Intimidation:
Linking Negotiation and Conflict

Sambuddha Ghosh, Gabriele Gratton, and Caixia Shen*

March 12, 2019

Abstract

A challenger wants a resource initially held by a defender, who can negotiate a settlement by offering to share the resource. If challenger rejects, conflict ensues. During conflict each player could be a tough type for whom fighting is costless. Therefore non-concession intimidates the opponent into conceding. Unlike in models where negotiations happen in the shadow of exogenously specified conflicts, offers made during negotiations determine how conflict unfolds if negotiations fail. In turn, how conflict is expected to unfold determines the players' negotiating positions. In equilibrium, negotiations always fail with positive probability, even if players face a high cost of conflict. Allowing multiple offers leads to brinkmanship—the only acceptable offer is the one made when conflict is imminent. If negotiations fail, conflict is prolonged and non-duration dependent.

Keywords: Intimidation, reputation, terrorism, negotiation, brinkmanship, costly war-of-attrition. **JEL Classification Numbers:** D74, D82.

---

*Ghosh: School of Economics, Shanghai University of Finance and Economics, 777 Guoding Road, Shanghai, China 200433, sghosh@mail.shufe.edu.cn. Gratton: School of Economics, UNSW Business School, UNSW Sydney, NSW 2052, Australia, g.gratton@unsw.edu.au. Shen: School of International Business Administration, Shanghai University of Finance and Economics, 777 Guoding Road, Shanghai, China 200433, shencaixia@gmail.com. We are grateful to three anonymous referees and Masaki Aoyagi for numerous detailed suggestions. We also thank Gaurav Aryal, Laurent Bouton, Alessandra Casella, Micael Castanheira, Jimmy Chan, Pauline Grosjean, Anton Kolotilin, Hongyi Li, Qingmin Liu, Youming Liu, Massimo Morelli, Santiago Oliveros, Carlos Pimienta, Ronny Razin, Santiago Sanchez-Pages, Balázs Szentes, John Tang, and Adam Wong. Ghosh is especially grateful to Costas Cavounidis and Juan Ortner for extended discussions, and to the Chinese National Science Foundation for support. We thank seminar participants at the Australian National University, Boston University, Columbia University, European University Institute, Georgetown University, University of Montreal, the 2014 ASSA Meetings, the 2014 Econometric Society Australasian Meetings, the 2013 Australasian Economic Theory Workshop, and the 2014 SAET Conference. Barton Lee provided excellent research assistance.

1 Introduction

Interstate conflicts begin when negotiations end. But why do negotiations fail to prevent conflict even when compromise solutions are available, commitment is possible, and conflict is likely to be long and painful?[1] In the crisis bargaining literature, two parties negotiate in the shadow of an exogenously given conflict.[2] If the parties' costs of fighting the conflict are private information, then each party has an incentive to misrepresent its real cost so to build a "reputation." This process leads to prolonged crises and, possibly, to war. What drives the choices made during negotiations is the type of conflict the parties expect, its expected length, and the probability of ultimate victory of each party.

However, how conflict unfolds (its length and the probability of ultimate victory of each party) is also determined by past negotiations. In fact, by rejecting a generous ultimatum offer, one party may signal that its cost of fighting is low, which in turn makes its opponent more cautious once the conflict begins. In this paper we argue that this possibility of signaling strength by rejecting a generous offer weakens the effectiveness of negotiations. In particular, it induces the parties to deliberately make meager offers which are rejected with strictly positive probability. For example, the Rambouillet Agreement offered by NATO to Yugoslavia before the onset of the Kosovo War was described as "a provocation, an excuse to start bombing"[3] that "deliberately set the bar higher than the Serbs could accept."[4]

To understand this two-way feedback relation between negotiation and conflict, we develop a model where negotiation and conflict are interlinked. In our model, Challenger (she) and Defender (he) want a resource that yields flow utility. Defender, who initially holds the resource, can try to negotiate peace by offering a share to Challenger. If all offers are rejected, conflict begins. In each period of conflict, Challenger chooses whether to attack. If she attacks, Defender chooses whether to concede the resource to Challenger. If Defender does not concede, conflict moves to the following period. Attacks are normally costly for both Challenger and Defender, but with an arbitrarily small probability either party is *tough*—does not experience the cost of attacks.

We fully characterize the set of equilibria of this model when the interval between

[1] In a seminal paper, Fearon (1995) poses this as a "central puzzle" that rationalist explanations fail to solve (see also Reiter, 2003). The key word 'rationalist' rules out explanations where the parties to the conflict are entirely or largely irrational.

[2] Notable examples include Esteban and Sàkovics (2008), Fearon (1992, 1994), Özyurt (2014), Powell (2004), and Sechser (2010).

[3] Henry Kissinger, *Daily Telegraph*, 28 June 1999.

[4] George Kenney, "Rolling Thunder: the Rerun," *The Nation*, 27 May 1999.

periods is sufficiently small. Because whether Challenger is tough is her private information, Challenger's attacks *intimidate* Defender: they scare Defender that Challenger will attack again if Defender does not concede. Similarly, because whether Defender is tough is his private information, not conceding intimidates Challenger: it scares her that further costly attacks are useless. This logic of intimidation draws from the literature on reputation building in bargaining and wars of attrition (Abreu and Gul, 2000; Chatterjee and Samuelson, 1988; Ponsatì and Sàkovics, 1995). The small uncertainty about the parties' payoffs is magnified by equilibrium play into a significant force that protracts conflict.

The same logic of intimidation also explains why the negotiation cannot succeed with certainty: even if Defender can make offers to Challenger before conflict begins, equilibrium offers are rejected with strictly positive probability by normal (i.e., not tough) Challenger. One might imagine that the negotiation fails because Defender is afraid to reveal whether he is tough or normal. We show that this intuition is incomplete by focusing on the case of Defender being *uninformed*, in the sense that he does not know his cost of pursuing conflict when he makes an offer, but will only discover it if and when Challenger first attacks. Even in this case where offers do not reveal whether Defender is tough or normal, the opportunity to make an offer is a double-edged sword for Defender. On the one hand, higher offers are better for Defender because they have a higher probability of being accepted. Indeed, if beliefs were held fixed both before and after the negotiation, Defender could completely avoid conflict with normal Challenger by offering slightly more than her expected value of entering conflict. On the other hand, a generous offer that has a high probability of being accepted increases Challenger's expected payoff from conflict because it increases Defender's belief that Challenger is tough if the negotiation fails. Therefore, rather than deterring conflict, more generous offers may encourage Challenger to seek conflict. We show that in equilibrium Defender always makes an offer that is both accepted and rejected with strictly positive probability by normal Challenger. Therefore, conflict begins with positive probability even if Challenger is normal.

This detrimental effect of generous offers is particularly evident when Defender has multiple chances to make offers. Indeed, suppose that before the last round Defender makes an offer that normal Challenger accepts with positive probability. Then Challenger strictly prefers to reject it, as this signals she is tough, thus intimidating Defender into making an even more generous offer in the last round. In other words, offers that could be successful (but weren't) make subsequent offers more costly. We show that in equilibrium all offers, except the one made in the last round, have vanishingly small probabilities of acceptance. Therefore, long negotiations resolve in *brinkmanship*: the parties make no progress towards a peaceful solution up until the last opportunity to negotiate.

This brinkmanship result provides an explanation for negotiation in the shadow of conflict not making any progress until the very last chance. A possible example are the negotiations leading up to the Treaty of Porthsmouth, which subsequently gained President Theodore Roosevelt the Nobel Peace Prize, in the Russo-Japanese War (1904–1905). Negotiations for the treaty were held while Russia brought further troops to Manchuria, a move that would have given Russia an advantage in case of conflict. Fredrik (2010) notes how the Japanese delegation demanded the southern part of the island of Sakhalin *and* war reparations throughout the negotiations. Only upon the arrival of further four Russian divisions, at what was conceivably their last chance to do so, did the Japanese drop their claim for reparations and avert conflict.

While in our benchmark model negotiations can only happen before conflict begins, in reality the parties to a conflict may wish to negotiate also once conflict has begun. In Section 7 we allow Defender to make offers in any period (and therefore after being informed about whether he is tough). We show that our main result holds even in this variation of the model: negotiations fail to avoid conflict with certainty.

A common finding in the empirical literature on both interstate (Bennet and Stam, 1996) and civil (Collier, Hoeffler, and Soderbom, 2004) conflicts is that conflicts are non-duration dependent: the probability of an additional period of conflict does not depend on the conflict's past length. Our framework allows us to explicitly derive the probability that conflict extends one more period. In equilibrium, if conflict does not end in the very first period, the probability that it extends from one period to the next is (i) independent of the probabilities that the players are tough and (ii) equal to a constant strictly less than 1, until a period we call the *conflict horizon* in which normal Defender concedes with certainty. Thus, our model predicts that, until the conflict horizon is reached, conflict is non-duration dependent.[5]

Our main contribution is to provide an integrated model of negotiation and conflict. Without such a model it would be impossible to study the two-way feedback relation by which negotiations affect conflict and conflict affects negotiations. Essentially, we connect two literatures: the crisis bargaining literature and the bargaining and reputation literature. In the crisis bargaining literature, conflict is an exogenously given outside option for the negotiating parties. Once conflict begins, the parties' relative military strengths de-

[5] In contrast, Reiter (2003) notes that existing models of conflict based on informational asymmetries fail to capture this feature as these asymmetries should fade with time. From a theoretical perspective, our result that conflict is non-duration dependent is equivalent to the constant concession rate in reputational models of bargaining (Abreu and Gul, 2000) and international crises in the shadow of conflict (Özyurt, 2014). In our model, armed conflict is itself a reputational game and therefore the parties capitulate at a constant rate.

termine the final outcome. This literature focuses on explaining why parties would delay reaching a settlement and ultimately reach conflict with positive probability. Since Gul, Sonnenschein, and Wilson (1986) clarified that private information alone cannot lead to delays in bargaining, a large literature has arisen. Fearon (1994) and Özyurt (2014) model "audience costs": further delays increase the cost of conceding to one's opponent. Therefore, delaying helps one commit to fighting. In our model, delay happens in spite of this motive being absent. Furthermore, in contrast with our result, flexible or strategic types in Özyurt (2014) always concede before war begins.

A different source of delay in bargaining is the one driven by reputation *à la* Kreps and Wilson (1982) and Milgrom and Roberts (1982).[6] Acharya and Grillo (2015) explicitly model this reputational motive in a model of crisis bargaining with irrational types. The option to engage in total war is never exercised in equilibrium, unlike in our model. Nevertheless, the same reputational motives drive the conflict part of our model. An approach similar to ours is the one by Lapan and Sandler (1988), who model terrorism as a repeated game between players who are tough with some probability. In their model, absent a concession, the public belief that a player is tough jumps up to an arbitrary and exogenously given quantity. Hodler and Rohner (2012) make this endogenous, but they have only two periods, which in turn means that they predict attacks only when the probability of the terrorist being tough is very large. Our model endogenously determines both the termination of the war and the evolution of beliefs about the degree of irrationality of one's opponent, and shows that prolonged conflict is compatible with very small degrees of irrationality.

Closer to us, Heifetz and Sagev (2005) study a model of negotiation during war. In their model, conflict is already happening, but one of the players, the aggressor, can choose to escalate—essentially increasing the flow cost of conflict. They show that there exist conditions such that, restricting attention to separating equilibria, the aggressor chooses to escalate. This is because, if escalation is more costly to the defender than to the aggressor, escalation both shortens conflict and induces further negotiations more in the aggressor's favor. The choice to escalate is similar to our choice to go to conflict, but in our model the parties go to conflict with strictly positive probability in any equilibrium even if this decision limits their ability to negotiate further (and for any relative cost of conflict for Challenger and Defender).

Our logic of brinkmanship shares a common feature with Sechser (2010) who shows that if conflict is potentially repeated, a player may incur the cost of rejecting an offer to

[6] Delay also arises in models with a non-common prior so that each player could be overly optimistic about her chances of being selected as the proposer (Yildiz 2004; Bénabou and Tirole, 2009).

avoid revealing its weakness in view of future negotiations. In Brito and Intriligator (1985) time is needed to screen among various types of opponents; in Sánchez-Páges (2009) time is needed to convey credible information to an uninformed party about the eventual outcome of rejecting agreements and triggering conflict. In contrast to these three papers, in our model brinkmanship arises because the uninformed party chooses to avoid giving the informed party a chance to signal she is tough.

Schelling (1956, 1960, 1966)[7] first developed the idea that bargaining parties can benefit if they convince their opponent that they are committed to their threat—hence the argument that governments should appear committed to hawkish positions when facing a terrorist threat. But in our model, as well as in Abreu and Gul (2000) and Özyurt (2014), once conflict begins, the expected payoff for (normal) Defender is independent of his probability of being tough. In fact, the entire advantage of being perceived as tough comes from the ability to induce a normal Challenger to attack with very low probability. But if the Challenger attacks nonetheless, then Defender must update his beliefs to assign a very high probability to Challenger being tough.

Our idea of intimidation is also related to Silverman (2004), a random-matching model where violence is instrumental in deterring future violence against oneself. If the fraction of agents who directly gain from violence is sufficiently large, then other agents can also engage in it to acquire a reputation for toughness. Yared (2009) considers a defender with private knowledge of his cost of conceding the flow resource in each period; in equilibrium the challenger attacks with positive probability when no concession is made, so that the defender has an incentive to concede often enough. Since costs are drawn independently across periods, there is no reputation at play, unlike in our model.

