

**Efficient Secrecy:  
Public versus Private Threats in Crisis Diplomacy**

Shuheï Kurizaki\*

Ph.D. Candidate  
Department of Political Science  
University of California, Los Angeles  
4289 Bunche Hall  
Los Angeles, CA 90095-1472  
kurizaki@ucla.edu

January 20, 2005

Work in progress  
Comments welcome

Abstract

This paper explores when and why political leaders choose to communicate secretly in international disputes. I examine a crisis bargaining game with two-sided domestic audiences, where a challenger can make a threat either in public or in private. When a threat is issued privately, the challenger cannot enhance its credibility by tying its hands because domestic audiences cannot observe the threat. I show that although private threats convey only limited credibility, it is rational to make them under certain conditions, because their rationality stems not from the informational efficacy but from their less provocative nature. Further, a private threat improves efficiency by expanding the range of peaceful outcomes when a crisis takes place before multiple audiences because its credibility is derived from the defender's sensitivity to audience costs. These results suggest that secrecy works in diplomacy despite its informational inefficacy because leaders can save face from diplomatic humiliations before their domestic audiences.

---

\* Earlier versions were presented at the annual meetings of the Peace Science Society, November 2003, the Midwest Political Science Association, April 2004, and the American Political Science Association, September 2004. I am grateful to Lisa Blaydes, Marisa Kellam, James Morrow, Barry O'Neill, Ken Schultz, Branislav Slantchev, Art Stein, Ahmer Tarar, and Leeat Yariv for valuable comments, and to the Ronald W. Burkle Center for International Relations (UCLA), the Institute on Global Conflict and Cooperation (University of California), and the Matsushita International Foundation (Panasonic) for financial support. First draft: July 16, 2003.

## 1. Introduction

During the Cuban Missile Crisis in October 1962, President Kennedy and Chairman Khrushchev communicated mostly through public media. But it was the *secret* deal to trade the Jupiter missiles in Turkey that eventually resolved the crisis, and this offer was made by Robert Kennedy to Ambassador Dobrynin at a *private* meeting. About a decade later, when the construction of a Soviet submarine base was discovered in 1970 in Cuba, Henry Kissinger and President Nixon pursued almost entirely private diplomacy to settle the Cienfuegos Submarine Base Crisis. To avoid “a dramatic confrontation on the order of 1962,” the Nixon administration “considered that quiet diplomacy was best suited to giving the USSR an opportunity to withdraw without humiliation” (Kissinger 1979, 651). Although this crisis could have been the second Cuban Crisis, possibly no one remembers this incident because it was managed quietly and secretly so that the public would not notice this as a crisis.

Uncertainty has long been claimed to be a fundamental cause of war (Schelling 1960; Blainey 1988; Morrow 1989; Fearon 1995). Accordingly, the literatures on deterrence and crisis bargaining have searched for credible information-revealing mechanisms that help state leaders overcome uncertainty and hence avoid inefficient fighting (Powell 2002). What eventually emerges from these studies is an audience cost story of the origins of war (Fearon 1994). A key result of the recent literature on audience costs is that state leaders can credibly communicate their resolve by going public with their demands so as to establish a tying-hands mechanism by creating audience costs that they would suffer if they fail to follow through on the threats (Fearon 1997; Schultz 1998, 2001; see also Ramsay 2004; Smith 1998). Consequently, this line of argument leads to the conclusion that normal form of diplomacy may be irrelevant during a crisis because incentives to misrepresent resolve undermines diplomatic signaling (e.g., Fearon 1995; see also Sartori 2002).

However, the two Cuban Crises in 1962 and 1970 pose a challenge to this conclusion. They suggest that state leaders go against the logic of audience costs by staying private and seeking normal diplomatic communications, rather than going public and thereby increasing the risk of inefficient

outcomes. In other words, the standard rationalist explanations leave unexplored this class of historical cases of crisis diplomacy. One might argue that these cases are merely “exceptions” and “quiet diplomatic exchanges may be insufficient to allow states to learn what concessions an adversary would ... be willing to make” (Fearon 1994, 586). Yet, the anecdotes above nonetheless raise questions as to when and why leaders rationally conclude that staying private is preferable to going public. Moreover, if private diplomacy were irrelevant to conflict resolution as the standard rationalist explanations suggest, it would be puzzling that secrecy has been an enduring feature of modern diplomatic institutions for centuries (Berridge 2002, 107; Nicolson 1954, 75).

The current literature does not address these questions. In fact, the conventional conclusions are logical, given their purpose to explain why leaders may rationally go to war despite its inefficiency (Morrow 1989; Fearon 1995). Appealing to the tying-hand mechanism, the standard rationalist logic concludes that private diplomacy is irrelevant because secrecy “unlocks” leaders from inefficient outcomes by allowing them to maintain their “escape clause” to disavow their threats.

But this logic brings us back to the original puzzle of war: why do leaders rationally choose to take actions in public that tie their hands to inefficient outcomes? If going public increases the risk of inefficiency, it should follow that the political risk of generating audience costs by going public should outweigh the benefits of enhanced credibility (Baum 2004). In principle, *ex post* inefficiency of going public opens up an *ex ante* bargaining range through private communications, which makes staying private preferable to going public in a crisis. In fact, very few international disputes become wars, and yet audience costs are rarely observed (Fearon 1997; Schultz 2001).

Hence, it must be shown when private channels can *and* cannot achieve peaceful settlements not only to account for diplomatic records of private tactics and quiet maneuvers in a crisis but also to support the conventional rationalist conclusions. The remainder of this paper is an attempt to explore such a condition and to look for the credibility of private signals in crisis bargaining. It takes a standard

audience cost model and adds simple twists to it by relaxing two crucial assumptions, upon which the conventional conclusion heavily depends.<sup>1</sup>

The first assumption is that international crises are public events played out in front of domestic audiences (Fearon 1994; Ramsay 2004; Schultz 2001a). While this assumption is reasonable since those models seek to explain why states take costly actions in public during a crisis, it also effectively rules out the possibility of making private threats. Early models assume away the possibility that leaders prefer to stay private in sending signals, with a notable exception. Although their analysis of public and private communication methods is ad hoc, Snyder and Diesinger (1977, 251-54) presume that a public threat is a better credibility-generating mechanism and conjecture that private communication best serves as a supplement to public threats.<sup>2</sup> I show the exact opposite.

The second assumption common to the previous audience costs models is that only the sender of signals can create domestic audience costs, while assuming away audience costs for the receiver.<sup>3</sup> This assumption is empirically less plausible, as there are numerous episodes suggesting that decision makers take into account the nature and magnitude of audience costs for the opponent in making decisions during crises. For example, Saddam Hussein refused the American demand of not invading Kuwait in 1990 based on his belief that the US had high audience costs. Immediately before his invasion of Kuwait, Hussein reportedly told the American ambassador that the US threats of intervention were not credible because, “Yours is a society which cannot accept 10,000 dead in one battle” (*New York Times*, September 22, 1990, A1). I show that the receiver’s audience costs drive an interesting mechanism that has been neglected in previous explanations of conflict.

---

<sup>1</sup> Besides these two assumptions, what is crucial to the conclusion that signals must be costly to convey information is the use of the framework of Spence’s (1973) costly signaling model. But cheap talk models has long shown that that signals need not be costly to be informative.

<sup>2</sup> Other exceptions include Leventoglu and Tarar’s (2004) analysis of public commitment during international negotiations and O’Neill’s (2003) analysis of mediation and national honor.

<sup>3</sup> Note that previous models by Fearon (1994) and Schultz (2001) also consider audience costs for both states in their models of international crises. Yet, the key difference here is that in Fearon’s model, audience costs are automatically raised by the onset of a crisis, but not by a rational choice of any states. In Schultz’s model, audience costs for both parties are signaling costs associated with the acts of signaling, but not the receiver’s political costs provoked by the adversary’s signals.

The core contribution of this paper is identifying the existence of an equilibrium in which a private threat works, and demonstrating its efficiency. In particular, I show that there exist “public” and “private” equilibria to the resulting game. The public equilibrium captures a conventional audience cost story, where only public threats can credibly reveal information. Conversely, in the private equilibrium, private threats are also rational. The merit of the present model is that while simple in its structure, small twists to a canonical crisis bargaining model generate both the well known conventional result as well as previously unexplored new result simultaneously.

There are several principal results. First, I show that a private threat conveys only limited credibility, while a public threat is fully credible. In fact, a private threat does not improve the defender’s beliefs about credibility. Nevertheless, a private threat is equally effective as a public threat in that the former achieves successful deterrence at the same rate as the latter, despite its lack of informational benefits. This is primarily because a private threat makes a private concession attainable. These results suggest that *the rationality of a private threat comes not from its informational role but from its less provocative nature.*

Second, the model shows that the private equilibrium is (Pareto) efficient, where peaceful outcomes are possible under broader conditions than in the public equilibrium, as a private threat expands the range of peaceful settlements that are mutually acceptable. This efficiency result is driven by two facts: 1) a private threat can produce the equilibrium outcome – a private concession – that cannot arise in the public equilibrium; and 2) resorting to a private threat significantly reduces the *ex ante* risk of war.

Third, I show that a private threat works because the states send partially credible private threats when they could go public to enhance the credibility of their threats. While forgoing a fully credible public threat *reduces* the credibility of the threat, doing so makes it easier for the receiver to concede, as secrecy avoids provoking audience costs. This suggests that *a private threat works only in the shadow of credible public threats.* This result may explain why Theodore Roosevelt carried a big stick when speaking softly.

The key to these new results is that signaling takes place in front of *multiple* audiences, meaning that not only the sender but also the receiver must incur audience costs *ex post* when they experience diplomatic humiliation (e.g., backing down from threats and public acquiescence in threats). In particular, in the absence of the receiver's sensitivity to its audience costs, the sender is not able to customize its signals and hence a private threat will not work. In other words, a standard audience cost claim that going public may be the only credible means of communication fails when audience costs for the receiver are not explicitly modeled.

This finding augments the literatures on crisis bargaining and audience costs, as it identifies when the conventional conclusion fails. While this model builds on a standard crisis game with audience costs, it identifies when and why leaders sometimes go public about their diplomatic positions *and* stay private at other times. In doing so, this model helps to explain why earlier models concluded that quiet normal diplomacy is irrelevant in a crisis, and when this conclusion does not hold.

What emerges from this study is a theoretical rationale for secret diplomacy. The literature often characterizes diplomacy as “cheap talk” because leaders can always afford to disavow diplomatic exchanges under the surface (Fearon 1995; Guisinger and Smith 2002; Sartori 2005). The model presented here might help explain why secrecy has been a central feature of diplomatic institutions ever since what we know as diplomacy today, or the “French system,” was established in the seventeenth century and perhaps earlier (Berridge 2002, 107; Nicolson 1954, 75).<sup>4</sup>

The paper proceeds as follows. The next section presents a model of international crises with two-sided domestic audiences and a private signaling option. Section 3 analyzes the model and presents the results of the equilibrium analysis. I also discuss what incentives public and private demands can generate for political leaders. Section 4 explores the implications of the equilibrium analysis and relates them to the importance of diplomacy. To highlight the central result concerning the roles of public and private threats, section 5 presents an illustrative case study. Some conclusions follow.

---

<sup>4</sup> For seminal work in this direction, see Guisinger and Smith 2002 and Sartori 2005, although their focus of diplomacy is exclusively on *reputations* for honesty.

## 2. The Model

The model studied here is a natural extension of the audience cost models, building on a canonical crisis bargaining game that shares basic strategic elements with previous models (e.g., Fearon 1997; Schultz 2001). The model is simple in its structure and one could easily add further complications, but this simplicity is designed to highlight the questions above, clarify some of the less intuitive consequences of the strategic problems, and facilitate comparison with previous studies.

A crisis game involves two states—the challenger ( $C$ ) and the defender ( $D$ )—in a dispute over some good or policy.  $C$  can alter the status quo by challenging  $D$ . For the sake of simplicity, we assume that the status quo represents  $D$ 's ideal position.  $D$  thus has no incentive to initiate a crisis, whereas  $C$ , hoping to change the current situation, has both an opportunity and an incentive to do so against  $D$ . Each state knows its own level of resolve, but does not know its adversary's resolve.

