War and Diplomacy on the World Stage: Crisis Bargaining Before Multiple Audiences*

Scott Wolford
swolford@austin.utexas.edu
University of Texas

Abstract

I analyze a three-actor model of crisis bargaining with two key features. First, an informed state can avoid diplomatic opposition, which would raise the costs of war, by conveying limited aims to a potential partner. Second, the means of signaling limited aims may fail to convince an enemy state tempted to risk war of the informed state's willingness to fight. I derive three results. First, war is more likely when third parties believe the informed state have limited aims. Second, the threat of opposition that modestly affects the costs of war discourages risky bluffing. Third, the threat of opposition that substantially raises the costs of war can lead resolute states to mask their willingness to fight, winning diplomatic support at the price of an elevated risk of war, despite the availability of a credible signal. Thus, building diplomatic coalitions can simultaneously make credible communication both easy and unattractive.

Keywords
bargaining, war, signaling, diplomacy

*A previous version of this paper was presented at the 2011 American Political Science Association Annual Conference, 2-5 September, Seattle WA. Thanks to Phil Arena, Jeff Carter, Andrew Coe, Moonhawk Kim, Hans-Inge Langø, Julie Phillips, Toby Rider, Emily Ritter, and Amy Yuen for numerous helpful comments and suggestions, and to Crown & Anchor Pub, Dog & Duck Pub, Pinthouse Pizza, Taco Flats, and Workhorse Bar in Austin, as well as Beer Hive Pub and Squatter's Pub in Salt Lake City, for excellent work environments.

States often struggle to convince others of their willingness to fight, because when communication is cheap nothing stands in the way of bluffing. Bidding up the political costs of backing down from a threat (Fearon 1994, Schultz 1998) or expending resources on military mobilization (Lai 2004, Slantchev 2005) can help states willing to fight to distinguish themselves from bluffers, yet fully separating signals appear rare even when they're available (Fearon 1997, Russett 1963). Leaders infrequently stake their political survival on the outcomes of crises or engage in substantial military mobilizations (see Sechser 2010, Snyder and Diesing 1977, Snyder and Borghard 2011), instead sending "halfhearted" signals that fail to clarify their willingness to fight short of going to war (Byman and Waxman 2002, Christensen 2011, Papayoanou 1997, Wolford 2014). But why would a state that truly is willing to fight over an issue send an ambiguous signal when a costlier, more effective signal is available? Why tolerate the risk of a bloody conflict when that risk could in principle be eliminated with a signal that, though costly, is still cheaper than war?

Explanations for ambiguous signals range from rapidly changing public tastes for war, the potentially provocative effects of strong signals, and the risk of emboldening one's allies (Fearon 1997, p. 84), to the desire to maintain flexibility (Snyder and Diesing 1977, Snyder and Borghard 2011) or surprise (Slantchev 2010), to the need to placate cost-sensitive coalition partners (Wolford 2015, Ch. 4). Yet these explanations overlook the incentives of states that must worry about not only domestic and coalition audiences but also third parties concerned about confronting today's belligerents in the future. Observers of Japan's rise to empire at the turn of the 19th and 20th centuries, major and minor powers alike, worried future confrontations after the First Sino-Japanese and Russo-Japanese Wars (see Paine 2003, 2017). Likewise, the great powers used diplomacy to limit Serbia's gains after twin victories in the First and Second Balkan Wars (Stevenson 1997). More recently, Iraq's neighbors in the wars of 1991 and 2003, Central and Southeastern Europe during NATO interventions in Yugoslavia, and much of East Asia during the Korean War, viewed great power military actions through their own fears of future victimization. States involved in

disputes and wars thus have incentives to convince third parties that their aims are limited. Yet the same actions that can signal a willingness to fight in one crisis—e.g., large military mobilizations—may also signal a willingness to fight in future crises, provoking the costly diplomatic opposition that states wish to avoid.

I explore this strategic problem in a three-player model of crisis bargaining in which an informed state faces two uncertain audiences: an enemy and a potential diplomatic partner. An enemy that believes the informed state willing to fight offers good terms, but a partner that believes it to have expansive aims—that is, to be generally willing to fight—may engage in diplomatic opposition that raises the informed state's costs for war. The informed state would like to convince its enemy of its willingness to fight today and the third party that said willingness doesn't extend into the future. But each audience draws inferences from the same actions during the crisis when the willingness to fight travels, to at least some degree, from present to future interactions. There are three key results. First, war is more likely when the informed state is believed *ex ante* to have limited aims. Second, when diplomatic opposition has a modest effect on the cost of war, the threat opposition can discourage risky bluffing. Third, when opposition has a larger impact on the cost of war, informed states use moderated threats to avoid it, regardless of their willingness to fight. Thus they fail to resolve enemies' uncertainty, sustaining an elevated risk of war despite the availability of a credible signal. Informed states in these equilibria *can* signal a willingness to fight, but they choose not to, preferring a low-cost war in a favorable diplomatic environment to sending a costlier signal that their costs for war are truly low.

The analysis shows that third-party beliefs about a state's foreign policy aims shapes interactions typically treated dyadically (see Braumoeller 2012, Croco and Teo 2005, Poast 2010). Some accounts link multilateral interactions to greater levels of uncertainty (Huth, Bennett and Gelpi 1992, Lake 2010/11), but the addition of third parties to a simple crisis bargaining model can also *reduce* uncertainty and lower the risk of war. Diplomatic coalition-building rarely figures into studies of extended deterrence (Werner 2000, Yuen

2009), alliance politics (Fang, Johnson and Leeds 2014, Johnson and Leeds 2011, Leeds 2003, Trager 2015) and military coalitions (Henke 2016, Wolford 2015, Wolford and Ritter 2016), but I show that the pursuit of diplomatic support affects both whether and how great powers credibly communicate their intentions.[1] Beliefs over whether a states's aims are limited or expansive shape patterns of diplomatic cooperation *and* violent conflict, but the tools that states wield to influence third-party beliefs can also shape the beliefs of their immediate adversaries. Military threats that are effective at the dyadic level can be counterproductive at the systemic level, and vice versa. The theory thus offers a rationale for linking otherwise disparate literatures on crisis bargaining, the balance of power, military and diplomatic multilateralism, and the elusive concept of world opinion.

Third-Party Diplomacy and Crisis Bargaining

When third parties observe others in a war or dispute, they use today's behavior to make guesses about future threats and decisions over whether and how to get involved in today's conflict (see, *inter alia*, Powell 1999, Ch. 5). China, for example, worried that Soviet guns would swivel east after the invasion of Czechoslovakia in 1968 (Westad 2012, p. 360-362), just like potential partners in the Third World, most notably in Southeast Asia and Africa, shied away from China after judging it too aggressive in its border war with India in 1962 (*ibid.*, p. 334). In the first case, the Soviet Union deepened the Sino-Soviet split by showing a willingness to invade Communist countries that deviated from Moscow's line, and in the second, China alienated potential supporters by striking an imperial pose, undermining its anti-colonial credentials through belligerence on its southern frontier. Third parties believe that some element of the willingness to use force travels across crises, which enables the cultivation of reputations.[2] A state's costs for war, in this formulation, have both an idiosyn-

[1] Kreps (2011) examines the construction of diplomatic coalitions in American interventions, but does not examine their impact on signaling and the escalation of disputes to war.

[2] Whether it inheres in governments or leaders (Dafoe 2012, Dafoe, Renshon and Huth 2014, Huth 1997, Weisiger and Yarhi-Milo 2015, Wolford 2007, Wu and Wolford n.d.), a state's reputation is another actor's

cratic, crisis-specific and a systematic, general component, where the latter applies to every crisis. As formalized below, states with low systematic costs for war have *expansive* aims, such that they're willing to use force in a wider range of circumstances than states with *limited* aims, whose higher systematic costs war render them willing to fight under a narrower set of circumstances—e.g., against today's enemy but not against a potentially fearful third party that may be tempted to align diplomatically against states with expansive aims.

Failing to secure diplomatic support can raise the costs of war if it means alienating allies or facing obstacles in international institutions. It can draw in great powers to limit one's gains (including compensatory indemnities) at peace conferences, which coalitions of powerful states did to Japan in 1895 (Paine 2017, Ch. 2) and to Serbia in 1913 (Stevenson 1997). It may even provoke a third party to support one's enemies, as Iran did during the civil war that followed the American-led invasion of Iraq in 2003 (Kilcullen 2006). Even less direct forms of opposition, or "soft balancing" (Kelley 2005, Paul 2005), can raise the costs of war if other states refuse to extend wartime credit (see Shea 2014), as the Western great powers did to Japan in 1905 (Paine 2017, Ch. 3), or cooperate to impose economic sanctions; the Arab oil embargo, launched in response to American support for Israel in the 1973 Yom Kippur War, seriously disrupted Western economies (Gaddis 2005, Tyler 2009). American president Nixon, aware of the threat as war loomed in 1973, cautioned that "we don't want to be so pro-Israel that the oil states—the Arabs that are not involved in the fighting—will break ranks. . . PR [public relations] is terribly important" (quoted in Tyler 2009, p. 211). States thus have strong incentives to care about "PR" in crisis diplomacy.

