Great Powers and Third-Party Diplomacy: The Logic of Signaling to Multiple Audiences∗

Scott Wolford†
swagenutexas.edu

September 22, 2011

Abstract

Great powers have incentives to demonstrate resolve to opponents in interstate crises, but their considerable military power also provokes fears of exploitation in outsiders, creating a competing incentive to signal restraint. I explore this dilemma in a model of crisis bargaining in which a third party can reward one belligerent’s restraint by imposing diplomatic costs on the other if the crisis escalates to war, which has two effects on the probability of war. First, winning outside support can impose sufficient costs on one’s opponent to discourage it from risking war by reducing the role of uncertainty. Second, the promise of imposing diplomatic costs on an opponent may encourage states to hide their resolve, refusing to reveal private information and maintaining a risk of war even when credible signaling is possible. Thus, while strategic restraint can reassure fearful observers, its exercise can also lead to otherwise avoidable wars.

∗Draft version. Comments welcome, but please do not cite without permission. Thanks to Phil Arena, David Bearce, Terry Chapman, Joe Jupille, Ashley Leeds, Amy Liu, Tig Reynolds, Toby Rider, and Emily Ritter for helpful comments and suggestions. A special thanks is reserved for Moonhawk Kim, who listened to more half-baked and ill-formed ideas during the gestation of this project than anyone should reasonably be expected to tolerate. A previous version of this paper was presented at the 2011 Annual Meeting of the American Political Science Association, September 2-5, Seattle, WA.

†Assistant Professor of Government, University of Texas at Austin
Demonstrating resolve can help states win concessions and avert war in international disputes, and sending credible signals of a willingness to fight is often viewed as essential for successful crisis management (Fearon 1997, Jervis 1970, Schelling 1966, Slantchev 2005, Snyder and Diesing 1977). Yet some states have a competing interest in signaling restraint, or a reluctance to use force, to third parties that fear exploitation at the hands of considerable military strength (Ikenberry 2001, Thompson 2006, Voeten 2005). These fearful outsiders can impose substantial diplomatic costs on great powers, or engage in “soft balancing” (Pape 2005), for failures to exercise restraint. As a result, powerful states face a dilemma between signaling resolve, which can secure a better crisis outcome at the risk of provoking third party diplomatic opposition, and restraint, which may prevent diplomatic opposition at the cost of a less favorable resolution to the crisis. How do great powers manage this tradeoff, and what does outside diplomatic support or opposition imply for their crisis bargaining behavior?

American decisionmakers faced just such a dilemma during the Berlin Crisis of 1961-1962, as they debated the merits of a strategy that would send ground forces across the intra-German border as a signal of resolve (Freedman 2000, pp. 62-3, 66-7) and an alternative built around lower levels of escalation and public signals of a willingness to negotiate (Aono 2010, p. 326, 334). Concerned with the reaction of both nervous allies and non-aligned states to further escalation, the US opted for the latter, hoping to “outtalk the Russians rather than outfight them” (Freedman 2000, pp. 65). Ambassador to the Soviet Union Llewellyn Thompson “believed that if the West failed to obtain the support of [world] public opinion, Khrushchev would exploit this and take far bolder actions” (Aono 2010, p. 333). For his part, Ambassador to the UN Adlai Stevenson saw “American restraint” as a way to “put the Soviet Union on the defensive ‘in the court of world opinion’ ” (Freedman 2000, pp. 80-1). The United States thus considered outside
diplomatic support as a way to impose costs on the Soviet Union, which might discourage risky behavior. However, this decision is at odds with the notion that the US’s goal should have signaling resolve over the status of Berlin. Can the promise of outside diplomatic support discourage escalation, as American decisionmakers hoped? On the other hand, might there be conditions under which efforts to reassure outsiders can increase the probability of war?

To answer these questions, I begin with a model of asymmetric information in which a state with private information over its resolve can take costly steps to escalate a crisis. Where similar models consider only two disputants (e.g., Fearon 1994, 1997, Slantchev 2005), I introduce a third party who is also uncertain over the first state’s resolve. This observer, fearful of possible exploitation in the future, prefers to support the first only if it is restrained, or generally reluctant to use force (see Ikenberry 2001), and it can impose diplomatic costs on the side it opposes. This captures a fundamental tension for the first state: convincing opponents of one’s willingness to fight can discourage those opponents from taking risky actions, but it can also incur diplomatic costs from outsiders who prefer to support military restraint. Great powers, then, face competing incentives to signal both resolve and restraint in crisis bargaining. During the Cold War, for example, nonaligned states sought to “tilt without tipping” (Gaddis 2005, p. 122), punishing one superpower’s lack of restraint by supporting the other, in the service of “resisting superpower hegemony” (p. 123-4). As a result, during the Berlin Crisis, third parties “had to be reassured that the United States would not be rash, [while] the Soviet Union had to be persuaded that it just might be” (Freedman 2000, p. 93).

In equilibrium, the promise of winning third-party diplomatic support, which imposes costs on one’s opponent, has two competing effects on the probability of war. First, as anticipated by American diplomats during the Berlin Crisis, diplomatic costs imposed on an opponent can
increase the costs of war, lessening the impact of uncertainty on its strategy and discouraging it from making aggressive demands that risk war. Second, and more surprisingly, the promise of winning outside support can increase the probability of war when it induces an otherwise resolute state to signal restraint rather than resolve. Specifically, it may choose not to send a costly, credible signal of its willingness to fight—which would avert war—because war with the support of outsiders is preferable to peace without it. This maintains uncertainty that a resolute state would otherwise pay to eliminate, preserving a possibility of war where, in the absence of third parties willing and able to impose diplomatic costs, it would otherwise not occur.

This signaling dilemma contributes to our understanding of international conflict, great power politics, and the role of third parties in three ways. First, it provides an account of how the diplomatic behavior of third parties, specifically soft balancing aimed to counter future threats, can affect both the signaling behavior of great powers and the probability of war. Most such diplomatic moves are excluded from models of crisis bargaining, where third parties take a more direct military role in the crisis if they take a role at all (Werner 2000, Yuen 2009, Zagare and Kilgour 2000). It also explains why, despite their considerable military power, great powers can rationally show concern for the reaction of world opinion to their crisis behavior, providing microfoundations for Fearon’s (1997) conjecture that the failure of some signals to convey resolve is less important “than the effect achieved on various other audiences” (p. 84).¹ Second, it shows that some of the very actions great powers take to reassure outsiders of their forbearance in the use of force, or “strategic restraint” (Ikenberry 2001), can lead them to adopt strategies that preserve uncertainty and, perversely, provoke their opponents into risking war.

¹This is also consistent with Russett’s (1963) contention that “pressures from the ‘attacker’s’ own allies and world opinion” (p. 102, f.n. 8) affected escalatory behavior in Cold War-era crises over the Bay of Pigs and the Suez Canal.
Thus, as great powers design policies to reassure some states, they run the risk of signaling information to others that leads to other undesirable outcomes, suggesting that the interaction of soft balancing and strategic restraint can have a tragic consequence—war—that each strategy in isolation is designed to avert. Finally, the model links the anticipated response of third parties to the incentive to misrepresent one’s resolve, without which private information alone is insufficient to cause war (Fearon 1995). While it provides a rationale for assuming signaling difficulties in models of crisis bargaining, it also suggests that overcoming them may involve not only political features of the state in question (Fearon 1994, Schultz 1998) but also the politics of order in the international system.