2 Benchmark model

In our benchmark model the parties to a resource dispute have a single chance to reach an agreement before conflict begins. In Section 5 we discuss how our results generalize to multiple rounds of negotiation.

Time is continuous and indexed by $\tau \geq 0$, but actions are in discrete time. There are two players: Challenger (C) and Defender (D). Both players discount future payoffs at rate $r > 0$. They contest a resource, which is initially held by Defender. Holding the resource gives a flow rent normalized to 1.

[7] See Crawford (1982) for a formal treatment of this idea.

The game played by Challenger and Defender is best described by dividing it into two phases: negotiation and conflict.

Negotiation. At time $\tau = 0$, Defender can offer a fraction x of the resource to Challenger. Upon observing the offer, Challenger decides to accept or reject the offer. If Challenger accepts the offer, then the game ends and thereafter Challenger and Defender enjoy flow rents x and $(1 - x)$ respectively. Otherwise, the game immediately enters the conflict phase.

Conflict. In the conflict phase, the following two-stage game is played out at each time $\tau_t = (t-1)\,\Delta$, $t \in \{1, 2, \dots\}$.

Stage 1: Challenger chooses whether to attack or concede;

Stage 2: Defender chooses whether to resist or concede.

We refer to the interval of time from τ_t to τ_{t+1} as period t. Notice that there is no time interval between the two stages.

As soon as one party concedes, the other party gets to enjoy the entire resource in that period and forever afterwards; thus conflict is less flexible than negotiation.

Types and payoffs. Each player can be of two types: *tough* or *normal*. Player $i \in \{C, D\}$ is tough with non-zero probability π^i. Normal types dislike conflict in the sense that in each period in which Challenger attacks, normal player i suffers losses

$$L^i := \left(1 - e^{-r\Delta}\right) \ell^i, \text{ with } \ell^i > 0.$$

Therefore, in any period in which neither Challenger nor Defender concede, normal Challenger's payoff is given by $-L^C$, while normal Defender's payoff is given by $V - L^D$, where

$$V := 1 - e^{-r\Delta}$$

is the value of holding the resource for one period. As soon as Defender (Challenger) concedes, Challenger (Defender) enjoys a payoff equal to V from that period onward. Notice that the cost of the attack is suffered by both players whenever Challenger attacks. If in any period Challenger attacks and Defender concedes, normal Challenger's payoff in that period is $V - L^C$ and normal Defender's payoff is $-L^D$.

In contrast, both tough Challenger and tough Defender suffer no loss from conflict. Tough Challenger does not accept any offer short of $x = 1$ and attacks until Defender concedes. Tough Defender never makes an offer greater than $x = 0$ and never concedes.

Challenger privately observes her type at the beginning of the game. Our aim is to show that negotiation fails even if Defender has no incentive to conceal his type. Therefore, in our benchmark model we assume that Defender privately observes his type only if and when conflict begins. While we do this because it better illustrates the logic behind our main results, we also note that this may be a realistic assumption in many conflicts. For example, when Challenger attacks using new technologies that have yet to be proven, both players may be uncertain about Challenger's ability to overcome Defender's defenses and cause him harm. Another example is when Defender is a democratic government who does not know a priori whether the electorate will be able to withstand Challenger's attacks. In both cases, Defender will be able to evaluate the losses caused by Challenger's first attack and therefore discover his type at the beginning of period 1. The timing at which players privately observe their types is common knowledge.

Obviously, if L^C is too large, then normal Challenger would never carry out an attack, even if Defender concedes for sure in period 1. Also, if L^D is too small, then normal Defender would never concede, even if Challenger attacks for sure at every period. We are interested in those cases in which it is at least conceivable for normal Challenger to attack and for normal Defender to concede. Therefore, we restrict our attention to those conflicts in which L^C is sufficiently small and L^D is sufficiently large.

Assumption 1. *Let $\delta \equiv e^{-r\Delta}$ be the discount factor between periods. We assume $\delta L^D > V$; $L^C < V(1-\delta)^{-1}$.*

Henceforth "equilibrium" refers to perfect Bayesian equilibrium as in Fudenberg and Tirole (1991), which requires *sequentially rational* strategies given *reasonable* beliefs. In what follows we characterize the equilibrium of the game when the opportunities to attack are frequent, i.e., when Δ is sufficiently small. Nonetheless, we solve the conflict phase for any $\Delta > 0$, as this allows us to derive precise comparative statics and study the negotiation phase without worrying about equilibrium selection.

2.1 Strategies and beliefs

Negotiation. In the negotiation phase, Defender's strategy is an offer x in $[0, 1]$. Normal Challenger's strategy α maps any offer to her probability of accepting it. Since tough

Challenger does not accept offers, Challenger's total probability of accepting an offer x is $\bar{\alpha}(x) = \left(1 - \pi^C\right) \alpha(x)$.

Conflict. Let $\mathcal{H}_t^i$ denote the set of all possible histories at which player $i \in \{C, D\}$ moves in period t. Notice that each history $h_t^C \in \mathcal{H}_t^C$ comprises an offer x, a rejection of the offer, and $t-1$ periods in which Challenger attacks and Defender resists; each history $h_t^D \in \mathcal{H}_t^D$ comprises an offer x, a rejection of the offer, $t-1$ periods in which Challenger attacks and Defender resists, and a further attack by Challenger in period t. Thus any history $h_t^i \in \mathcal{H}_t^i$ at which player i moves is uniquely identified by the offer x and the period t.

In the conflict phase, a (behavior) strategy for the normal type of player $i \in \{C, D\}$ is a sequence of mappings $(\sigma_t^i : [0,1] \to [0,1])_{t \in \mathbb{N}}$, where $\sigma_t^i(x)$ is the probability that normal player i concedes at the unique history h_t^i that follows offer x. Notice that σ_t^i is the probability with which normal player i concedes in period t, conditional on no previous concession.

Beliefs. For each stage in each period $t \in \{1, 2, \dots\}$ we define the public (i.e, the opponent's) belief that a player is tough at the unique history that follows offer x. At the start of stage 1 of period t, the public beliefs that Challenger and Defender are tough are given by $\pi_{t-1}^C(x)$ and $\pi_{t-1}^D(x)$ respectively; at the start of stage 2, these public beliefs are $\pi_t^C(x)$ and $\pi_{t-1}^D(x)$ respectively.[8] Of particular importance are the *post-negotiation beliefs* $\pi_0^i(x)$ for $i \in \{C, D\}$ with which conflict begins if Challenger rejects the offer made during the negotiation phase. These beliefs may differ from the priors depending on the offer made.

Total probabilities of concession. Since tough players never concede, the *total* (conditional) probability of concession of player i in period t is obtained by multiplying the probability of normal player i by the probability that normal player i concedes:

$$\bar{\sigma}_t^i(x) = \left(1 - \pi_{t-1}^i(x)\right) \sigma_t^i(x), \; i \in \{C, D\}. \tag{1}$$

3 Preview of results

In this section, we preview our main results. We will show that in equilibrium:

[8]Notice that in stage 2 the period index differs for Challenger and Defender. This is because Defender's belief that Challenger is tough is updated in light of Challenger's action at stage 1; Challenger's belief that Defender is tough is not updated because Defender does not move at stage 1.

Negotiation fails. Defender makes an offer x^* that is both accepted and rejected by normal Challenger with strictly positive probability. Therefore, conflict begins with probability strictly greater than π^C, but strictly smaller than 1.

Even when negotiation fails, the offers made during negotiation affect the belief $\pi_0^C(x^*)$ Defender holds when conflict begins, in turn affecting how conflict is played. In particular, a higher probability that negotiation succeeds leads, in case of failure, to a higher belief that Challenger is tough.

Conflict. In period 1 Challenger attacks with probability 1. Following the first attack, both normal players mix between attacking (resisting) and conceding for a number n of periods that depends on $\pi_0^C(x^*)$. After each attack, Defender raises his belief that Challenger is tough. Similarly, after each time Defender resists, Challenger raises her belief that Defender is tough. In period $n+1$, normal Challenger mixes between attacking and conceding. Thereafter, both normal players concede with probability 1.

In the following sections, we develop a precise analysis of the equilibrium play, explain the intuition behind these results, and offer some useful comparative statics. We solve the game backwards, beginning with the conflict phase.

4 Equilibrium of the conflict phase

We now study equilibrium behavior in the continuation game of conflict after negotiation has failed: Defender has made an offer x that Challenger rejected. Any such continuation game begins with post-negotiation beliefs $\left(\pi_0^C(x), \pi_0^D(x)\right)$.

In our model, each normal player concedes if he or she believes the enemy to be tough with probability 1. Since payoffs are continuous in beliefs, so are the optimal strategies. Therefore, in any period t, if Defender believes Challenger to be tough with sufficiently high probability, he will concede immediately even if he knows that normal Challenger will concede in stage 1 of period $t+1$ (at her next chance to concede). Similarly, if Challenger believes Defender to be tough with sufficiently high probability, then she will prefer to concede immediately, even if she knows that normal Defender will concede with certainty in stage 2 of period t itself (at his next chance to concede). The following lemma finds these thresholds as $\bar{\pi}^C$ and $\bar{\pi}^D$ defined in (2) and (3).

Lemma 1 (Threshold beliefs)**.** *Let*

$$\bar{\pi}^C := \frac{1}{\delta}\left[1+L^D\right]^{-1}, \text{ and} \tag{2}$$
$$\bar{\pi}^D := 1-L^C. \tag{3}$$

In any continuation equilibrium after Challenger rejects an offer x,

(i) *if the strategy of normal Defender is to concede immediately after Challenger attacks in period $t+1$, then normal Challenger strictly prefers to attack in period $t+1$ if $\pi_t^D(x) < \bar{\pi}^D$, is just indifferent if $\pi_t^D(x) = \bar{\pi}^D$, and strictly prefers to concede otherwise;*

(ii) *if the strategy of normal Challenger is to concede in period $t+1$, then normal Defender strictly prefers to resist in period t if $\pi_t^C(x) < \bar{\pi}^C$, is just indifferent if $\pi_t^C(x) = \bar{\pi}^C$, and strictly prefers to concede otherwise.*

Proof. In Appendix A.2. □

Proposition 1 below completely characterizes the unique equilibrium of the conflict phase for any $\Delta > 0$, but under an additional assumption. We show in Appendix A.5 that when this assumption fails the set of equilibria can still be identified using Proposition 1. Furthermore, as Δ becomes small, the expected payoff of all these equilibria converge to the expected payoff of the equilibrium in Proposition 1. Therefore, as Δ becomes small, this assumption plays no role in the determination of the expected payoffs in the conflict game and the unique equilibrium of the whole model.

Assumption 2. *The quantities $\ln \pi_0^D(x) / \ln \bar{\pi}^D$ and $\ln \pi_0^C(x) / \ln \bar{\pi}^C$ are not integers.*

Equilibrium play after negotiation fails crucially depends on how the post-negotiation public beliefs, $\pi_0^C(x)$ and $\pi_0^D(x)$, compare to their respective thresholds, $\bar{\pi}^C$ and $\bar{\pi}^D$. First, the number of periods for which normal Challenger attacks and normal Defender resists with strictly positive probability depends on the *conflict horizon.*

Definition 1. The *conflict horizon* is the largest non-negative integer n such that $\pi_0^C(x) < \left(\bar{\pi}^C\right)^n$ and $\pi_0^D(x) < \left(\bar{\pi}^D\right)^n$.

Second, the continuation equilibrium once negotiation has failed is unique, but it can be of one of two types, depending on the post-negotiation public beliefs $\pi_0^C(x)$ and $\pi_0^D(x)$. What matters is whether Challenger or Defender are perceived to be more or less tough, compared to what would be needed to induce the other player to concede immediately. If

$\pi_0^C(x)$ is sufficiently closer to $\bar{\pi}^C$ than $\pi_0^D(x)$ is to $\bar{\pi}^D$, then we say that Challenger is more *intimidating*. Otherwise we say that Defender is more intimidating. As we shall see, when negotiation fails, then conflict always begins with Challenger being more intimidating.

Proposition 1. *Any continuation game beginning with beliefs $\left(\pi_0^C(x), \pi_0^D(x)\right)$, conflict horizon n, and thresholds $\bar{\pi}^C$ and $\bar{\pi}^D$ defined in (2) and (3) admits a unique continuation equilibrium. In this equilibrium:*

1. *If $\left(\bar{\pi}^C\right)^{n+1} < \pi_0^C(x) < \left(\bar{\pi}^C\right)^n$ and $\pi_0^D(x) < \left(\bar{\pi}^D\right)^{n+1}$ (Challenger is more intimidating), in period $t = 1$, Challenger attacks with probability 1, Defender concedes with total probability*

$$\bar{\sigma}_1^D(x) = 1 - \frac{\pi_0^D(x)}{\left(\bar{\pi}^D\right)^n}, \tag{4}$$

and $\pi_1^C(x) = \pi_0^C(x)$ and $\pi_1^D(x) = \left(\bar{\pi}^D\right)^n$. From stage 1, period $t = 2$ to stage 1, period $t = n+1$, Challenger and Defender concede with total probability $\bar{\sigma}_t^i(x) = 1 - \bar{\pi}^i$ and beliefs evolve according to

$$\pi_t^i(x) = \frac{\pi_{t-1}^i(x)}{\bar{\pi}^C},\ i \in \{C, D\}. \tag{5}$$

2. *If $\left(\bar{\pi}^D\right)^{n+1} < \pi_0^D(x) < \left(\bar{\pi}^D\right)^n$ and $\pi_0^C(x) < \left(\bar{\pi}^C\right)^n$ (Defender is more intimidating), in period $t = 1$, Challenger concedes with total probability*

$$\bar{\sigma}_1^C(x) = 1 - \frac{\pi_0^C(x)}{\left(\bar{\pi}^C\right)^n}, \tag{6}$$

and $\pi_1^C(x) = \left(\bar{\pi}^C\right)^n$. From stage 2, period $t = 1$ to stage 2, period $t = n$, Challenger and Defender concede with total probabilities $\bar{\sigma}_t^i(x) = 1 - \bar{\pi}^i$, $i \in \{C, D\}$ and beliefs evolve according to (5).