States can secure diplomatic support in one of two general, often incompatible, ways. First, they may offer sufficient spoils from victory (Schweller 1994) or pose such grave threats (Mearsheimer 2001) that some states choose to bandwagon, either profiting from or hoping to avoid exploitation at the hands of an aggressive great power in the future. Second, they can attempt reassurance by signaling a willingness to use force in only a limited

subjective belief about the value of some enduring quality, such as the willingness to use force.

number of circumstances. But achieving reassurance is neither easy nor cheap (Kydd 2005, Kydd and McManus 2017), because states with expansive aims would prefer that potential victims believe their aims to be limited. The approval of international institutions can sometimes signal limited aims (Chapman 2011, Kreps 2011, Thompson 2006), helping third parties coordinate on expectations that great powers are worth supporting (Voeten 2005), though this sometimes requires a moderation of aims in order to win the cooperation of potential partners (Wolford 2015, Ch. 5).[3] Military threats are central to crisis bargaining both before and after the emergence of global security institutions in the middle of the 20th century (see Slantchev 2011, Stevenson 1997), so in the analysis that follows I focus on how an informed state's threat-making behavior influences both an opponents' and a potential diplomatic supporter's beliefs about the expansiveness of its aims.[4]

However valuable it is in reducing the costs of war, the pursuit of diplomatic support rarely enters directly into the study of crisis bargaining. To the extent that theories of crisis and war focus on multilateral processes, they do so in the context of extended deterrence and alliance commitments (Johnson and Leeds 2011, Leeds 2003, Trager 2015) or military coalitions (Wolford 2015), where coalition-building involves the aggregation of capabilities aimed at shaping the military balance. The cultivation of diplomatic support, though, shapes the expected outcome of war by altering its costs. Diplomatic coalitions form not only when states can offer partners sufficient rewards in return for assistance but also when potential partners can be confident that the state seeking their support won't go on to reward that support by posing a subsequent threat. States win diplomatic support by convincing potential partners of their limited aims, but the means of conveying that—in the analysis

[3]Institutional support, if it lowers the costs of war, can also embolden uninformed states to risk war when they otherwise wouldn't (Chapman and Wolford 2010).

[4]If states worry that potential supporters will draw undesirable inferences from public threats against enemies, they may also attempt reassurance through secret diplomacy (Brown 2014*a*,*b*, Carson 2016, Kurizaki 2007, Yarhi-Milo 2013). Some non-crisis negotiations are easy to keep secret (Crall and Martin 2013), but international crises are often public events, where states often have no choice but to act in public view. Further, secret reassurance might be possible with respect to *some* third parties, but it's not possible for *all* third parties. Military threats and bargaining positions are visible to more states than can be offered specific, private signals of reassurance, such that third parties not privy to secret reassurance may still draw undesirable, diplomatically consequential inferences about a great power's restraint (or lack thereof).

below, moderated military threats—may have unintended consequences for states' ability (and desire) to communicate their intentions during crises.

Model

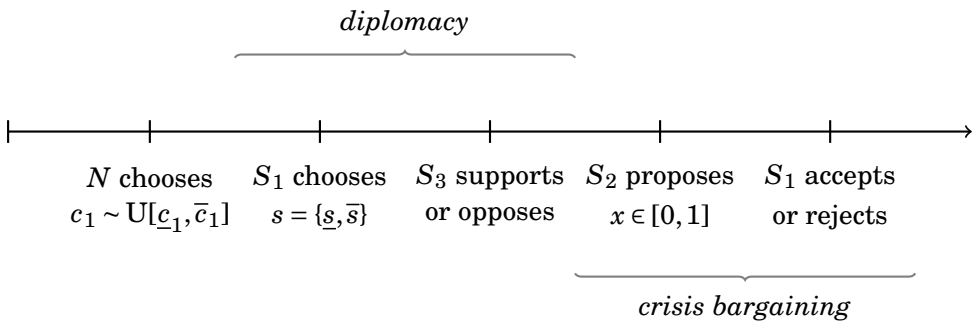
Suppose that two states, S_1 and S_2 (collectively, the "primary disputants"), disagree over how to divide a pie of unit value in the shadow of war, while a third-party state (S_3) commits during the crisis to supporting or opposing S_1 diplomatically should war break out. The other players begin the game uncertain over the systematic component of S_1's costs for war, which travels across crises, in contrast to its idiosyncratic, or crisis-specific, costs for war. When S_1's systematic costs are high, its aims are limited, but when its systematic costs are low, its aims are expansive—or as Powell (1999, p. 194) as it, "generally aggressive." This renders S_2 uncertain over how much it can successfully demand of S_1 short of war and S_3 uncertain over the value of S_1 opposing diplomatically in today's crisis. These other states may, though, be able to improve their estimates after S_1 makes a costly military threat. Both S_2 and S_3 prefer that S_1 has limited aims, leading to a better deal for the former and lower risks of exploitation for the latter, but S_1 would like the other players to hold divergent beliefs. It may get a better bargain from an enemy yet provoke the opposition of a third party that believes its costs for fighting are low, where S_3's opposition raises S_1's costs for war today, draining resources that could be used in the future. I refer to S_3 as a single state to keep the things simple, but it can also be viewed as a pivotal third party or coalition whose diplomatic alignment determines whether S_1's costs of war see a net rise or fall.

The informed state, S_1, pays costs for fighting ($c_1 > 0$) that entail idiosyncratic, or crisis-specific ($\omega_1 > 0$), and systematic, or enduring ($\sigma_1 > 0$), components, or

$$c_1 = \omega_1 + \sigma_1,$$

where σ_1 is the key source of uncertainty for the other players. Figure 1 shows that the

Figure 1: The sequence of play



game begins as Nature chooses σ_1, which defines S_1's type, from the uniform distribution $\sigma_1 \sim \mathrm{U}[\underline{\sigma}_1, \overline{\sigma}_1]$ and reveals them only to S_1.[5] When σ_1 is high, S_1 has limited aims, and its aims are expansive when σ_1 is low. An expansive type pays low costs for war in crises both today and tomorrow, such that it's more willing to use force in subsequent disputes than a limited type, who pays higher costs for war both today and into the future. S_2 cares about S_1's willingness to fight today, which depends on both systematic (σ_1) and idiosyncratic ($\omega_1 > 0$) costs, but S_3 cares only about S_1's willingness to fight tomorrow, which depends on S_1's systematic costs (see also Powell 1999, Ch. 5). The commonly known distribution from which σ_1 is drawn constitutes other players' prior beliefs over S_1's type, i.e. the information available to the uninformed players at the beginning of the crisis. S_1's type distribution is commonly known, so it has a natural interpretation as other players' shared beliefs about S_1's systematic willingness to fight in this stylized three-actor international system.

After learning its type, S_1 chooses a military threat ($s > 0$), an action like mobilization or arming that's costly up front but improves S_1's chances of winning a war (see Slantchev 2011). The threat, for which S_1 pays marginal cost ($a > 0$), can be small ($s = \underline{s}$) or large ($s = \overline{s}$), where $0 < \underline{s} < \overline{s}$. The larger the mobilization, the better S_1's military prospects, such that it defeats S_2 with probability $p(s) \in (0,1)$, where $\partial p(s)/\partial s > 0$, and S_2 wins with the com-

[5] I choose the uniform distribution to pin down a clear analytical solution, but the key results—in particular, the existence of each equilibrium—require only that σ_1 be drawn from a differentiable distribution with strictly positive support.

plementary probability. After observing S_1's mobilization, S_3 chooses whether to support or oppose S_1 diplomatically. Support amounts to tacit acceptance, but if S_3 opposes, then S_1 pays an additional cost ($d > 0$) for war. When d is high, S_3's opposition is consequential, and when d is low, opposition is inconsequential. If S_3 opposes, S_1's total costs of war are $(c_1 + d)$, where it pays $c_1 > 0$ regardless of S_3's choice. S_2 pays $c_2 > 0$ for war. Letting $j = 1$ when S_3 supports and $j = 0$ when it opposes S_1, the primary disputants' war payoffs are

$$EU_1(\text{war}) = p(s) - c_1 - d(1-j) \quad \text{and} \quad EU_2(\text{war}) = 1 - p(s) - c_2,$$

where war is a costly lottery that allocates the whole prize to the winner.[6] Following S_3's choice, S_2 proposes a division of the prize in which S_1 receives x and S_2 receives $1-x$ if S_1 accepts. If S_1 rejects, a war ensues, with payoffs as defined above.

For S_3's preferences I abstract away from the distributive outcome of the current crisis to focus on the future implications of S_3's alignment choice.[7] Suppose that, in some future interaction, S_3's payoffs depend on S_1's willingness and ability to pose a threat.[8] First, S_3 expects to do better against an S_1 with limited aims than an S_1 with expansive aims. This means $u_3(\sigma_1) > 0$, where $\partial u_3(\sigma_1)/\partial\sigma_1 > 0$, such S_3's future payoffs increase in σ_1. To keep the analysis simple, let $u_3(\sigma_1) = \sigma_1$. Second, S_1's ability to pose a future threat depends on the resources left over from today's crisis. When S_3 opposes S_1, it drains a share $r \in (0,1)$ of the resources S_1 can muster in future crises (cf. Treisman 2004). Finally, S_3 pays nothing to support S_1, but opposition entails a cost $k > 0$ in foregone side payments (see Wolford 2015, Wolford and Ritter 2016) and damaged relations with S_1 and its allies. Therefore, support

[6]The model's solution is not substantively different when S_3's opposition to (support for) S_1 implies support for (opposition to) S_2. I use the simpler specification here, but the alternative might be useful for analyzing superpower competition during the Cold War or crises between two great powers.