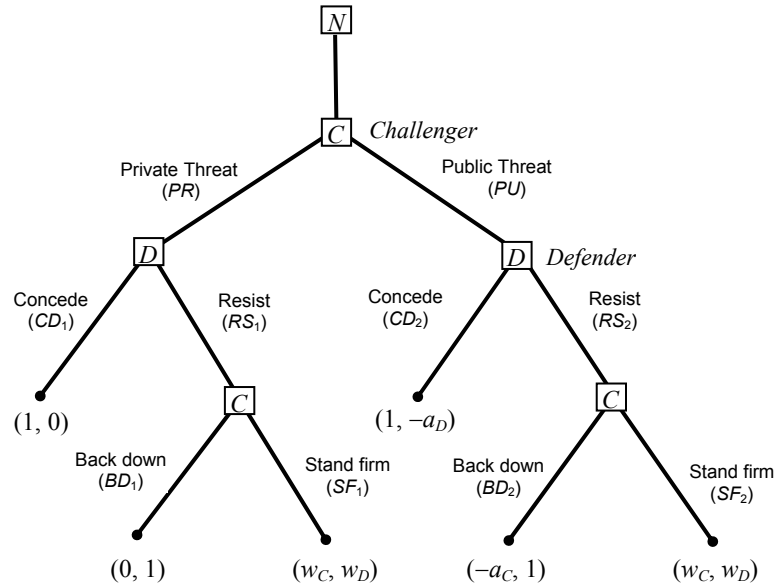
I introduce two twists to this crisis game. First, crisis bargaining is played out before a two-sided domestic audience. There are audiences both in  $C$  and  $D$ , rather than one audience belonging to the sender of a signal. The main implication of this innovation is that  $C$ 's challenge can provoke  $D$ 's audience costs. Second, I allow  $C$  to decide whether to go public or stay private when it comes to sending signals. This twist allows for the possibility of crisis diplomacy being carried out in private, where domestic audiences may know of the crisis but know nothing about the course of events.

***Sequence of Moves.*** The crisis game begins with nature's move and involves four stages. First, at the onset of a crisis, nature informs both  $C$  and  $D$  of their values for fighting over the issue at stake,  $w_C$  and  $w_D$ , respectively. Second,  $C$  makes a challenge to  $D$  to alter the status quo by threatening to use force.  $C$  can do so either in public ( $PU$ ) or in private ( $PR$ ). Because the focus of this paper is on  $C$ 's incentives to make a private threat and the associated credibility condition, he has no option of retaining the status quo at the outset of the game.<sup>5</sup> Instead,  $C$  must choose whether to make a public or private threat.<sup>6</sup>

---

<sup>5</sup> I refer to the challenger,  $C$ , as “he” and the defender,  $D$ , as “she” throughout this paper.

**Figure 1: Crisis Bargaining Game with Public versus Private Threats**



Third, once  $C$  makes a threat, whether publicly or privately,  $D$  updates her beliefs about  $C$ 's value of  $w_C$  according to Bayes' rule, and then either concedes ( $CD$ ) or resists ( $RS$ ). Once a private threat is issued,  $D$  does not have an option to make it public. As such, conceding to a public (private) threat constitutes a public (private) concession. If  $D$  concedes, whether publicly or privately, then  $C$  changes the status quo peacefully to his favored policy position, and the game ends. If, on the other hand,  $D$  resists, the crisis escalates to the fourth stage, where  $C$  either backs down ( $BD$ ) or carries out his threat ( $SF$ ). Backing away from a public (private) threat is (un)observable for domestic audiences. If the challenger stands firm, war occurs. The extensive form game is depicted in Figure 1.

**Outcomes and Payoffs.** Without loss of generality, the payoffs for the status quo are normalized to be 1 for  $D$  and 0 for  $C$ . If  $C$  makes a public threat and  $D$  concedes, the payoffs are 1 for  $C$  and  $-a_D \leq 0$  for  $D$ , indicating that  $C$  obtains his value for the issue at stake and  $D$  not only loses her value for the status quo but also incurs audience costs from suffering from “diplomatic humiliation” (Fearon 1994; see also

---

<sup>6</sup> I assume the two states are already in a confrontation, but do not ask how they got there, following previous crisis bargaining models (O'Neill 1992, 466). We could allow the challenger to do nothing and retain the status quo, but as Corollary 1 demonstrates, doing so does not change the results.

O'Neill 1999). If, on the other hand,  $D$  resists the public threat and  $C$  backs down,  $D$  keeps the status quo with the payoff of 1, and  $C$  gains nothing but pays audience costs  $-a_C \leq 0$ .

However, if  $C$  makes a private threat, neither side incurs audience costs by backing down or conceding, since the audiences are not aware of  $C$ 's threats. In such situations, both states are able to back down without being caught by domestic audiences. So when  $D$  makes a *private concession*, she gets 0 and  $C$  gets 1. If  $C$  backs off from his private threat, the status quo prevails with the status quo payoffs, 0 for  $C$  and 1 for  $D$ .

War occurs when  $D$  resists—whether publicly or privately—and  $C$  stands firm in response. In the event of war, the payoffs are the expected values for war  $w_C$  and  $w_D$  for  $C$  and  $D$ , respectively, which, I assume, are bounded:  $w_i \in [\underline{w}_i, 1]$ ,  $i = C, D$ , where the lower bound is a negative and arbitrarily small real number:  $\underline{w}_i < 0$ . So neither side can gain from fighting more than the value of the good, which is worth 1.

**Information and Beliefs.** I assume that the crisis game involves two-sided uncertainty: Each state is uncertain about the other's true value for war  $w_i$ . I generate this uncertainty by assuming that nature randomly selects  $w_C$  and  $w_D$  from the cumulative distribution  $F_i(x) = \Pr(w_i \leq x)$ , where support of  $F_i(\bullet)$  lies in the interval  $[\underline{w}_i, 1]$  for  $i = C, D$ . The distribution functions  $F_C(\bullet)$  and  $F_D(\bullet)$  are assumed to be independent, identical, and common knowledge.  $C$  observes the selection of  $w_C$ , and  $D$  observes her  $w_D$ , but neither state observes the other's draw, so each state forms pre-crisis beliefs about  $w_C$  and  $w_D$ , respectively. In particular, define  $p_{public}$  and  $p_{private}$  as  $D$ 's pre-crisis beliefs that  $C$ 's public and private threats, respectively, are genuine, and  $q_{public}$  and  $q_{private}$  as her updated beliefs when she receives a threat.

### 3. Results

There exist two perfect Bayesian equilibria: one *public*, one *private*. The public equilibrium is the case where only a public threat can reveal  $C$ 's private information, which captures a well-known standard audience cost story (e.g., Fearon 1997; Schultz 2001). This equilibrium may help explain why states sometimes go public and invoke dramatic confrontations that may lead to inefficient outcomes such

as costly backing-down and costly fighting. The private equilibrium, in contrast, demonstrates a new result, in which a private threat not only conveys (limited) credibility but also improves (*Pareto efficiency*). This equilibrium suggests a rational underpinning for private signaling as an immediate consequence of a standard audience cost story.

All equilibria to this game can be described by a set of cutpoints defined as follows. First, consider the equilibrium strategy for  $C$ . At the final decision node, by subgame perfection, there exists a unique type with  $w_C = -a_C$ , who is indifferent between carrying out a public threat ( $SF_2$ ) and backing down in public ( $BD_2$ ), provided that he has issued a public threat ( $PU$ ). Hence, all challenger-types in the range  $(-a_C, 1]$  strictly prefer  $SF_2$  to  $BD_2$  and all other types in  $[\underline{w}_C, -a_C)$  strictly prefer  $BD_2$ , where  $\underline{w}_C < -a_C < 0$ . Similarly, having made a private threat ( $PR$ ), the critical type with  $w_C = 0$  such that  $C$  stands firm in private ( $SF_1$ ) if  $w_C \geq 0$  and backs down ( $BD_1$ ) otherwise.

Now consider  $C$ 's initial decision. Suppose a unique challenger-type who is indifferent between going public ( $PU$ ) and staying private ( $PR$ ). Let  $k$  denote this unique type's value for war such that all challenger-types with  $w_C \in (k, 1]$  strictly prefer a public threat ( $PU$ ) to a private threat ( $PR$ ), and the types with  $w_C \in [\underline{w}_C, k)$  strictly prefer private threats.

Second, consider  $D$ 's strategy. There exist two critical types  $w_{D,private}^*$  and  $w_{D,public}^*$  who are indifferent between refusing and conceding, upon seeing private and public threats, respectively. Then, higher types with  $w_D \geq w_{D,private}^*$  resist ( $RS_1$ ) and lower types concede ( $CD_1$ ); likewise, higher types of  $D$  ( $w_D \geq w_{D,public}^*$ ) resist ( $RS_2$ ) and lower types ( $w_D < w_{D,public}^*$ ) concede ( $CD_2$ ).

As such, these cutpoints partition  $C$ 's possible types into four ranges, and  $D$ 's into two. Note, however, that all ranges of types need not exist for all possible combinations of parameters. Below, I discuss the public and private equilibria in turn.

### 3.1. The Public Equilibrium: A Standard Audience Costs Story

I first characterize the *public equilibrium* in which a private threat is irrelevant in the course of crisis bargaining: the status quo always prevails as a result of a private threat. Upon seeing a private threat,  $D$  detects that the threat is a bluff and, in response,  $C$  backs down without being caught by his domestic audience. As a result, a challenging leader must go public in order to make his threats credible. Behavior in this equilibrium turns out to be equivalent to the crisis dynamics generally captured by a standard audience cost story where signals are sent only in public before one-sided domestic audience (e.g., Fearon 1997; Ramsay 2004; Schultz 1998; Smith 1998). Proposition 1 establishes the condition under which this is the case.

***Proposition 1 (Going Public)*** *If  $C$  values fighting war less than retaining the status quo, then he must go public to communicate information in the game specified above.*

All proofs are in the Appendix. Proposition 1 shows the sufficient condition for the public equilibrium to exist.<sup>7</sup> If the issue at stake is not worth fighting over, then it may not be possible to convey private information through private channels and hence  $C$  must appeal to the public and expose himself to domestic political costs in order to demonstrate his resolve through making a threat.

**Inconsequential private threats:** One natural corollary of Proposition 1 is that  $C$  can never benefit from making a private threat in this equilibrium. This result may raise skepticism about my model specification that excludes the possibility that  $C$  retains the status quo at the onset of the crisis game. Corollary 1.1, however, shows that the types who would retain the status quo are a subset of those types who back down from a private threat in the public equilibrium.

***Corollary 1.1.*** *Going private (PR) weakly dominates retaining the status quo (SQ) at the onset of crisis in the public equilibrium.*

---

<sup>7</sup> Note, however, that the proof of Proposition 1 establishes necessity for the public equilibrium:  $k < 0$ . Because  $k$  is a cutpoint on  $w_C$ , when the upper bound on  $w_C$  is 0,  $k$  must also be bounded above by 0. This implies that the public equilibrium can exist even if  $w_C \geq 0$  as long as  $k < 0$ .