3. *Thereafter, normal Challenger and Defender concede with probability 1 and, if player $i \in \{C, D\}$ does not concede, $\pi_t^i(x) = 1$.*

We give the full proof of Proposition 1 and further details about its intuition in Appendix A.2. Here we offer only the basic thrust behind its logic to show how negotiation and conflict are linked.

First, notice that equilibrium beliefs are derived from the equilibrium strategies and the initial post-negotiation beliefs $\left(\pi_0^C(x), \pi_0^D(x)\right)$ using Bayes' rule:

$$\pi_t^i(x) = \frac{\pi_{t-1}^i(x)}{1 - \bar{\sigma}_t^i(x)}, \text{ for all } i \in \{C, D\}, t \in \{1, 2, \dots\}$$

We can now turn to why the equilibrium strategies are optimal given beliefs. Intuitively, normal players weigh the costs and benefits of protracting conflict to the next period. On the one hand, refusing to concede may lead to further losses. On the other hand, it opens up the possibility that the opponent will concede.

When Defender is more intimidating, Challenger's (Defender's) total probability of concession $\bar{\sigma}^C(x) = 1 - \bar{\pi}^C$ ($\bar{\sigma}^D(x) = 1 - \bar{\pi}^D$) is such that normal Defender (Challenger) is indifferent between conceding and resisting (attacking).[9] After each time Challenger attacks, Defender raises his belief that Challenger is tough. Similarly, after each time Defender resists, Challenger raises her belief that Defender is tough. This is compatible with a constant total probability of concession because normal Challenger (Defender) increases the probability of concession over time according to:

$$\sigma_t^i(x) = \frac{1 - \bar{\pi}^i}{1 - \pi_{t-1}^i(x)}, \text{ for all } i \in \{C, D\}, \ t \in \{2, \dots, n\}. \tag{7}$$

After n periods, $\pi_n^D(x) > \bar{\pi}^D$ (while $\pi_n^C(x) = \bar{\pi}^C$) so that, by Lemma 1, Challenger strictly prefers to concede in period $n+1$.

When Challenger is more intimidating, the equilibrium is slightly different because in period 1 Defender concedes with a higher probability. Therefore, in period 1 Challenger strictly prefers to attack. Thereafter play proceeds as when Defender is more intimidating, but now $\pi_{n+1}^C(x) > \pi^C$ (while $\pi_n^D(x) = \bar{\pi}^D$) so that, by Lemma 1, Defender strictly prefers to concede in period $n+1$.

We now give an intuitive explanation for why normal Challenger mixes between attacking and conceding in any period $t > 1$ in which $\pi_{t-1}^D(x) \leq \bar{\pi}^D$. Suppose that normal Challenger concedes with certainty in period t. By Bayes' rule, normal Defender is then so intimidated by an attack that he would concede immediately after it. But then normal Challenger strictly prefers to attack and receive the whole resource rather than avoiding the loss L^C. If instead she attacks with certainty in equilibrium, then Defender's belief that Challenger is tough after the attack, $\pi_{t+1}^D(x)$, equals $\pi_t^D(x)$, and therefore his propensity to concede does not change from period $t-1$ to t. But then normal Challenger strictly prefers to concede as attacking brings no benefits while she incurs the loss L^C. A similar logic explains why normal Defender also mixes between resisting and conceding in any period $t > 1$ in which $\pi_t^C(x) \leq \bar{\pi}^C$.

The key difference between the two cases in Proposition 1 is that when Challenger is more intimidating, she attacks with probability 1 at the start, whereas she mixes when

[9]See Lemma 4 in Appendix A.2.

Defender is more intimidating.[10] That is, when Challenger is perceived to be more likely to be tough, her expected equilibrium payoff from conflict is greater. In fact, when Defender is more intimidating, in period 1 Challenger is indifferent between attacking and conceding—payoff equals 0. Instead, when she is more intimidating, Challenger strictly prefers to attack.

Remark 1. If at the beginning of conflict Challenger is more intimidating, then normal Challenger's expected payoff $u^C\left(\pi_0^C(x), \pi_0^D(x)\right)$ is given by

$$u^C\left(\pi_0^C(x), \pi_0^D(x)\right) = \left(1 - \frac{\pi_0^D(x)}{(\bar{\pi}^D)^n}\right) - L^C \tag{8}$$

and normal Defender's expected payoff is $-L^D$. The unconditional probability of an attack in period $t \in \{2, \ldots, n+1\}$ is given by

$$\Pr(\text{attack at } t) = \frac{\pi_0^D(x)}{(\bar{\pi}^D)^n}\left(\bar{\pi}^C\right)^{t-1}\left(\bar{\pi}^D\right)^{t-2}.$$

Remark 2. If at the beginning of conflict Defender is more intimidating, then normal Challenger's expected payoff is 0 and normal Defender's expected payoff $u^D\left(\pi_0^C(x), \pi_0^D(x)\right)$ is given by

$$u^D\left(\pi_0^C(x), \pi_0^D(x)\right) = \left(1 - \frac{\pi_0^C(x)}{(\bar{\pi}^C)^n}\right) - \frac{\pi_0^C(x)}{(\bar{\pi}^C)^n}L^D. \tag{9}$$

The unconditional probability of an attack in period $t \in \{1, \ldots, n\}$ is given by

$$\Pr(\text{attack at } t) = \frac{\pi_0^C(x)}{(\bar{\pi}^C)^n}\left(\bar{\pi}^C\bar{\pi}^D\right)^{t-1}.$$

4.1 Comparative statics of conflict

We now provide some comparative statics of the conflict phase and link them to empirical regularities.

As in Abreu and Gul (2000), the conflict horizon is finite and it is shorter when players are believed to be tough with greater probability.

Corollary 1. *Unless both players are tough, the maximum length of a conflict is determined by the conflict horizon n. If Challenger is more intimidating, normal Challenger does not attack after period $n+1$. If Defender is more intimidating, normal Challenger does not attack after period n.*

[10] In reputation models, asymmetries in prior probabilities for being tough are leveled up by adjusting initial concession rates (see Abreu and Gul, 2000; and Özyurt, 2014 for a continuous-time version of this result).

The next result determines the probability that a conflict that has lasted until period t, $1 < t \leq n$ survives to period $t+1$. This probability is independent of which player is more intimidating, how much the players are likely to be tough, or the period t. That is, the hazard rate of the conflict depends only on the threshold values $\bar{\pi}^C$ and $\bar{\pi}^D$, which do not depend on the initial beliefs, $\pi_0^C(x)$ and $\pi_0^D(x)$, or on t. Thus, until period $t = n$ is reached, conflict is non-duration dependent.

Corollary 2. *In each period $t : 1 < t \leq n$, if conflict has not yet ended, Challenger attacks with constant probability $\bar{\pi}^C$ and Defender resists with constant probability $\bar{\pi}^D$. Therefore, if conflict has not yet ended by the end of period $t-1$, the probability that conflict will not end by the end of period t equals $\bar{\pi}^C\bar{\pi}^D$.*

The non-duration dependence of conflict is driven by the reputational nature of our model of conflict. In fact, as we noted in the introduction, in models of bargaining with reputation (e.g., Abreu and Gul, 2000; Özyurt, 2014), the conditional probability that bargaining continues is constant in time. Notice that as $\pi_0^C(x)$ and $\pi_0^D(x)$ become small, n grows and the length of conflict (conditional on there being a conflict at the end of period 1) is therefore approximated by a geometric distribution with hazard rate $\bar{\pi}^C\bar{\pi}^D$.

Although the probability that conflict extends to the next period does not depend on time, it nonetheless depends in intuitive ways on other primitives of the model, in particular on the players' costs of fighting. Thus conflict duration is determined by the opponents' military capabilities, rather than their ability to intimidate.

Corollary 3. *Conditional on there being a conflict at time $t > 1$, the probability of an attack in period $t' > t$ is decreasing in the cost of fighting (L^C and L^D).*

For $\pi_0^C(x)$ and $\pi_0^D(x)$ sufficiently small, such that conflict length is approximated by a geometric distribution, the same comparative statics apply to the expected length of conflict. Furthermore, our normalization of the value of the resource to 1 does not allow us to derive explicit comparative statics with respect to it. Nonetheless, it is easy to show that if the value of the resource is given by V', then the probability of an attack in period t' is increasing in V'.

The evidence concerning the effect of institutional characteristics on the duration of conflict is ambiguous. On the one hand, Bennet and Stam (2009), Langlois and Langlois (2009), and Henderson and Bayer (2013) find that the relation between democracy and conflict duration is not significant once the military capabilities of the parties and the physical characteristics of conflict (common boundaries, terrain, etc.) are taken into account. On the other hand, models of co-determination of war duration and outcome find that wars that are initiated by democracies are shorter (Bennet and Stam, 1998; Clark and Reed, 2003; Slantchev, 2004).[11] Our results suggest that that the probability of continuation of conflict depends indeed only on the costs and benefits of war, and only to a lesser extent on the probability of being tough. Thus, in our model physical and technological characteristics matter more than political ones.

We now turn to the question of when an armed conflict is more likely to begin, i.e. there is a first attack if there is no chance to negotiate. The following corollary describes how this probability depends on the public beliefs that the players are tough.

Corollary 4. *Fix the likelihood $\pi_0^D(x)$ that Defender is tough. The probability that Challenger begins to attack is increasing in the belief that Challenger is tough $\pi_0^C(x)$. It is* strictly *increasing if and only if Defender is more intimidating than Challenger.*

For Defender, an image of toughness can pay: if Defender is more intimidating than Challenger, then the probability of a first attack is strictly less than 1. In this case, the probability of a first attack is $\pi_0^C(x) / \left(\bar{\pi}^C\right)^n$, where n is the largest natural number such that $\pi^D \leq \left(\bar{\pi}^D\right)^n$. Thus, if $\pi_0^D(x)$ increases, the probability of a first attack decreases.

Corollary 5. *Let Defender be more intimidating than Challenger. Then, the probability that Challenger begins to attack is decreasing in $\pi_0^D(x)$.*

Nonetheless, the advantage of being perceived as tough should not be overstated. After the first attack is carried out, Challenger levels the playing field with Defender and the expected payoff for Defender is $-L^C$, independently of $\pi_0^D(x)$. Indeed, in equilibrium, Defender is indifferent between conceding and resisting whenever he plays.

One of the few empirical regularities about conflict is that pairs of democracies are less likely to fight each other.[12] Our results suggest that violence begins when there is greater imbalance between the parties' probability of being tough. We argue that democratic leaders are kept in check by their citizens and that therefore democracies tend to have similar probabilities of being tough or irrational. On the contrary, autocrats of the like of Kim Jung-un are known for their unpredictable behavior. Thus, a pair of autocracies or one democracy and one autocracy are more likely to have unbalanced probabilities of being tough and are therefore more likely to engage in conflict.

[11] Lyall (2010) finds that democracies fight longer wars in a sample of 286 counterinsurgency wars.
[12] See Bueno de Mesquita and Smith (2012) for a survey.

5 Why negotiations fail to eliminate conflict

We now turn to the negotiation phase of our model to show why conflict may be unavoidable even if Defender and Challenger can commit to a peaceful division of the resource. An important feature of our model is that negotiation and conflict are linked: actions taken during negotiation affect public beliefs at the onset of conflict. These beliefs, in turn, determine the expected payoff from conflict, affecting the relative appeal of negotiating peace. In this section we show that this link between negotiation and conflict induces Defender to make an offer that is rejected with strictly positive probability by normal Challenger.

5.1 Preliminaries

We begin by showing how more generous offers, which are accepted by Challenger with greater probability, also make conflict more valuable to Challenger.

By Remarks 1 and 2, if post-negotiation beliefs are such that Challenger is more intimidating, then normal Challenger's expected payoff of conflict $u^C\left(\pi_0^C(x), \pi_0^D(x)\right)$ is given by $\bar{\sigma}_1^D(x) - L^C$, where

$$\bar{\sigma}_1^D(x) = 1 - \frac{\pi_0^D(x)}{\left(\bar{\pi}^D\right)^n};$$

otherwise $u^C\left(\pi_0^C(x), \pi_0^D(x)\right) = 0$.