[7]This is without loss of generality as long as S_3 contends with S_1 on an issue whose resolution can be shaped by the outcome of the first crisis. In practice, it simply rules out bandwagoning as a potential motive for support.

[8]A similar game would represent a subsequent interaction as a game in which S_1 and S_3 bargain in the shadow of war such that bargains struck reflect each side's costs for fighting (see Powell 1999, Ch. 4). Results aren't substantively different from the reduced form version, which I present here to keep solution and explication space to a minimum.

yields σ_1, and opposition yields $-k+(1-r)\sigma_1$. Opposition makes future aggression more difficult, but S_3 would rather save the costs of opposition if S_1 has truly limited aims.

The following functions define payoffs over the game's terminal nodes of settlement and war, both with and without diplomatic support. For S_1,

$$u_1 = \begin{cases} -as + x & \text{if settlement } \forall j \\ -as + p(s) - c_1 & \text{if war and support} \\ -as + p(s) - c_1 - d & \text{if war and opposition.} \end{cases}$$

For S_2,

$$u_2 = \begin{cases} 1 - x & \text{if settlement } \forall j \\ 1 - p(s) - c_2 & \text{if war } \forall j, \end{cases}$$

and for S_3,

$$u_3 = \begin{cases} \sigma_1 & \text{if support} \\ -k + (1-r)\sigma_1 & \text{if oppose.} \end{cases}$$

The model shares several features with standard two-player crisis bargaining games, in that an informed side can signal a willingness to fight with costly actions (e.g. Arena n.d., Fearon 1997, Schultz 1998, Slantchev 2005). Other models introduce third parties with different information and roles in the conflict: Werner's (2000) foundational three-player game occurs under complete information, Yuen (2009) treats third-party preferences as a source of uncertainty, Favretto (2009) shows how the commonly-known preferences of third parties can reduce the impact of bilateral uncertainty, and Trager (2015) shows how third-party commitments can enable costless communication between adversaries.[9] In the most

[9] Johns (2007) examines agents who hold information private from two principals in an international-bureaucratic, as opposed to crisis bargaining, framework.

similar model, Wolford (2014) introduces a potential military partner for the informed player with preferences over both today's crisis and the costs of participating in the war; however, the potential partner is also informed of S_1's type and can shape the military balance with its alignment choice. In the present formulation, the third party's alignment has no effect on the distribution of military power, and it begins the game as uncertain as the enemy over S_1's willingness to fight. This creates a signaling dilemma derived from third-party beliefs over a state's general willingness to use force, a factor thus far unexplored in the literatures on crisis bargaining, multilateralism, and war.

Analysis

How do states manage multiple audiences in crisis bargaining? The answer depends on (a) the extent to which S_3's opposition increases S_1's costs for war and (b) whether S_3 conditions its alignment on S_1's military threat in today's crisis. These dimensions define two types of Perfect Bayesian Equilibrium (PBE).[10] First, in a pooling equilibrium, S_1 makes the same threat (large or small) regardless of its type, preventing other players from updating their beliefs. Second, in a semi-separating equilibrium, S_1 plays a cutpoint strategy, such that all types below a certain cost threshold make the large threat and all types above the threshold make the small threat, allowing other players to truncate their beliefs to smaller, still uniformly-distributed, partitions. After detailing how S_2 and S_3 condition their strategies on S_1's threat, I discuss the conditions under which each type of equilibrium exists.

The analysis rests on two substantive assumptions. First, S_1's total costs for war are never believed to be so large as to make fighting incredible, so $\overline{c}_1 < p(\underline{s}) - d$. This ensures that S_2 faces a genuine risk-return tradeoff in choosing its proposal and that its proposals are always interior, or $x^* \in (0,1)$, which simplifies the analysis at no loss of generality. Second,

[10] In PBE, strategy profiles are sequentially rational and weakly consistent with beliefs updated according to Bayes' Rule wherever possible (Fudenberg and Tirole 1991, Ch. 8).

S_1's use of military threats is not prohibitively costly, or

$$\left(p(\overline{s}) - p(\underline{s})\right) - a\left(\overline{s} - \underline{s}\right) \equiv M > 0 \Leftrightarrow a < \frac{p(\overline{s}) - p(\underline{s})}{\overline{s} - \underline{s}}, \tag{1}$$

such that the net military gain (M) from making the large threat is positive. This (i) ensures that any choice not to make the large threat doesn't depend on conditions under which such a threat might be *ex ante* too costly and (ii) stacks the deck against the revelation of information, increasing the analyst's confidence in the effects of the proposed mechanisms to lead to either credible signaling or pooling on small military threats.

The Multiple Audience Problem

S_1's challenge is to convince S_2 of its willingness to fight today and S_3 of its limited aims into the future. In this section, I derive general features of S_2's and S_3's equilibrium strategies, then show how they shape S_1's initial choice over large or small military threats. First, the more S_2 offers, the more likely is S_1 to accept, but the less favorable are S_2's terms of settlement. Thus, S_2's offer trades an acceptable chance of war for a better bargain if S_1 accepts. Lemma 1 characterizes S_2's strategy for any combination of S_1's threat (s), S_2's beliefs when it makes its proposal ($\sigma_1 \sim \mathrm{U} \in [\underline{\sigma}_1', \overline{\sigma}_1']$), and S_3's diplomatic alignment (j).

Lemma 1. *For any s, $\sigma_1 \sim U \in [\underline{\sigma}_1', \overline{\sigma}_1']$ and j, S_2's equilibrium proposal is*

$$x = p(s) + \frac{c_2 - \omega_1 - \overline{\sigma}_1' - d(1-j)}{2}, \tag{2}$$

which all types $\sigma_1 < p(s) - \omega_1 - d(1-j) - x^*$ reject.

Equation (2) shows that S_2 makes more generous offers, reducing the chances of rejection, when its own costs of war (c_2) are high, its estimate of S_1's maximum possible costs of war ($c_1 = \omega_1 + \overline{\sigma}_1'$) is low, and when S_3 supports S_1 ($j = 1$). The last two elements reflect S_1's dilemma: larger threats, by virtue of their cost, may convince S_2 that S_1's costs of war

are low, producing more generous terms, but convincing S_3 of the same fact can provoke diplomatic opposition that leads S_2 to make a less generous offer, since S_2 can propose to keep more for itself if it knows that S_3's choice has increased S_1's costs of war.

S_1's strategy may provoke third-party opposition, which creates a dilemma when S_3 supports small threats and opposes large ones. If S_3 is sure to support or oppose S_1 regardless of the size of its military threat, the model collapses to a standard two-player signaling game. Lemma 2 characterizes the conditions under which S_3 conditions support on the size of the threat, where $\mathbf{E}(\sigma_1|s)$ is S_3's updated estimate of S_1's systematic costs for fighting.

Lemma 2. *S_3 plays a conditional strategy, supporting iff $s = \underline{s}$ when $-ru_3\left(\mathbf{E}(\sigma_1|\underline{s})\right) \leq k < -ru_3(\mathbf{E}(\sigma_1|\overline{s}))$, where $\mathbf{E}(\sigma_1|\underline{s}) > \mathbf{E}(\sigma_1|\overline{s})$ ensures that $u_3(\mathbf{E}(\sigma_1|\overline{s})) < u_3\left(\mathbf{E}(\sigma_1|\underline{s})\right)$.*

S_3 conditions its strategy on S_1's threat when the latter is more likely to make a large threat if its aims are unlimited than if its aims are limited. This is precisely the condition needed for S_1 to demonstrate a willingness to fight to S_2 and to secure a more generous offer in standard two-player models, but it activates S_1's signaling dilemma once third-party alignments are at stake. Lemma 2 expresses this condition in terms of a middling range of the costs of opposition (k). S_3 plays a conditional strategy when opposition is neither too expensive nor too cheap relative to the consequences of the information revealed by S_1's threat. This condition becomes easier to satisfy as S_3's opposition drains more resources for future use (r). Thus, S_3's ability and desire to restrain S_1's willingness to wage war depends on how prior beliefs over S_1's restraint interact with S_1's choice of signal.

The Signaling Dilemma

Defining S_2's and S_3's strategies entails several equilibria. I first characterize PBE in which S_3 supports or opposes S_1 unconditionally, where the only equilibria entail large threats ($s^* = \overline{s}$) regardless of S_1's type. In these equilibria, S_2's information problem remains unaffected by either S_1's threat or S_3's alignment. This facilitates comparison to PBE that

exist when S_3 plays a conditional strategy, which induces pooling on large threats under some conditions, as well as two additional PBE: (a) a semi-separating equilibrium in which unlimited types make a large threat while limited types issue the small threat and (b) a distinct pooling equilibrium in which S_1 makes a small threat ($s^* = \underline{s}$) regardless of its type, securing S_3's support at the cost of leaving S_2's information problem unsolved.

S_3 supports unconditionally when the costs of opposition are sufficiently high, opposing unconditionally when the costs of opposition are sufficiently low.[11] In such equilibria, no barrier—here, the threat of losing diplomatic support—stands in the way of of S_1's limited types making large threats to bluff about their willingness to fight. Thus, only pooling equilibria, in which S_1 makes the large military threat regardless of its type, exist.