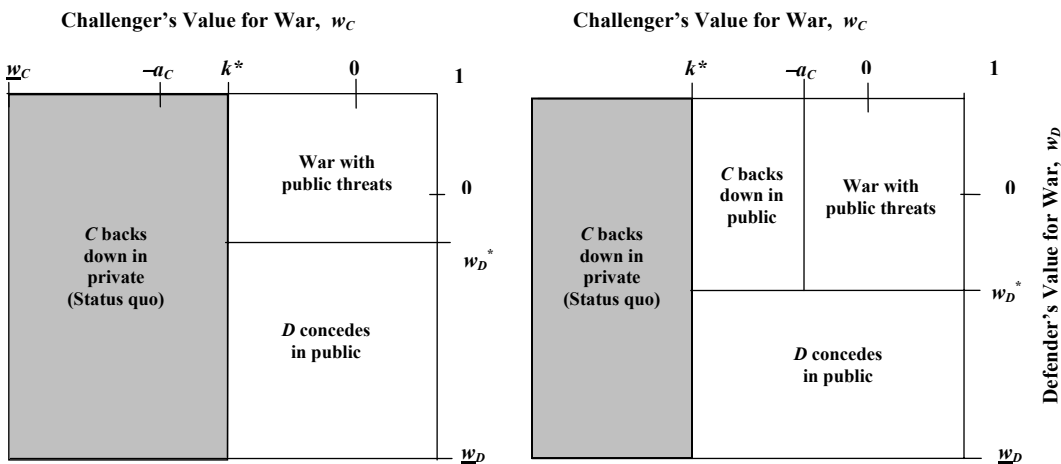
The behavioral implication is that all types who make a private threat will back down eventually because these types have low values for war  $w_C \in [\underline{w}_C, k^*]$ , where  $k < 0$  — values for war are less than the status quo. Anticipating this, whenever  $C$  issues a private threat,  $D$  detects that the threat is not genuine ( $w_C < 0$ ), and therefore she will always resist, to which  $C$  responds by backing down quietly without incurring audience costs. Consequently, as is evident from Figure 2, in the public equilibrium the status quo always prevails as a result of a private threat without affecting both states' status quo utilities. Hence, making private threats in the public equilibrium is inconsequential because doing so is essentially equivalent to retaining the status quo.

On the other hand, once  $C$  has chosen to *go public* about his threat, the crisis dynamics generally remain the same as a conventional audience cost story (e.g., Fearon 1997; Schultz 2001). By resorting to public threats,  $C$  can tie his hands and thereby enhance the credibility of his threats since doing so attracts the attention of his domestic audiences. But going public may have some perverse side effects that increase the *ex ante* risk of inefficient outcomes such as fighting unwanted wars and public concessions, because public threats create audience costs that  $C$  would suffer *ex post* if he backed down. As such, the next result establishes the well known effect of audience costs on the crisis dynamics.

**Corollary 1.2.** *In the public equilibrium, bluffing may occur when  $a_C < F_D(-a_D)/(1 - F_D(-a_D))$ .*

When audience costs for  $C$  are high, a half-hearted challenger will shy away from public commitment, and hence a public threat allows full communication of information. When audience costs are low, however, weaker types may bluff and run a risk of costly backing down in public, because low audience costs generate an incentive to gamble for irresolute types with  $w_C \in [k^*, -a_C]$ . Hence, as  $a_C$  decreases, irresolute types are more likely to make public threats and thereby public threats convey less information, leaving greater uncertainty about  $C$ 's true type.

**Figure 2. Public Equilibrium**



(a) No bluffing case when  $-a_C < k$

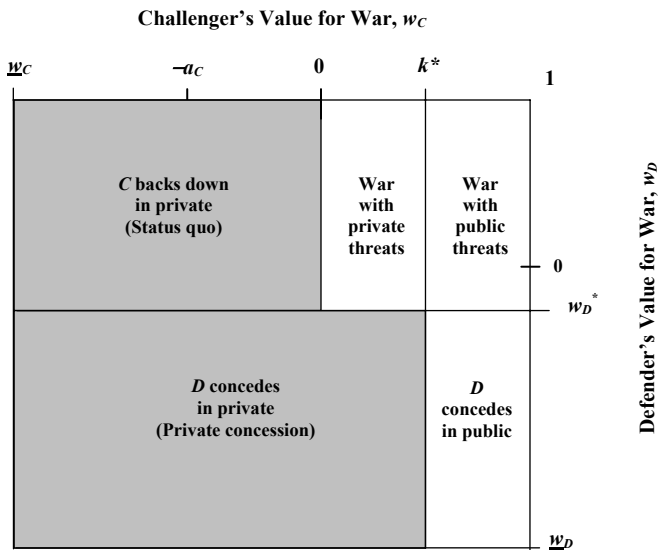
(b) Bluffing case when  $k < -a_C$

**The second audience and its informational role:** It should be stressed that the introduction of audience costs for  $D$  gives rise to a new result in a conventional audience cost story regarding the informational role of public threats. A threat is *informative* if it reduces  $D$ 's uncertainty as to whether  $C$  will follow through on his threat.

**Corollary 1.3.** *A public threat becomes more informative as audience costs for  $D$  get larger. Further, a public threat is fully informative when  $a_D \geq -F_D^{-1}(a_C/(1 + a_C))$ .*

When audience costs of public concessions are invoked,  $D$  finds it difficult to concede. Higher audience costs make it harder for a larger range of  $D$ 's types to resist a public threat. Facing a higher probability of refusal, only resolute types of  $C$  can afford to make public threats. Accordingly, imposing audience costs on  $D$  improves the ability of higher types of  $C$ , for which  $w_C \geq \max\{k^*, -a_C\}$ , to distinguish themselves from lower types. This will cause less bluffing than if there were no audience costs for  $D$  (i.e.,  $a_D = 0$ ) as is the case with the existing models with one-sided audience costs. In fact, Corollary 1.3 shows that bluffing will never occur when  $D$ 's audience costs are large enough (i.e.,  $a_D \geq -F_D^{-1}(a_C/(1 + a_C))$ ). Hence, when  $D$  suffers from greater audience costs *ex post*, resolved challengers can convey greater credibility.

**Figure 3. Private Equilibrium**



Note: The shaded area represents peaceful outcomes due to private threats.

#### 4. The Private Equilibrium: “Speak Softly and Carry a Big Stick”

Going public about a threat and creating *ex post* domestic audience costs is a well recognized signaling technique that helps leaders turn their threats into credible signals. Consequently, when signals are sent privately without appealing to audiences, it would be difficult for *C* to convey private information because the absence of domestic consequences of their foreign policy choices complicates the credibility problem, as signals become costless and non-binding.

However, there exists another equilibrium to the crisis game, in which under certain conditions a non-binding, private threat can reliably communicate information without invoking audience costs to improve its credibility. I shall call this case the *private equilibrium*. In this equilibrium, a private threat can expand the range of peaceful settlements that are mutually acceptable because it does not involve costly public contests that may lead to costly backing down or costly fighting of an unwanted war, as is the case with a public threat.

##### 4.1. How Private Signals Work

We first characterize the private equilibrium to show how private threats work. Proposition 2 suggests that the private equilibrium is pooling over “private” types (i.e.,  $w_C \in [\underline{w}_C, k^*]$ ), but separating over “public” types (i.e.,  $w_C \in [k^*, 1]$ ).<sup>8</sup> While all public threats are genuine, there exists a range of challenger types who make a credible private threat. Figure 3 summarizes the equilibrium outcomes.

**Proposition 2 (Going Private).** *Suppose (i)  $-a_C \leq 0$ , and (ii)  $k > 0$ . Then in equilibrium  $C$  makes a public threat if  $w_C \geq k^*$  and a private threat otherwise. Upon receiving a threat, whether it be public or private,  $D$  resists if  $w_D \geq w_{D,private}^* = w_{D,public}^*$ . In response to  $D$ 's refusal, all types who have made a public threat will stand firm; the types who have issued a private threat will stand firm if  $w_C > 0$  and will back down otherwise.*

**Challenger-types and corresponding strategies:** I first describe  $C$ 's initial decision. Conditions (i) and (ii) of the proposition generate four sets of challenger types, three of which send a private threat. These four sets of types and their corresponding strategies are summarized in Figure 4. First, there is a critical type  $w_C = k^*$  who is indifferent between going public and staying private. Challenger-types who are more resolved than this critical type ( $w_C \geq k^*$ ) always go public and never back down from their public threats. These types are so willing to fight that they have no interest in diplomatic solutions whatsoever. I label this set of types “hard-liner.” When Iraqi officials (reportedly) offered a concession through a private contact in an attempt to avert war on the eve of the American invasion in March 2003, the Bush administration surely was not willing to back-channel a deal (*New York Times*, November 6, 2003, A1). The Bush administration's resolve was so high that they “didn't view it [a private contact] as a credible opportunity or credible communication ... because ... [t]he front door was wide open.”<sup>9</sup>

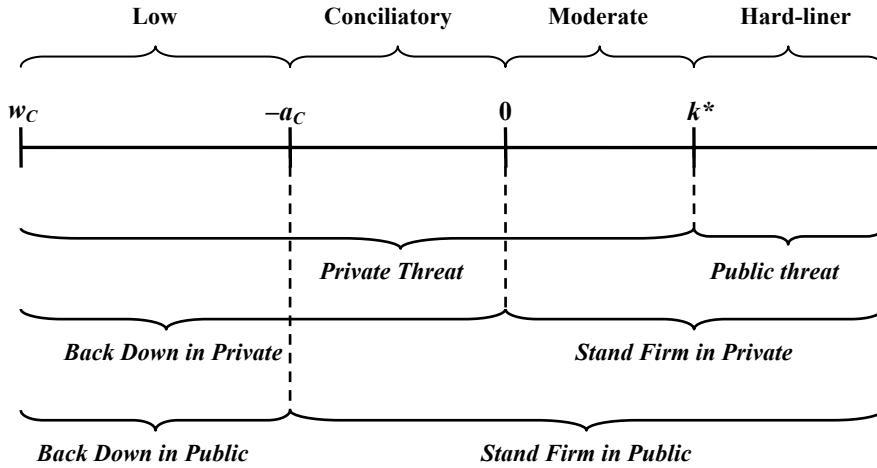
Second, consider challenger-types with  $w_C \in [0, k^*]$  who are less resolved than this critical type

<sup>8</sup> The public equilibrium never pools over “private” types.

<sup>9</sup> Although the Bush administration officially does not either deny or confirm the existence of such a contact, it apparently deemed this backdoor channel not credible.

<http://www.whitehouse.gov/news/releases/2003/11/20031106-5.html>

**Figure 4: Challenger-Types and Equilibrium Strategies from Proposition 2**



( $w_C = k^*$ ) but more resolved than a higher critical type  $w_C = 0$ . Challenger-types in this range make a private threat but stand firm ( $SF_1$ ) if  $D$  fails to make a concession. I call types in this range “moderate” because while their resolve is high enough to stand firm both in public and in private, they will seek a peaceful settlement through a back-door channel that is easier for  $D$  to accept. Given condition (ii), no types in this range would do better by sending a public threat. A historical example of a private threat made by a highly resolved challenger can be seen in the “Melian Dialogue” during the Peloponnesus War. When the Athenians delivered an ultimatum to the Melians, they did so at a private meeting, despite the fact that it was common practice in ancient Greece that diplomatic envoys negotiate at public assemblies (Adcock and Mosley 1975). Despite the norm of public debate, the Melians, out of fear of invoking public outcry, “did not invite [the Athenian] representatives to speak before the people.” When the Melians refused the Athenian demand, the Athenians carried out their private ultimatum and killed the entire male population of the Melians (Thucydides 1972, 400).

The third set of types are in the range  $[-a_C, 0]$  and will make a private threat from which they back down ( $BD_1$ ) in response to a refusal. Recall that the type  $w_C = -a_C$  is indifferent between standing firm ( $SF_2$ ) and backing down ( $BD_2$ ), having issued a public threat. Hence, though types with  $w_C \in [-a_C, 0]$  never issue a public threat, their off-the-equilibrium-path strategy is to stand firm ( $SF_2$ ), if they had to

make a choice between fighting and backing down in front of the audience. I label types in this range “consiliatory.” As the Appendix shows, no types with  $w_C \in [-a_C, 0]$  can profitably deviate by making public threats.

Finally, the fourth set is the “low” types with  $w_C \in [\underline{w}_C, -a_C]$  and they behave in exactly the same manner as the “consiliatory” type in equilibrium: stay private and back down. But these types behave differently off the equilibrium path: they would back down if they had issued a public threat. Thus, all types whose value for war is less than  $k^*$  will send a private threat.