Great Powers, Soft Balancing, and Restraint

In models of signaling in dyadic crises, peaceful outcomes often rest on a resolute type’s ability to distinguish itself from an irresolute type with a costly signal, given the irresolute type’s incentive to bluff (see Jervis 1970). They may tie their hands or sink costs (Fearon 1994, 1997), make military moves that perform both functions (Slantchev 2005), or use brinkmanship to manipulate the autonomous risk of war (Powell 1990, Schelling 1966), but regardless of the policy instrument the logic of crisis signaling remains the same: separation. A resolute type must pay some price that an irresolute type is unwilling to pay, influencing its opponent’s estimate of its resolve such that the opponent offers acceptable terms; otherwise, an opponent may run a risk of war if it is optimistic enough that the state in question truly is irresolute.

However, signals sent in crises can transmit information not only to one’s opponent but also to third parties. Russett (1963, p. 102, f.n. 8) and Fearon (1997, p. 84) discuss the possibility that
the dynamics of threat and signaling can be influenced by states not involved in the immediate crisis, but no work as yet develops microfoundations to explain how and under what conditions we might see dyadic incentives to signal resolve confounded by competing incentives to send other information to third parties. To be sure, states with direct distributive interests in a crisis or formal commitments may intervene—or threaten to—militarily (e.g. Werner 2000, Zagare and Kilgour 2000). Yet many observers are more concerned about what a great power’s behavior says about its general willingness to use force and what it means for their future safety, and they may adjust their diplomatic positions accordingly. To explore this possibility, I focus on a specific type of third party: observers who fear exploitation at the hands of unchecked military power and who seek to reward the exercise of restraint with their diplomatic behavior (Ikenberry 2001, Thompson 2006, Voeten 2005).  

If states balance against not only power but also threat (Walt 1987), then great powers may be especially attuned to the responses of outsiders to their signals, especially when signals reveal something about their underlying willingness to use military force. Threatened states, even minor powers, can engage in a variety of behaviors, from realignment to the imposition of diplomatic costs via “soft balancing” (Pape 2005), designed to punish great powers for a lack of restraint in the use of force. Traditional conceptions of responding to threats range from altering alliance commitments to intervening (or not) in ongoing wars (Christensen and Snyder 1990, Powell 1999, Schweller 1994, Wagner 2007, Walt 1987), but the full spectrum of options available to third parties includes a variety of nonmilitary activity short of direct intervention that

---

2 There may, for instance, be third parties to which a great power wishes to send the same signal of resolve that it sends to its opponent—say, West Germany and France in the Berlin Crisis, who most wanted to see the US prove its willingness to fight (Aono 2010, Freedman 2000)—but this merely reinforces the desire to signal resolve. While it may encourage irresolute types to bluff about their resolve, it does not create the kind of dilemma in which I am interested here.
can nonetheless impose serious costs on the great powers.\footnote{See Trager (2010) for examples of nonmilitary actions that the targets of threats may take in a dyadic context.}

The literature on soft balancing (e.g. Pape 2005, Paul 2005) identifies several indirect strategies that third parties can adopt to affect a great power’s payoffs from using force. Pape (2005) discusses “nonmilitary tools [used] to delay, frustrate, and undermine” (p. 10) the foreign policies of aggressive or unrestrained great powers, including “territorial denial, entangling diplomacy, economic strengthening, and signaling of resolve to participate in a balancing coalition” (p. 36). To this list Paul (2005, p. 69) adds obstruction in international institutions, including the UN, where threatened powers can deny support for the current conflict or deny approval for multilateral actions, neither of which represents direct military intervention but can be considered “low-cost diplomatic strategies to…constrain” (p. 58) the exercise of power. To the extent that several observers can coordinate their actions during the crisis in response to a perceived lack of restraint (Thompson 2006, Voeten 2005), great powers must weigh consequences of signaling resolve beyond the direct need to influence their opponents’ behavior.

For example, the Cold War saw both the United States and the Soviet Union attempt to curry the favor of largely unaligned and “anti-bloc” Third World countries, especially during crises when their diplomatic support—or opposition—might prove decisive in changing an opponent’s strategic calculations (Lane 2003, p. 152). American diplomats sought “to champion ‘neutralism’ in order to prevent the Afro-Asian states from orienting towards the Soviet sphere” (ibid., p. 151). Kennedy saw in the Non-Aligned Movement “a possibility to achieve some additional censure of the Soviet Union” (ibid., p. 166, f.n. 62) in the Berlin Crisis, hoping that winning its support would entail sufficient short-term diplomatic costs for the Soviets that they might back down. The president was also concerned that Khrushchev, for his part, might try to “de-
velop UN sentiment as a source of pressure” on the United States (quoted in Freedman 2000, p. 120). Finaly, a great power’s own allies may engage in soft balancing, as the United States learned before the 2003 Iraq War when Turkey refused to allow the use of its territory to open a northern front into Iraq (Gordon and Trainor 2006, pp. 42,115) and when France and Germany guided NATO’s refusal to commit to the defense Turkey during the invasion (Paul 2005, p. 69), both due to a perceived lack of American restraint.

Managing power asymmetries between the great powers and those that fear their military strength is a fundamental problem of world order, and a variety of strategies exist for addressing it, including domination, the diffusion of power, or self-binding through international institutions (Ikenberry 2001), where securing multilateral approval can signal benign intent (Thompson 2006) or policy-relevant information (Chapman 2011, Voeten 2005) to observers. Ikenberry (2001) terms the sending of costly signals of one’s forbearance in the use of force “strategic restraint,” and it is this very incentive to reassure outsiders that comes into tension with the desire to signal resolve. However, to date the literature on strategic restraint has not considered its implications for its use in bargaining and war. On the other hand, while research on soft balancing sheds light on why great powers may care about ensuring outside support in otherwise bilateral crises, it does not explore the consequences of pursuing such support for the dynamics of signaling and the outcomes of crisis bargaining. Soft balancing may be designed to contain

---

4Pressure from the rank of middle powers can indeed be quite severe: the Arab oil embargo, launched in response to American support for Israel in the 1973 Yom Kippur war, damaged Western economies for nearly a decade and remains among the most prominent examples of minor powers imposing significant economic costs on more powerful states (Gaddis 2005, Tyler 2009). Tragically, President Richard Nixon was aware of this threat as war approached between Israel and its neighbors in 1973, cautioning, “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 is terribly important” (quoted in Tyler 2009, p. 211).

5The need to reassure allies also appeared in the crisis over Quemoy and Matsu in 1858, where “the U.S. government felt the need to be explicitly firm to deter the Communist Chinese…and vague enough to calm the timid British” (Snyder and Diesing 1977, p. 223).
great powers, and strategic restraint is intended to reduce international opposition, but absent an understanding of how the two interact in the strategic setting of crisis bargaining, the net effect of the two on the probability of great power war is not obvious. Specifically, when will great powers signal resolve, when might they signal restraint, and what does this mean for the chances of war? In the next section, I specify a model designed to explore these possibilities.

The Model

The model features two belligerents, states 1 and 2, who dispute the division of some benefits worth 1 to each side. 1 begins the game uncertain over 2’s resolve, or the costs of escalating the crisis and fighting a war, and should they fail to divide the benefits peacefully, the disputants fight a costly war. The crisis unfolds in two stages. First, in the diplomatic stage, 1 and 2 choose levels of costly escalation, such as mobilizing troops or increasing military readiness, that can improve their chances of prevailing in war. 2’s escalatory decision in the diplomatic stage may also allow 1 to update its beliefs over 2’s resolve heading into the bargaining stage, in which 1 makes a take-it-or-leave-it demand to divide the benefits and where 2’s rejection ends the crisis in war. The model differs from two-player treatments, however, by introducing a third party, $O$, that is also uncertain over state 2’s resolve. $O$ can be a single state or a coalition of several, and it may be drawn from 2’s allies, enemies, or nonaligned states with reason to fear 2’s military power—what matters is that its diplomacy can alter a belligerent’s costs for war. $O$ prefers to support 2 in the crisis only if it is truly restrained in the use of force, because a state reluctant to use military force poses a smaller threat to its interests than a resolute one. Given their reluctance to use force, I also refer to states with higher costs for war, or irresolute states
in standard terminology, as “restrained.”