Proposition 1. *PBE with the following strategies exist when S_3's strategy is unconditional.*

(a.) When $k < -r(\overline{\sigma}_1 + \underline{\sigma}'_1)/2$, where $\underline{\sigma}'_1 > \underline{\sigma}_1$ off the equilibrium path, there exists a pooling equilibrium in which $s^* = \overline{s}$ for all σ_1 and S_3 opposes S_1 unconditionally.

(b.) When $k \geq -r(\overline{\sigma}_1 + \underline{\sigma}_1)/2$, there exists a pooling equilibrium in which $s^* = \overline{s}$ for all σ_1 and S_3 supports S_1 unconditionally.

Proposition 1 confirms that the model generates standard crisis bargaining patterns when S_3 supports S_1 unconditionally. S_2 is more willing to risk war when its own costs for war are low, less willing when it believes that S_1's costs for war are low. With no disincentive to bluff, low-cost types of S_1 stand nothing to gain by revealing their type with a small threat, so they pool with lower-cost types and leave S_2's risk-return tradeoff unaffected. But as shown in Equation (3), which characterizes the probability of war in any PBE with pooling on large threats, the probability of war is lower when S_3 opposes unconditionally.

[11] Unconditional strategies can also exist for the same values of k that support conditional strategies, and the constraints supporting either unconditional strategy can overlap in middling ranges of k as well. These strategies are unique, however, at extreme values of k.

Letting $j = 1$ for support and $j = 0$ for opposition, the equilibrium probability of war is

$$\Pr(\text{war}|j) = 1 - \frac{d(1-j) + c_2 + \omega_1 + \overline{\sigma}_1}{2(\overline{\sigma}_1 - \underline{\sigma}_1)}, \tag{3}$$

which falls when S_3 opposes ($j = 0$) and as the consequences of its opposition (d) grow more severe. Therefore, *war is more likely when third parties believe S_1 to have limited aims*, because S_1 secures S_3's support by default. On the other hand, war is *less* likely when third parties believe S_1 to have expansive aims, because it is sure to provoke diplomatic opposition and face costlier wars. Setting aside any potential effects on S_1's willingness to use threats as signals, S_3's opposition reduces the probability of war by increasing the range of settlements S_1 accepts in lieu of war, regardless of its type. This reduction in the probability of war comes at S_1's expense, however, because S_1 receives less favorable offers. This creates a strategic incentive to seek out S_3's diplomatic support when possible.

If S_3's costs of opposition fall in a middling range, or $-ru_3\left(\mathbf{E}(c_1|\underline{s})\right) \leq k < -ru_3(\mathbf{E}(c_1|\overline{s}))$, new information about S_1's type allows S_3 to support only small mobilizations. This can alter S_1's strategy when S_3's opposition causes a sufficiently large increase in the costs of war, i.e. when $d \geq M$. Proposition 2 states that the large-threat pooling equilibrium exists, despite S_3's conditional strategy, when opposition only minimally raises S_1's costs of war ($d \leq 2M$); however, as d increases, other PBE emerge with different implications for the credibility of signals, the willingness to send them, and the probability of war.

Proposition 2. *PBE with the following strategies exist when S_3 supports iff $s^* = \underline{s}$.*

(a.) When $d \leq 2M$ and

$$-r\left(\frac{\overline{\sigma}_1 + \underline{\sigma}_1}{2}\right) \leq k < -r\left(\frac{\overline{\sigma}_1 + \underline{\sigma}'_1}{2}\right),$$

where $\underline{\sigma}'_1 > \underline{\sigma}_1$ off the equilibrium path, there exists a pooling equilibrium in which $s^* = \overline{s}$ for all σ_1.

(b.) When $\left(2M + c_2 + \omega_1 + \underline{\sigma}_1\right) < d < (2M + c_2 + \omega_1 + \overline{\sigma}_1)$ *and*

$$-r\left(\frac{\overline{\sigma}_1 + \underline{\sigma}_1}{2}\right) \leq k < -r\left(\frac{\overline{\sigma}_1 + \underline{\sigma}_1'}{2}\right),$$

there exists a semi-separating equilibrium in which $s^* = \overline{s}$ *for* $c_1 < \hat{c}_1$ *and* $s^* = \underline{s}$ *for* $c_1 \geq \hat{c}_1$.

(c.) When $d \geq M$ *and*

$$-r\left(\frac{\overline{\sigma}_1 + \underline{\sigma}_1}{2}\right) \leq k < -r\left(\frac{\overline{\sigma}_1 + \underline{\sigma}_1'}{2}\right),$$

where $\overline{c}_1' < \overline{c}_1$ *off the equilibrium path, there exists a pooling equilibrium in which* $s^* = \underline{s}$ *for all* c_1.

See appendix for proof.

When S_3's opposition has little impact on S_1's costs of war, even a conditional strategy cannot dissuade limited types of S_1 from bluffing with large threats. When $d \leq 2M$, S_1 tolerates diplomatic opposition and makes a large threat regardless of its type, because limited types are happy to take advantage of S_2's uncertainty and bluff, just as they are in the equilibrium with unconditional opposition in Proposition 1 (where, in effect, $d = 0$). But when diplomatic opposition is more consequential for S_1's costs for fighting, S_1 has a correspondingly weaker incentive to convey a willingness to fight, and limited types are discouraged from using a large threat to mimic expansive types.

When S_3 conditions diplomatic support on the size of military threats, S_1 strategically chooses the size of its threat. If limited S_1 are unwilling to provoke opposition to convey resolve, the threat of diplomatic opposition sometimes redounds to S_3's advantage; when d is large (but not too large), such that $\left(2M + c_2 + \omega_1 + \underline{\sigma}_1\right) < d < (2M + c_2 + \omega_1 + \overline{\sigma}_1)$, a semi-separating equilibrium exists in which S_1 issues large threats only when its aims are unlimited. Limited types of S_1 are happy to reveal themselves as such, because the expected gains

from conveying low costs for fighting are offset by the savings of earning S_3's diplomatic support. The threat of provoking third-party opposition disciplines the incentive to bluff—an additional cost to large mobilizations that makes communication credible, whether S_1 avoids the large mobilization or not. The third party's promise of support can discourage bluffing that might otherwise occur, allowing both limited and expansive types to win concessions peacefully, the latter revealing their willingness to fight by provoking diplomatic opposition.

Ikenberry (2001) terms the pursuit of limited aims in order to win diplomatic support "strategic restraint," and this strategy succeeds in the semi-separating equilibrium; states with limited aims build diplomatic coalitions by moderating their military threats. Proposition 2 shows that this sanguine view of strategic restraint is only valid, however, when S_3 can only modestly raise S_1's costs for war. When losing diplomatic support affects the costs of war substantially, S_3's conditional strategy loses its ability to discipline S_1's threats and instead *discourages* the revelation of information. When $d \geq M$, S_1 plays a pooling strategy that turns the logic of bluffing on its head: expansive types mimic limited types, securing diplomatic support by choosing a small mobilization. S_1's most expansive types, which S_3 would like to oppose, refuse to reveal themselves, accepting a low-cost war over a peaceful settlement reflecting the substantial costs of diplomatic opposition. This is the dark side of strategic restraint. States may work hard to limit their threats to secure diplomatic support, but in the process they (a) prevent third parties from opposing the most expansive states and (b) sustain an elevated probability of war, despite the availability of a credible signal that would convince S_2 of S_1's willingness to fight—if only it were sent. And in contrast to a similar pooling PBE in Wolford (2014)'s model with only one uncertain audience, the foregone large threat is credible off the equilibrium path; if S_1 were to make the large threat, S_2 would believe it to be higher-cost and make an accordingly increased offer.

Table 1 shows that when S_1 secures S_3's support by pooling on a small threat, the probability of war equals that found in Equation (3) when $j = 1$. S_2's prior beliefs are unaltered,