We move a step back and study the continuation game beginning just after Defender makes any offer x. Recall that $\alpha(x)$ is the probability that normal Challenger accepts offer x, so that $\bar{\alpha}(x) = \left(1 - \pi^C\right)\alpha(x)$ is her total probability of acceptance. Then, in any continuation equilibrium, $\pi_0^C(x)$ is determined by $\bar{\alpha}(x)$ using Bayes' rule:[13]

$$\pi_0^C(x) = \frac{\pi^C}{1 - \bar{\alpha}(x)}. \tag{10}$$

Therefore, an offer x that is accepted with a greater probability $\bar{\alpha}(x)$ implies a higher public belief $\pi_0^C(x)$ that Challenger is tough if x is rejected and conflict begins.

For a given post-negotiation belief π_0^D that Defender is tough, increasing Defender's post-negotiation belief that Challenger is tough from π_0^C to $\pi_0'^C$ increases normal Challenger's expected payoff from conflict in one of two ways. First, if $\left(\pi_0^C, \pi_0^D\right)$ is such

[13]By Property 1 in Definition 3.1 in Fudenberg and Tirole (1991), belief $\pi_0^C(x)$ must be derived from π^C and $\alpha(x)$ by Bayes' rule, even if the equilibrium of the whole game assigns probability zero to the offer x.

that Defender is more intimidating, the change to $\pi_0'^C$ may induce a conflict where Challenger is more intimidating, thereby increasing Challenger's expected payoff from 0 to $\bar{\sigma}_1^D(x) - L^C > 0$. Second, if $\left(\pi_0^C, \pi_0^D\right)$ is such that Defender is more intimidating, the increase to $\pi_0'^C$ induces a (weakly) greater probability $\bar{\sigma}_1^D(x)$ that Defender concedes in period 1, because the conflict horizon n is (weakly) decreasing in $\pi_0^C(x)$. This immediately gives the following.

Remark 3. Let x and x' be two offers and fix the post-negotiation belief π_0^D that Defender is tough. In any equilibrium, for any acceptance strategy for Challenger α, if $\alpha(x) > \alpha(x')$, then $u^C\left(\pi_0^C(x), \pi_0^D\right) \geq u^C\left(\pi_0^C(x'), \pi_0^D\right)$.

We can now study the first move in our game: Defender chooses an offer. Our equilibrium concept implies that Defender's offer cannot signal what he does not know (Fudenberg and Tirole, 1991).[14] Therefore, $\pi_0^D(x) = \pi^D$ for all offers $x \in [0, 1]$.

We consider two extreme offers that Defender may choose to make. At one extreme, Defender may choose an offer x so meager that Challenger would surely reject: $\bar{\alpha}(x) = 0$. By Bayes' rule, then conflict begins with $\pi_0^C(x) = \pi^C$. But for Challenger to reject x, then x must not exceed normal Challenger's expected payoff from conflict, $u^C\left(\pi^C, \pi^D\right)$. Thus, the maximum offer that Challenger could surely reject is $\underline{x} = u^C\left(\pi^C, \pi^D\right)$. Furthermore, suppose that Challenger were to accept an offer $x < \underline{x}$ with strictly positive probability. By Bayes' rule, conflict would begin with belief $\pi_0^C(x) > \pi^C$. But then Challenger would strictly prefer to reject x as, by Remark 3, $u^C\left(\pi_0^C(x), \pi_0^D\right) \geq u^C\left(\pi^C, \pi_0^D\right) = \underline{x} > x$. This immediately gives the following.

Remark 4. In any equilibrium, normal Challenger surely rejects all offers $x < \underline{x} = u^C\left(\pi^C, \pi^D\right)$ and accepts with strictly positive probability all offers $x > \underline{x}$.

At the other extreme, Defender may choose an offer x so generous that normal Challenger would surely accept: $\bar{\alpha}(x) = 1 - \pi^C$. By Bayes' rule, if the offer is rejected then conflict begins with $\pi_0^C(x) = 1$. But for normal Challenger to accept x, then x must be no less than normal Challenger's expected payoff from conflict, $u^C\left(1, \pi^D\right)$. Thus, the minimum offer that Challenger would surely accept is $\bar{x} = u^C\left(1, \pi^D\right)$. Furthermore, suppose that normal Challenger were to reject an offer $x > \bar{x}$ with strictly positive probability. By Bayes' rule, conflict would begin with belief $\pi_0^C(x) < 1$. But then normal Challenger

[14] Properties 2 and 3 in Definition 3.1. Formally, Fudenberg and Tirole's (1991) definition applies to "multi-stage games with observed actions." In particular, each player is informed about his or her type at the start of the game. In our benchmark model Defender does not know his own type when making offers. Nonetheless, the definition still applies if we interpret our game as a game with three players: uninformed Defender, informed Defender, and Challenger. The offer is made by uninformed Defender, whose payoff equals the expected payoff of informed Defender.

would strictly prefer to accept x as, by Remark 3, $u^C\left(\pi_0^C(x), \pi^D\right) \leq u^C\left(1, \pi^D\right) = \bar{x} < x$. This immediately gives the following.

Remark 5. In any equilibrium, normal Challenger surely accepts all offers $x > \bar{x} = u^C\left(1, \pi^D\right)$ and rejects with strictly positive probability all offers $x < \bar{x}$.

5.2 Equilibrium

Proposition 2 says that Defender's equilibrium offer is bounded by the extreme offers $\underline{x}$ and $\bar{x}$. The equilibrium offer is neither surely rejected nor surely accepted by normal Challenger; in fact, Challenger mixes in response to off-path offers that are neither below $\underline{x}$ nor above $\bar{x}$. Obviously, normal Challenger's expected payoff from conflict must then be strictly positive. By Remarks 1 and 2, this happens if and only if public beliefs at the beginning of conflict (and therefore after negotiation has failed) are such that rejecting an offer induces equilibrium beliefs where Challenger is more intimidating.

Proposition 2. *For any Δ sufficiently small, any equilibrium features a pair $(x_\Delta, \alpha_\Delta)$ such that:*

1. *Defender makes an offer $x_\Delta \in [\underline{x}, \bar{x}]$, where $\underline{x} = u^C\left(\pi^C, \pi^D\right)$ and $\bar{x} = u^C\left(1, \pi^D\right)$.*
2. *Normal Challenger accepts offer x with probability $\alpha_\Delta(x)$ such that $\alpha_\Delta(x_\Delta) \in (0,1)$ and*

$$\alpha_\Delta(x) \begin{cases} = 0 & \text{if } x < \underline{x}; \\ \in [0,1) & \text{if } x = \underline{x}; \\ \in (0,1) & \text{if } x \in (\underline{x}, x); \\ \in (0,1] & \text{if } x = \bar{x}; \\ = 1 & \text{if } x > \bar{x}. \end{cases}$$

3. *Post-negotiation beliefs are determined by*

$$\pi_0^C(x) = \frac{\pi^C}{1 - (1 - \pi^C)\,\alpha_\Delta(x)};$$
$$\pi_0^D(x) = \pi^D.$$

4. *For any $x \in [\underline{x}, \bar{x}]$, if $\alpha_\Delta(x) > 0$, then there exists $n \in \mathbb{N}$ such that $\left(\bar{\pi}^C\right)^{n+1} \leq \pi_0^C(x) < \left(\bar{\pi}^C\right)^n$ and $\pi_0^D(x) = \pi^D < \left(\bar{\pi}^D\right)^{n+1}$.*

Furthermore, for every sequence of Δ going to zero, every sequence of equilibrium pairs $(x_\Delta, \alpha_\Delta(x_\Delta))$ converges to the same limit $(x^*, \alpha^*) \in (0,1)^2$.

To prove Proposition 2, we first study Defender's choice of an offer when Challenger is more intimidating if conflict were to begin with pre-negotiation beliefs $\left(\pi^C, \pi^D\right)$. If an offer x is then made and rejected, post-negotiation beliefs are such that $\pi_0^C(x) \geq \pi^C$ and $\pi_0^D(x) = \pi^D$. Therefore, every continuation conflict game begins with Challenger being more intimidating. By Remark 1, Challenger accepts an offer x with positive probability only if[15]

$$x \geq u^C\left(\pi_0^C(x), \pi^D\right) = \left(1 - \frac{\pi^D}{\left(\bar{\pi}^D\right)^n}\right) - \left(1 - e^{-r\Delta}\right)\ell^C. \tag{11}$$

Let $U_\Delta^D(x, \bar{\alpha}(x))$ denote uninformed Defender's expected payoff from making an offer x that is accepted with total probability $\bar{\alpha}(x)$:

$$U_\Delta^D(x, \bar{\alpha}(x)) := \bar{\alpha}(x)(1-x) + (1-\bar{\alpha}(x))\left\{\pi^D - \left(1-\pi^D\right)\left(1-e^{-r\Delta}\right)\ell^D\right\}. \tag{12}$$

The factor multiplying $1 - \bar{\alpha}(x)$ in the second term shows that the Defender correctly anticipates that, if negotiation fails, he will be tough with probability π^D (and get the full utility of the resource) and with the remainder probability he will be normal (and get no utility from the resource if conflict begins). An offer x_Δ^* is optimal if it maximizes (12) subject to (11) and (10). Since there is no point for Defender to offer more than what normal Challenger would accept, the first constraint binds.

Thus far we have considered our benchmark model with $\Delta > 0$. In the remainder of this section we derive the limit of optimal offers as Δ approaches zero, when it is easier to quantify the intuition. Appendix A.3 uses a continuity argument to guarantee the result for all Δ sufficiently small. The limit of Defender's payoff as $\Delta \downarrow 0$ is given by

$$U^D(x, \bar{\alpha}(x)) = \bar{\alpha}(x)(1-x) + (1-\bar{\alpha}(x))\pi^D. \tag{13}$$

The limit of Challenger's payoff as $\Delta \downarrow 0$ is given by the following lemma.

Lemma 2. *Assume that Challenger is more intimidating if conflict begins with beliefs* $\left(\pi^C, \pi^D\right)$. *Then* $\lim_{\Delta\downarrow 0} u^C\left(\pi_0^C(x), \pi^D\right) = U^C\left(\pi_0^C(x), \pi^D\right) := 1 - \pi^D\left(\pi_0^C(x)\right)^{\frac{\ell^C}{1-\ell^D}}$.

Proof. In Appendix A.3. □

Therefore, in the limit as $\Delta \downarrow 0$, Defender's problem reduces to

[15] Recall from Proposition 1 that $\bar{\sigma}_1^D(x) = 1 - \pi_0^D(x) / \left(\bar{\pi}^D\right)^n$ in any continuation equilibrium and $L^i \equiv \left(1 - e^{-r\Delta}\right)\ell^i$, $i \in \{C, D\}$.

$$\max_{(x,\bar{\alpha})\in[0,1]\times[0,1-\pi^C]} U^D(x,\bar{\alpha}) \text{ s.t. } x = U^C\left(\pi_0^C(x),\pi^D\right) \text{ and } \pi_0^C(x) = \frac{\pi^C}{1-\bar{\alpha}}$$

Using Lemma 2 and the second constraint, we can write x as a function of $\bar{\alpha} \in \left[0, 1-\pi^C\right]$ and, after substituting this expression in U^D we have a maximization problem in one variable, $\bar{\alpha}$. The first-order condition is:

$$\frac{dU^D}{d\bar{\alpha}} = \pi^D\left(\frac{\pi^C}{1-\bar{\alpha}}\right)^{\frac{\ell^C}{1-\ell^D}} - \bar{\alpha}\frac{dx}{d\bar{\alpha}} - \pi^D = 0. \tag{14}$$

Substituting

$$\frac{dx}{d\bar{\alpha}} = -\pi^D\left(\frac{\pi^C}{1-\bar{\alpha}}\right)^{\frac{\ell^C}{1-\ell^D}}\frac{1}{1-\bar{\alpha}}\frac{\ell^C}{1-\ell^D}.$$

into (14) yields

$$\frac{1}{\pi^D}\frac{dU^D}{d\bar{\alpha}} = \left(\frac{\pi^C}{1-\bar{\alpha}}\right)^{\frac{\ell^C}{1-\ell^D}}\left(1+\frac{\bar{\alpha}}{1-\bar{\alpha}}\frac{\ell^C}{1-\ell^D}\right) - 1 = 0. \tag{15}$$

It is easy to see that $dU^D/d\bar{\alpha}$ is strictly positive at $\bar{\alpha} = 0$, strictly negative at $\bar{\alpha} = 1-\pi^C$, and U^D is single-peaked over its domain. This implies that the function U^D attains a maximum at a unique interior value $\bar{\alpha} \in \left(0, 1-\pi^C\right)$, which corresponds to a unique value x^* since $\bar{\alpha}$ is strictly increasing in x whenever $\bar{\alpha} \in \left(0, 1-\pi^C\right)$. Therefore, when Challenger is more intimidating if conflict were to begin with beliefs $\left(\pi^C, \pi^D\right)$, Defender's optimal offer is rejected with strictly positive probability by normal Challenger.