**Defender’s equalization:** Now consider  $D$ ’s response to  $C$ ’s decision on whether to go public or stay private.  $D$ ’s equilibrium strategy is a function of her beliefs, which in turn depends upon  $C$ ’s signal. When it comes to deciding whether to refuse  $C$ ’s demand,  $D$ ’s critical concern is the probability that  $C$  will stand firm in response to her refusal. When  $D$  sees public threats, she knows for certain that  $C$  will stand firm ( $SF_2$ ) because the only types that issue public threats in this case are those with  $w_C > k^*$ . Hence,  $D$ ’s posterior belief that the threat is genuine is one, as long as the threat is public. On the other hand, when  $D$  sees a private threat, she is uncertain about whether the threat is genuine, because, as condition (ii) of Proposition 2 suggests, there exist some types who prefer fighting over the goods to giving them up ( $w_C > 0$ ). Hence, upon seeing a private threat,  $D$  revises her assessment of  $C$ ’s type and believes it is distributed in  $[\underline{w}_C, k^*]$ .

Given these beliefs,  $D$  resists if and only if doing so yields a higher expected payoff than conceding. This condition generates a unique cutpoint strategy for  $D$  as defined earlier. Given this cutpoint structure,  $D$  picks her (public and private) refusal rates so that challenger-type with  $w_C = k^*$  is indifferent between public and private threats.

Interestingly, this yields  $D$ ’s optimal rates of refusal such that  $r_{public}^* = r_{private}^* = 1 - F_D(-a_D)$ , meaning that in equilibrium she refuses a public and private challenge at the same rate. The force driving  $D$ ’s equalization behavior is simply her optimal reaction to the threats given her audience-cost constraints. And yet, as I will next show, the implication of this equalization is not as trivial as it seems.

## 4.2. Why Private Signals Work

Having observed that  $D$  concedes to some private threats, we now address the next question: *Why* do private threats work? The credibility problem is a major obstacle to convey information through private communication. It is tempting to conjecture that some conditions render private threats more informative and hence credible. To the contrary, however, I show that credibility is not necessary for private threats to work. Indeed,  $D$  revises downward her beliefs that  $C$  will follow through on his private threats. Rather, the private equilibrium is governed by the curious behavior of the “moderate” and “consiliatory” types of  $C$ , both on and off the equilibrium path, which induces  $D$  to equalize her rates of refusal in public and in private because of her audience costs.

**Reduced credibility and private concessions:** Recall from Figure 4 that although the “moderate” and “consiliatory” types of  $C$  in the range  $[-a_C, k^*]$  will never have to choose between public backing-down and public standing-firm in equilibrium, their off-the-equilibrium-path strategy is to stand firm in public, if they have gone public. That is, the types in this range could have gone public to enhance the credibility of their threats by tying their hands, as these types are willing enough to stand firm in public (because  $-a_C < 0$ ). Nonetheless, neither “moderate” nor “consiliatory” types tie their hands in equilibrium; in fact, they forgo the credibility enhancing device and seek a more difficult communication medium: a private threat.

The conscious choice of a more difficult communication method by challenger-types in this range leaves  $D$  puzzled regarding  $C$ 's true type. This is because the “moderate” and “consiliatory” types reduce the credibility of their threats by forgoing a public threat and instead going private, and this curious signaling behavior compels  $D$  to revise downward her prior beliefs. When  $D$  sees a private threat, she believes with a greater probability that the threat is *not* genuine, because many of the resolved types with  $w_C > -a_C$  would not have made a private threat. The Appendix shows that the posterior belief that the private threat is genuine is indeed lower than the prior  $q_{private} = (F_C(k^*) - F_C(0))/F_C(k^*) < p_{private} = 1 - F_C(0)$ .

However, this loss of credibility is compensated by the fact that more types of  $D$  will concede to a private threat. This result is driven by the fact that  $C$ 's choice to forgo public signaling makes a private concession possible, and making private concessions incurs no audience cost for  $D$ . Also note that there is a positive *ex post* probability that private threats are credible (i.e.,  $q_{private} > 0$ ). Hence, when  $D$  receives a private threat, she can either concede privately for a certain payoff of 0 or resist and gamble that  $C$  is bluffing. When the threat is a bluff ( $w_C < 0$ ), this gamble pays off; when the private threat is genuine ( $w_C \in [0, k^*]$ ), the refusal results in costly fighting. This strategic calculus then forces unresolved types of  $D$  with  $w_D < w_{D, private}^*$  to concede to some private threats because they cannot afford to gamble.

Accordingly  $D$  finds it optimal to resist at the same rates in public and in private ( $r_{public}^* = r_{private}^*$ ), so that the critical type of  $C$  ( $w_C = k^*$ ) is indifferent between making public and private threats. Unless equalized, some types of  $C$  may have an incentive to deviate from a public threat to a private threat or vice versa, and consequently, complete pooling or complete separating may occur and the private equilibrium will collapse. As a result,  $D$  ends up making a concession to non-credible private threats issued by “consiliatory” or “low” types, as well as credible private threats by the “moderate” types.

**Effective deterrence without informational efficacy:** A private threat can be thought of as “cheap talk” in the sense that it has no immediate domestic consequences (or payoffs) on either side. But the source of rationality lies *not* in its informational role affecting  $D$ 's beliefs, as it does in the Crawford-Sobel (1982) tradition of cheap talk models. Indeed, private threats have only limited capacity to convey  $C$ 's resolve. In the private equilibrium when  $D$  observes a public threat she knows for certain that  $C$  is resolved, whereas when she observes a private threat she assigns a fairly low probability that  $C$  is resolved. As such, private threats are less *informative* than public signals. To summarize this claim, following Sartori (2005) I now define two properties of a threat.

*Informational efficacy:* A threat has informational *efficacy* if  $D$ 's posterior belief about credibility of the threat is greater than her prior, upon receiving the threat. Formally, a threat  $j$  has informational efficacy if  $q_j > p_j$ .

*Effectiveness:* A threat is more *effective* if, upon receiving a threat,  $D$  concedes at a higher rate than otherwise. Formally, a threat  $j$  is more effective than threat  $i$  if  $r_j > r_i$ .

The next propositions establish that a private threat achieves effective deterrence despite its lack of informational efficacy.

**Proposition 3.1 (Informational Efficacy).** *A private threat does not have informational efficacy either in the public or private equilibrium, while a public threat is always efficacious in both equilibria.*

**Proposition 3.2 (Effectiveness).** *A private threat is equally effective as a public threat ( $r_{private}^* = r_{public}^*$ ) in the private equilibrium, while it can never be effective ( $r_{private}^* = 1$ ) in the public equilibrium.*

Propositions 3.1 and 3.2 state that, despite the informational inefficacy, a private threat is no less effective than a public threat in the private equilibrium. This is because  $D$  finds it optimal to resist at the same rate given both public and private threats, as Proposition 2 demonstrates. However, if  $C$  wanted to convey greater credibility of his threats, he could go public because public threats always have informational efficacy, while private threats never do. This result suggests that *credibility may not at all be necessary for successful deterrence*. But how is this possible?

**Second audience and customized signals:** The driving force behind these dynamics is the audience cost for  $D$ . Recall that public threats invoke audience costs not only for  $C$  but also for  $D$ . Then, the “moderate” and “consiliatory” types of  $C$  ( $w_C \in [-a_C, k^*]$ ) find that while public threats convey greater information, this informational benefit of public threats is counteracted by the greater probability that  $D$  will have to resist in order to avoid suffering audience costs from making a public concession.

This is proven in the next section (Proposition 4) by the fact that the private equilibrium does not exist when  $D$  has no audience costs ( $a_D = 0$ ).

$C$ 's choice to forgo the credibility-enhancing device by going private implies  $C$ 's ability to tailor a signal depending on his own type as well as the identity of the audience.<sup>10</sup> This ability, in turn, not only allows the lower type of  $C$  to back down quietly without incurring audience costs, but also allows  $D$  to realize that some private threats are indeed credible, and eventually allows  $D$  to make concessions without losing her face with the domestic audience.

Because the “moderate” and “consiliatory” types of  $C$  will not pay audience costs in equilibrium, the only reason for these types to create audience costs *ex post* by going public is to convey credibility. Then the only reason for these types not to go public but instead to stay private is to avoid provoking  $D$ 's domestic audience. This is because provoking  $D$ 's domestic audience with a public threat will lock her into resisting, which in turn will lock the higher types of  $C$  into costly fighting. Understanding this *implicit* signal of  $C$ 's likely intention in a manner of “forward induction,”  $D$  assigns a positive probability to the possibility that  $C$  is willing to follow through on his threat even when she receives a private threat. Consequently, private threats have some limited credibility. Despite this limited credibility, a private threat precludes some wars that would be caused because  $D$ 's honor has been raised due to her audience costs. This is the mechanism that makes private threats work.

#### **4.3. When Private Signals Work.**

Having described how and why a private threat works, we now consider *when* this is the case. Proposition 4 establishes the conditions under which the private equilibrium exists.

**Proposition 4.** *The private equilibrium exists if and only if the following conditions are met: (i) the public option is also credible ( $k^* > -a_C$ ); (ii) a viable domestic audience exists on  $D$ 's side, while it need not exist*

---

<sup>10</sup> Work in economics obtains analogous results that cheap talk messages cannot be informative in a single-audience situation (Farrell and Gibbons 1989, 1216-17; see also Gertner et al. 1988; Ramsay 2003).

on  $C$ 's side ( $a_C \geq 0$ ,  $a_D > 0$ ); and (iii) the defending leader is reasonably accountable to her domestic audience ( $a_D \geq F_C(0)/(1 - F_C(0))$ ).

**Speak Softly and Carry a Big Stick:** The first condition suggests that a public threat must also be credible always. Technically, it must be the case that  $k^* > -a_C$ , which is a natural corollary of conditions (i) and (ii) of Proposition 2. This result guarantees that it is only “hard-liner” types with  $w_C \in [k^*, 1]$  who make public threats, and therefore every public threat in the private equilibrium is always credible. This condition implies that a private threat can be effective only *in the shadow* of a credible public signal.<sup>11</sup> This is why President Theodore Roosevelt had to carry a big stick when he spoke softly.

**Two-sided Audiences and the Hidden Equilibrium:** The second condition of Proposition 4 states that a domestic audience must exist in the defending country ( $a_D > 0$ ) for a private threat to work, while it is not necessary on  $C$ 's side ( $a_C \geq 0$ ). On the one hand, even if  $C$ 's leader does not incur audience costs ( $a_C = 0$ ) when he backs down in public, the private equilibrium does *not* collapse. On the other hand, however, if the defending leader does not incur audience costs ( $a_D = 0$ ) in the event of public backing-down, the private equilibrium collapses. This is because  $a_D = 0$  implies  $F_C(0) = 0$ . In words, if there is no audience cost for  $D$ , “consiliatory” and “low” types of  $C$  with  $w_C \in [\underline{w}_C, 0]$  will be washed out. But this, in turn, generates unfeasible beliefs about  $C$ 's types that cannot be sustained in equilibrium. Hence, there must be multiple audiences for a private threat to work; otherwise  $C$  is unable to customize his signals.

The observation that the private equilibrium collapses when  $a_D = 0$  explains why existing audience cost models conclude that private signals can never make a difference (e.g., Fearon 1997; Schultz 1998; Ramsay 2003). Setting  $D$ 's audience costs equal to zero effectively changes the present assumption of two-sided domestic audiences to that of a one-sided audience, and as a result the private equilibrium collapses. Thus, the private equilibrium proposed here cannot be found in the earlier models that assume away audience costs for the receiver of signals, while imposing them only on the signaler.

---

<sup>11</sup> This result is analogous to Austen-Smith and Banks' (2000) result that the availability of costly signals (i.e., burned money in their model) makes cheap talk informative.

**Accountable leaders and private concessions:** The last condition states that the defending leader must pay reasonably high audience costs in the event of a public concession in front of her audience. Technically, it must be the case that  $a_D \geq F_C(0)/(1 - F_C(0))$ , which is implied by the fact that  $D$ 's value for war  $w_D$  is bounded above by 1. Substantively, this result suggests that  $C$  does not have leverage in private diplomacy unless  $D$  is politically accountable to her domestic audience (at least to some degree), while  $C$  himself need not be politically accountable to acquire that leverage.