Table 1: Probability of war by equilibrium strategies

3's strategy	1's strategy	Probability of war
When $k \geq -ru_3(\mathbf{E}(\sigma_1 \mid \overline{s}))$,		
Support $\forall s$	Pool on $s^* = \overline{s}$	$1-(c_2+\omega_1+\overline{\sigma}_1)/2(\overline{\sigma}_1-\underline{\sigma}_1)$
When $-ru_3\left(\mathbf{E}(\sigma_1 \mid \underline{s})\right) \leq k < -ru_3(\mathbf{E}(\sigma_1 \mid \overline{s}))$,		
Support iff $s^* = \underline{s}$	Pool on $s^* = \overline{s}$	$1-(d+c_2+\omega_1+\overline{\sigma}_1)/2(\overline{\sigma}_1-\underline{\sigma}_1)$
Support iff $s^* = \underline{s}$	$s^* = \overline{s}$ if $\sigma_1 < \hat{\sigma}_1$,	$1-(d+c_2+\omega_1+\hat{\sigma}_1)/2(\hat{\sigma}_1-\underline{\sigma}_1)$
	$s^* = \underline{s}$ if $\sigma_1 \geq \hat{\sigma}_1$	$1-(c_2+\omega_1+\overline{\sigma}_1)/2(\overline{\sigma}_1-\hat{\sigma}_1)$
Support iff $s^* = \underline{s}$	Pool on $s^* = \underline{s}$	$1-(c_2+\omega_1+\overline{\sigma}_1)/2(\overline{\sigma}_1-\underline{\sigma}_1)$
When $k < -ru_3\left(\mathbf{E}(\sigma_1 \mid \underline{s})\right)$,		
Oppose $\forall s$	Pool on $s^* = \overline{s}$	$1-(d+c_2+\omega_1+\overline{\sigma}_1)/2(\overline{\sigma}_1-\underline{\sigma}_1)$

and since S_1 pays no additional costs for fighting, the probability of war is just as high as it is in a PBE where S_3 supports unconditionally. S_3 supports some types of S_1 that it regrets after the fact, but the probability of war in this pooling on low-threats equilibrium is *not* the highest in the game. When S_3's impact on the costs of war is most consequential (when d is sufficiently large), war is most likely in precisely those cases where S_3 is happiest offering S_1 its diplomatic support—that is, when S_1 plays a semi-separating strategy and reveals its aims to be relatively limited ($s^* = \underline{s}$)—when it emboldens state S_2 to risk war. On the other hand, war is less likely in the same semi-separating equilibrium when S_3 opposes after a large threat ($s^* = \overline{s}$), because S_2 can secure ever more favorable bargains as S_1's costs of war increase.[12] Therefore, when its opposition is most consequential, S_3 achieves its goal of opposing states with unlimited aims at the cost of increasing the probability that today's dispute ends in war; but when S_1 *does* provoke opposition, the probability of war falls to its lowest level in any equilibrium so long as d is sufficiently large.

The model thus offers an explanation for why states can build large diplomatic coalitions only to see their targets still doubt their willingness to fight: the process of signaling limited

[12]Recall that $\hat{c}_1$ is the threshold (defined in Proposition 2) separating S_1's types that issue small ($c_1 \leq \hat{c}_1$) and large ($c_1 > \hat{c}_1$) threats.

aims that ensures diplomatic support can also preserve doubts in an opponent's mind that the state building the coalition is truly committed to fighting. Building coalitions requires concessions (Kreps 2011, Riker 1962, Wolford 2015), and winning support may require compromise in the form of small or moderated military threats. In aid of securing both local and broader regional support for the reversal of Iraq's annexation of Kuwait, for example, the United States targeted its 1990-91 buildup narrowly at Kuwait and southeastern Iraq (Atkinson 1993, Bush and Scowcroft 1998). This secured wide diplomatic support for the preservation of recognized international borders (Kreps 2011), but it had the unintended consequence of preserving some optimism in the Iraqi leadership that it could bog the coalition down in a costly, narrow-front war in rough terrain (on this point, see Lindsey 2015). Only after the successful execution of the armored "left hook" strategy was Iraq disabused of the notion that the coalition faced significant costs for prosecuting the war. Likewise, the United States consented to limited military threats—i.e., an air campaign in lieu of a ground invasion—against Serbia before the Kosovo War of 1999 (Clark 2001, Henriksen 2007, Richardson 2000), which preserved both sufficient diplomatic support *and* Serbian doubts about NATO's willingness to wage a lengthy war (Wolford 2015, Ch. 4).

Conclusion

States often send ambiguous signals of their willingness to fight when more credible signals are available, prompting their opponents to risk war. Prevailing explanations for this pattern, focused on leader psychology and domestic politics, overlook a systemic imperative that encourages states to moderate their threats even when they're truly willing to fight: preventing diplomatic opposition that can raise the costs of war. When third-party alignment decisions have even a modest impact on an informed state's costs for fighting, the desire to secure support can simultaneously make signaling a willingness to fight in today's crisis easier *and* less attractive. Therefore, whether states can and wish to signal credibly

a willingness to use force during particular crises depends on beliefs among third parties about the expansiveness of their foreign policy aims.

To the extent that states value broad diplomatic support, the model identifies an endogenous and reciprocal link between a systemic variable, i.e. prevailing beliefs over the aims of particular states, and the risk of war in individual crises—in stark contrast to Waltz's (1959) claim that the structure of the system gives only permissive conditions for war. Third-party beliefs determine the ease with which support is gained, while securing it (or failing to) also changes prevailing assessments of a state's aims for future interactions. Particular constellations of beliefs dictate whether third parties play conditional diplomatic strategies, but in the aggregate, the probability of war in a given crisis should be *greater* when third parties more trusting of that state's foreign policy ambitions. System-level variables—configurations of military power (Mansfield 1994, Monteiro 2011/12), distributions of ideology (Braumoeller 2012, Kadera, Crescenzi and Shannon 2003, Mitchell 2002), and the offense-defense balance (Jervis 1978)—are often linked to the risk of war, and this model suggests another potential avenue for the system to shape dyadic interactions: third-party beliefs about a disputant's aims, which are shaped by the same actions in today's crisis as the beliefs held by a disputant's opponent. Thus, a specific feature of the international system can induce patterns of third-party diplomatic alignment, opponent bargaining positions, and great power signaling, associated with varying levels of information transmission and risks of war, absent any meaningful variation in either international institutions, the distribution of power, or domestic politics. The presence of diplomatic support, especially when it forces states into making limited threats, may be a crucial omitted variable in dyadic studies of the success of coercive threats in crisis bargaining.

Finally, the theory suggests that the concept of "world opinion" plays a more important role in international politics than generally accepted; shared beliefs about state preferences can be viewed as a meaningful element of the structure of the international system (cf. Wendt 1999), one that can change more frequently than material elements like the dis-

tributions of power or political regimes and ideologies. During the Cold War, for example, nonaligned states sought to "tilt without tipping" (Gaddis 2005, p. 122), punishing a superpower's lack of restraint by opposing it in particular crises, all in the service of "resisting superpower hegemony" (pp. 123-4). These decisions are not without consequence when diplomatic alignments can alter the costs of war; both the United States and the Soviet Union tried to win the diplomatic support of "anti-bloc" third-party states in hopes of lowering their own costs of war (Lane 2003, pp. 151-2, 166). When cultivating diplomatic coalition partners requires adjustments to military threats—as happened in moments as diverse as the July Crisis of 1914 (Clark 2012, Hastings 2013), the Berlin Crisis of 1961 (Aono 2010, Freedman 2000), the Persian Gulf War of 1991 (Atkinson 1993, Kreps 2011), and the Kosovo Crisis of 1999 (Clark 2001, Wolford 2015)—and when failures of reassurance result in third-party opposition—as it did in the Iraq War of 2003—diplomatic coalition-building is not mere window-dressing. It may determine whether particular great power crises are settled peacefully or end in war. Strategic restraint, useful for great powers in creating a sustainable international order in the long run (see Ikenberry 2001), may come at the cost of an increased risk of war in the short run, depending on the ideational structure of the international system.

Appendix

Proof of Lemma 1. Begin with S_1's acceptance rule, which stipulates that it accepts some x when $x \geq p(s) - c_1 - d(1-j)$, such that S_1 accepts iff $\sigma_1 \geq p(s) - \omega_1 - d(1-j) - x$. Given beliefs $\sigma_1 \sim \mathrm{U}[\underline{\sigma}_1', \overline{\sigma}_1']$, S_2 sets x to solve $\max_x\{EU_2(x)\}$, or

$$\max_x \left\{ \int_{\underline{\sigma}_1'}^{p(s)-\omega_1-d(1-j)-x} (1 - p(s) - c_2)\, \mathrm{d}\sigma_1 + \int_{p(s)-\omega_1-d(1-j)-x}^{\overline{\sigma}_1'} (1-x)\, \mathrm{d}\sigma_1 \right\}.$$

The first order condition $2(p(s)-x)+c_2-\omega_1-\overline{\sigma}_1'-d(1-j)=0$ yields an optimum at

$$x^* = p(s) + \frac{c_2-\omega_1-\overline{\sigma}_1'-d(1-j)}{2},$$

and since $\partial^2 EU_2(x)/\partial x^2 = -2$, x^* is sure to yield a maximum. □

Proof of Lemma 2. Two inequalities must be true. First, S_3 supports if $s=\underline{s}$ when

$$u_3\left(\mathbf{E}(\sigma_1|\underline{s})\right) \geq -k+(1-r)u_3\left(\mathbf{E}(\sigma_1|\underline{s})\right),$$

or when $k \geq -r\left(\mathbf{E}(\sigma_1|\underline{s})\right)$. Second, S_3 opposes if $s=\overline{s}$ when

$$-k+(1-r)u_3(\mathbf{E}(\sigma_1|\overline{s})) > u_3(\mathbf{E}(\sigma_1|\overline{s})),$$

or when $k < -ru_3(\mathbf{E}(\sigma_1|\overline{s}))$. Therefore, S_3 plays a conditional strategy when

$$-ru_3\left(\mathbf{E}(\sigma_1|\underline{s})\right) < k \leq -ru_3(\mathbf{E}(\sigma_1|\overline{s})),$$

opposing unconditionally when k falls below this range and supporting unconditionally when k falls above. □

Proof of Proposition 1. For the equilibrium in which S_3 opposes unconditionally, strategies and beliefs are as follows. S_1 sets $s^*=\overline{s}$, accepts iff $c_1 \geq p(\overline{s})-d-x$ on the equilibrium path and accepts iff $\sigma_1 \geq p(\underline{s})-\omega_1-d-x$ off the equilibrium path. On the equilibrium path, S_2 and S_3 retain their prior beliefs, and off the equilibrium path they believe $\sigma_1 \sim [\underline{\sigma}_1', \overline{\sigma}_1]$, where $\underline{\sigma}_1 < \underline{\sigma}_1'$ such that higher-cost types are believed to be more tempted to save the costs of large mobilization than lower-cost types. By Lemma 1, S_2 proposes $x^* = p(\overline{s})+(c_2-\omega_1-\overline{\sigma}_1-d)/2$ on the equilibrium path and $x' = p(\underline{s})+(c_2-\omega_1-\overline{\sigma}_1-d)/2$ off the equilibrium path. S_3 opposes on and off the equilibrium path.