We now turn to the case when Defender is more intimidating if conflict were to begin with beliefs $\left(\pi^C, \pi^D\right)$: there exists $m \in \mathbb{N}$ such that $\left(\bar{\pi}^D\right)^{m+1} < \pi^D < \left(\bar{\pi}^D\right)^m$ and $\pi^C < \left(\bar{\pi}^C\right)^m$. We rule out by contradiction the following possibility: Defender makes an offer that is accepted with so small a probability that, even if Challenger rejects it, the resulting post-negotiation beliefs still induce a conflict in which Defender is more intimidating. Suppose that Defender makes such an offer x. Then (10) implies that

$$\pi_0^C(x) = \frac{\pi^C}{1-\bar{\alpha}(x)} < \left(\bar{\pi}^C\right)^m. \tag{16}$$

Since (16) implies $\bar{\alpha}(x) < 1-\pi^C$, normal Challenger must reject x with strictly positive probability.[16] From the point of view of normal Challenger, though, accepting such an offer gives a payoff of $x > 0$. Instead, rejecting it induces a conflict with Defender being more intimidating, and thus gives an expected payoff of 0. Therefore, normal Challenger prefers to accept this offer with certainty, contradicting $\bar{\alpha}(x) < 1 - \pi^C$. Therefore, even if Defender is more intimidating before negotiation begins, Challenger must be more intimidating after negotiations fail. The optimal offer must then satisfy the first-order condition in (14), and the rest of the logic is unchanged. Hence, Defender's optimal offer is rejected with strictly positive probability by normal Challenger.

6 Brinkmanship (multiple offers)

Previous sections established why a single round of negotiation may fail to avoid conflict. But with multiple rounds, Defender could potentially 'screen' Challenger, i.e., take out a proportion of normal types at each round, lowering the probability of conflict, which could conceivably go to 0 as the number of rounds increases. We now show that this is not in fact the case: our results on the failure of negotiation are robust to an arbitrary number of offers.

The model is modified as follows to accommodate K rounds of negotiation. At each round $k = -(K-1), \ldots, -1, 0$, Defender can offer any fraction of the resource to Challenger. The conflict phase is triggered if all K offers are rejected. Defender learns his type only if conflict starts and Challenger attacks, while Challenger knows her type throughout. Decision histories extend the benchmark model in a straightforward way. A decision history in the negotiation phase is uniquely identified by a vector of offers and the player who moves. If the player who moves is Defender, then the action chosen is another offer; if it is Challenger, it is acceptance or rejection of the last offer in the sequence. A decision history in the conflict phase is uniquely identified by the vector of rejected offers, the number of periods of conflict that have elapsed, and the identity of the player whose moves.

Notice that Defender can still attain the same expected payoff he would get with only one round of negotiation. In fact, he can choose to make offers that no Challenger would accept until the last round ($k = 0$), and then make the same offer he would make with one-round negotiation. Therefore, the question is whether or not Defender can do better than this by making an acceptable offer before the last round of negotiation. Proposition 3

[16] As we have argued before, if normal Challenger accepts with certainty ($\bar{\alpha}(x) = 1 - \pi^C$) then a rejection forces Defender to conclude that Challenger is tough with certainty ($\pi_0^C(x) = 1 > \left(\bar{\pi}^C\right)^m$ for any $m \in \mathbb{N}$).

says that the answer to this question is negative: in equilibrium offers are accepted with vanishingly small probability except on the brink of conflict.

Proposition 3. *For any* $\Delta > 0$, *let* $\bar{\alpha}_k^*(\Delta)$, $k = -(K-1), \ldots, -1, 0$ *be the equilibrium probability that Challenger accepts the offer in round* k *of the game with* K *rounds of negotiation. Let* $\bar{\alpha}^* \in \left(0, 1 - \pi^C\right)$ *be the solution to the first-order condition (15). For any sequence of* Δ *going to zero, every sequence* $\left(\bar{\alpha}^*_{-(K-1)}(\Delta), \ldots, \bar{\alpha}^*_{-1}(\Delta), \bar{\alpha}^*_0(\Delta)\right)$ *converges to* $(0, \ldots, 0, \bar{\alpha}^*)$.

The proof in Appendix A.4 explicitly allows for varying non-negative intervals of time between rounds of negotiation, and thus implies the above result with all intervals being equal to Δ.

Once again, the intuition behind our brinkmanship result lies in the ability of Challenger to intimidate Defender—in this case, into making larger offers in the future. To illustrate this intuition, we focus here on the simple case of $K = 2$ as Δ goes to 0 (in Appendix A.4 we show how the intuition goes through for $\Delta > 0$). Let x^* be the optimal offer in our limiting benchmark model. Notice that if an agreement is not reached in the first round, then the continuation game is identical to our benchmark model. Yet, Defender's belief that Challenger is tough when the last round comes along may have been affected by actions taken in the previous round. Suppose that, in the first round, Defender makes an offer that Challenger accepts with probability $\bar{\alpha}' > 0$. If Challenger rejects, then Defender's belief that Challenger is tough rises to

$$\pi^{C\prime} = \frac{\pi^C}{1 - \bar{\alpha}'} > \pi^C.$$

But then, when the second and last round comes along, Defender would make an offer larger than x^*. Since Challenger must be indifferent between accepting the first and the second offer, then the earlier offer too must be larger than x^*. Thus, the cost for Defender of buying Challenger's agreement is the same in the two rounds and it is greater than x^*—the cost it pays when there is only one round of negotiation. Since Defender can always attain the same expected payoff he would get if there was only one round of negotiation (for example by making a zero offer in all rounds until the final one), making an acceptable offer in the first round is not optimal for him.

7 Negotiating during conflict, and informed Defender

Our benchmark model makes a stark distinction between the actions available to the parties before and after conflict begins: before conflict begins, the parties can negotiate a peace that involves sharing the resource between them; after conflict begins, they can only concede the whole resource. This is arguably realistic. As noted by Langlois and Langlois (2009), negotiations *during* conflict between states are uncommon. For example, Pillar (1983) finds that in only nineteen of one-hundred and forty two interstate wars parties negotiated during conflict and before an armistice. However, there are situations in which negotiations may happen after conflict begins. We now discuss a variant of the model in which Defender can make any offer to share the resource to Challenger in any period $t \in \{1, 2, \ldots\}$. Since Defender learns his type once conflict begins, this variant of the model also features offers from an informed Defender. The main message is that neither the possibility of negotiating during conflict nor Defender being informed at the time he makes an offer change our main result: initial negotiations fail with strictly positive probability and the ensuing conflict is prolonged.

Consider the following variation of our model. The only difference with our benchmark model is that if Challenger rejects the (initial) negotiation offer x, the following three-stage game is played at each period $t \in N$:

Stage 1: Challenger attacks with certainty unless she has previously accepted an offer;

Stage 2: Defender makes an offer $y_t \in [0, 1]$;

Stage 3: Challenger chooses whether to accept or reject the offer. If she accepts, the game ends and Challenger and Defender enjoy flow rents y_t and $(1 - y_t)$, respectively, thereafter. If she rejects, the game moves to the next period.

As before, Defender learns his own type immediately after the first attack by Challenger.

Implicit in our model is the assumption that Challenger's decision to reject an offer and continue the conflict is taken at the end of a period rather than at the beginning of the next.[17] This makes this model readily comparable to our benchmark, where, after Challenger attacks, Defender can choose to concede the whole resource, in which case the transfer happens immediately. Similarly here, after Challenger attacks, Defender can choose to concede a fraction of the resource. If Challenger accepts, then the transfer happens immediately.

Let $\mathcal{H}_t^i$ denote the set of all possible histories at which player $i \in \{C, D\}$ moves in period t. A history $h_t^C \in \mathcal{H}_t^C$ comprises a rejected initial offer x, $t - 1$ rejected offers y_s, $s = 1, \ldots, t - 1$, and the offer y_t Defender has made in period t. A history $h_t^D \in \mathcal{H}_t^D$

[17]This is why Challenger attacks with certainty at stage 1 after rejecting an offer in the previous period: accepting an offer strictly dominates rejecting the same offer and then conceding in the next period.

comprises a rejected initial offer x, and $t-1$ rejected offers y_s, $s=1,\dots,t-1$. Thus any history $h_t^i \in \mathcal{H}_t^i$ at which player i moves is uniquely identified by the period t and a vector of offers—$\boldsymbol{y}_t = (x, y_1, \dots, y_t)$, if $i = C$; $\boldsymbol{y}_{t-1}$, if $i = D$. Let $\pi_t^D(\boldsymbol{y}_t)$ denote Challenger's belief that Defender is tough at the history comprising offers $\boldsymbol{y}_t$ at which Challenger moves in period t. Furthermore, let $\pi_t^C\left(\boldsymbol{y}_{t-1}\right)$ denote Defender's belief that Challenger is tough when Defender moves in period t.

Since Defender is informed about her type at the moment she makes an offer in period $t \geq 1$, then making a strictly positive offer reveals that he is normal. Here we focus on equilibria satisfying the following 'degeneracy' property: if Defender is publicly known to be normal and Challenger is publicly believed to be tough with strictly positive probability, then Defender offers the whole resource. Arguments mirroring those in the reputational bargaining literature (Abreu and Gul, 2000) show that all equilibria must exhibit degeneracy as Δ goes to zero. When Δ is sufficiently small, any equilibrium of the entire game with the degeneracy property is of the form below.

(Initial) negotiation fails. Defender makes an initial offer that is both accepted and rejected by normal Challenger with strictly positive probability. Therefore, conflict begins with probability strictly greater than π^C, but strictly smaller than 1.

If negotiation fails, conflict begins with Challenger being more intimidating.

Conflict. There is a threshold level $\pi'^C < \bar{\pi}^C$ such that in the continuation game of conflict:

1. if $0 < \pi_t^D\left(\boldsymbol{y}_{t-1}\right) \leq \bar{\pi}^D$ and $0 < \pi_{t-1}^C\left(\boldsymbol{y}_{t-1}\right) \leq \pi'^C$, normal Defender mixes between offering $y_t = \delta\bar{\pi}^D$ and $y_t = 0$; normal Challenger accepts any offer $y_t \geq \delta\bar{\pi}^D$ with certainty; mixes between accepting and rejecting $y_t = 0$; and rejects $y_t \in \left(0, \delta\bar{\pi}^D\right)$ with certainty.

2. If $0 < \pi_t^D\left(\boldsymbol{y}_{t-1}\right) \leq \bar{\pi}^D$ and $\pi'^C < \pi_{t-1}^C\left(\boldsymbol{y}_{t-1}\right) < \bar{\pi}^C$, normal Defender mixes between offering $y_t = 1$ and $y_t = 0$; normal Challenger accepts any offer $y_t \geq \delta\bar{\pi}^D$ with certainty; mixes between accepting and rejecting $y_t = 0$; and rejects $y_t \in \left(0, \delta\bar{\pi}^D\right)$ with certainty.

3. If $\pi_{t-1}^C\left(\boldsymbol{y}_{t-1}\right) > \bar{\pi}^C$, normal Defender offers $y_t = 1$.

4. If $\pi_{t-1}^D\left(\boldsymbol{y}_{t-1}\right) > \bar{\pi}^D$, normal Challenger accepts zero offers.

5. Beliefs are updated using Bayes' rule and equilibrium strategies as in our benchmark model.

The logic behind this equilibrium follows that in our benchmark model. Challenger's (Defender's) total probability of acceptance (making a strictly positive offer) is such that normal Defender (Challenger) is indifferent between making a strictly positive offer and offering nothing (accepting and rejecting). After each time Challenger rejects, Defender raises his belief that Challenger is tough. Similarly, after each time Defender offers nothing, Challenger raises her belief that Defender is tough. This model differs from the benchmark one because it allows *non-zero* offers short of everything. Such an offer reveals immediately that Defender is normal, and thus will not be accepted unless it is so generous that Challenger would rather not reject it, attack once, and force Defender to offer the whole resource in the next period. This latter option gives $\delta\left(1-L^C\right)=\delta\bar{\pi}^D$.[18] Anything below $\delta\bar{\pi}^D$ is counterproductive (because it will invite an extra attack before Defender anyway concedes), and anything higher than this is unnecessary if the Defender puts reasonably high probability that Challenger is normal and will accept it; this leads to point 1 above. If Challenger is more likely to be tough, but still less than the threshold $\bar{\pi}^C$, Defender prefers to end conflict with both normal and tough Challenger if he is going to reveal that he is normal; this leads to point 2 above.

It remains to explicitly calculate π'^C, the belief at which Defender's offer switches from $\delta\bar{\pi}^D$ to 1. If Defender offers $\delta\bar{\pi}^D$, with probability $1-\pi^C_{t-1}\left(\boldsymbol{y}_{t-1}\right)$ Challenger is normal and will accept; otherwise, Challenger is tough and Defender will concede the whole resource with certainty after incurring an additional attack. Therefore Defender's expected payoff equals

$$\left(1-\pi^C_{t-1}\left(\boldsymbol{y}_{t-1}\right)\right)\left(1-\delta\bar{\pi}^D\right)+\pi^C_{t-1}\left(\boldsymbol{y}_{t-1}\right)\left(V-\delta L^D\right).$$

Offering the whole resource right away gives Defender the certainty to avoid any further conflict and a payoff of 0. So Defender prefers to offer $\delta\bar{\pi}^D$ rather than the entire resource only if

$$\pi^C_{t-1}\left(\boldsymbol{y}_{t-1}\right)<\frac{\left(1-\delta\bar{\pi}^D\right)}{(1-\delta\bar{\pi}^D)+(\delta L^D-V)}=:\pi'^C<1.$$

We now turn to the question of why pre-conflict negotiation fails with strictly positive probability when Δ is sufficiently small. Notice that $\delta\bar{\pi}^D\uparrow 1$ as $\Delta\downarrow 0$. That is, as Δ approaches 0, the equilibrium of the conflict phase in this version of the model approaches the equilibrium of our benchmark model. Therefore, Challenger's and Defender's ex-

[18]Recall that $V=\left(1-e^{-r\Delta}\right)=1-\delta$. Therefore $\delta\left(\frac{V}{1-\delta}-L^C\right)=\delta\left(1-L^C\right)$.

pected payoffs from conflict also approach their value in our benchmark model. It follows that, in the limit as Δ approaches 0, the equilibrium of the whole game (inclusive of pre-conflict negotiation) converges to that in our benchmark model. As for Proposition 2, a continuity argument works for any $\Delta > 0$ sufficiently small.