This is why President Theodore Roosevelt won a concession from Canada through private letters. In response to Canadian Prime Minister Laurier's plea to save face with his domestic audience during the Alaskan Boundary dispute in 1903, Roosevelt agreed to appoint an international tribunal to camouflage the apparent surrender to American threats, while he sent troops quietly and sent private letters containing an ultimatum. This made it easier for Laurier to concede the territory to the United States, as he was sensitive to domestic costs of a public concession (Penlington 1972, 62-63).

Conditions (ii) and (iii) highlights the key to the private equilibrium:  $D$ 's sensitivity to her audience costs. As long as  $D$  incurs audience costs when conceding in public, a private concession becomes preferable for  $D$ , and consequently  $C$  can influence  $D$ 's beliefs by customizing his signals.

### 5.1. Efficient Secrecy

I have shown that there exist two distinctive equilibrium mechanisms, through which leaders signal their private information in international crises: going public and staying private. The multiplicity of the equilibria suggests that state leaders may have freedom to choose one mechanism over another as their crisis management strategy. Given this, the question to ask is this: Which mechanism should leaders choose? I address this question by comparing the welfare values of the two equilibria.

Proposition 5 presents a simple *efficiency* result. From both an *ex ante* and an *interim* perspective (Holmström and Myerson 1983), the private equilibrium is *Pareto superior* to the public equilibrium. In particular, the private equilibrium weakly *interim dominates* the public equilibrium, in which all challenger types, except "hard-liner" types with  $w_C \in [k^*, 1]$ , strictly prefer the private equilibrium. On

the other hand, the “hard-liner” types are indifferent between the public and private equilibria, because their values for war always exceed their audience costs so that the “hard-line” challenger has no concern of public backing-down.

***Proposition 5 (Efficient Secrecy).*** *For all types, the private equilibrium is ex ante efficient. Furthermore, for all challenger-types with  $w_C \leq k^*$  and for any defender-types, the private equilibrium (strictly) interim dominates the public equilibrium.*

This efficiency result demonstrates that both  $C$  and  $D$  are always better off with the private equilibrium than the public equilibrium (at least weakly) no matter what types they are. Intuitively, a private threat can expand the range of peaceful settlements that are mutually acceptable. Comparing Figures 2 and 3, it is obvious that peaceful outcomes are possible under broader conditions in the private equilibrium than in the public equilibrium where only public threats are credible.

This result is driven by two facts. First, private threats can produce the equilibrium outcome—a private concession—that cannot arise in the public equilibrium. A private concession is possible only in the private equilibrium (Figure 3) and is impossible in the public equilibrium (Figure 2). Second, this result is also driven by the reduced probability of war due to public threats. The next result establishes the effect of private threats on the probability of war.

***Corollary 5.1 (Risk of War).*** *The ex ante probability of war is strictly greater in the public equilibrium than in the private equilibrium.*

This result implies that resorting to private threats decreases the *ex ante* probability of war—that is, the probability of war prior to nature’s draw of  $w_C$  and  $w_D$ . This result is driven by the fact that  $D$ ’s probabilities of refusal are the same in the public and private equilibria. While in the private equilibrium war could result not only from public threats but also from private threats, as Figures 2 and 3 indicate, the

risk of war is limited to the higher range of  $w_C$  in the private equilibrium (i.e.,  $w_C \in [0, 1]$ ), whereas the war outcome can be realized in the lower range of  $w_C$  in the public equilibrium (i.e.,  $w_C \in [k^*, 1]$ , where  $k^* < 0$ ).<sup>12</sup> Given this, the expected probability of war across the entire range also falls.

There are clear advantages to making private threats for both  $C$  and  $D$ . On the one hand, private threats help  $C$  escape from costly backing-down and costly fighting. Because staying private allows  $C$  to pretend as if nothing had happened, secret diplomacy can secure leeway to disavow the threat if  $D$  refuses to concede. Consequently, this leeway enables  $C$  to avoid domestic political costs if he has no choice but not back down, and thus avoid being “locked into” fighting when the domestic audience costs are too high.

On the other hand, private threats help  $D$  escape from costly public concessions and therefore it may help her avoid fighting an unwanted war in order to defend her honor from public humiliation. A *private concession* is attractive for  $D$  mainly because it in fact lowers the costs of conceding by concealing the identity of the state offering a concession (O’Neill 2003) or the compromise itself (Fearon 1992, 127). Thus, a private threat allows  $D$  to rationally make a concession, which can never be attainable in public diplomacy where a concession may involve domestic political costs. Historical norms are that state leaders employ secret threats as a means of face-saving tactics.<sup>13</sup> For example, during the final phase of the Cuban Missile Crisis, President Kennedy wanted to make sure that “Every opportunity was to be given to the Russians to find a peaceful settlement which would not diminish their national security or be a political humiliation” (Kennedy 1969, 81).

Contrary to the popular perception that transparency or “open diplomacy” carries beneficial effects in the age of democracy (Nicolson 1963; Finel and Lord 1999), *Efficient Secrecy* posits that the private equilibrium is a more valuable mechanism for all types of any players, as it can lead to better bargaining outcomes than the public equilibrium can. State leaders, hence, cannot rationally ignore a

---

<sup>12</sup> For the public equilibrium, the expected probability of war is calculated in the separating case, because the risk of war is strictly higher in the semi-separating case.

<sup>13</sup> See Snyder and Diesing (1977, 257) for other instances that involve face-saving gestures.

private threat simply because it is “cheap” talk. Secrecy in diplomacy may not only be rational but also efficient.

## 5.2. Rational Diplomacy

The contemporary literature of international relations has downplayed the role of diplomacy in shaping international outcomes until recent years (Sartori 2002; Guisinger and Smith 2002). In particular, as Sartori (2002, 121) points out, the standard rationalist explanations of war implies that “normal forms of diplomatic communication may be worthless” in international relations, because they are cheap talk and non-binding (Fearon 1994, 578). However, this conclusion contradicts with the fact that for centuries states have invested much time and energy into diplomacy.

Historically, modern diplomatic institutions were created in response to the security dilemma caused by uncertainty, as a stable communication system between states (Mattingly 1955, 51-76).<sup>14</sup> Secret communication has been the norm of diplomacy ever since ancient times<sup>15</sup> Even with the advent of mass democracy in the 19th century or the post World War I Wilsonian demand for “open” diplomacy, secrecy has not ceased from being the central feature of diplomacy (Jönsson and Hall 2003; Nicolson 1954). The model suggests that such secret communication can be rational in crisis diplomacy.

Yet the rationality of diplomacy stems not so much from its informational benefits as from its less provocative nature, even though its historical development suggests that one of the principal functions of diplomacy is communication among states (e.g., Jönsson and Hall 2003, 195-96). As Propositions 3.1 and 3.2 suggest, private diplomacy is as equally effective as public provocative measures in convincing the adversary to concede, despite the fact that private diplomacy never has informational efficacy, while public signals always does.

---

<sup>14</sup> Modern diplomatic institutions were formulated as the “Italian system” during the Renaissance and established as the “French system” during the reign of Louis XIV (Berridge 2002, 107; Nicolson 1954, 53-61). The rudiments of what we now know as diplomacy can be traced back to Ancient Greece (Adcock and Mosley 1975) as well as the Ancient Near East (Cohen and Westbrook 2000).

<sup>15</sup> The only exception is Ancient Greece, where diplomatic envoys had to report to public assemblies and argue in public (Jönsson and Hall 2003; also see Adcock and Mosley 1975).

Precisely for this reason, state leaders would always have to rely on costly signals such as troop mobilization and public threats that generate a real risk of inefficient war, if their sole purpose were to enhance the credibility of their threats (Fearon 1994, 579).<sup>16</sup> In this sense the result here is consistent with the existing rationalist view in that “quiet diplomatic exchanges may be insufficient to allow states to learn what concessions an adversary would ... be willing to make” (Fearon 1994, 586).

I argue, however, that the conventional conclusion about diplomacy overlooks the very nature of diplomatic institutions: by definition, the primary objective of diplomacy is “the promotion of the national interest *by peaceful means*” (Morgenthau 1956, 505). Indeed, the model offers such a view on diplomacy: private diplomacy can achieve the same policy goal with equal effectiveness but without increasing the risks of costly fighting or costly diplomatic humiliation.

Private diplomacy can be rational because audience costs are not provoked for the adversary as long as the demands and threats remain private, and such secrecy does not lock the adversary into a situation where she has no choice but to stand firm. This non-provocative nature helps to overcome private diplomacy’s limited ability to convey information.

The mechanism of private diplomacy hinges on the adversary’s sensitivity to audience costs that leaders would suffer *ex post* in the event of a public concession. Indeed, as Proposition 4 demonstrates, the assertion that quiet diplomacy does not make a difference holds only if  $D$  suffers no political costs from diplomatic humiliation ( $a_D = 0$ ), as the existing rationalist explanations assume. The presence of a viable domestic audience for  $D$  disciplines  $C$ ’s communication so as to customize the signals to save face. An important implications here is that the benefit from face-saving afforded by private concessions is a source of rationality for private threats and secret diplomacy more generally.

Furthermore, this logic may explain why secrecy is a persistent feature of the so-called “French system” of diplomacy (Berridge 2002, 107; Nicolson 1954, 75). As Proposition 5 suggests, private

---

<sup>16</sup> Diplomatic historians seem to concur on the informational benefits of military threats in coercive diplomacy. For example, Paul Gordon Lauren (1994, 25) argues that “despite its inherent dangers, this extreme variant of coercive diplomacy [an ultimatum] conveys resolve and urgency better than, say, an ambiguously worded diplomatic protest.”

diplomacy is not only rational but also efficient. As Louis XIV observed, “Open negotiations ... incline negotiators to consider their own prestige and to maintain the dignity ... with undue obstinacy and prevent them from giving way to the frequently superior arguments of the occasion” (quoted in Nicolson 1954, 61).<sup>17</sup>

## **5. Alaska Boundary Dispute, 1903: A Successful Private Threat**

Before concluding this paper, it is useful to see how the logic of *efficient secrecy* operates in an actual case. To illustrate my theoretical claims about when, why, and how a private threat works, I present a historical episode from the Alaska Boundary Dispute in 1903, where President Theodore Roosevelt’s “speak softly and carry a big stick” foreign policy was most evident.

The dispute occurred in 1902 when the United States claimed the Canadian territory adjacent to Alaska along the Pacific coast. The original boundary was nothing more than a convenient construct when Britain and Russia concluded a treaty in 1825. The border became of strategic importance when gold was presumed to be discovered in the area. This dispute was eventually resolved in October 1903 in the favor of the U.S.—and the U.S. gained a town now known as Juneau—with an appearance of reasonable compromises, although in fact Canada conceded in the face of Roosevelt’s private threats of waging war.

In March 1902, when Roosevelt was warned of the consequence of the discovery of gold, he decided to send in troops “as quietly and unostentatiously as possible ... to prevent any possible disturbance along the disputed boundary line” (Collin 1985, 174-178). In the meantime, Roosevelt sent a message to Ottawa about the possibility of violence, implicitly challenging Canada with a territorial demand. Knowing that Roosevelt would not pull back the troops from the disputed area, Prime Minister Laurier wanted to make a private concession so as to avoid an apparent surrender of territory to Roosevelt’s threats. So Laurier’s government proposed an arbitrational settlement so that his government

---

<sup>17</sup> This may also help to make sense of why “protocole”—the undue ceremonial diplomatic procedure designed to burnish honor and prestige—is another feature of the French system of diplomacy (Nicolson 1977, 43; Berridge 2002, 107).

could at least save face with Canadians. Laurier reportedly “pleaded to Henry White, the head of the American Embassy, that he would like to ‘save his face’ with Canadians by an arbitration” (Nevins 1930, 192-193; Penlington 1972, 62-63).

In response, the U.S. agreed to appoint a tribunal to review the disputed case in the courts, rather than arbitration as the Canadians hoped (Francis et al. 1992, 116). However, the tribunal was just meant to be a face-saving device to gesture a compromise and to conceal the identity of the party submitted. Roosevelt from the beginning refused to arbitrate this dispute or to consider any decision whatsoever short of complete victory in this case, but he was willing to gesture a compromise if the Canadian side wished the settlement to appear to be a tribunal settlement (Collin 1985, 174-76; Penlington 1972, 62-63). In fact, the composition of the tribunal was designed so that the U.S. could never lose the case.