After observing belligerents’ escalatory decisions, $O$ chooses in the final move of the diplomatic stage which side to support, thereby imposing diplomatic costs on the side it opposes if the crisis escalates to war. The costs $O$ imposes can range from economic to military to political, from embargoes to changed alignments to obstruction in international institutions and the denial of territorial privileges. Thus, the side $O$ supports enters the bargaining stage with the knowledge that its opponent will pay these additional costs for the use of force. This ensures that $2$ has competing incentives for the revelation of its private information. Escalating improves its military position and may convince $1$ of its resolve, yet this comes at the risk of failing to win $O$’s diplomatic support. On the other hand, choosing not to escalate can win $O$’s support, yet $2$ foregoes the military and signaling benefits of escalation.

To specify the game formally, let Nature draw a cost term $c_i > 0$, or Bayesian type, for each state $i = 1, 2$, which represents its costs for military action, both for escalation during the crisis and for a possible war. Using a common cost term across both stages of the crisis ensures that it represents how each side evaluates the costs of coercive action in general, independently of the current crisis. Thus, low values of $c_i$ indicate a resolute state and higher values a restrained state. A resolute state, for example, views itself as relatively unbound by norms of peaceful conflict resolution and assesses lower costs of using the military instrument in a given crisis; a restrained
state, on the other hand, finds the use of force as *prima facie* more costly, either because of the acceptance of peaceful norms or greater concern for its reputation for restraint. While Nature reveals state 1’s type, $c_1$, to the other players, only 2 knows its own type. Specifically, 2 can be either resolute, $c_2 = c_2$, or restrained, $c_2 = \bar{c}_2$, where $c_2 < \bar{c}_2$. 1 and $O$ know only the distribution from which 2’s type is drawn: restrained with probability $\phi$ and resolute with probability $1 - \phi$. This ensures that 1 is unsure of which demands 2 will accept and reject, while $O$ is unsure of whether it is more profitable to support state 1 or 2 diplomatically.

The sequence of moves, described in Figure 1, is designed to capture 2’s incentives to reveal or misrepresent its private information through escalatory behavior that can also affect the choices of the other players. 2 makes the first move, choosing a level of escalation, $e_2 \geq 0$, that increases its military prospects at some cost. Escalation may involve the mobilization or positioning of troops—e.g., the proposed deployment into East Germany—increases in readiness or alert status, loosening rules of engagement, and even the delegation of authority over nuclear weapons to commanders in the field (see also Slantchev 2005). 2’s sensitivity to the costs of escalation is determined by its resolve, such that, for a given level of escalation, it pays $c_2 e_2$, such that escalation is costlier as $c_2$ increases. Next, using whatever information may have been revealed by 2’s decision, 1 chooses its own level of escalation $e_1 \geq 0$, for which it pays $c_1 e_1$.

In the final move of the diplomatic stage, $O$ chooses to support 1 or 2 for the bargaining stage, where supporting one side entails opposing the other. The side $O$ opposes pays additional diplomatic costs $c_D > 0$ in the event that the crisis escalates to war. $O$’s diplomatic sup-

---

6Allowing 2 the first move at the escalation stage and 1 the first move at the bargaining stage is similar to the sequence used in similar games of cooperation under mistrust (Kydd 2005, Schultz 2005), and it captures, with a minimum of machinery, the essentials of 2’s incentives for signaling in front of two audiences.

7While it is possible to give $O$ another option or neutrality, imposing costs on neither side, doing so would merely add to the number of equilibria without adding any additional insight into the research question, which turns on the effects of winning the diplomatic support of third parties.
port or opposition can involve decisions to grant or revoke economic and logistical cooperation, support in international institutions, overflight or transit privileges, as well as intelligence and information sharing. Each decision affects the exercise of power but fails to meet the threshold of direct intervention or traditional balancing (Pape 2005). Thus, I restrict attention to those actions that belligerents can take advantage of in the short term as they enter the bargaining stage, which introduces the possibility that a resolute 2 might use O’s uncertainty strategically to win diplomatic support, when O would not support it under complete information.

In the bargaining stage, 1 proposes a division of the benefits, \((x, 1-x)\), in which it receives \(x\) and 2 receives \(1-x\). If 2 accepts, the deal is implemented with each side receiving the proposed share of benefits, less any costs of escalation incurred in the diplomatic stage, or

\[
  u_1(x) = x - c_1 e_1 \quad \text{and} \quad u_2(x) = 1 - x - c_2 e_2.
\]

If the crisis ends peacefully, then the side that O has opposed does not pay the diplomatic costs of O’s opposition, since the crisis does not escalate to war. However, as shown next, diplomatic costs can affect the credibility of a state’s threat to resort to war, which helps determine which proposals are made and accepted in the bargaining stage.

War is a costly lottery in which each side pays a cost regardless of outcome and the victor receives all the benefits. The probability that a state wins is a function of each side’s military capabilities prior to the crisis, \(m_i\), and escalation levels, \(e_i\). For example, 2 wins a war with probability \((m_2 + e_2)/(m_1 + e_1 + m_2 + e_2)\), pays costs \(c_2\) for fighting and, if O supported the other
side, diplomatic costs $c_D$. Therefore, its payoff for the game if the crisis ends in war is

$$EU_2 = \frac{m_2 + e_2}{m_1 + e_1 + m_2 + e_2} \times 1 - c_2 - c_2e_2 - O_2c_D,$$

(1)

where \(c_2e_2\) captures the escalatory costs paid in the diplomatic stage, and \(O_2\) is an indicator such that \(O_2 = 1\) if \(O\) opposes \(O_2\) in the diplomatic stage and \(O_2 = 0\) if \(O\) supports it. State 1’s payoffs for a crisis ending in war are defined similarly by substituting \(i = 1\) for \(i = 2\) and vice versa into Equation (1).

Finally, \(O\) receives its best payoff when it supports a restrained \(O_2\). Specifically, it receives 1 if it supports \(O_2\) when \(O_2\) is restrained, or \(c_2 = c_2\), and 0 if \(O_2\) is resolute, or \(c_2 = c_2\). Likewise, if it supports 1, it receives 1 if \(O_2\) is resolute and 0 if \(O_2\) is restrained. Note that \(O\) has no distributive interest in the outcome of the crisis, caring only that it has the opportunity to impose diplomatic costs on the side it chooses to oppose. However, \(O\) may be uncertain over \(O_2\)'s type when it makes its decision, and, unless \(O_2\) reveals its type through escalatory behavior, \(O\)'s uncertainty is not resolved until the game ends. This allows for \(O\)'s diplomatic decisions to occur under conditions of uncertainty, such that it may choose to support \(O_2\) in the present crisis when, against expectations, it turns out to be resolute. While \(O\) may regret a decision \textit{ex post}—and may therefore wish to switch its support later—I do not model such future choices, because even short-term diplomatic support can have an effect on the belligerents’ ability to exercise military power.\textsuperscript{8}

I also place two technical restrictions on parameters in order to simplify the analysis. First, since the results do not depend on particulars of the pre-crisis military balance, I assume that \(m_1 = m_2 > 0\), or that the primary belligerents begin with equal and nonzero military capabili-

\textsuperscript{8}In any event, should \(O\) change its mind in the future, it would be reasonable to discount such payoffs, which further increases the short term effects of its initial support decision.
ties, denoted $m_i$. This ensures choosing not to escalate, $e_i = 0$, leaves a state with some chance of prevailing in a war. Second, I assume that neither side is so strong \textit{ex ante} as to make escalation a pointless expense, or $m_i \leq \hat{m}_i$, where $\hat{m}_i$ is defined in the appendix. Technically, this ensures that optimal escalation levels are nonzero, $e_i > 0$, or that choosing not to escalate foregoes a clear military benefit, as was the case in the Berlin Crisis.