8 Discussion

We have seen how negotiations and conflict are linked by a two-way feedback. On the one hand, when choosing how to negotiate, the parties to a dispute take into account how they expect an eventual conflict to unfold. On the other hand, the way conflict unfolds is also determined by how and why negotiations fail. Our model of this two-way feedback between negotiations and conflict uncovers a novel reason why negotiations may fail to avoid conflict: offers that have a higher probability of being accepted also increase the incentive for the aggressor to initiate conflict. Thus, negotiations can mitigate but not prevent conflict. This result should not be seen as a claim that negotiations are useless: a neutral observer who seeks to reduce conflict will gain from bringing the parties to the negotiation table. If negotiations succeed, then conflict is avoided; if they fail, conflict will be shorter. In fact, in our model, the result of failed negotiations is to increase the belief π_0^C that Challenger is tough, and therefore decrease the conflict horizon n. Furthermore, the ability to negotiate increases the expected payoff of both players.

One possible caveat to our results is that we do not give Defender a chance to afford peace with a tough Challenger unless he offers the entire resource to Challenger. In many realistic applications, even a committed Challenger would admit that before conflict ensues she is willing to accept a smaller offer $X < 1$. Nevertheless, such a scenario would not change the main insight from our model unless X is sufficiently small. In fact, suppose that $X > \left(1 - \pi^D\right) V$. As we discussed in Section 5, an offer equal or greater to $\left(1 - \pi^D\right) V$ would convince normal Challenger to accept for sure. We can then see whether Defender would prefer to strike a deal only with normal Challenger or with both normal and tough Challenger. In this case, as $\Delta \downarrow 0$, Defender's expected payoff of offering N equals

$$\frac{1}{r}\left[\left(1 - \pi^C\right)(1 - N) + \pi^C \pi^D\right]$$

where $N = \lim_{\Delta \downarrow 0} \left(1 - \pi^D\right) V$. Instead, Defender's expected payoff of offering X equals $(1 - X)/r$. Therefore, Defender prefers to strike a deal with only normal Challenger as long as

$$X > \left(1 - \pi^C\right) N + \pi^C \left(1 - \pi^D\right).$$

From Section 5, we know that if Defender prefers to strike a deal with only normal Challenger, then it strictly prefers to make an offer that normal Challenger both accepts and rejects with strictly positive probability.

In our model, we also do not allow Challenger to offer peace deals. But our key insight about the failure of negotiations is also true if both Defender and Challenger can make offers. In fact, if normal Defender accepts the equilibrium offer for sure, then post-negotiation beliefs would be such that normal Challenger would never attack (that is, $\pi_0^D = 1$). But then normal Defender would strictly prefer to reject such an offer. Similarly, consider a variant of the model in which there are K rounds of two-step negotiations. At each stage, Challenger makes a demand first. If Defender accepts the offer, the game ends; otherwise, Defender makes an offer. If Challenger accepts the offer, the game ends; otherwise, the game moves to the next round of negotiations (or to conflict, if the last round of negotiations has already been reached). If Challenger ever demands less than unity, she reveals herself to be the normal type and, in the ensuing game, Defender has no incentive to concede anything to her and Challenger would never attack. Thus Challenger will simply demand 1 at each round and Defender will refuse and make an unacceptably low offer (say, zero) until the last round—this recovers our brinkmanship result.

Our brinkmanship result shows that there is little point in insisting on multiple rounds of negotiations. If acceptable offers are to be made, they will only be made at the last opportunity to avert conflict. In a sense, neutral observers should take advantage of any ultimatum imposed by the parties, rather than pressuring them to have softer deadlines. An interesting extension of our model would be to allow for there to be multiple commitment types of Challenger, from softer ones that would accept all offers above a threshold less than 1 to the toughest ones who would accept only $x = 1$. Even in this model, the forces that lead to delay would be present, but an additional force would be at work—an uninformed Defender could potentially screen some of the softest types in the early rounds of negotiations. Furthermore, the main insights from our model will still drive the analysis if we were to allow softer types to strategically mimic the toughest ones.

References

[1] **Abreu, Dilip and Faruk Gul (2000).** Bargaining and Reputation, *Econometrica,* 68(1): 85–117.

[2] **Acharya, Avidit and Edoardo Grillo (2015).** War with Crazy Types. *Political Science Research and Methods*, 3(2): 281–307.

[3] **Bénabou, Roland and Jean Tirole (2009).** Over my Dead Body: Bargaining and the price of dignity. *The American Economic Review* 99(2): 459–465.

[4] **Bennet, Scott and Allan C. Stam (1998)**. The declining advantages of democracy: A combined model of war outcomes and duration. *Journal of Conflict Resolution*, 42(3): 344–366.

[5] **Bennet, Scott and Allan C. Stam (1996).** The Duration of Interstate Wars, 1816-1985, *The American Political Science Review*, 90(2): 239–257.

[6] **Bennet, Scott and Allan C. Stam (2009).** Revisiting Predictions of War Duration, *Conflict Management and Peace Science*, 26(3): 256–267.

[7] **Brito, Dagobert L. and Michael D. Intriligator (1985).** Conflict, War and Redistribution, *The American Political Science Review*, 79(4): 943–957.

[8] **Bueno de Mesquita, Bruce and Alastair Smith (2012).** Domestic Explanations of International Relations, *Annual Review of Political Science*, 14: 1–21.

[9] **Chatterjee Kalyan and Larry Samuelson (1988).** Bargaining Under Two-sided Incomplete Information: The Unrestricted Offers Case. *Operations Research*, 36(4): 605–618.

[10] **Clark, David H. and William Reed (2003).** A Unified Model of War Onset and Outcome. *The Journal of Politics*, 65(1): 69–91.

[11] **Collier, Paul, Anke Hoeffler, and Mans Soderbom (2004).** On the Duration of Civil War, *Journal of Peace Research*, 41(3): 253–273.

[12] **Crawford, Vincent P. (1982).** A Theory of Disagreement in Bargaining, *Econometrica*, 50(3): 607–637.

[13] **Esteban, Joan and József Sákovics (2008).** A Theory of Agreements in the Shadow of Conflict: The Genesis of Bargaining Power, *Theory and Decision*, 65(3): 227–252.

[14] **Fearon, James D. (1995).** Rationalist Explanations for War, *International Organization*, 49(3): 379–414.

[15] **Fudenberg, Drew, and Jean Tirole (1991).** Perfect Bayesian equilibrium and sequential equilibrium, *Journal of Economic Theory* 53(2): 236–260.

[16] **Gul, Faruk, Hugo Sonnenschein and Robert Wilson (1986).** Foundations of Dynamic Monopoly and the Coase Conjecture, *Journal of Economic Theory* 39: 155–190.

[17] **Heifetz, Aviad and Ella Segev (2005).** Escalation and Delay in Protracted International Conflicts, *Mathematical Social Sciences*, 49: 17–37.

[18] **Henderson, Errol A. and Reşat Bayer (2013).** Wallets, Ballots, or Bullets: Does Wealth, Democracy, or Military Capabilities Determine War Outcomes?, *International Studies Quarterly*, 57(2): 303–317.

[19] **Hodler, Roland and Dominic Rohner (2012).** Electoral Terms and Terrorism, *Public Choice*, 150(1): 181–193.

[20] **Kreps, David and Robert Wilson (1982).** Reputation and Imperfect Information. *Journal of Economic Theory*, 27: 253–79.

[21] **Langlois, Catherine C. and Jean-Pierre P. Langlois (2009).** Does Attrition Behavior Help Explain the Duration of Interstate Wars? A Game Theoretic and Empirical Analysis, *International Studies Quarterly*, 53(4): 1051–1073.

[22] **Lapan, Harvey E. and Todd Sandler (1988).** To Bargain or Not to Bargain: That Is the Question, *The American Economic Review*, 78(2): 16–21.

[23] **Lyall, Jason (2010).** "Do Democracies Make Inferior Counterinsurgents? Reassessing Democracy's Impact on War Outcomes and Duration." *International Organization*, 64(1): 167-192.

[24] **Milgrom, Paul and John Roberts (1982).** Predation, Reputation, and Entry Deterrence, *Journal of Economic Theory*, 27: 280–312.

[25] **Özyurt, Selçuk (2014).** Audience Costs and Reputation in Crisis Bargaining, *Games and Economic Behavior*, 88: 250–259.

[26] **Pillar, Paul R. (1983).** *Negotiating Peace: War Termination as a Bargaining Process*, Princeton: Princeton University Press.

[27] **Ponsatì Clara and József Sákovics (1995).** The War of Attrition with Incomplete Information, *Mathematical Social Sciences*, 29(3): 239-254.

[28] **Powell, Robert (1996).** Bargaining in the Shadow of Power, *Games and Economic Behaviour*, 15: 255–289.

[29] **Powell, Robert (2004).** Bargaining and Learning While Fighting, *American Political Science Review*, 48(2): 344–361.

[30] **Reiter, Dan (2003).** Exploring the Bargaining Model of War, *Perspectives on Politics*, 1(1): 27–43.

[31] **Sanchez-Pages, Santiago (2009).** Conflict as a Part of the Bargaining Process, *Economic Journal*, 119: 1189–1207.

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

[33] **Schelling, Thomas C. (1956).** An Essay on Bargaining, *American Economic Review*, 6: 281–306.

[34] **Schelling, Thomas C. (1960).** *The Strategy of Conflict*, Cambridge, MA: Harvard University Press.

[35] **Schelling, Thomas C. (1966).** *Arms and Influence*, New Haven, CT: Yale University Press.

[36] **Silverman, Dan (2004).** Street Crime and Street Culture, *International Economic Review*, 45(3): 761–786.

[37] **Slantchev, Branislav L. (2004).** How initiators end their wars: The duration of warfare and the terms of peace, *American Journal of Political Science*, 48(4): 813–829.

[38] **Stanton, Fredrik (2010).** *Great Negotiations: Agreements that Changed the Modern World*, Yardley, Pennsylvania: Westholme Publishing.

[39] **Yared, Pierre (2009).** A Dynamic Theory of War and Peace, *Journal of Economic Theory*, 145: 1921–1950.

[40] **Yildiz, Muhamet (2005).** Waiting to Persuade. *The Quarterly Journal of Economics*, 119(1): 223–248.

A Appendix

A.1 Proof of Lemma 1

Proof. (i) Assume that normal Defender concedes immediately after Challenger attacks in period $t+1$. That is $\sigma^D_{t+1}(x) = 1$. By Bayes's rule, if Defender does not concede, then $\pi^D_{t+1}(x) = 1$. Therefore, normal Challenger would concede at $t+2$. Then the expected payoff of attacking in period $t+1$ equals

$$\left(1 - \pi^D_t(x)\right) \frac{V}{1-\delta} - L^C = \left(1 - \pi^D_t(x)\right) - L^C.$$

Since not attacking yields a payoff of 0, Challenger strictly prefers to attack if $\pi^D_t(x) < \bar{\pi}^D$, is indifferent if $\pi^D_t(x) = \bar{\pi}^D$, and strictly prefers to concede if $\pi^D_t(x) > \bar{\pi}^D$.

(ii) Assume that normal Challenger concedes in period $t+1$. That is, $\sigma_{t+1}(x) = 1$. By Bayes's rule, if Challenger does not concede, then $\pi^C_{t+1}(x) = 1$. Therefore, normal Defender would concede at $t+1$. Then the expected payoff of resisting in period t equals

$$V + \delta\left[\left(1 - \pi^C_t(x)\right)\frac{V}{1-\delta} - \pi^C_t(x) L^D\right] = (1-\delta) + \delta\left[\left(1 - \pi^C_t(x)\right) - \pi^C_t(x) L^D\right].$$

Since resisting yields a payoff of 0, Defender strictly prefers to resist if $\pi^C_t(x) < \bar{\pi}^C$, is indifferent if $\pi^C_t(x) = \bar{\pi}^C$, and strictly prefers to concede if $\pi^C_t(x) > \bar{\pi}^C$. □

A.2 Equilibrium in the game of conflict (Proposition 1)

We now present a few lemmas that identify necessary conditions for equilibrium and thereby pin down the unique one in the game of conflict.

We first ask if a normal player ever mimics the tough type. Lemma 3 says that only Challenger in period 1 can mimic the tough type (i.e. attack) with probability 1.

Lemma 3. *In any equilibrium, normal types of both players concede with strictly positive probability in all periods, except possibly Challenger in period* 1*.*

Proof. The *concession sequence* $\langle \kappa_i \rangle_{i \in \mathbb{N}}$ of any strategy profile is a sequence in $[0,1]$, where each odd (even) term is the probability that Challenger (respectively, Defender) concedes at that time conditional on no player having conceded yet. A concession sequence arising from an equilibrium profile is called an *equilibrium concession sequence*.