In June 1903, Roosevelt established a contingency plan to dispatch troops if the tribunal failed to reach a settlement in favor of the U.S. In November, Roosevelt issued a public statement to justify his seizure of Panama if the Panama revolution had not taken place. This statement demonstrated to Britain and Canada that he was highly resolved to fight over the Alaska boundary and his threat was not merely bluffing (Beale 1956, 130). Consistent with the equilibrium behavior of  $C$  of “moderate” types ( $w_C \in [0, k^*]$ ), Roosevelt signaled successfully through these actions that he would try out a private maneuver first but he would also stand firm if the Britain-Canada side refused his demand

As for Britain, the disputed territory was not a great concern as long as a settlement did not deteriorate her relations with the U.S., as her international situation with France, Germany, Japan, and Russia had developed unfavorably by 1903. Thus Britain’s decision was primarily based not on Alaska but on the essential need to maintain friendship and *détente* with the U.S. (Collin 1985, 183; Francis et al. 1992, 115-116; Penlington 1972, 92). As for the Canadian leaders, they really had only two choices, given Roosevelt’s high resolve: either to conclude the tribunal favorably for American case and save face with fellow Canadians, or to surrender territory to the U.S. forces in public and be humiliated.

Roosevelt chose to threaten Britain rather than Canada, because the British representative, Lord Alverstone—Chief Justice of England and President of the Alaska Boundary Tribunal—was a pivotal

voter on the tribunal. The six tribunal members consisted of three Americans, two Canadians, and one Briton. The U.S. kept sending messages to the British leaders to convince them that Britain's "self-interest would be better served by aligning with America rather than Canada" (Collin 1985, 182), and threatened that it would draw the boundary using military force, should they fail to win the American case (Francis et al. 1992, 116; Penlington 1972, 89-90).

Roosevelt notes in retrospect that his personal letter handed to British foreign minister Joseph Chamberlain from Supreme Court Justice Oliver W. Holmes at a private meeting "was one of the decisive elements in the eventual American victory" (Collin 1985, 179-180; see also Munro 1970, 58). Justice Holmes privately met Chamberlain on behalf of Roosevelt to communicate his message: settle or fight. Roosevelt instructed Judge Holmes in a personal letter dated July 25, 1903 that:

"... if you happen to meet Chamberlain ... you are entirely at liberty to tell him what I say, although of course it must be privately and unofficially. ... [I]f there is a disagreement I wish it distinctly understood, not only that there will be no arbitration of the matter, but that ... I shall take a position ... which will render it necessary for Congress to give me the authority to run ... the boundary on my own hook ... as we claim it" (Munro 1970, 56-57).

In the end, the British representative Alverstone sided with the Americans and accepted their position of the boundary and territorial control as the U.S. claimed. Alverstone was reportedly instructed either by the British Prime Minister or the Foreign Secretary to side with the American demands (Penlington 1972, 90-99). In addition, Roosevelt gave Senator Lodge not an 'official and authoritative' letter but a private message to be conveyed to British leaders. Senator Lodge showed this personal letter to Alverstone whose vote was decisive in settling the case with a complete American victory.

On the surface, the dispute appeared to be resolved through a tribunal settlement. But it was actually Britain's acquiescing of Canadian territory under the private threats Roosevelt repeatedly issued. During the course of the Alaska Boundary dispute, Roosevelt spoke softly by publicly holding a tribunal, but he carried a big stick by quietly dispatching troops and privately blackmailing them. That way, Roosevelt made it easier for Laurier to surrender the territory.

## 6. Conclusions

This study is a natural extension of the audience costs story (Fearon 1994, 1997; Smith 1998; Ramsay 2004; Schelling 1960; Schultz 1998, 2001). As noted at the outset, much of the literature on international crises and disputes has developed to explain why states take costly actions in public during a crisis. With a few exceptions (Baum 2004; Leventoğlu and Tarar 2004; O'Neill 2003), research in this area has not addressed questions as to when and why state leaders sometimes go private in the course of international bargaining. This paper addresses this shortcoming by extending the audience costs logic beyond its original concern. In particular, it relaxes two assumptions commonly held by standard models of crisis bargaining with audience costs (e.g., Fearon 1997; Schultz 2001). The resulting game shows that there exists a previously unknown equilibrium in which a private signal carries some limited credibility. It also shows that a private threat can win a concession from the adversary equally effectively as a public signal, and it also enhances the states' overall ability to resolve disputes peacefully.

The main contribution of this paper is in identifying the mechanism of private threats and demonstrating its rationality. The rationality stems not from its informational advantage but from its less provocative nature. Private threats help the adversary to make a concession by permitting the defender to save face with her public when doing so. The fact that private threats can be rational as well as efficient under reasonable conditions may provide a logical underpinning for apparent predominance of secrecy in diplomacy in general despite the popular perception that secrecy is socially inefficient.

On the other hand, the standard audience cost logic is at work for public threats. The rationality of public threats comes from the fact that the very action of signaling increases the risk of inefficient outcomes such as war. This is due to a dual role that a public threat plays: it enhances the states' ability to communicate their resolve with the adversary in a crisis; but it also makes it harder for the defender to concede. The driving force behind this logic is the fact that going public with military threats provokes domestic audience costs for both states in a crisis.

Contrasting the mechanisms of private threats to that of public ones demonstrates that the audience cost story can be extended to explain a much wider range of state behavior than originally

envisioned. While a standard logic was developed to explain publicly demonstrated military coercion (Fearon 1994; 1997), it can also be extended to explain privately conducted diplomatic maneuvers.

What eventually emerges from this insight is a rationale for “normal” diplomacy as opposed to “coercive” diplomacy. The equilibrium logic developed here may account for why state leaders have maintained the current form of diplomatic institutions for (at least) three centuries. As such, this paper is part of a growing set of formal models that investigate the role of diplomacy in conflict resolution in order to fill the gap between the empirical facts and theoretical implications of diplomacy (Guisinger and Smith 2002; Sartori 2002). Although the current diplomatic institutions were formulated as a stable communication system among city-states in Renaissance Italy, the rationality of private diplomacy resides not so much in its informational benefits but in its secrecy and its face-saving function.

Perhaps more interesting is that I derive the rationale for diplomacy from a standard rationalist framework (i.e., the audience cost logic) that previously downplayed the role of quiet diplomatic communication. I agree with the rationalists (e.g., Fearon 1995) and diplomatic historians (e.g., Lauren 1994) that quiet diplomatic communication is less informative than provocative public confrontation. I argue, however, that informational inefficacy of private signals does not directly translate into the irrelevancy of diplomacy. This paper establishes this claim by identifying the private equilibrium where informational efficacy is not necessary for a private threat to work, and by demonstrating the efficiency of secrecy in crisis diplomacy. Diplomacy is not irrelevant in conflict resolution, but is rational. States are always better off with quiet normal diplomacy than public and provocative coercion. State leaders, hence, cannot rationally ignore a private threat simply because it is “cheap” talk.

## **A. Appendix**

This appendix presents proofs of the propositions and corollaries. The solution concept is perfect Bayesian equilibrium (PBE), which requires that the strategies of  $C$  and  $D$  must maximize their utility, given the other’s strategy and their beliefs. The beliefs must be consistent with equilibrium strategies of  $C$  and  $D$ , and determined by Bayes’s rule if possible.

**Proof of Proposition 1.** Proposition 1 establishes sufficiency. To prove necessity, consider  $C$ 's cutpoint-strategy. Then, it is obvious that if  $k < 0$ ,  $w_C \in [\underline{w}_C, k] \cap w_C \in [0, 1] = \emptyset$ . Hence, having made a private threat ( $PR$ ), there is no type who will stand firm in private ( $SF_1$ ) in equilibrium. From this, if  $k < 0$ , it must be the case that  $q_{private} = 0$ , where  $q_{private}$  denotes  $D$ 's posterior belief that  $C$ 's private threat is genuine (i.e.,  $w_C \geq 0$ ). Since  $q_{private} = 0$ ,  $PR$  cannot be credible in equilibrium.

It remains to be shown that a public threat ( $PU$ ) conveys information. Since  $w_C \in [k, 1] \cap w_C \in [0, 1] \neq \emptyset$ , there must exist some types who issue  $PU$  and stand firm in public. Thus,  $D$  is uncertain about whether a threat is genuine when she receives  $PU$ . Let  $q_{public}$  denote  $D$ 's posterior belief that  $PU$  is genuine ( $w_C \geq -a_C$ ) such that  $q_{public} \equiv \Pr(w_C \geq -a_C \mid w_C \geq k)$ . There are two cases, depending on the relative magnitudes of  $k$  and  $-a_C$ .

Case 1:  $k < -a_C$ . Using Bayes's rule,  $q_{public}$  is given by  $q_{public} = \frac{\Pr(w_C \geq -a_C)}{\Pr(w_C \geq -a_C) + \Pr(k \leq w_C < -a_C) + \Pr(w_C < k)}$

$\Leftrightarrow \frac{1 - F_C(-a_C)}{1 - F_C(-a_C) + [F_C(-a_C) - F_C(k)]}$ , which reduces to

$$\frac{1 - F_C(-a_C)}{1 - F_C(k)}. \quad (1)$$

Since  $k < -a_C < 0$ , the prior probability that  $PU$  is genuine is  $p_{public} \equiv \Pr(w_C \geq -a_C) = 1 - F_C(-a_C)$ . Also since  $F_C(\bullet) \in [0, 1]$ , it follows that  $(1 - F_C(-a_C)) / (1 - F_C(k)) > 1 - F_C(-a_C)$ . Hence, when  $k < -a_C$ ,  $PU$  conveys information.

Case 2:  $k > -a_C$ . Given the cutpoint-structure specified above, when  $-a_C < k < 0$ , it follows that  $w_C \in [k, 1] \cap w_C \in [\underline{w}_C, -a_C] = \emptyset$ . This means that all public threats are genuine, which implies

$$q_{public} = 1 \quad (2)$$

Hence, when  $k > -a_C$ , a public threat always conveys information. **Q.E.D.**

**Proof of Corollary 1.1.** Suppose  $C$  retains the status quo ( $SQ$ ), and the game ends. Then  $C$ 's certain payoff is zero  $U_C(SQ) = 0$ . Note that it is a dominant strategy to back down ( $BD_1$ ) from a private threat

( $PR$ ) if  $w_C < 0$ . Also note that in the public equilibrium, when  $C$  makes  $PR$ ,  $w_C$  must be in the interval  $[\underline{w}_C, k^*]$ , where  $k^* < 0$ . So, given the types in this range, the expected utility of  $C$  from making  $PR$  is given by  $EU_C(PR | w_C \in [\underline{w}_C, k^*]) = 0$ . Thus,  $U_C(SQ) = EU_C(PR | w_C \in [\underline{w}_C, k^*]) = 0$ , and hence in the public equilibrium making  $PR$  weakly dominates retaining  $SQ$  for any value of  $k < 0$ . **Q.E.D.**

**Proof of Corollary 1.2.** To show that  $a_C = F_D(-a_D)/(1 - F_D(-a_D))$  is the cutpoint, below which bluffing may occur, I need to find the equilibrium strategies  $r_{public}^*$  and  $k^*$  for  $D$  and  $C$ , respectively. There are two cases, depending on the relative magnitudes of  $-a_C$ , and  $k$ .

*Case 1:  $-a_C < k$ .* Recall from (2) that  $q_{public} = 1$  when  $-a_C < k$ . Hence, given a public threat ( $PU$ ), sequential rationality requires that  $D$  resist if and only if  $EU_D(RS_2 | PU) \geq EU_D(CD_2 | PU) \Rightarrow w_D \geq -a_D$ . Then, the probability that  $D$  resists in public is given by  $r_{public} \equiv \Pr(w_D > -a_D) = 1 - F_D(-a_D)$ .