This model both shares some features with others and builds on them in notable ways. First, it is similar to Slantchev’s (2010) and Trager’s (2010) models of bilateral crisis bargaining in that signaling resolve may have undesirable consequences, though it locates these consequences in the strategic behavior of a third party as opposed to the choice set of the receiver. Next, three-player models of deterrence and intervention often involve asymmetric information (e.g. Morrow 1994, Quackenbush 2006, Yuen 2009, Zagare and Kilgour 2000), yet states do not face the challenge of wishing to signal different information to multiple audiences, as 2 does here. Finally, Chapman and Wolford (2010) analyze a model in which an uninformed state can seek the support of a biased international organization able to alter its costs for war. However, the organization can affect the war payoffs of only the side that appeals to it, which is also the uninformed side. In contrast to previous work, then, the present model considers multiple uncertain audiences with divergent incentives to respond to new information, and it allows a third party to affect the costs of war for \textit{either} side in the crisis.

\textbf{Analysis}

How might the diplomatic behavior of third parties affect the crisis behavior of great powers? In this section, I discuss three equilibria that bear on this question. I first isolate the conditions
under which a credible signal is available to a resolute type, then describe when and why a resolute type might refuse to send it, mimicking a restrained type in order to win outside diplomatic support. I then consider the effect of this support on the probability of war, showing that, while diplomatic support can induce caution on the part of one's opponent by raising its costs for war, the promise of winning O's support can encourage a resolute state to knowingly tolerate war when, in the absence of a third party, it would not do so.

Given the game's information structure, I solve for Perfect Bayesian Equilibria (PBE), which requires that strategies be sequentially rational and that beliefs are updated according to Bayes' Rule wherever possible. Signaling games can exhibit multiple equilibria, especially when players are allowed to randomize their actions, so I discuss only pure strategy equilibria in order to isolate the dynamics of the most substantive interest. Even in pure strategies there can be many equilibria, but since I am interested in the existence of particular equilibria, I focus on only three that are substantively plausible and nonoverlapping. First, I discuss a separating equilibrium in which resolute and irresolute types of 2 take different actions, revealing their types to both 1 and O, followed by two pooling equilibria in which both types take the same action, ensuring that the other players retain their uncertainty.

### Diplomacy and Bargaining

Before discussing equilibria in detail, it is worth highlighting the basic strategic tensions that link the diplomatic and bargaining stages of the crisis. Begin with the bargaining stage. If 2

---

9 An equilibrium consists of (a) 2's escalation level and response rule for 1's proposal, (b) 1's escalation level and equilibrium proposal, (c) O's choice of which side to support, and (d) posterior beliefs for 1 and O following 2's moves, potentially reflecting new information over 2's type.

10 Nonoverlapping equilibria do not exist for the same parameter values, which ensures that when one is possible, the other is not, and vice versa.
has revealed its type, which occurs if the resolute type escalates (choosing $e_2 > 0$) while the restrained does not ($e_2 = 0$), the costliness of war ensures that the crisis will be resolved peacefully; as 1 knows exactly what it can demand of 2 without provoking a war. However, if 1 remains uncertain over 2’s type, as it would if types were to pool on the same action in the diplomatic stage (choosing identical values of $e_2$), it faces a risk-return tradeoff in which it may trade a risk of war for larger concessions (Powell 1999). When 1 is uncertain but sufficiently optimistic over 2’s type, it makes a proposal that only the restrained type will accept, risking war if 2 turns out to be the resolute type, who rejects. Finally, should the crisis escalate to war, the side that drew $O$’s opposition pays additional diplomatic costs $c_D$.

Anticipating behavior in the bargaining stage, 2 chooses its actions in the diplomatic stage in order to influence the other players’ beliefs. In the typical signaling story where there is no added value to signaling restraint, the challenge for a resolute type is to reveal itself by sending a signal so costly—here, by escalating sufficiently—that the restrained type prefers not to send it. This resolves 1’s uncertainty, ensuring that it will find it unprofitable to risk war when setting its demands. However, the promise of diplomatic support and the imposition of additional costs on 1 creates a signaling dilemma for state 2. Revealing itself as resolute guarantees that $O$ will oppose it, forcing 2 to pay additional diplomatic costs, leading under some conditions to the resolute type preferring to hide its type, pooling with the restrained type on no escalation in order to win $O$’s favor. I show below that this occurs when 2 enters the crisis with an *ex ante* favorable reputation for restraint and $O$’s opposition is sufficiently costly.

---

11I show in the Appendix that 1 will never make a demand that both types reject, and that it will make only one of two unique proposals designed to increase gains from settlements as much as possible without (1) provoking needless rejection and (2) demanding less than what would be agreed to by the other side.
Separating Equilibrium

In this section, I identify the conditions conducive to peaceful settlement via costly signaling by finding a separating equilibrium in which each type of 2 takes a unique action that reveals its type. Establishing the existence of a separating equilibrium requires identifying a signal that the resolute type can send that will be too costly for a restrained type to mimic, revealing 2’s type and convincing 1 not to risk war. Proposition 1 states that such an equilibrium exists when the restrained type has sufficiently high costs for war and the diplomatic costs of O’s opposition are sufficiently low.

Proposition 1 (Separating Equilibrium). When \( \bar{c}_2 \geq c_1 + c_\ell \) and \( c_D \leq \bar{c}_D \), there exists a PBE in which 2 escalates if resolute and does not escalate if restrained, O supports 2 iff it does not escalate, and the crisis is resolved peacefully. See appendix for proof.

In the separating equilibrium, which occurs on the left side of Figure 2, the resolute type escalates, setting \( e^*_2 > 0 \) (defined in the appendix), while the restrained type does not, choosing \( e^*_2 = 0 \), which allows both 1 and O to update their beliefs over 2’s type. 1 chooses its own level of escalation in response, and O supports 2 if it does not escalate and supports 1 if 2 does escalate. Finally, as expected under conditions of complete information, the crisis is resolved peacefully in the bargaining stage, where 1 demands only so much as 2 is willing to concede short of war, given its revealed type.

The existence of a separating equilibrium here depends on two factors. First, the restrained type must find escalation so costly, i.e. \( \bar{c}_2 \geq c_1 + c_\ell \), that it will not find it profitable to pay these costs in order to mimic the resolute type. The restrained type thus foregoes the military benefits of escalation, and each type of 2 is content to reveal its type by its action and see the
Figure 2: The Possibility of War in Equilibrium when $\phi \geq 1/2$ and $\bar{c}_2 \geq c_1 + c_2$

The Possibility of War in Equilibrium

Crisis resolved peacefully. The second condition, $c_D \leq \bar{c}_D$, requires that the diplomatic costs of opposition not be so high that the resolute type finds them an unacceptable price for revealing its type. In other words, the resolute type knows that revealing its private information will cost it diplomatically, but when those who would impose these costs are either too weak or too poorly coordinated to affect its well-being, the resolute type is not deterred from sending a revelatory signal, improving its military prospects against 1, and striking a peaceful bargain.