Lemma 3 then says that in any equilibrium concession sequence, all terms (except possibly the first) must be strictly positive.

Step 1. The proof is based on the key idea that if the string $(\kappa_i, 0, \kappa_{i+2})$ appears in an equilibrium concession sequence and $\kappa_{i+2} > 0$, then $\kappa_i = 1$: if the opponent is not conceding in the interim the value of concession can only go down because the positive cost to fighting strictly exceeds the flow utility derived from the resource; therefore concession should have been strictly better at the step before.

Step 2. We now show that, along any concession sequence, adjacent terms cannot be 0. Let $\kappa_i = 0 = \kappa_{i+1}$; if $\kappa_{i+2} > 0$, it would contradict Step 1. Induction implies that if two adjacent terms of the concession sequence are 0, all subsequent terms are 0 too. But since there is a positive probability of the tough type, it cannot be an equilibrium to never concede, knowing that your opponent will not. Therefore, no equilibrium concession sequence contains adjacent 0's.

Step 3. Suppose $\kappa_i = 0$ for some $i > 1$. By Step 2, we must have $\kappa_{i+1} > 0$; from Step 1 it means that $\kappa_{i-1} = 1$. If the player who is supposed to concede with probability 1 does not do so, his/her reputation immediately jumps to 1 and the normal opponent must concede immediately thereafter, i.e. $\kappa_i = 1$—a contradiction! □

Lemma 4 characterizes the players' strategies if they are indifferent for two consecutive periods.

Lemma 4. *If Challenger is indifferent between conceding at times t and $t+1$ in any equilibrium, then normal Defender's equilibrium concession probability and the public beliefs about him are*

$$\sigma_t^D(x) = \tilde{\sigma}^D\left(\pi_{t-1}^D(x)\right) := \frac{1-\bar{\pi}^D}{1-\pi_{t-1}^D(x)}; \text{ and } \pi_t^D(x) = \frac{\pi_{t-1}^D(x)}{\bar{\pi}^D} \tag{17}$$

respectively. Similarly, if Defender is indifferent between conceding at times t and $t+1$ in any equilibrium, then normal Challenger's probability of conceding and the public beliefs about her type are

$$\sigma_{t+1}^C(x) = \tilde{\sigma}^C\left(\pi_t^C(x)\right) := \frac{1-\bar{\pi}^C}{1-\pi_t^C(x)}, \text{ and } \pi_{t+1}^C(x) = \frac{\pi_t^C(x)}{\bar{\pi}^C}. \tag{18}$$

Proof. Let V_t^i:$[0,1] \times [0,1] \to \mathbb{R}$ be the mapping from public beliefs at a history at which player $i \in \{C, D\}$ moves in period t to the expected payoff of not conceding. For ease of exposition, we drop the history qualifier x from beliefs such that $\pi_t^i(x) = \pi_t^i$.

Challenger is indifferent if and only if $V_t^C\left(\pi_{t-1}^C, \pi_{t-1}^D\right) = 0$, which is the payoff of Challenger from conceding. Suppose that Challenger is indifferent for two consecutive periods: $V_t^C\left(\pi_{t-1}^C, \pi_{t-1}^D\right) = V_{t+1}^C\left(\pi_t^C, \pi_t^D\right) = 0$. By Lemma 3, $\sigma_{t+1}^C(x) \neq 1$. Therefore

$$\bar{\sigma}_t^D(x) = 1 - \bar{\pi}^D.$$

This corresponds to a strategy for Defender such that $\sigma_t^D(x) = \tilde{\sigma}^D\left(\pi_{t-1}^D\right)$ in (17) if $\pi_{t-1}^D \leq \bar{\pi}^C$.

Defender is indifferent if and only if $V_t^D\left(\pi_t^C, \pi_{t-1}^D\right) = -L^D$, the payoff Defender gets if he concedes.[19] Suppose that Defender is indifferent for two consecutive periods (or is indifferent at time t and concedes at time $t+1$): $V_{t+1}^D\left(\pi_{t+1}^C, \pi_t^D\right) = -L^D = V_t^D\left(\pi_t^C, \pi_{t-1}^D\right)$. By Lemma 3, $\sigma_t^D(x) \neq 1$. Therefore

$$\bar{\sigma}_{t+1}^C(x) = 1 - \bar{\pi}^C.$$

This corresponds to a strategy for Challenger such that $\sigma_{t+1}^C(x) = \tilde{\sigma}^C\left(\bar{\pi}_t^C\right)$ in (18) if $\pi_t^C \leq \bar{\pi}^C$. □

Remark 6. Challenger's mixing probability at $t = 1$ need not equal $\tilde{\sigma}^C$; Defender's mixing at $t = 1$ can be different from $\tilde{\sigma}^D$ only if Challenger strictly prefers to attack at $t = 1$.

Combining Lemmas 3 and 4, in equilibrium both players concede with the probabilities in Lemma 4 above, except possibly at $t = 1$. Beliefs evolve according to the above lemma, except possibly at $t = 1$ and until they hit $\bar{\pi}^C$ or $\bar{\pi}^D$.

Conflict continues as long as no player has conceded. If players mix, then beliefs about their type increase until a threshold is crossed.

Lemma 5 says that if both beliefs are strictly below their threshold, no belief leaps over the corresponding threshold at the next step without touching the corresponding threshold exactly.

Lemma 5. *In equilibrium (i)* $\pi_t^D(x) < \bar{\pi}^D$ *and* $\pi_{t+1}^C(x) < \bar{\pi}^C$ *implies* $\pi_{t+1}^D(x) \leq \bar{\pi}^D$ *; (ii)* $\pi_t^C(x) < \bar{\pi}^C$ *and* $\pi_t^D(x) < \bar{\pi}^D$ *implies* $\pi_{t+1}^C(x) \leq \bar{\pi}^D$.

Proof. For ease of exposition, we drop the history qualifier x from beliefs and write π_t^i in place of $\pi_t^i(x)$.

We proceed by contradiction. Let $\pi_t^D < \bar{\pi}^D$, $\pi_{t+1}^C < \bar{\pi}^C$ but $\pi_{t+1}^D > \bar{\pi}^D$. Lemma 1 implies that normal Challenger will concede with probability 1 at time $t+2$ if Defender does not

[19] This payoff is not 0 but $-L^D$ because when it is Defender's turn to decide if he wants to concede or resist, Challenger has already attacked and the loss will be experienced by Defender in the current period regardless of his choice of move.

concede at $t+1$. So if Defender does not concede at time $t+1$ he gets a continuation payoff of 1 from $t+2$ onward if Challenger is the normal type. Since Challenger is normal with probability $1-\pi_{t+1}^{C}$, Defender's payoff from $t+1$ (the current period) onward is

$$(1-\delta)\left(V-L^{D}\right)+\delta\left[\left(1-\pi_{t+1}^{C}\right)V-\pi_{t+1}^{C}L^{D}(1-\delta)\right].$$

Defender strictly prefers to not concede if the above exceeds the payoff $-(1-\delta)L^{D}$ from conceding immediately at $t+1$:

$$(1-\delta)V+\delta\left[\left(1-\pi_{t+1}^{C}\right)V-\pi_{t+1}^{C}L^{D}(1-\delta)\right]>0. \tag{19}$$

Inequality (19) reduces to $\pi_{t+1}^{C}<\bar{\pi}^{C}$, which is true by assumption. Therefore Defender strictly prefers to fight at $t+1$, i.e. $\sigma_{t+1}^{D}(x)=0$—this contradicts Lemma 3, implying that $\pi_{t}^{D}<\bar{\pi}^{D}$ and $\pi_{t+1}^{C}<\bar{\pi}^{C}$ cannot lead to $\pi_{t+1}^{D}>\bar{\pi}^{D}$.

Now let $\pi_{t}^{C}<\bar{\pi}^{C}$ and $\pi_{t}^{D}<\bar{\pi}^{D}$, but $\pi_{t+1}^{C}>\bar{\pi}^{C}$. By a similar logic Challenger strictly prefers to fight at $t+1$ if

$$-(1-\delta)L^{C}+\left(1-\pi_{t}^{D}\right)V>0.$$

The expression above reduces to $\pi_{t}^{D}<\bar{\pi}^{D}$. So Challenger strictly prefers to fight at $t+1$, i.e. $\sigma_{t+1}^{C}(x)=0$—this contradicts Lemma 3. □

What do our previous results imply about period 1's probability of attack? By Lemma 4, from period 2 onward beliefs must grow by a factor $\left(\bar{\pi}^{C}\right)^{-1}$ and $\left(\bar{\pi}^{D}\right)^{-1}$, respectively. The solid line in Figure 1 depicts the equilibrium evolution of beliefs in a conflict of horizon 2 with Challenger being more intimidating. The dashed line represents the evolution of beliefs if it is common knowledge that both Defender and Challenger play the strategies in Lemma 4 from period 1 onward. In this case, $\pi_{1}^{D}(x)<\bar{\pi}^{D}$ and $\pi_{2}^{C}(x)>\bar{\pi}^{C}$, violating Lemma 5. In equilibrium, Defender must concede with sufficiently large probability in period 1 so as to 'level the playing field' with Challenger and guarantee $\pi_{2}^{D}(x)=\bar{\pi}^{D}$. Since Defender is conceding with a higher probability than what would make Challenger indifferent, in period 1 Challenger strictly prefers to attack.

Remark 7. If Challenger is more intimidating, in period 1 Challenger attacks with probability 1 and Defender concedes with probability $1-\pi_{0}^{D}(x)/\left(\bar{\pi}^{D}\right)^{n}>1-\bar{\pi}^{D}$.

A similar logic applies to the case when Defender is at least as committed as Challenger. In this case, Challenger must concede with sufficiently high probability in period 1 so as to 'level the playing field' with Defender and guarantee $\pi_{n}^{C}=\bar{\pi}^{C}$.

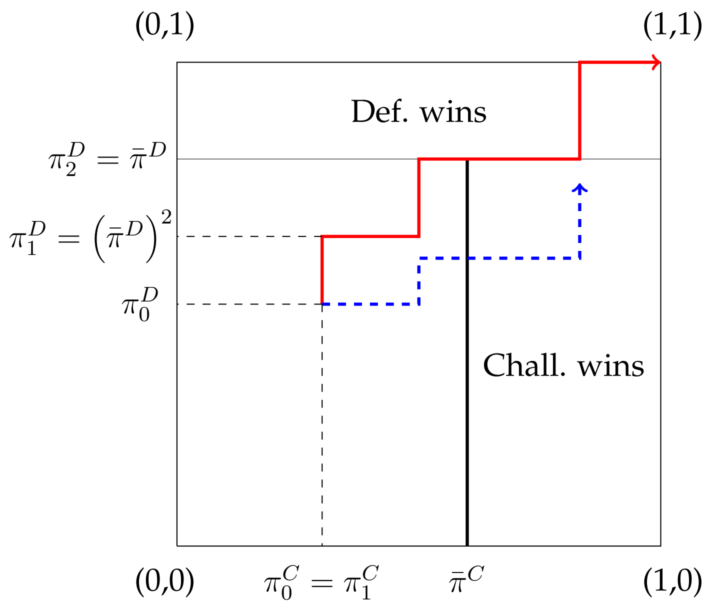

Figure 1: How Beliefs Evolve

Remark 8. If Defender is more intimidating, in period 1 Challenger attacks with probability $\pi_0^C(x) / \left(\bar{\pi}^C\right)^n$.

The next lemma shows that along the equilibrium path, provided no one concedes, both reputations grow according to (5) from period 2 onward, until a time t when either $\pi_t^C = \bar{\pi}^C$ or $\pi_t^D = \bar{\pi}^D$.

Lemma 6. *For any period $t \geq 2$, if $\pi_{t-1}^D(x) \leq \bar{\pi}^D$ and $\pi_t^C(x) \leq \bar{\pi}^C$, then Challenger plays $\tilde{\sigma}^C\left(\pi_t^C(x)\right)$ and Defender plays $\tilde{\sigma}^D\left(\pi_{t-1}^D(x)\right)$.*

Proof. We show the result for Defender. The result for Challenger follows a symmetric argument.

Proceed by contradiction. If $\sigma_t^D(x) \neq \tilde{\sigma}^D\left(\pi_t^D(x)\right)$, by Lemma 4, Challenger is not indifferent at either t or at $t+1$. There are two possibilities. First, she strictly prefers to concede. But then Defender would concede with probability 0 in the previous period, contradicting Lemma 3. Second, she strictly prefers to fight. But then by Lemma 3 she is Challenger in period 0 and $t = 1 < 2$. □

The lemma above is useful in proving Proposition 1. We divide the proof in two cases: when Challenger is more intimidating, and when Defender is more intimidating.

A.2.1 Proof of Proposition 1 when Challenger is more intimidating

Proof. Let $\left(\bar{\pi}^{C}\right)^{n+1} < \pi_0^C(x) < \left(\bar{\pi}^{C}\right)^{n}$ and $\pi_0^D(x) < \left(\bar{\pi}^{D}\right)^{n+1}$ for some $n \in \mathbb{N}$.

