( m_2 \), announce \( a_s = \chi_2 \) following \( m_2 \), \( a_s = \chi_3 \) following \( m_3 \), \( a_s = \chi_1 \) following any signal other than \( m = 2, 3 \), and choose \( r = 1 \) iff either \( a_t \neq a_s \) or \( a_t < u_s^2 - 1(w_s) \); for \( u_s^3 \), send \( m_3 \), announce \( \chi_3 \) following \( m_3 \), \( a_s = \chi_2 \) following \( m_2 \), \( a_s = \chi_1 \) following any signal other than \( m = 2, 3 \), and choose \( r = 1 \) iff either \( a_t \neq a_s \) or \( a_t < u_s^3 - 1(w_s) \).

Take the Target’s strategy to be: for \( u_t^1 \), announce \( a_t = \chi_2 \) following \( m_2 \), \( a_t = \chi_3 \) following \( m_3 \), \( a_t = \chi_1 \) following any signal other than \( m = 2, 3 \); for \( u_t^2 \), announce \( a_t = \chi_2 \) following \( m_2 \), \( a_t = \chi_1 \) following any signal other than \( m = 2 \); for \( u_t^3 \), announce \( a_t = \chi_1 \) following any signal.

The Target’s updated beliefs following the signal are: for all \( y \neq 2, 3 \), \( \mu(u_s^1 | m_y) = 1 \), \( \mu(u_s^2 | m_y) = \mu(u_s^3 | m_y) = 0 \); \( \mu(u_s^1 | m_2) = \frac{h_s(u_s^1)}{h_s(u_s^1) + h_s(u_s^3)} \), \( \mu(u_s^2 | m_2) = \frac{h_s(u_s^2)}{h_s(u_s^1) + h_s(u_s^3)} \), \( \mu(u_s^3 | m_2) = 0 \); \( \mu(u_s^1 | m_3) = \mu(u_s^2 | m_3) = 0 \), and \( \mu(u_s^3 | m_3) = 1 \). Note that these beliefs are
consistent with Bayes’ rule and the Signaler’s strategy. We need not specify the Signaler’s updated beliefs at its final move except to say that these beliefs must also accord with Bayes’ rule.

The sequential rationality of the Signaler’s war choice follows directly from the Signaler’s payoffs. Consider the optimality of the Signaler’s strategy at information sets at the announcement stage. For type $u^1_s$, at the information set following $m_2$, the Signaler’s strategy results in an expected utility of $(h_t(u^1_t) + h_t(u^2_t))u^1_s(\chi_2) + (1 - h_t(u^1_t) - h_t(u^2_t))u^1_s(\chi_1)$ given the players’ strategies and beliefs, whereas any other announcement results in $u^1_s(\chi_1)$ for sure. Similarly, following $m \neq 2, 3$, matching the Target’s announcement of $\chi_1$ yields the same outcome and payoff as any other announcement. Following $m = 3$, the Signaler’s equilibrium strategy yields an expected payoff of $h_t(u^1_t)u^1_s(\chi_3) + (1 - h_t(u^1_t))u^1_s(\chi_1)$ whereas any other announcement yields $u^1_s(\chi_1)$ for sure.

For type $u^2_s$, at the information set following $m_2$, the Signaler’s strategy results in an expected utility of $(h_t(u^1_t) + h_t(u^1_t))u^2_s(\chi_2) + (1 - h_t(u^1_t) - h_t(u^1_t))w_s$ given the players’ strategies and beliefs, whereas any other announcement results in $\chi_1$ for sure as the status quo and thus the payoff of $w_s$ in the game because the Signaler’s strategy implies that it chooses war in the next stage. Following $m \neq 2, 3$, matching the Target’s announcement of $\chi_1$ yields $\chi_1$ as the status quo and again the war payoff in the game, which is the same result as any other announcement the Signaler might make. Following $m = 3$, the action assigned by the Signaler’s equilibrium strategy yields an expected payoff of $h_t(u^1_t)u^2_s(\chi_3) + (1 - h_t(u^1_t))w_s$ whereas any other announcement again yields $w_s$ for sure.
For type $u^3_s$, at the information sets following $m \neq 3$, the Signaler’s assigned action is again optimal because doing so results in $w_s$, the same expected utility as any other announcement in such an information set. Following $m = 3$, the Signaler’s equilibrium strategy yields an expected payoff of $h_t(u^1_t)u^3_s(\chi_3) + (1 - h_t(u^1_t))w_s$ whereas any other announcement yields $w_s$ for sure.

I now turn to the sequential rationality of the message component of the Signaler’s strategy given the players’ strategies and beliefs in the equilibrium. For type $u^3_s$, the Signaler’s strategy yields an expected utility of $h_t(u^1_t)u^3_s(\chi_3) + (1 - h_t(u^1_t))w_s$ whereas any other message $m$ yields $w_s$ for sure. For type $u^2_s$, the Signaler’s expected utility from $m_2$ is $(h_t(u^1_t) + h_t(u^2_t))u^2_s(\chi_2) + (1 - h_t(u^1_t) - h_t(u^2_t))w_s$ whereas it’s expected utility from $m_3$ is $h_t(u^1_t)u^2_s(\chi_3) + (1 - h_t(u^1_t))w_s$ and $w_s$ for sure for any other message. For type $u^1_s$, the Signaler’s strategy yields expected utility $(h_t(u^1_t) + h_t(u^2_t))u^1_s(\chi_2) + (1 - h_t(u^1_t) - h_t(u^2_t))u^1_s(\chi_1)$ whereas it’s utility from $m_3$ is $h_t(u^1_t)u^1_s(\chi_3) + (1 - h_t(u^1_t))u^1_s(\chi_1)$ and $u^1_s(\chi_1)$ for sure for any other message. Thus, for types $u^1_s$ and $u^2_s$, the Signaler’s actions at these information sets are optimal for $h_t(u^1_t)$ sufficiently low.

Now consider the Target’s strategy given the Target’s beliefs and the Signaler’s strategy. Following $m_3$, the Target’s strategy is optimal for the $u^1_t$ type because it yields an expected utility of $u^1_t(\chi_3)$ and any other announcement yields an expected utility of $w_t$. The Target’s strategy for types $u^2_t$ and $u^3_t$ of announcing $\chi_1$ is optimal because these types prefer war to $\chi_3$ and any announcement other than $\chi_3$ is expected to result in war given the other components of the equilibrium. Following $m_2$, the Target’s strategy for the $u^1_t$ type yields
\( u_t^1(\chi_2) \) whereas any other announcement yields expected utility \( h_s(u_s^1)u_t^1(\chi_1) + (1 - h_s(u_s^1))w_t \).

Similarly, the Target’s strategy for the \( u_t^2 \) type yields \( u_t^2(\chi_2) \) and any other action for this type yields \( h_s(u_s^1)u_t^2(\chi_1) + (1 - h_s(u_s^1))w_t \). Thus, the Target’s strategy is optimal for \( u_t^1 \) and \( u_t^2 \) for sufficiently low \( h_s(u_s^1) \). For the \( u_t^3 \), the Target’s strategy yields \( h_s(u_s^1)u_t^3(\chi_1) + (1 - h_s(u_s^1))w_t > u_t^3(\chi_2) \), and any other choice at that information set yields the same or \( u_t^3(\chi_2) \). Following a signal other than \( m_2 \) or \( m_3 \), the Target’s strategy yields expected utility \( u_t^{z_t}(\chi_1) \forall z_t \) and any other choice is expected to yield the same. \( \blacksquare \)
References


Figure 1. Example of Signaler Utilities
Figure 2. Probability of War