Two other points are of note. First, the condition $c_D \leq \bar{c}_D$ also applies to a world in which third party diplomatic costs are nonexistent or irrelevant, i.e. $c_D = 0$. If third party diplomacy has no effect on the course of disputes between the great powers, then as long as the restrained type's costs for war are so high that it will not mimic the resolute type, war will not occur when the resolute 2 can credibly signal its resolve. To the extent that war occurs in the equilibria
discussed below, the conditions supporting them can be said to cause war, given a value of $c_2$ that supports the existence of a credible signal. Finally, the uninformed players’ prior beliefs $\phi$ over 2’s type, or its \textit{ex ante} reputation for restraint, play no role in determining the possibility of a separating equilibrium. Straightforwardly, this is because 2 has an incentive to reveal its type, which eliminates the effects of uncertainty on 1’s demand behavior. However, as shown in the next section, a favorable reputation for restraint and sufficiently large diplomatic costs can radically change 2’s escalatory incentives.

**Pooling Equilibria**

If costly signaling averts war when third-party diplomatic costs are sufficiently low, how do 2’s incentives change when $O$ poses a more serious threat of diplomatic punishment? When the diplomatic costs of opposition become large enough that the resolute type no longer finds it tolerable to provoke $O$’s opposition, it opts instead not to escalate, masking its type by acting like a restrained type that finds escalation too costly. In other words, state 2 may engage in strategic restraint designed to manipulate $O$’s beliefs about 2’s willingness to use force. Proposition 2 states that such pooling equilibria occur when $c_D$ is sufficiently high and when $O$’s prior belief is such that it believes that 2 is likely to be restrained.

**Proposition 2** (Pooling Equilibria). \textit{When $c_2 \geq c_1 + \epsilon_2$, $c_D > \overline{c}_D$, and $\phi \geq 1/2$, there exist PBE in which neither type of 2 escalates, $O$ supports 2, and 1 risks war iff $\phi > \hat{\phi}$. See appendix for proof.}

As the diplomatic costs of opposition increase, passing the threshold defined by $\overline{c}_D$ in Figure 2, the separating equilibrium defined in Proposition 1 is no longer sustainable, and two strategic restraint equilibria emerge in which neither type of 2 escalates, setting $e_2 = 0$ instead. Note
that since $c_2 \geq c_1 + c_2$, the restrained 2 still finds it too costly to send an escalatory signal, meaning that the resolute type can send a perfectly revelatory signal if it wishes to do so. However, when revealing itself as resolute brings sufficiently high diplomatic costs for escalation, the associated decrease in the resolute type’s expected payoff from war means that even a peaceful resolution to the conflict comes at a heavy diplomatic price. Accordingly, when $c_D > e_D$, the resolute type has an incentive to mask its type by acting like a restrained type.

However, simply pooling with the restrained type does not guarantee that O will support 2, because with no new information O can rely only on its prior beliefs over 2’s restraint. Therefore, non-escalatory pooling (or strategic restraint) equilibria are only possible when $\phi > 1/2$, or when 2 enters the crisis with a sufficiently favorable reputation for exercising restraint. Since O will support 2 under these circumstances only if $\phi(1) + (1-\phi)(0) \geq \phi(0) + (1-\phi)(1)$, or $\phi \geq (1-\phi)$, the relevant constraint is simply $\phi \geq 1/2$. Otherwise, if $\phi < 1/2$ and O has prior reason to suspect that 2 is not the restrained type, O will default on supporting 1 instead. Therefore, the resolute type should pool with a non-escalating restrained type only when it begins the crisis an ax ante favorable reputation among observers seeking to reward restraint.

When 2 uses strategic restraint to win the support of world opinion, 1 faces the risk-return tradeoff discussed above, choosing between a moderate demand that both types accept and a risky one that the resolute type rejects. As shown in Figure 2, 1 makes the risky demand only when it is sufficiently optimistic that 2 is restrained, i.e. $\phi > \hat{\phi}$—or, as it is rendered in Figure 2, when 2 is sufficiently unlikely to be resolute, $1 - \phi < 1 - \hat{\phi}$. Therefore, pooling on restraint most favors a peaceful settlement when O believes that 2 is sufficiently likely to be restrained while 1 believes that there is a large enough risk that it is truly resolute, or $1/2 \leq \phi \leq \hat{\phi}$. In fact,

As shown in the appendix, the existence of this no-risk equilibrium requires that $\hat{\phi} > 1/2$, which is true when
the restrained type receives a very favorable bargain here by virtue of the resolute type's pooling behavior, since 1 opts to make a proposal generous enough that the resolute type accepts. However, when both 1 and O believe that 2 is sufficiently likely to be restrained, or \( \phi > \max\{1/2, \hat{\phi}\} \), 1 is assured to risk war even as O is increasingly disposed to favor 2. Thus, even when it is sure to fight a war by refusing to reveal its private information, the resolute 2 prefers a war with O's support to a peaceful settlement of the crisis without O's support.\(^{13}\)

What can these equilibria tell us about American behavior in the Berlin Crisis of 1961-1962? With Kennedy's recent humiliation at the Vienna summit in June 1961 (Freedman 2000, p. 64), West Germany and France loudly asserting the need to deter further Soviet action (Aono 2010), and President Eisenhower having successfully managed a similar challenge in 1958 with his own signals of resolve (Fursenko and Naftali 2006, Taubman 2003), the United States would seem to have ample reason to signal a willingness to fight to protect the status quo in Germany.\(^{14}\)

However, the US opted against a strategy of escalation. Why? First, to the extent that the United States was truly willing to risk war to protect Allied rights in Berlin (see Dallek 2003, Taubman 2003), it seems to have weighed the costs of losing diplomatic support more heavily than the immediate benefits to be gained from improving the military situation in Berlin through the proposed mobilization (Freedman 2000, pp. 62-63,67). Indeed, Paul Nitze asserted that it was “essential not to scare people to death with our buildup” (quoted in Aono 2010, p. 336), and there was broad consensus on the existence of “positive benefits of being seen to make the effort” to negotiate (Freedman 2000, p. 69). Second, the United States must have be-

---

\(^{13}\) As specified in the appendix, 1 and O respond to any unexpected, i.e. out-of-equilibrium, escalation by believing that 2 is resolute, since it pays lower costs for escalation than the restrained type.

\(^{14}\) For extensive discussions of Kennedy's reputational desire to signal resolve over Berlin, see Freedman (2000), Dallek (2003), and Harrison (2003).
lieved that its reputation for restraint was sufficiently strong that world opinion would support it against Soviet attempts to revise the status quo. At the very least, Kennedy’s performance during the Bay of Pigs incident suggested that the he was less than resolute in his willingness to use force, and he had already stated his desire to show restraint in negotiations over a ban on nuclear testing, where he believed that a resumption of testing would be a “public relations disaster” (Freedman 2000, p. 52). Finally, the Soviet response, to back off the deadline and instead build the Berlin Wall (see Fursenko and Naftali 2006, Harrison 2003), suggests that while third parties may have had a sufficiently favorable view of the United States, the chances that it was truly willing to use force over Berlin were sufficiently large that, for the USSR, following through on its aggressive bargaining posture would pose too great a risk: hence the construction of the Berlin Wall rather than a transfer of authority in the city to East Germany (this places the Berlin Crisis in the peaceful pooling equilibrium in Figure 2). While the model provides a plausible account of the otherwise puzzling American decision to signal restraint rather than resolve, we can still ask if there is reason to believe that American decisionmakers were correct in their assessment that winning the support of world opinion would lower the chances of war, a question to which I turn in the following section.

**Signals of Restraint and the Probability of War**

If third-party diplomacy can affect great powers’ escalatory behavior, how does it affect the escalation of crises to war? During the Berlin Crisis, American decisionmakers believed that increasing the diplomatic costs of escalation would deter the Soviet Union from risking war, but

---

15This is consistent with Stevenson’s above-quoted claim that refusing to escalate would place the disapprobation of world opinion on the Soviet Union rather than the West.
Proposition 2 shows that resolute types may be tempted to forego revealing their private information, provoking war where they otherwise would not if third party diplomacy were inconsequential. In this section, I explore both effects: specifically, the promise of winning outside diplomatic support can lead to wars that would otherwise be avoided, yet once gained, such support can render one’s opponent less willing to risk war.