Now, I derive an expression for  $k^*$  such that it is sequential rational. In equilibrium  $C$ 's type with  $w_C = k^*$  must be indifferent between  $PU$  and  $PR$ . This condition holds when  $EU_C(PU | w_C = k^*) =$

$$EU_C(PR | w_C = k^*) \Rightarrow (1 - F_D(-a_D)) k^* + F_D(-a_D) = 0, \text{ which implies } k^* = \frac{-F_D(-a_D)}{1 - F_D(-a_D)}. \text{ Since } -a_C < k,$$

for the cutpoint  $k^*$  to constitute a PBE, it must be that  $a_C > \frac{F_D(-a_D)}{1 - F_D(-a_D)} \Leftrightarrow a_D > -F_D^{-1}\left(\frac{a_C}{1 + a_C}\right)$ .

*Case 2:  $k < -a_C$ .* Given  $PU$ ,  $D$  resists if and only if  $EU_D(RS_2 | PU) \geq EU_D(CD_2 | PU) \Rightarrow q_{public} \cdot w_D + (1 - q_{public}) \geq -a_D$ , which implies  $w_D \geq \frac{q_{public} - 1 - a_D}{q_{public}}$ . Then, the probability that  $D$  resists, given  $PU$  is

$$r_{public} \equiv (RS_2 | PU) = \Pr\left(w_D \geq \frac{q_{public} - a_D - 1}{q_{public}}\right) = 1 - F_D\left(\frac{q_{public} - a_D - 1}{q_{public}}\right). \quad (3)$$

Plugging (1) into (3) yields:

$$r_{public} = 1 - F_D\left(\frac{F_C(k^*) \cdot (1 + a_D) - F_C(-a_C) - a_D}{1 - F_C(-a_C)}\right). \quad (4)$$

Now, I derive an expression for  $k^*$ . Sequential rationality requires that  $C$  with  $w_C = k^*$  be indifferent between  $PU$  and  $PR$ . This condition holds when  $EU_C(PU | w_C = k^*) = EU_C(PR | w_C = k^*) \Rightarrow r_{public}(-a_C) + (1 - r_{public}) = 0$ , which generates the optimal rate of refusal such that

$$r_{public}^* = \frac{1}{1 + a_C}. \quad (5)$$

Since  $r_{public} = r_{public}^*$  in equilibrium,  $k^*$  must be such that (4) and (5) hold simultaneously. Hence,

substituting  $r_{public}$  with  $r_{public}^*$ , and solving for  $k^*$  gives us the cutoff value such that

$$k^* = F_C^{-1} \left\{ \frac{F_D^{-1} \left( \frac{a_C}{1+a_C} \right) \cdot [1 - F_C(-a_C)] + F_C(-a_C) + a_D}{1 + a_D} \right\}. \quad (6)$$

Since  $k < -a_C$ , for the cutoff value of  $k^*$  to be a part of a PBE, it must be the case that

$$\frac{F_D^{-1} \left( \frac{a_C}{1+a_C} \right) \cdot [1 - F_C(-a_C)] + F_C(-a_C) + a_D}{1 + a_D} < F_C(-a_C). \quad (7)$$

Solving (7) for  $a_C$  gives us  $a_C < F_D(-a_D)/(1 - F_D(-a_D))$ . **Q.E.D.**

**Proof of Corollary 1.3.** It follows immediately from Corollary 1.2 that in the public equilibrium, bluffing never occurs if  $a_D \geq -F_D^{-1}(a_C/(1 + a_C))$ . Since no bluffing occurs,  $C$ 's types are fully separated, so a public threat ( $PU$ ) is fully informative. By (1), when  $a_D < -F_D^{-1}(a_C/(1 + a_C))$ ,  $D$ 's equilibrium belief that  $PU$  is credible, upon seeing it, is given by  $q_{public}^* = (1 - F_C(-a_C))/(1 - F_C(k^*))$ , where  $F_C(k^*) =$

$$\frac{F_D^{-1} \left( \frac{a_C}{1+a_C} \right) [1 - F_C(-a_C)] + F_C(-a_C) + a_D}{1 + a_D} \text{ from (6). Thus, since } \frac{\partial}{\partial a_D} (1 - F_C(k^*)) < 0, \text{ differentiating } q_{public}^* \text{ w.r.t. } a_D$$

yields  $\partial q_{public}^* / \partial a_D > 0$ . **Q.E.D.**

**Proof of Proposition 2.** The private equilibrium is characterized by a series of cutpoints established by Preliminaries in the text. Then, provided that (i)  $-a_C \leq 0$  and (ii)  $k > 0$ , the proposed equilibrium strategy for  $C$  follows immediately from this cutpoint structure as follows.

$$\begin{aligned}
& \{PU, SF_1, SF_2\} \text{ if } w_C \geq k^*, \\
& \{PR, SF_1, SF_2\} \text{ if } w_C \in [0, k^*], \\
& \{PR, BD_1, SF_2\} \text{ if } w_C \in [k^*, -a_C], \text{ and} \\
& \{PR, BD_1, BD_2\} \text{ if } w_C \leq -a_C.
\end{aligned}$$

Next, consider  $D$ 's beliefs about  $C$ 's type. Upon receiving a private threat ( $PR$ ),  $D$  forms a posterior belief that  $PR$  is genuine, according to Bayes's rule, such that  $q_{\text{private}} \equiv \Pr(w_C \geq 0 \mid w_C < k) = \frac{\Pr(0 < w_C \leq k)}{\Pr(w_C \leq 0) + \Pr(0 < w_C \leq k)} = \frac{F_C(k) - F_C(0)}{F_C(k)}$ . Conversely, when  $D$  receives public threats ( $PU$ ), she knows for certain that  $C$  will stand firm in response to a refusal, because the only types that make  $PU$  are the types for which  $w_C \geq k^*$ , where  $k > 0 > -a_C$ . Hence, it follows that  $q_{\text{public}} = 1$ .

Given the beliefs, sequential rationality requires that  $D$  resist if and only if her expected utility from resisting is greater than that from conceding. When  $PR$  is made, this condition holds when  $EU_D(RS_1 \mid PR) \geq EU_D(CD_1 \mid PR)$ , which implies  $w_D \geq (q_{\text{private}} - 1) / q_{\text{private}} \equiv w_{D,\text{public}}^*$ . Likewise, when  $PU$  is made, this condition holds when  $EU_D(RS_2 \mid PU) \geq EU_D(CD_2 \mid PU)$ , which implies  $w_D \geq -a_D \equiv w_{D,\text{private}}^*$ .

Let  $r_{\text{private}}$  and  $r_{\text{public}}$  denote the probabilities that  $D$  resists, upon receiving  $PR$  and  $PU$ , respectively. Then,  $D$ 's equilibrium strategy is characterized by  $w_{D,\text{private}}^*$  and  $w_{D,\text{public}}^*$  as follows.

$$r_{\text{private}} \equiv \Pr\left(w_D > \frac{q_{\text{private}} - 1}{q_{\text{private}}}\right) = 1 - F_D\left(-\frac{F_C(0)}{F_C(k) - F_C(0)}\right) \quad (8)$$

$$r_{\text{public}} \equiv \Pr(w_D > -a_D) = 1 - F_D(-a_D). \quad (9)$$

In equilibrium  $D$  mixes her strategy so that  $C$ 's type with  $w_C = k^*$  is indifferent between  $PU$  and  $PR$ .

Recall also that when  $w_C = k^*$  he stands firm since  $-a_C < k$  by conditions (i) and (ii). Then, this condition holds when  $EU_C(PR \mid w_C = k) = EU_C(PU \mid w_C = k)$ , or

$$r_{\text{private}}^* = r_{\text{public}}^* \quad (10)$$

Given (8) and (9), I can rewrite (10) such that  $1 - F_D(-F_C(0)/(F_C(k) - F_C(0))) = 1 - F_D(-a_D)$ . Solving this resulting equation for  $F_C(k)$  and taking the inverse yields:

$$k^* = F_C^{-1} \left( F_C(0) + \frac{F_C(0)}{a_D} \right). \quad (11)$$

To complete the proof, it remains to show that the proposed cutoff strategy for  $C$  is sequentially rational. First, consider the “moderate” types with  $w_C \in [0, k^*]$ . It is incentive compatible for these types to make  $PR$  if and only if no types within this range would do better by making  $PU$ . This condition holds when  $EU_C(PR | w_C \in [0, k^*]) \geq EU_C(PU | w_C \in [0, k^*]) \Rightarrow r_{private}^*(w_C) + (1 - r_{private}^*) \geq r_{public}^*(w_C) + (1 - r_{public}^*)$ . Substituting for  $r_{private}^*$  and  $r_{public}^*$ , and simplifying the resulting equation, I find that  $(q_{private} - 1)/q_{private} \geq -a_D$ . Substituting  $q_{private}$  in this inequality gives us

$$F_C(k^*) \geq F_C(0) + F_C(0)/a_D. \quad (12)$$

Recall that  $F_C(0) > 0$  and  $a_D \geq 0$ . Then, this incentive compatibility argument implies that  $k^* \geq 0$  in equilibrium, so sending  $PR$  is sequential rational for types  $w_C \in [0, k^*]$ .

Finally, for the “consiliatory” and “low” types with  $w_C < 0$ , it is incentive compatible to make  $PR$  if and only if  $EU_C(PR | w_C < 0) \geq EU_C(PU | w_C < 0) \Rightarrow r_{private} + (1 - r_{private}) \geq r_{public} \cdot (-a_C) + (1 - r_{public})$ .

Plugging (8) and (9) into this inequality yields  $F_D \left( -\frac{F_C(0)}{F_C(k) - F_C(0)} \right) - 1 \geq (F_D(-a_D) - 1) \cdot (-a_C + 1)$ .

Plugging (11) into this inequality yields  $a_C \geq 0$ , which is consistent with our assumption. Therefore, it is sequential rational for the types  $w_C < 0$  to make  $PR$ . **Q.E.D.**

**Proof of Proposition 3.1.** The priors that a threat is credible are given by  $p_{private} = 1 - F_C(0)$  and  $p_{public} = 1 - F_C(-a_C)$ , for a private threat ( $PR$ ) and a public threat ( $PU$ ), respectively. Then, in the public equilibrium the posterior beliefs, upon seeing  $PR$  and  $PU$ , respectively, are given by:

$$q_{private} = 0 \quad \text{and} \quad q_{public} = \begin{cases} \frac{1 - F_C(-a_C)}{1 - F_C(k^*)} & \text{if } -a_C > k^* \\ 1 & \text{if } -a_C \leq k^* \end{cases},$$

where  $k < 0$ . It immediately follows that  $PU$  has efficacy, while  $PR$  does not.

Similarly, the posterior beliefs in the private equilibrium, upon seeing  $PR$  and  $PU$ , are  $q_{private} = (F_C(k^*) - F_C(0))/F_C(k^*)$  and  $q_{public} = 1$ , respectively. By contradiction. Suppose  $PR$  has efficacy in the private equilibrium. Then, it must follow that

$$\frac{F_C(k^*) - F_C(0)}{F_C(k^*)} > 1 - F_C(0). \quad (13)$$

Substituting  $F_C(k^*)$  with  $F_C(0) + F_C(0)/a_D$  and re-arranging (13) yields  $1 < F_C(k^*)$ . However, this inequality never holds since  $F_C(\bullet) \in [0, 1]$ . Thus (13) does not hold either. **Q.E.D.**

**Proof of Proposition 3.2.** The claim that  $D$ 's two refusal rates,  $r_{public}^*$  and  $r_{private}^*$ , are identical is given by (10) in Proposition 2. **Q.E.D.**

**Proof of Proposition 4.** I first show condition (iii)  $a_D > \frac{F_C(0)}{1-F_C(0)}$ . This result follows directly from two assumptions:  $w_C < 1$  and  $k < 1$ . Then, it must be that  $F_C(k^*) < 1$ . Given (11), this condition holds when

$$1 > F_C(0) + \frac{F_C(0)}{a_D} \Rightarrow a_D > \frac{F_C(0)}{1-F_C(0)}. \quad (14)$$

Therefore, for the private equilibrium to exist  $D$ 's audience cost  $a_D$  must be greater than  $F_C(0)/(1 - F_C(0))$ . This result is useful for the next result.