Since $\underline{\sigma}_1 < \underline{\sigma}_1'$ and only relatively high cost types are believed to have deviated to the low

threat, showing that S_3 opposes off the equilibrium path is sufficient to show that it will oppose on the path as well. Therefore,

$$-k+(1-r)\int_{\underline{\sigma}_1'}^{\overline{\sigma}_1}\left(\sigma_1\times\frac{1}{\overline{\sigma}_1-\underline{\sigma}_1'}\right)\mathrm{d}\sigma_1>\int_{\underline{\sigma}_1'}^{\overline{\sigma}_1}\left(\sigma_1\times\frac{1}{\overline{\sigma}_1-\underline{\sigma}_1'}\right)\mathrm{d}\sigma_1,$$

which reduces to $k<-r(\overline{\sigma}_1+\underline{\sigma}_1')/2$, establishes S_3's behavior in and out of equilibrium. Finally, I establish that the highest-cost type of S_1 ($\sigma_1=\overline{\sigma}_1$) issues the large threat rather than deviate to the small threat and induce the other players to believe that it is relatively high cost; if this type will not deviate to a smaller threat, then no types will. This is a type that accepts S_2's offer in and out of equilibrium, so it sets $s=\overline{s}$ when $-a\overline{s}+x^*\geq -a\underline{s}+x'$, or when $\left(p(\overline{s})-p(\underline{s})\right)-a\left(\overline{s}-\underline{s}\right)\geq 0$, which is true by construction.

For the equilibrium in which S_3 supports unconditionally, strategies and beliefs are as follows. S_1 sets $s^*=\overline{s}$, accepts iff $\sigma_1\geq p(\overline{s})-\omega_1-x$ on the equilibrium path and accepts iff $\sigma_1\geq p(\underline{s})-\omega_1-x$ off the equilibrium path. On the equilibrium path, S_2 and S_3 retain their prior beliefs, and off the equilibrium path they believe $\sigma_1\sim[\underline{\sigma}_1',\overline{\sigma}_1]$, where $\underline{\sigma}_1<\underline{\sigma}_1'$ such that higher-cost types are believed to be more tempted to save the costs of large mobilization than lower-cost types. By Lemma 1, S_2 proposes $x^*=p(\overline{s})+(c_2-\omega_1-\overline{\sigma}_1)/2$ on the equilibrium path and $x'=p(\underline{s})+(c_2-\omega_1-\overline{\sigma}_1)/2$ off the equilibrium path. S_3 supports on and off the equilibrium path.

Since $\underline{\sigma}_1<\underline{\sigma}_1'$ and only relatively high cost types are believed to have deviated to the small threat, showing that S_3 supports on the equilibrium path is sufficient to show that it will support off the path as well. Therefore,

$$\int_{\underline{\sigma}_1}^{\overline{\sigma}_1}\left(\sigma_1\times\frac{1}{\overline{\sigma}_1-\underline{\sigma}_1}\right)\mathrm{d}\sigma_1\geq -k+(1-r)\int_{\underline{\sigma}_1}^{\overline{\sigma}_1}\left(\sigma_1\times\frac{1}{\overline{\sigma}_1-\underline{\sigma}_1}\right)\mathrm{d}\sigma_1,$$

which reduces to $k\geq -r(\overline{\sigma}_1+\underline{\sigma}_1)/2$, establishes S_3's behavior in and out of equilibrium. Finally, I establish that the highest-cost type of S_1 ($\sigma_1=\overline{\sigma}_1$) issues the large threat rather than deviate to the small threat and induce the other players to believe that it is relatively

high cost; if this type will not deviate to a smaller threat, then no types will. This is a type that accepts 2's offer in and out of equilibrium, so it sets $s = \overline{s}$ when $-a\overline{s} + x^* \geq -a\underline{s} + x'$, or when $\left(p(\overline{s}) - p(\underline{s})\right) - a\left(\overline{s} - \underline{s}\right) \geq 0$, which is true by construction. □

Proof of Proposition 2. For the equilibrium in which S_1 pools on the high threat, strategies and beliefs are as follows. S_1 sets $s^* = \overline{s}$, accepts iff $\sigma_1 \geq p(\overline{s}) - \omega_1 - d - x$ on the equilibrium path and accepts iff $\sigma_1 \geq p(\underline{s}) - \omega_1 - x$ off the equilibrium path. On the equilibrium path, S_2 and S_3 retain their prior beliefs, and off the equilibrium path they believe $\sigma_1 \sim [\underline{\sigma}_1', \overline{\sigma}_1]$, where $\underline{\sigma}_1 < \underline{\sigma}_1'$ such that only relatively restrained types would be tempted to deviate to the low threat, saving the costs of a large mobilization and avoiding S_3's opposition. By Lemma 1, S_2 proposes $x^* = p(\overline{s}) + (c_2 - omega_1 - \overline{\sigma}_1 - d)/2$ on the equilibrium path and $x' = p(\underline{s}) + (c_2 - \omega_1 - \overline{\sigma}_1)/2$ off the equilibrium path. S_3 opposes on the equilibrium path and supports off the equilibrium path.

Two inequalities must be satisfied for S_3's strategy to be sequentially rational. On the equilibrium path, S_3 opposes when

$$-k + (1-r)\int_{\underline{\sigma}_1}^{\overline{\sigma}_1}\left(\sigma_1 \times \frac{1}{\overline{\sigma}_1 - \underline{\sigma}_1}\right)\mathrm{d}\sigma_1 > \phi\int_{\underline{\sigma}_1}^{\overline{\sigma}_1}\left(\sigma_1 \times \frac{1}{\overline{\sigma}_1 - \underline{\sigma}_1}\right)\mathrm{d}\sigma_1,$$

or when $k < -r(\underline{\sigma}_1 + \overline{\sigma}_1)/2$. Off the equilibrium path, it supports when

$$\int_{\underline{\sigma}_1'}^{\overline{\sigma}_1}\left(\sigma_1 \times \frac{1}{\overline{c}_1 - \underline{c}_1'}\right)\mathrm{d}\sigma_1 \geq -k + (1-r)\int_{\underline{\sigma}_1'}^{\overline{\sigma}_1}\left(\sigma_1 \times \frac{1}{\overline{\sigma}_1 - \underline{\sigma}_1'}\right)\mathrm{d}\sigma_1,$$

or when $k \geq (\underline{\sigma}_1' + \overline{\sigma}_1)/2$. Therefore, S_3's strategy is sequentially rational when $-r(\underline{\sigma}_1 + \overline{\sigma}_1)/2 < k \leq -r(\underline{\sigma}_1' + \overline{\sigma}_1)/2$. Finally, I show that the highest-cost type of S_1 ($\sigma_1 = \overline{\sigma}_1$) issues the large threat rather than deviate to the small threat and induce the other players to believe that it is relatively high cost; if this type will not deviate to a small threat, then no types will. This is a type that accepts S_2's offer in and out of equilibrium, so it sets $s^* = \overline{s}$ when $-a\overline{s} + x^* \geq -a\underline{s} + x'$, or when $d \leq 2\left(p(\overline{s}) - p(\underline{s})\right) - a\left(\overline{s} - \underline{s}\right) \equiv 2M$.

For the semi-separating equilibrium, strategies and beliefs are as follows. S_1 sets $s^* = \underline{s}$

and accepts iff $\sigma_1 \geq p(\underline{s}) - \omega_1 - x$ when $\sigma_1 \geq \hat{\sigma}_1$, and it sets $s^* = \overline{s}$ and accepts iff $\sigma_1 \geq p(\overline{s})\omega_1 - d - x$ when $\sigma_1 < \hat{\sigma}_1$. S_2 and S_3 believe that $\sigma_1 \sim [\hat{\sigma}_1, \overline{\sigma}_1]$ if $s^* = \underline{s}$ and that $\sigma_1 \sim [\underline{\sigma}_1, \hat{\sigma}_1)$ if $s^* = \overline{s}$; there are no out of equilibrium beliefs. By Lemma 1, 2 proposes $x^* = p(\overline{s}) + (c_2\omega_1 - \hat{\sigma}_1 - d)/2$ if $s^* = \overline{s}$ and $x' = p(\underline{s}) + (c_2 - \omega_1 - \overline{\sigma}_1)/2$ if $s^* = \underline{s}$. S_3 opposes if $s^* = \overline{s}$ and supports if $s^* = \underline{s}$.