Return to Figure 2, which shades those parts of the equilibrium space in which war occurs with positive probability, as a function of 1’s beliefs over 2’s resolve and the diplomatic costs of opposition, given that 2 begins the game with a favorable reputation for restraint, or \( \phi \geq 1/2 \). Each of the three equilibria described in Propositions 1 and 2 appear, with the separating equilibrium occurring for low diplomatic costs and the two pooling equilibria when diplomatic costs are higher. Finally, within the space characterized by pooling equilibria, there is a risk of war only when 1 is sufficiently optimistic that it faces a restrained opponent. As noted above, the occurrence of a peaceful strategic restraint (i.e. pooling) equilibrium requires a delicate balance: O must be confident enough that 2 is restrained to support it in a pooling equilibrium, but if 2 is too likely to be restrained, 1 cannot be deterred from risking war. Perversely, then, the very conditions that lead O to give 2 the benefit of the doubt when it comes to rewarding restraint are those that also encourage 1 to risk war against it.

When pooling equilibria exist, 1’s willingness to risk war decreases in the diplomatic costs of O’s opposition, which it pays in the event of war, since O supports 2 in the non-escalatory pooling equilibria in question.\(^{16}\) The threshold below which 1 risks war is

\[
1 - \hat{\phi} = \frac{\bar{c}_2 - c_2}{c_1 + c_D + \bar{c}_2},
\]

\(^{16}\)I show in the appendix that no non-escalatory pooling equilibria exist in which would O support 1.
which decreases in the diplomatic costs of opposition. Why is this the case? Though O’s behavior does not affect 1’s information about 2’s type, it does directly affect 1’s expected utility for war, as shown in Equation (1). By increasing the costs of war, the prospective opposition of third parties makes risking war less attractive, regardless of 1’s beliefs over 2’s resolve. By rendering 1’s outside option of war less attractive, these additional costs ensure that 1 must be ever more optimistic in order to risk war, rendering the threshold in Equation (2) that supports the risk-of-war equilibrium ever harder to satisfy. Thus, it would appear that, conditional on winning the support of outsiders, American decisionmakers were right about its potential to discourage further Soviet escalation over the Berlin question. However, the very promise of winning outside diplomatic support can encourage war where it otherwise would not occur.

When the diplomatic costs of O’s opposition are small or nonexistent, then there exists a separating equilibrium in which the resolute type finds it profitable to send a costly signal of its resolve by escalating the crisis. However, as $c_D$ increases, the resolute type finds it less and less attractive to provoke diplomatic opposition until $c_D \geq \bar{c}_D$, and it then chooses to forego the opportunity to reveal its private information and engages in strategic restraint. Further, when both 1 and O are sufficiently optimistic over 2’s restraint, the resolute type masks its resolve in the full knowledge that it will end up in a war against 1. Despite the costs of war, a violent resolution to the crisis if preferable to a peaceful one in which 2 faces third-party diplomatic opposition; put differently, peace with international opposition can be worse than war with its support, and when this is the case then increasing diplomatic costs of escalation can encourage 2 to tolerate a war that it would otherwise pay to avoid if third parties were inconsequential.

Third-party diplomatic support thus has competing effects on the course of great power disputes. When a state has outside support, its opponent can be deterred from risking war,
but the very promise of winning that support can, perversely, lead a resolute state to mask its
type and provoke a war that it has the means, in principle, to avoid. However, the promise
of third-party diplomatic support encourages war only when a state enters the crisis with a
sufficiently favorable reputation for restraint, because it both encourages $O$ to support it and $I$
to risk war against it, which brings into stark relief the fundamental dilemma for great powers
between signaling resolve and restraint. This suggests, of course, that those great powers most
widely believed to be restrained are precisely those that are most likely to find themselves in
wars as the result of their ability to take advantage of a favorable reputation. Great powers with
unfavorable reputations for restraint, on the other hand, do not face the temptation of winning
support through pooling and will find themselves more often in separating equilibria that favor
peaceful resolutions to their crises.

**Conclusion**

Great powers face a signaling dilemma in interstate crises. Demonstrating resolve through esca-
lation may deter an opponent, but it may also provoke diplomatic opposition from third parties
fearful of unrestrained military power; on the other hand, failing to signal resolve may reas-
sure observers but result in either unfavorable settlements or avoidable wars. Where previous
studies examine the incentive to signal either resolve or restraint in isolation (e.g. Fearon 1997,
Ikenberry 2001), I examine tradeoffs between them in a three-player model of crisis bargaining
under asymmetric information. Even when great powers can send credible signals of resolve,
they may instead engage in strategic restraint, refusing to signal resolve and generating a risk
of war, in order to win outside diplomatic support. While such support, once gained, induces
caution in a state's adversaries and reduces the probability of war, the very promise of winning it can prevent that state from revealing its private information, generating a positive probability of war where none would otherwise exist.

By drawing explicit links between third party diplomacy, or soft balancing, and crisis bargaining behavior, the model makes predictions over the probability of war that depend on both the anticipated costs of incurring diplomatic opposition and the reputation for restraint that a great power brings to the crisis. Specifically, war is possible when a great power with a reputation for reluctance to use force exercises strategic restraint, refusing to escalate in a way that reveals its private information, maintaining diplomatic support yet provoking war in the process. Far from refusing to admit weakness, then, great powers may refuse to admit their strength, with tragic consequences for the outbreak of war. Thus, the very reputation that curries the favor of third party observers also encourages a great power's opponent to risk war in making demands during the crisis. On the other hand, the strength of outside diplomatic support, once gained, can discourage risk-taking and decrease the probability of war. Perhaps tragically, then, two strategies designed to limit the occurrence of great power war, soft balancing and strategic restraint, can interact to increase the chances that it occurs.

This may explain why great powers may send ambiguous or uninformative signals about their resolve (see Fearon 1997, Russett 1963), and it provides a substantive example for the crucial and often-overlooked second element of Fearon's (1995) rationalist explanation for war: (a) private information with (b) incentives to misrepresent it. Absent such a reason to dissemble, bargains ought to be easy to reach as long as players are sufficiently patient (Leventoglu and Tarar 2008), and explaining observed conflicts in terms of this mechanism requires an account of why belligerents were unwilling or unable to share critical war-avoiding information. Other
incentives to misrepresent, be they secrecy about military strength (Meirowitz and Sartori 2008) or bolstering a reputation for resolve (Wolford 2007), manifest as irresolute types mimicking resolute types. However, the story presented here involves a resolute type acting like a restrained type, refusing to reveal its private information even when the dyadic characteristics of the crisis are conducive to credible signaling. Strategic restraint, then, while potentially useful for great powers in the maintenance of international order (Ikenberry 2001), may come at the cost of otherwise avoidable wars when the benefits of engaging in it discourage costly signaling.

More broadly, the analysis suggests that multilateral processes, including the diplomatic behavior of non-participants and signaling to multiple audiences, generally excluded from dyadic models of crisis bargaining can improve our understanding of the dynamics of crisis bargaining and the outbreak of war. Unambiguous signals of resolve or commitment tend to be rare in crisis bargaining, and while this may stem from leaders’ pragmatic desires to keep their future options open or the need to appease domestic audiences (Snyder and Diesing 1977), the logic developed above indicates that another force discouraging separating signals may emerge from basic characteristics of the international system—specifically, a great power’s need to signal restraint in order to reduce the costs of exercising power. More states watch, fear, and respond to great powers than other states, and one further implication of this analysis is that, all else equal, powerful states should find themselves more prone than weaker states to masking their resolve in order to win the short-term gains of outside diplomatic support.