Next, I show condition (ii). Note that  $a_D \geq 0$  by assumption. Then, the claim that  $a_D > 0$  immediately follows from (14) by contradiction. Suppose  $a_D = 0$ . Then, it must be the case that  $F_C(0) = 0$ . Had this been the case,  $w_C \in [\underline{w}_C, 0] = \emptyset$ . Given this, unless  $a_D > 0$  and thereby  $F_C(0) \neq 0$ , the proposed cutoff strategy for  $C$  in Proposition 2 cannot constitute a PBE. Finally, condition (ii)  $a_C \geq 0$  has been already shown as  $C$ 's incentive compatibility constraint in (13).

Lastly, to show condition (i) that the public option is also credible ( $k^* > -a_C$ ), it suffices to show that  $k^* \geq 0$  since  $-a_C \leq 0$  by assumption. It is trivial that claim  $k^* \geq 0$  follows immediately from the incentive compatibility for the ‘‘moderate’’ types with  $w_C \in [0, k^*]$ . Recall that for a proposed strategy to

constitute a PBE, the inequality (12) must be true:  $F_C(k^*) \geq F_C(0) + F_C(0)/a_D$ . Also recall the results derived above that  $F_C(0) > 0$  and  $a_D > 0$ . Then, this inequality implies that  $k^* \geq 0$  in equilibrium. **Q.E.D.**

**Proof of Proposition 5.** First I show the *ex ante* efficiency result. In analogy with the Myerson-Satterthwaite (1983) result, the expected payoffs of the public equilibrium are greater if  $a_C < F_D(-a_D)/(1 - F_D(-a_D))$ , as implied by Corollary 1.2. Then, we only need to compare the expected payoffs of the bluffing public equilibrium with those of the private equilibrium.  $C$ 's *ex ante* expected payoffs of the public and private equilibrium, respectively, are:

$$EU_C(\text{public}) = (1 - F_C(k^*))(1 - F_D(-a_D))w_C + (1 - F_C(k^*))F_D(-a_D), \text{ where } k^* < 0; \text{ and}$$

$$EU_C(\text{private}) = (1 - F_C(0))(1 - F_D(-a_D))w_C + F_D(-a_D), \text{ where } k^* > 0.$$

Since the war outcome is realized only when  $w_C > 0$  in the private equilibrium, while in the public equilibrium war occurs when  $w_C \in [k^*, 1]$ , whereby  $k^* < 0$ , it follows that  $EU_C(\text{public}) < EU_C(\text{private})$ .

For  $D$ , an analogous argument shows that  $EU_D(\text{public}) < EU_D(\text{private})$ .

Similarly, for the types with  $w_C \leq k^*$ , the *interim* efficiency result can be obtained by comparing  $C$ 's expected values in the two equilibria as follows:  $EU_C(\text{public} | w_C \leq k^*) = 0$  and  $EU_C(\text{private} | w_C \leq k^*) = (1 - F_C(k^*))F_D(-a_D) + (F_C(k^*) - F_C(0))(1 - F_D(-a_D))w_C$ . Since the war payoff in the private equilibrium is strictly positive,  $EU_C(\text{private} | w_C \leq k^*) > 0$ . Thus,  $EU_C(\text{public} | w_C \leq k^*) < EU_C(\text{private} | w_C \leq k^*)$ . **Q.E.D.**

**Proof of Corollary 5.1.** Let  $\pi^*$  denote the expected probability of war. The *ex ante* probability of war in the private equilibrium is given by

$$\pi^* = (1 - F_C(0))(1 - F_D(-a_D)). \quad (15)$$

The *ex ante* probabilities of war in the public equilibrium are given by

$$\pi^* = \begin{cases} (1 - F_C(k^*))(1 - F_D(-a_D)) & \text{if } -a_C < k^* \\ (1 - F_C(-a_C))(1 - F_D\left(\frac{q_{\text{public}}^{-1} - a_D}{q_{\text{public}}}\right)) & \text{if } -a_C > k^* \end{cases} \quad (16)$$

When  $-a_C < k^*$ , the claim that (16) is strictly greater than (15) follows immediately from the result that  $k^* < 0$  for the public equilibrium. To prove this claim when  $-a_C > k^*$ , first note that since  $-a_D < 0$  by assumption, it is sufficient to show that  $1 - F_D(-a_D) < 1 - F_D((q_{public} - 1 - a_D)/q_{public})$ , where  $q_{public} = (F_C(k^*)(1 + a_D) - F_C(-a_C) - a_D)/(1 - F_C(-a_C))$ . Rearranging this inequality yields  $F_C(-a_C) > F_C(k^*)$ , which always holds because  $-a_C > k^*$ . Thus, (16) is strictly greater than (15). **Q.E.D.**

## References

- Adcock, Frank and D.J. Mosley. 1975. *Diplomacy in Ancient Greece*. New York : St. Martin's Press.
- Cohen, Raymond and Raymond Westbrook.
- Austen-Smith, David and Jeffrey S. Banks. 2000. "Cheap Talk and Burned Money." *Journal of Economic Theory*, 91, pp. 1-16.
- Baum, Matthew A. 2003. "Going Private: Public Opinion, Presidential Rhetoric, and the Domestic Politics of Audience Costs in U.S. Foreign Policy Crises." *Journal of Conflict Resolution*, Vol. 48, No. 5, pp. 603-631.
- Beale, Howard K. 1956. *Theodore Roosevelt and the Rise of America to World Power*. Baltimore: Johns Hopkins University Press.
- Berridge, G. R. 2002. *Diplomacy: Theory and Practice*, Second Edition. Palgrave Macmillan
- Callières, François de. 1716. *The Art of Negotiating with Sovereign Princes*. London: Geo. Strahan.
- Cohen, Raymond and Raymond Westbrook 2000. *Amarna Diplomacy: the Beginnings of International Relations*. Baltimore, Md.: Johns Hopkins University Press.
- Collin, Richard H. 1985. *Theodore Roosevelt, Culture, Diplomacy, and Expansion: A New View of American Imperialism*. Baton Rouge: Louisiana State University Press.
- Crawford, Vincent P. and Joel Sobel. 1982. "Strategic Information Transmission." *Econometrica*, Vol. 50, No. 6, pp. 1431-1451.

- Farrell, Joseph and Robert Gibbons. 1989. "Cheap Talk with Two Audiences." *American Economic Review*, Vol. 79, No. 5, pp. 1214-1223.
- Fearon, James D. 1992. "Threats to Use Force: The Role of Costly Signals in International Crises." Ph.D. Dissertation, University of California, Berkeley.
- Fearon, James D. 1994. "Domestic Political Audiences and the Escalation of International Disputes." *American Political Science Review*, Vol. 88, No. 3 (September), pp. 577-592.
- Fearon, James D. 1995. "Rationalist Explanations for War." *International Organization*, Vol. 49, No. 3 (Summer), pp. 379-414.
- Fearon, James D. 1997. "Signaling Foreign Policy Interests: Tying Hands versus Sinking Costs." *Journal of Conflict Resolution*, Vol. 41, No. 1 (February), pp. 68-90.
- Finel, Bernard I. and Kristin M. Lord. 1999. "The Surprising Logic of Transparency." *International Studies Quarterly*, Vol. 43, No. 2, pp. 315-339.
- Francis, R. Douglas, Richard Jones, and Donald B. Smith. 1992. *Destinies: Canadian History since Confederation*, Second edition. Toronto: Holt, Rinehart and Winston of Canada.
- Gertner, Robert, Robert Gibbons, and David Scharfstein. 1988. "Simultaneous Signalling to the Capital and Product Markets." *Rand Journal of Economics*, Vol. 19, No. 2 (Summer), pp. 173-190.
- Guisinger, Alexandra and Alastair Smith. 2002. "Honest Threats: The Interaction of Reputation and Political Institutions in International Crises." *Journal of Conflict Resolution*, Vol. 46, No. 2, (April), pp. 175-200.
- Holmström, Bengt and Roger B. Myerson. 1983. "Efficient and Durable Decision Rules with Incomplete Information." *Econometrica*, Vol. 51, No. 6, pp. 1799-1819.
- Jönsson, Christer and Martin Hall. 2003. "Communication: An Essential Aspect of Diplomacy" *International Studies Perspectives*, Vol. 4, No. 2 (May), pp. 195-210.
- Kennedy, Robert F. 1969. *Thirteen Days: A Memoir of the Cuban Missile Crisis*. New York: W.W. Norton.
- Kissinger, Henry. 1979. *White House Years*. Boston: Little Brown.

- Lauren, Paul Gordon. 1994. "Coercive Diplomacy and Ultimata: Theory and Practice in History." in Alexander L. George and William E. Simons, eds. *The Limits of Coercive Diplomacy*, second edition. Boulder: Westview Press.
- Leventoğlu, Bahar and Ahmer Tarar. 2004. "Pre-negotiation Public Commitment in Domestic and International Bargaining." Mimeo. Texas A&M and SUNY Stony Brook.
- Mattingly, Garrett. 1955. *Renaissance Diplomacy*. Boston: Houghton Mifflin.
- Morgenthau, Hans J. 1956. *Politics Among Nations*, Second Edition. New York: Alfred A. Knopf
- Morrow, James D. 1989. "Capabilities, Uncertainty, and Resolve: A Limited Information Model of Crisis Bargaining." *American Journal of Political Science*, Vol. 33, No. 4, pp. 941-972.
- Munro, John A. 1970. *The Alaska Boundary Dispute*. Toronto: Copp Clark.
- Myerson, Roger B. and Mark A. Satterthwaite. 1983. "Efficient Mechanism for Bilateral Trading." *Journal of Economic Theory*, Vol. 29, No. 2, pp. 265-281.
- Nevins, Allan, *Henry White, Thirty Years of American Diplomacy* (New York, Harper, 1930).
- Nicolson, Harold. 1963. *Diplomacy*, third edition. London: Oxford University Press.
- Nicolson, Harold. 1954. *The Evolution of Diplomatic Method*. London: Constable.
- O'Neill, Barry. 1999. *Symbols, Honor, and War*. Ann Arbor, MI: University of Michigan Press.
- O'Neill, Barry. 2003. "Mediating National Honour: Lessons from the Era of Dueling." *Journal of Institutional and Theoretical Economics*, Vol. 159, No. 1, pp. 1-19.
- Powell, Robert. 1999. *In the Shadow of Power*. Princeton: Princeton University Press.
- Powell, Robert. 2002. "Bargaining Theory and International Conflict." *Annual Review of Political Science*, 5: 1-30.
- Penlington, Norman. 1972. *The Alaska Boundary Dispute: A Critical Reappraisal*. McGraw-Hill Ryerson.
- Ramsay, Kristopher W. 2004. "Politics at the Water's Edge: Crisis Bargaining and Electoral Competition." *Journal of Conflict Resolution*, Vol. 48, No. 4 (August), pp. 459-486.
- Sartori, Anne E. 2005. *Deterrence by Diplomacy*. Princeton: Princeton University Press.

- Schelling, Thomas C. 1960. *The Strategy of Conflict*. Cambridge: Harvard University Press.
- Schultz, Kenneth A. 1998. "Domestic Opposition and Signaling in International Crises." *American Political Science Review*, Vol. 92, No. 4, (December), pp. 829-844.
- Schultz, Kenneth A. 2001. "Looking for Audience Costs." *Journal of Conflict Resolution*, Vol. 45, No. 1, pp. 32-60.
- Smith, Alastair. 1998. "International Crises and Domestic Politics." *American Political Science Review*. Vol. 92, No. 3 (September), pp. 623-638.
- Snyder, Glenn H. and Paul Diesing. 1977. *Conflict among Nations: Bargaining, Decision Making, and System Structure in International Crises*. Princeton: Princeton University Press.
- Spence, Michael. 1973. "Job Market Signaling." *Quarterly Journal of Economics*, Vol. 87, No. 3, pp. 355-74.
- Thucydides. 1972. *History of the Peloponnesian War*. Translated by Rex Warner. London: Penguin.