Two inequalities must be satisfied for S_3's strategy to be sequentially rational. If $s^* = \overline{s}$, S_3 opposes when

$$-k + (1-r)\int_{\underline{\sigma}_1}^{\hat{\sigma}_1} \left(\sigma_1 \times \frac{1}{\hat{\sigma}_1 - \underline{\sigma}_1}\right) \mathrm{d}\sigma_1 > \int_{\underline{\sigma}_1}^{\hat{\sigma}_1} \left(\sigma_1 \times \frac{1}{\hat{\sigma}_1 - \underline{\sigma}_1}\right) \mathrm{d}\sigma_1,$$

or when $k < -r(\hat{\sigma}_1 + \underline{\sigma}_1)/2$. If $s^* = \underline{s}$, S_3 supports when

$$\int_{\hat{\sigma}_1}^{\overline{\sigma}_1} \left(\sigma_1 \times \frac{1}{\overline{\sigma}_1 - \hat{\sigma}_1}\right) \mathrm{d}\sigma_1 \geq -k + (1-r)\int_{\hat{\sigma}_1}^{\overline{\sigma}_1} \left(\sigma_1 \times \frac{1}{\overline{\sigma}_1 - \hat{\sigma}_1}\right) \mathrm{d}\sigma_1,$$

or when $k \geq -r(\overline{\sigma}_1 + \hat{\sigma})/2$. Therefore, S_3's strategy is sequentially rational when $-r(\hat{\sigma}_1 + \underline{\sigma}_1)/2 < k \leq -r(\overline{\sigma}_1 + \hat{\sigma})/2$. It remains to show that S_1's signaling strategy is incentive-compatible. Begin by identifying $\hat{\sigma}_1$, or the type that is indifferent over issuing the small threat (which leads to rejection and S_3's support) and the large threat (which leads to acceptance and S_3's opposition). This type satisfies $-a\overline{s} + x^* = -a\underline{s} + p(\underline{s}) - \hat{c}_1$, such that

$$\hat{\sigma}_1 = \omega_1 + d - c_2 - 2\left(p(\overline{s}) - p(\underline{s})\right) + 2a\left(\overline{s} - \underline{s}\right).$$

Next, to ensure that this is a plausible type, it must be the case that $\underline{\sigma}_1 < \hat{\sigma} < \overline{\sigma}_1$. This is the case when

$$2\left(\left(p(\overline{s}) - p(\underline{s})\right) - a\left(\overline{s} - \underline{s}\right)\right) + c_2 + \omega_1 + \underline{\sigma}_1 < d < 2\left(\left(p(\overline{s}) - p(\underline{s})\right) - a\left(\overline{s} - \underline{s}\right)\right) + c_2 + \omega_1 + \overline{\sigma}_1,$$

or, equivalently, when $\left(2M + c_2 + \omega_1 + \underline{\sigma}_1\right) < d < (2M + c_2 + \omega_1 + \overline{\sigma}_1)$.

For the equilibrium in which S_1 pools on the low threat, strategies and beliefs are as

follows. S_1 sets $s^* = \underline{s}$, accepts iff $\sigma_1 \geq p(\underline{s}) - \omega_1 - x$ on the equilibrium path and accepts iff $\sigma_1 \geq p(\overline{s}) - \omega_1 - d - x$ off the equilibrium path. On the equilibrium path, S_2 and S_3 retain their prior beliefs, and off the equilibrium path they believe $\sigma_1 \sim [\underline{\sigma}_1, \overline{\sigma}'_1]$, where $\overline{\sigma}'_1 < \overline{\sigma}_1$ such that only relatively low cost types would be tempted deviate to the large threat, more willing to tolerate the costs of mobilization and diplomatic opposition. By Lemma 1, S_2 proposes $x^* = p(\underline{s}) + (c_2 - \omega_1 - \overline{\sigma}_1)/2$ on the equilibrium path and $x' = p(\overline{s}) + (c_2 - \omega_1 - \overline{\sigma}_1 - d)/2$ off the equilibrium path. S_3 supports on the equilibrium path and opposes off the equilibrium path.

Two inequalities must be satisfied for S_3's strategy to be sequentially rational. On the equilibrium path, S_3 supports when

$$\int_{\underline{\sigma}_1}^{\overline{\sigma}_1} \left(\sigma_1 \times \frac{1}{\overline{\sigma}_1 - \underline{\sigma}_1} \right) \mathrm{d}\sigma_1 \geq -k + (1-r) \int_{\underline{\sigma}_1}^{\overline{\sigma}_1} \left(\sigma_1 \times \frac{1}{\overline{\sigma}_1 - \underline{\sigma}_1} \right) \mathrm{d}\sigma_1,$$

or when $k \geq -r(\overline{\sigma}_1 + \underline{\sigma}_1)/2$. Off the equilibrium path, S_3 opposes when

$$-k + (1-r) \int_{\underline{\sigma}_1}^{\overline{\sigma}'_1} \left(\sigma_1 \times \frac{1}{\overline{\sigma}'_1 - \underline{\sigma}_1} \right) \mathrm{d}\sigma_1 > \int_{\underline{\sigma}_1}^{\overline{\sigma}'_1} \left(\sigma_1 \times \frac{1}{\overline{c}'_1 - \underline{c}_1} \right) \mathrm{d}\sigma_1,$$

or when $k < -r(\overline{\sigma}'_1 + \underline{\sigma}_1)/2$. Therefore, S_3's strategy is sequentially rational when $-r(\overline{\sigma}'_1 + \underline{\sigma}_1)/2 < k \leq -r(\overline{\sigma}_1 + \underline{\sigma}_1)/2$. Finally, I show that the lowest-cost type of S_1 ($\sigma_1 = \underline{\sigma}_1$) issues the small threat rather than deviate to the large threat and induce the other players to believe that it is relatively low cost; if this type will not deviate to a larger threat, then no types will. This is a type that rejects S_2's offer in and out of equilibrium, so it sets $s^* = \underline{s}$ when $-a\underline{s} + p(\underline{s}) - \omega_1 - \underline{\sigma}_1 \geq -a\overline{s} + p(\overline{s}) - \omega_1 - \underline{\sigma}_1 - d$, or when $d \geq \left(p(\overline{s}) - p(\underline{s})\right) - a\left(\overline{s} - \underline{s}\right) \equiv M$. □

References

Aono, Toshihiko. 2010. "'It Is Not Easy for the United States to Carry the Whole Load': Anglo-American Relations During the Berlin Crisis, 1961-1962." *Diplomatic History* 34(2):325–356.

Arena, Philip. n.d. "Costly Signaling, Resolve, and Martial Effectiveness." Typescript, University of Essex.

Atkinson, Rick. 1993. *Crusade: The Untold Story of the Persian Gulf War*. New York: Houghton Mifflin.

Braumoeller, Bear F. 2012. *The Great Powers and the International System: Systemic Theory in Empirical Perspective*. New York: Cambridge University Press.

Brown, Jonathan N. 2014*a*. "Immovable Positions: Public Acknowledgment and Bargaining in Military Basing Negotiations." *Security Studies* 23(2):258–292.

Brown, Jonathan N. 2014*b*. "The sound of silence: Power, secrecy, and international dueinces in US military basing negotiations." *Conflict Management and Peace Science* 31(4):406–431.

Bush, George H.W. and Brent Scowcroft. 1998. *A World Transformed*. New York: Knopf.

Byman, Daniel and Matthew Waxman. 2002. *The Dynamics of Coercion*. Cambridge: Cambridge University Press.

Carson, Austin. 2016. "Facing Off and Saving Face: Covert Intervention and Escalation Management in the Korean War." *International Organization* 70(1):103–131.

Chapman, Terrence L. 2011. *Securing Approval: Domestic Politics and Multilateral Authorization for War*. Chicago, IL: University of Chicago Press.

Chapman, Terrence L. and Scott Wolford. 2010. "International Organizations, Strategy, and Crisis Bargaining." *Journal of Politics* 72(1):227–242.

Christensen, Thomas J. 2011. *Worse Than a Monolith: Alliance Politics and Problems of Coercive Diplomacy in Asia*. Princeton: Princeton University Press.

Clark, Christopher. 2012. *The Sleepwalkers: How Europe Went to War in 1914*. New York: HarperCollins.

Clark, Wesley K. 2001. *Waging Modern War*. New York: PublicAffairs.

Crall, Dennis A. and Thomas M. Martin. 2013. "Cool Hand Nuke: Lessons From the Quiet Diplomacy of the Cienfuegos Non-Crisis." *Foreign Policy Analysis* 9(2):189–201.

Croco, Sarah E. and Tze Kwang Teo. 2005. "Assessing the Dyadic Approach to Interstate Conflict Processes: A.K.A. "Dangerous" Dyad-Years." *Conflict Management and Peace Science* 22(1):5–18.

Dafoe, Allan. 2012. Resolve, Reputation, and War: Cultures of Honor and Leaders' Time-in-Office PhD thesis Universita of Calirfornia, Berkeley.

Dafoe, Allan, Jonathan Renshon and Paul Huth. 2014. "Reputation and Status as Motives for War." *Annual Review of Political Science* 17(371-393).

Fang, Songying, Jesse C. Johnson and Brett Ashley Leeds. 2014. "To Conede or Resist? The Restraining Effect of Military Alliances." *International Organization* 68(4):775–809.

Favretto, Katja. 2009. "Should Peacemakers Take Sides? Major Power Mediation, Coercion, and Bias." *American Political Science Review* 103(2):248–263.