17 For an empirical approach to this problem, see Poast (2010).
Appendix

Recall that all proofs rest on the following restrictions. First, costs for war are never so large as to undermine the threat to fight, which ensures that all proposals $x$ are interior; this simplifies the analysis but does not alter the substantive force of the results. Second, pre-crisis military endowments are equal, or $m_1 = m_2 = m_i$. Finally, as derived below, $m_i \leq \hat{m}_i = \min\{c_2/(c_1 + c_2)^2, 1/4c_i\}$, which ensures that optimal escalation levels are positive, i.e. $e^*_i > 0$.

Proof of Proposition 1. Strategies in the separating equilibrium are as follows. First, 2 sets $e_2 = e^*_2 > 0$ if resolute, $e_2 = 0$ if restrained, and accepts $x$ iff $u_2(\text{accept}) \geq EU_2(\text{reject}|c_2)$. Second, 1 sets $e^*_1(e^*_2) > 0$ if $e_2 = e^*_2$, $e^*_1(0) > 0$ if $e_2 = 0$, and proposes the largest $x$ such that 2 accepts. Third, $O$ supports 2 if $e_2 = 0$ and supports 1 if $e_2 = e^*_2$. Posterior beliefs after 2’s escalation choice for both 1 and $O$ are $\phi' = 0$ if $e_2 = e^*_2$, and $\phi' = 1$ if $e_2 < e^*_2$. Since types reveal themselves immediately, end-of-game posteriors are trivial, and pure separation ensures that uninformed players update their beliefs to certainty, confirming that beliefs are consistent with strategies according to Bayes’ Rule.

To verify that strategies are sequentially rational, begin with $O$’s strategy. It supports 2 if $e_2 = 0$, which is a best reply given its beliefs since $1 > 0$, and that it supports 1 if $e_2 = e^*_2$, which is also a best reply given its beliefs since $1 > 0$. Now consider 1’s escalation decision, which it takes after updating its beliefs over 2’s type and in full knowledge of $O$’s strategy. If $e_2 = 0$, 1 believes that 2 is restrained and anticipates that $O$ will support 2. Therefore, its optimal level of escalation in preparation for war must satisfy

$$\max_{e_1} \left\{ \frac{m_i + e_1}{m_i + e_1 + m_i} - c_1 (1 + e_1) - c_D \right\}.$$
The first-order condition is \((m_i - c_1 (e_1 + 2m_i)^2)/((e_1 + 2m_i)^2) = 0\), and solving for \(e_1\) yields the optimal escalation

\[ e_1^*(0) = \frac{\sqrt{m_i}}{\sqrt{c_1}} - 2m_i. \]  

(3)

Since the second-order condition is \(-2m_i/(e_1 + 2m_i)^3\), any \(e_1 \geq 0\) as defined by Equation (3) represents a maximum. Ensuring that \(e_1^*(0) \geq 0\) requires \(m_i \leq 1/4c_1\), as stipulated above.

If \(e_2 = e_2^*\), 1 believes that 2 is restrained and anticipates that \(O\) will support 1. Therefore, its optimal level of escalation in preparation for war must satisfy

\[ \max_{e_1} \left\{ \frac{m_i + e_1}{m_i + e_1 + m_i + e_2} - c_1 (1 + e_1) \right\}, \]  

(4)

and 2's must satisfy

\[ \max_{e_2} \left\{ \frac{m_i + e_2}{m_i + e_1 + m_i + e_2} - c_2 (1 + e_2 - c_2) \right\}. \]  

(5)

This implies solving a system of first-order conditions,

\[ \frac{e_2 + m_i - c_1 (e_1 + e_2 + 2m_i)^2}{(e_1 + e_2 + 2m_i)^2} = \frac{e_1 + m_i}{(e_1 + e_2 + 2m_i)^2} - c_2 = 0, \]

yielding the following optima:

\[ e_1^*(e_2^*) = \frac{c_2 - m_i (c_1 + c_2)^2}{(c_1 + c_2)^2} \quad \text{and} \quad e_2^* = \frac{c_1}{(c_1 + c_2)^2} - m_i. \]  

(6)
Since the second-order conditions are $-2(e_2 + m_i)/(e_1 + e_2 + 2m_i)^3$ and $-2(e_1 + m_i)/(e_1 + e_2 + 2m_i)^3$ for Equations (4) and (5), respectively, any nonnegative values of $e_i$ are guaranteed to represent maxima, and the expressions in (6) are guaranteed to be nonnegative as long as $m_i \leq c_2/(c_1 + c_2)^2$. Therefore, to ensure that all escalatory choices are nonnegative, I impose the global restriction that $m_i \leq \hat{m}_i$, where

$$\hat{m}_i = \min \left\{ \frac{c_2}{(c_1 + c_2)^2}, \frac{1}{4c_i} \right\}.$$ 

Now consider behavior in the bargaining stage. The restrained 2 accepts a proposal $x$ iff

$$1 - x \geq \frac{m_2}{m_1 + e_1^*(0) + m_2} - \bar{c}_2 \Leftrightarrow x \leq 1 - \frac{m_2}{m_1 + e_1^*(0) + m_2} + \bar{c}_2. \quad (7)$$

If 1 wishes to induce 2’s acceptance, it sets $x^*(0)$ to satisfy Inequality (7) at equality, since demanding more provokes rejection and demanding less wins acceptance but at less favorable terms. Therefore, 1 proposes $x^*(0)$, rather than provoke war by demanding more, iff

$$x^*(0) \geq \frac{m_1 + e_1^*(0)}{m_1 + e_1^*(0) + m_2} - c_1 - c_D,$$

which is strictly true since $\bar{c}_2, c_1, c_D > 0$. The resolute 2 accepts a proposal $x$ iff

$$1 - x \geq \frac{m_2 + e_2^*}{m_1 + e_1^*(e_2^*) + m_2} - \bar{c}_2 - c_D \Leftrightarrow x \leq 1 - \frac{m_2 + e_2^*}{m_1 + e_1^*(e_2^*) + m_2} + \bar{c}_2 + c_D. \quad (8)$$

If 1 wishes to induce 2’s acceptance, it sets $x^*(e_2^*)$ to satisfy Inequality (8) at equality, since demanding more provokes rejection and demanding less wins acceptance but at less favorable
terms. Therefore, 1 proposes $x^*(e_2^*)$, rather than provoke war by demanding more, iff

$$x^*(e_2^*) \geq \frac{m_1 + e_1^*(e_2^*)}{m_1 + e_1^*(e_2^*) + m_2 + e_2^*} - c_1,$$

which is strictly true since $\bar{c}_2, c_1, c_D > 0$.