Fearon, James D. 1994. "Domestic Political Audiences and the Escalation of International Disputes." *American Political Science Review* 88(3):577–92.

Fearon, James D. 1997. "Signaling Foreign Policy Interests: Tying Hands versus Sinking Costs." *Journal of Conflict Resolution* 41(1):68–90.

Freedman, Lawrence. 2000. *Kennedy's Wars: Berlin, Cuba, Laos and Vietnam*. New York: Oxford University Press.

Fudenberg, Drew and Jean Tirole. 1991. *Game Theory*. Cambridge: MIT Press.

Gaddis, John Lewis. 2005. *The Cold War: A New History*. New York: Penguin.

Hastings, Max. 2013. *Catastrophe 1914: Europe Goes to War*. New York: Knopf.

Henke, Marina E. 2016. "The Politics of Diplomacy: How the United States Builds Multilateral Military Coalitions." *International Studies Quarterly* 61(2):410–424.

Henriksen, Dag. 2007. *NATO's Gamble: Combining Diplomacy and Airpower in the Kosovo Crisis 1998-1999*. Annapolis, MD: Naval Institute Press.

Huth, Paul, D. Scott Bennett and Christopher Gelpi. 1992. "System Uncertainty, Risk Propensity, and International Conflict Among the Great Powers." *Journal of Conflict Resolution* 36(3):478–517.

Huth, Paul K. 1997. "Reputations and Deterrence." *Security Studies* 7(1):72–99.

Ikenberry, G. John. 2001. *After Victory: Institutions, Strategic Restraint, and the Rebuilding of Order After Major Wars*. Princeton: Princeton University Press.

Jervis, Robert. 1978. "Cooperation under the Security Dilemma." *World Politics* 30(2):167–214.

Johns, Leslie. 2007. "A Servant of Two Masters: Communication and the Selection of International Bureaucrats." *International Organization* 61(2):245–275.

Johnson, Jesse C. and Brett Ashley Leeds. 2011. "Defense Pacts: A Prescription for Peace?" *Foreign Policy Analysis* 7(1):45–65.

Kadera, Kelly M., Mark J.C. Crescenzi and Megan L. Shannon. 2003. "Democratic Survival, Peace, and War in the International System." *American Journal of Political Science* 47(2):234–247.

Kelley, Judith. 2005. "Strategic Non-cooperation as Soft Balancing: Why Iraq was not Just about Iraq." *International Politics* 42(2):153–173.

Kilcullen, David. 2006. "Counter-insurgency *Redux*." *Survival* 48(4):111–130.

Kreps, Sarah E. 2011. *Coalitions of Convenience: United States Military Interventions After the Cold War*. New York: Oxford University Press.

Kurizaki, Shuhei. 2007. "Efficient Secrecy: Public versus Private Threats in Crisis Diplomacy." *American Political Science Review* 101(3):543–558.

Kydd, Andrew H. 2005. *Trust and Mistrust in International Relations*. Princeton: Princeton University Press.

Kydd, Andrew H. and Roseanne W. McManus. 2017. "Threats and Assurances in Crisis Bargaining." *Journal of Conflict Resolution* 61(2):325–348.

Lai, Brian. 2004. "The Effects of DIfferent Types of Military Mobilization on the Outcome of International Crises." *Journal of Conflict Resolution* 48(2):211–229.

Lake, David A. 2010/11. "Two Cheers for Bargaining Theory: Assessing Rationalist Explanations of the Iraq War." *International Security* 35(3):7–52.

Lane, Ann. 2003. "Third world neutralism and British Cold War strategy, 1960-62." *Diplomacy and Statecraft* 14(3):151–174.

Leeds, Brett Ashley. 2003. "Do Alliances Deter Aggression? The Influence of Military Alliances on the Initiation of Militarized Interstate Disputes." *American Journal of Political Science* 47(3):427–439.

Lindsey, David. 2015. "Military Strategy, Private Information, and War." *International Studies Quarterly* 59(4):629–640.

Mansfield, Edward D. 1994. *Power, Trade, and War*. Princeton: Princeton University Press.

Mearsheimer, John J. 2001. *The Tragedy of Great Power Politics*. New York: W.W. Norton and Company.

Mitchell, Sara McLaughlin. 2002. "A Kantian System? Democracy and Third-Party Conflict Resolution." *American Journal of Political Science* 46(4):749–759.

Monteiro, Nuno P. 2011/12. "Unrest Assured: Why Unipolarity Is Not Peaceful." *International Security* 36(3):9–40.

Paine, S.C.M. 2003. *The Sino-Japanese War of 1894-1895: Perceptions, Power, and Primacy*. Cambridge: Cambridge University Press.

Paine, S.C.M. 2017. *The Japanese Empire: Grand Strategy from the Meiji Restoration to the Pacific War*. Cambridge, UK: Cambridge University Press.

Papayoanou, Paul A. 1997. "Intra-Alliance Bargaining and U.S. Bosnia Policy." *Journal of Conflict Resolution* 41(1):91–116.

Paul, T.V. 2005. "Soft Balancing in the Age of U.S. Primacy." *International Security* 30(1):46–71.

Poast, Paul. 2010. "(Mis)Using Dyadic Data to Analyze Multilateral Events." *Political Analysis* 18(4):403–425.

Powell, Robert. 1999. *In the Shadow of Power*. Princeton: Princeton University Press.

Richardson, Louise. 2000. A Force for Good in the World? Britain's Role in the Kosovo Crisis. In *Alliance Politics, Kosovo, and NATO's War: Allied Force or Forced Allies?*, ed. Pierre Martin and Mark R. Brawley. New York: Palgrave.

Riker, William H. 1962. *The Theory of Political Coalitions*. New Haven: Yale University Press.

Russett, Bruce M. 1963. "The Calculus of Deterrence." *Journal of Conflict Resolution* 7(2):97–109.

Schultz, Kenneth A. 1998. "Domestic Opposition and Signaling in International Crises." *American Political Science Review* 92(4):829–844.

Schweller, Randall L. 1994. "Bandwagoning for Profit: Bringing the Revisionist State Back In." *International Security* 19(1):72–107.

Sechser, Todd S. 2010. "Goliath's Curse: Coercive Threats and Asymmetric Power." *International Organization* 64(4):627–660.

Shea, Patrick E. 2014. "Financing Victory: Credit, Democracy, and War Outcomes." *Journal of Conflict Resolution* 58(5):771–795.

Slantchev, Branislav. 2005. "Military Coercion in Interstate Crises." *American Political Science Review* 99(4):533–547.

Slantchev, Branislav. 2010. "Feigning Weakness." *International Organization* 64(3):357–388.

Slantchev, Branislav. 2011. *Military Threats: The Costs of Coercion and the Price of Peace*. Cambridge: Cambridge University Press.

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.

Stevenson, David. 1997. "Militarization and Diplomacy in Europe before 1914." *International Security* 22(1):125–161.

Thompson, Alexander. 2006. "Coercion Through IOs: The Security Council and the Logic of Information Transmission." *International Organization* 60(1):1–34.

Trager, Robert F. 2015. "Diplomatic Signaling among Multiple States." *Journal of Politics* 77(3):635–647.

Treisman, Daniel. 2004. "Rational Appeasement." *International Organization* 58(2):345–373.

Tyler, Patrick. 2009. *A World of Trouble: The White House and the Middle East—from the Cold War to the War on Terror*. New York: Farra, Straus and Giroux.

Voeten, Erik. 2005. "The Political Origins of the UN Security Council's Ability to Legitimize the Use of Force." *International Organization* 59:527–557.

Waltz, Kenneth. 1959. *Man, The State, and War: A Theoretical Analysis*. New York: Columbia University Press.

Weisiger, Alex and Keren Yarhi-Milo. 2015. "Revisiting Reputation: How Past Actions Matter in International Politics." *International Organization* 69(02):473–495.

Wendt, Alexander. 1999. *Social Theory of International Politics*. Cambridge: Cambridge University Press.

Werner, Suzanne. 2000. "Deterring Intervention: The Stakes of War and Third-Party Involvement." *American Journal of Political Science* 44(4):720–732.

Westad, Odd Arne. 2012. *Restless Empire: China and the World since 1750*. New York: Basic Books.

Wolford, Scott. 2007. "The Turnover Trap: New Leaders, Reputation, and International Conflict." *American Journal of Political Science* 51(4):772–788.

Wolford, Scott. 2014. "Showing Restraint, Signaling Resolve: Coalitions, Cooperation, and Crisis Bargaining." *American Journal of Political Science* 58(1):144–156.

Wolford, Scott. 2015. *The Politics of Military Coalitions*. New York: Cambridge University Press.

Wolford, Scott and Emily Hencken Ritter. 2016. "National Leaders, Political Security, and the Formation of Military Coalitions." *International Studies Quarterly* 60(3):540–551.

Wu, Cathy Xuanxuan and Scott Wolford. n.d. "Leaders, States, and Reputations." *Journal of Conflict Resolution* forthcoming.

Yarhi-Milo, Keren. 2013. "Tying Hands Behind Closed Doors: The Logic and Practice of Secret Reassurance." *Security Studies* 22(3):405–435.

Yuen, Amy. 2009. "Target Concessions in the Shadow of Intervention." *Journal of Conflict Resolution* 53(5):727–744.