Finally, consider 2’s strategy in the first move of the diplomatic stage. The restrained 2 sets $e_2 = 0$, which reveals its type to the other players, and thus wins $O$’s support. If it deviates, setting $e_2 = e_2^*$ like the resolute type, it will provoke $O$’s opposition. 2 plays its equilibrium strategy, honestly revealing its type, iff $1 - x^*(0) \geq 1 - x^*(e_2^*) - \bar{c}_2 e_2^*$, which is true as long as $
\bar{c}_2 \geq c_1 + c_2$ and $m_i < c_2/(c_1 + c_2)^2$, or $m_i \leq \hat{m}_i$. The resolute 2, in equilibrium, escalates and reveals its type; should it deviate by choosing not to escalate, 1 makes a proposal sufficient to win the acceptance only of the restrained type, which means that the resolute 2 will reject in its return to optimal play. Therefore, the resolute type’s strategy is a best reply iff

$$1 - x^*(e_2^*) - c_2 e_2^* \geq \frac{m_2}{m_1 + e_1^*(0) + m_2} - c_2 \Leftrightarrow c_D \leq -\sqrt{c_1 m_2 + m_2 c_2} + \frac{c_1^2}{(c_1 + c_2)^2} = \bar{c}_D. \quad (9)$$

Note that 1’s rejection is out-of-equilibrium, since both 1 and O expect 1’s proposal, given posterior beliefs $\phi' = 1$, to be accepted. Such beliefs are trivial here since they end the game, but it is reasonable to expect that the uninformed players believe $\phi'' = 1$ in response to an unexpected rejection.

\[\Box\]

**Proof of Proposition 2.** Strategies in the pooling equilibria are as follows. First, 2 sets $e_2 = 0$, and accepts $x$ iff $u_2(\text{accept}) \geq EU_2(\text{reject}|c_2)$. Second, on the equilibrium path, 1 sets $e_1^*(0) > 0$, but sets $e_1^*(e_2^*) > 0$ (as defined in Proposition 1) if 2 escalates; on the equilibrium path, 1 proposes
x = \bar{x}^*(0) if \phi \leq \hat{\phi} and x = \bar{x}^*(0) if \phi > \hat{\phi}, but if 2 has escalated it proposes x = x^*(e_2^*) (as defined in Proposition 1). Finally, on the equilibrium path O supports 2 but supports 1 if 2 escalates. Posterior beliefs after 2’s escalation choice for both 1 and O are \phi' = \phi on the equilibrium path and \phi' = 1 in response to an out-of-equilibrium escalation.

To verify that strategies are sequentially rational, begin with O’s strategy. It supports 2 iff \phi \geq (1 - \phi) \Leftrightarrow \phi \geq 1/2, which is a best reply given that it retains its priors, or \phi' = \phi. Now consider 1’s optimal escalation level in the diplomatic stage. If 2 does not escalate, 1’s optimal escalation is again e_1^*(0), as defined in Equation (3). Should 2 escalate by setting e_2 = e_2^*, as defined in (6), 1 chooses e_1^*(e_2^*), also defined in (6).

There are two equilibria distinguished by 1’s proposal in the bargaining stage. First, in the no-risk equilibrium, it proposes x = \bar{x}^*(0), which both types of 2 accept. Next, in the risk-of-war equilibrium, it proposes x = \bar{x}^*(0), which only the weak type accepts. I begin by establishing sequentially rational behavior in the no-risk equilibrium and follow by doing the same for the risk-of-war equilibrium.

In the bargaining stage, a player-type of 2, or c_2, accepts a proposal x iff

\[
1 - x \geq \frac{m_2}{m_1 + e_1^*(0) + m_2} - c_2 \Leftrightarrow x \leq 1 - \frac{m_2}{m_1 + e_1^*(0) + m_2} + c_2. \tag{10}
\]

Since c_2 takes on two different values, there are two proposals 1 can make in equilibrium. First, if it wishes to win the acceptance of both types, it satisfies Inequality (10) at equality where c_2 = c_2, denoted x = \bar{x}^*(0). Next, if it wishes to win the acceptance of only the restrained type, provoking the resolute’s rejection, it satisfies Inequality (10) at equality where c_2 = \bar{c}_2, denoted x = \bar{x}^*(0). It is straightforward to show that will never make a proposal that both types reject,
because doing so foregoes the opportunity to win an acceptance that transfers the bargaining surplus, \( c_1 + c_2 \), to it with a positive probability. Therefore, these are the only two proposals it makes in equilibrium. 1 thus proposes \( x = x^*(0) \) when \( u_1(x^*(0)) \geq u_1(\bar{x}^*(0)) \), or

\[
\bar{x}^*(0) \geq \phi \bar{x}^*(0) + (1 - \phi) \left( \frac{m_1 + e_1^*(0)}{m_1 + e_1^*(0) + m_2} - c_1 - c_D \right) \iff \phi \leq \frac{c_1 + c_D + c_2}{c_1 + c_D + \bar{c}_2} = \hat{\phi}.
\]

Combined with the constraint that \( \phi \geq 1/2 \), the existence of the no-risk equilibrium also requires \( 1/2 \leq \phi \leq \hat{\phi} \), which is true iff \( \bar{c}_2 < 2c_l \). Otherwise, when \( \phi > \max\{1/2, \hat{\phi}, \phi' \} \), 1 proposes \( x^*(0) \) and risks war.

First, consider 2’s strategies when 1 will make the no-risk proposal. Given that setting \( e_2 = e_2^* \) will lead the uninformed players to believe that it is resolute (with those strategies defined in Proposition 1), or \( \phi' = 0 \), the restrained type sets \( e_2 = 0 \) iff

\[
1 - x^*(0) \geq 1 - x^*(e_2^*) - \bar{c}_2 e_2^* \iff c_D \geq -\sqrt{c_1 \sqrt{m_2} - (1 - m_2) \bar{c}_2 + c_2 + \frac{c_1 (c_1 - \bar{c}_2 + c_2)}{(c_1 + c_2)^2}}.
\]

The resolute type chooses not to reveal itself through escalation iff

\[
1 - x^*(0) > 1 - x^*(e_2^*) - \bar{c}_2 e_2^* \iff c_D > \bar{c}_D,
\]

where \( \bar{c}_D \) is the same constraint identified in Inequality (9). Next,

\[
\bar{c}_D > -\sqrt{c_1 \sqrt{m_2} - (1 - m_2) \bar{c}_2 + c_2 + \frac{c_1 (c_1 - \bar{c}_2 + c_2)}{(c_1 + c_2)^2}}
\]

ensuring that \( \bar{c}_D \) is the binding constraint, when \( m_i < 1 + c_1 / (c_1 + c_2)^2 \), which must be true as
long as $m_i < \hat{m}_i$ is also true, as it is by construction. Therefore, the no-risk pooling equilibrium exists when $1/2 \leq \phi \leq \hat{\phi}$ and when $c_D > \overline{c}_D$.

Finally, consider 2’s strategy when 1 will make the risk-of-war proposal. Given that setting $e_2 = e_2^*$ will lead the uninformed players to believe that it is resolute (with those strategies defined in Proposition 1), or $\phi' = 0$, the restrained type sets $e_2 = 0$ iff

$$1 - \overline{x}^*(0) \geq 1 - x^*(e_2^*) - \overline{c}_2 e_2^* \Leftrightarrow c_D \geq -\sqrt{c_1 \sqrt{m_2} + m_2 \overline{c}_2} + \frac{c_1 (c_1 - \overline{c}_2 + \overline{c}_2)}{(c_1 + c_2)^2}.$$ 

The resolute type, which expects war on the equilibrium path, chooses not to reveal its type iff

$$\frac{m_2}{m_1 + e_1^*(0) + m_2} - c_2 > 1 - x^*(e_2^*) - \overline{c}_2 e_2^* \Leftrightarrow c_D > \overline{c}_D,$$

where $\overline{c}_D$ is the same constraint identified in Inequality (9). Next,

$$\overline{c}_D > -\sqrt{c_1 \sqrt{m_2} + m_2 \overline{c}_2} + \frac{c_1 (c_1 - \overline{c}_2 + \overline{c}_2)}{(c_1 + c_2)^2}$$

ensuring that $\overline{c}_D$ is the binding constraint, when $m_i < c_1 / (c_1 + c_2)^2$, which must be true as long as $m_i < \hat{m}_i$ is also true, as it is by construction. Therefore, the risk-of-war pooling equilibrium exists when $\phi > \max\{1/2, \hat{\phi}\}$ and $c_D > \overline{c}_D$. \hfill \qed

References