Existence. We first show that the strategies defined in Proposition 1 constitute an equilibrium. For ease of notation, we call this sequence of profiles of strategies $\sigma_t^* = \left(\sigma_t^{C*}, \sigma_t^{D*}\right)$, $t \in \mathbb{N}$. From Lemma 4 it is clear that after the first move by Challenger in period 1 players are indifferent and therefore willing to mix. Since normal players concede in * once the thresholds are crossed, this is consistent with Lemma 1. Since Defender concedes with a *larger* probability than $\tilde{\sigma}^D$ in the first period, Lemma 4 implies that Challenger strictly prefers to fight at $t = 1$. Also note that by Bayes' rule the equilibrium belief about Challenger's type after non-concession at $t = 1$ is given by $\left(\bar{\pi}^{C}\right)^{n}$.

Uniqueness. If $\pi_0^C(x) \geq \bar{\pi}^C$, then Lemma 1 implies that the above is the only equilibrium; similarly for the case $\pi_0^D(x) \geq \bar{\pi}^D$. Therefore let $\left(\pi_0^C(x), \pi_0^D(x)\right) < \left(\bar{\pi}^C, \bar{\pi}^D\right)$, so that $n \geq 1$. If normal types follow $\tilde{\sigma}^C, \tilde{\sigma}^D$ defined in equations (18) and (17) up to and including time n, then beliefs will cross their respective threshold in a way that violates Lemma 5, since $\pi_0^C(x) / \left(\bar{\pi}^{C}\right)^{n} > \bar{\pi}^C$ By Lemmas 4 and 6, the only freedom we have is in choosing different strategies for $t = 1$.

By contradiction, suppose that Challenger concedes with positive probability in period 1. This implies she expects Defender to concede with probability at least $\tilde{\sigma}^D$. But this implies that there is $m \leq n$ such that beliefs are $\left(\pi_{m+1}^C(x), \pi_m^D(x)\right)$ with $\pi_{m+1}^C(x) > \bar{\pi}^C$ and $\pi_m^D(x) < \bar{\pi}^D$, contradicting Lemma 5.

Last, since Challenger cannot concede with probability less than 0, we have $\pi_{n+1}^C(x) > \bar{\pi}^C$. Thus, by Lemma 5, Defender must concede in period 1 with probability exactly σ_1^{D*}. □

A.2.2 Proof of Proposition 1 when Defender is more intimidating

Proof. Let $\left(\bar{\pi}^{D}\right)^{n+1} < \pi_0^D(x) < \left(\bar{\pi}^{D}\right)^{n}$ and $\pi_0^C(x) < \left(\bar{\pi}^{C}\right)^{n}$ for some $n \in \mathbb{N}$.

Existence. Lemmas 1 and 4 imply that the strategy σ^* defined in Proposition 1 constitute an equilibrium. In particular, σ_1^{C*} and Bayes's rule imply that the equilibrium belief about Challenger's type after non-concession at $t = 1$ is given by $\left(\bar{\pi}^{C}\right)^{n}$.

Uniqueness. If $\pi_0^C(x) \geq \bar{\pi}^C$, then Lemma 1 implies that the above is the only equilibrium; similarly for the case $\pi_0^D(x) \geq \bar{\pi}^D$. Therefore let $\left(\pi_0^C(x), \pi_0^D(x)\right) < \left(\bar{\pi}^C, \bar{\pi}^D\right)$, so that $n \geq 1$. If normal types follow $\tilde{\sigma}^C, \tilde{\sigma}^D$ defined in equations (18) and (17) up to and including time n, then beliefs will cross their respective threshold in a way that violates

Lemma 5, since $\pi_0^D(x)/\left(\bar{\pi}^D\right)^n > \bar{\pi}^D$. By Lemmas 4 and 6, the only freedom we have is in choosing different strategies for $t = 1$.

Case 1: $\sigma_1^C < \sigma_1^{C*}$. Suppose that $\sigma_1^C < \sigma_1^{C*}$. The inequality $\sigma_1^C < \sigma_1^{C*}$ implies that Challenger's reputation increases at a slower rate such that $\pi_n^C(x) < \bar{\pi}^C$. If $\sigma_1^D < \tilde{\sigma}_1^D$, then Challenger prefers to concede immediately ($\sigma_1^C = 1$) since Challenger is just indifferent at $\tilde{\sigma}^D$; this contradiction implies that $\sigma_1^D \geq \tilde{\sigma}^D$, which in turn gives $\pi_1^D(x) \geq \pi_0^D(x)/\bar{\pi}^D$ and therefore $\pi_n^D(x) > \bar{\pi}^D$. I.e. there exists $m \leq n$ such that belief profile is $\left(\pi_m^C(x), \pi_m^D(x)\right)$ with $\pi_m^C(x) < \bar{\pi}^C$ and $\pi_m^D(x) > \bar{\pi}^D$, contradicting Lemma 5. Therefore, $\sigma_1^C \geq \sigma_1^{C*}$ is the only possibility in equilibrium.

Case 2: $\sigma_1^C > \sigma_1^{C*}$. Suppose that $\sigma_1^C > \sigma_1^{C*}$. Now $\pi_1^C(x) > \pi_0^C(x)/\bar{\pi}^C$, $\pi_2^C(x) > \pi_0^C(x)/\left(\bar{\pi}^C\right)^2$, etc. Since Defender's reputation is growing at rate $1/\bar{\pi}^D$ it follows from Defender being more intimidating and $\left(\bar{\pi}^C\right)^{n+1} < \pi_0^C(x)$ that $\pi_n^C(x) > \bar{\pi}^C$, i.e.,a violation of Lemma 5. Therefore, $\sigma_1^C \leq \sigma_1^{C*}$ is the only possibility in equilibrium.

Finally, since Challenger must be indifferent at $t = 0$ to play σ_1^{C*}, then $\sigma_1^D = \tilde{\sigma}^D = \sigma_1^{D*}$. □

A.3 Proof of Proposition 2

A.3.1 Proof of Lemma 2

Proof. When Challenger is more intimidating, the conflict horizon n satisfies $\left(\bar{\pi}^C\right)^n \leq \pi_0^C(x) \leq \left(\bar{\pi}^C\right)^{n+1}$. As $\Delta \downarrow 0$, both $\bar{\pi}^C \to 1$ and $\bar{\pi}^D \to 1$. Therefore the conflict horizon n tends to infinity. Taking (natural) logarithm and rearranging we get

$$1 \leq \frac{1}{n}\left(\frac{\ln \pi_0^C(x)}{\ln \bar{\pi}^C}\right) \leq 1 + \frac{1}{n} \;\Rightarrow \lim_{\Delta \downarrow 0} \frac{1}{n}\left(\frac{\ln \pi_0^C(x)}{\ln \bar{\pi}^C}\right) = 1.$$

This implies the first equality below:

$$\lim_{\Delta \downarrow 0} \ln\left(\bar{\pi}^D\right)^n = \lim_{\Delta \downarrow 0} n\left(\ln \bar{\pi}^D\right)\frac{1}{n}\left(\frac{\ln \pi_0^C(x)}{\ln \bar{\pi}^C}\right) = \left(\ln \pi_0^C(x)\right)\lim_{\Delta \downarrow 0}\left(\frac{\ln \bar{\pi}^D}{\ln \bar{\pi}^C}\right). \tag{20}$$

Noting that

$$\lim_{\Delta \downarrow 0}\left(-\frac{\ln \bar{\pi}^D}{\ln \bar{\pi}^C}\right) = \lim_{\Delta \downarrow 0}\left(\frac{\ln\left(1 - \left(1 - e^{-r\Delta}\right)\ell^C\right)}{\ln\left(e^{-r\Delta}\left[1 + \left(1 - e^{-r\Delta}\right)\ell^D\right]\right)}\right),$$

using L'Hôpital's rule, and taking the limit as $\Delta \downarrow 0$ yields

$$\lim_{\Delta \downarrow 0} \left(-\frac{\ln \bar{\pi}^D}{\ln \bar{\pi}^C} \right) = \frac{\ell^C}{1 - \ell^D}. \tag{21}$$

Note that the last expression is strictly negative because Assumption 1 implies that $\bar{\pi}^D, \bar{\pi}^C \in (0, 1)$. Substituting the limit in (21) into equation (20) gives the probability with which Defender concedes at the beginning of the limit game:

$$\lim_{\Delta \downarrow 0} \left(1 - \frac{\pi^D}{\left(\bar{\pi}^D\right)^n} \right) = 1 - \pi^D / \left\{ \lim_{\Delta \downarrow 0} \left(\bar{\pi}^D\right)^n \right\} = 1 - \pi^D \left(\pi_0^C (x) \right)^{\frac{\ell^C}{1 - \ell^D}}.$$

Finally, by Remark 1, $\lim_{\Delta \downarrow 0} u^C \left(\pi_0^C (x), \pi^D \right) = 1 - \pi^D \left(\pi_0^C (x) \right)^{\frac{\ell^C}{1-\ell^D}} - \lim_{\Delta \downarrow 0} L^C = 1 - \pi^D \left(\pi_0^C (x) \right)^{\frac{\ell^C}{1-\ell^D}}$. □

A.3.2 Proof of Proposition 2 for small $\Delta > 0$

Proof. We now show that that the solution to the game with discrete (i.e. $\Delta > 0$) but frequent enough opportunities to concede approaches the solution of the limiting problem we solve in the body of the paper. A value of Δ determines $\delta = e^{-r\Delta}$ and thus the thresholds $\bar{\pi}^C$ and $\bar{\pi}^D$ according to (2) and (3), which in turn determine the conflict horizon $n(\Delta, \bar{\alpha})$ as

$$\frac{\pi^C / (1 - \bar{\alpha})}{\left(\bar{\pi}^C\right)^{n(\Delta, \bar{\alpha}) - 1}} < \bar{\pi}^C \leq \frac{\pi^C / (1 - \bar{\alpha})}{\left(\bar{\pi}^C\right)^{n(\Delta, \bar{\alpha})}}. \tag{22}$$

This value of $n(\Delta, \bar{\alpha})$ in turn determines possible values of the equilibrium probability $\bar{\sigma}_1^D(x) = P^D(\Delta, \bar{\alpha})$ with which Defender concedes in period 1 if Challenger is more intimidating:

$$P^D(\Delta, \bar{\alpha}) = 1 - \frac{\pi^D}{\left(\bar{\pi}^D\right)^{n(\Delta, \bar{\alpha})}} \tag{23}$$

if both inequalities in (22) are strict, and

$$\underline{P}^D(\Delta, \bar{\alpha}) := 1 - \frac{\pi^D}{\left(\bar{\pi}^D\right)^{n(\Delta, \bar{\alpha}) + 1}} \leq P^D(\Delta, \bar{\alpha}) \leq 1 - \frac{\pi^D}{\left(\bar{\pi}^D\right)^{n(\Delta, \bar{\alpha})}} =: \bar{P}^D(\Delta, \bar{\alpha}) \tag{24}$$

if the inequality in (22) holds as an equality.

Defender's problem is

$$\mathcal{P}_\Delta : \quad \max\ U_\Delta^D(x, \bar{\alpha}) \text{ s.t. } (x, \bar{\alpha}) \in F(\Delta), \tag{25}$$

where $F : [0,1] \to [0,1] \times \left[0, 1 - \pi_0^C\right]$ gives the set of feasible pairs of $(x, \bar{\alpha})$ at Δ. A pair $(x, \bar{\alpha}) \in F(\Delta)$ if the following indifference condition holds, with $n(\Delta, \bar{\alpha})$ defined as in (22), and $P^D(\Delta, \bar{\alpha})$ defined by (23) or (24) as the case may be:

$$P^D(\Delta, \bar{\alpha}) - \ell^C \left(1 - e^{-r\Delta}\right) = x. \tag{26}$$

The limit problem studied in the body of the paper (see Section 5) is

$$\mathcal{P}_0 : \quad \max\ U_0^D(x, \bar{\alpha}) \text{ s.t. } (x, \bar{\alpha}) \in F(0), \tag{27}$$

where a pair $(x, \bar{\alpha}) \in F(0)$ if the following indifference condition holds

$$x = 1 - \pi^D \left(\pi^C / (1 - \bar{\alpha})\right)^{\frac{\ell^C}{1-\ell^D}} =: \ P^D(0, \bar{\alpha}). \tag{28}$$

We show in two steps that F is a compact-valued correspondence continuous at $\Delta = 0$.

<u>*Step 1.*</u> Let $\left(\Delta^k\right)_{k \geq 1}$ be a sequence decreasing to 0; let $\left(x^k, \bar{\alpha}^k\right)_k \in F\left(\Delta^k\right)$ be a sequence of feasible points. Consider a subsequence along which $\bar{\alpha}^k \to \bar{\alpha}$, drop the terms that are not in this convergent subsequence, and renumber the sequence. It is easy to see from (22), (23), and (24) that the correspondence mapping from $\bar{\alpha}$ values to x values is a step function for any Δ. Let the maximum and minimum values of x such that $\left(x^k, \bar{\alpha}^k\right) \in F\left(\Delta^k\right)$ be denoted by $\left[\underline{x}\left(\Delta^k, \bar{\alpha}^k\right), \bar{x}\left(\Delta^k, \bar{\alpha}^k\right)\right]$. Fix any $\eta > 0$. Since a step function is upper hemicontinuous, there exists an integer K_1 such that for all values of $k \geq K_1$ we have

$$d_H\left(x^k, \left[\underline{x}\left(\Delta^k, \bar{\alpha}^k\right), \bar{x}\left(\Delta^k, \bar{\alpha}^k\right)\right]\right) < \eta/2,$$

where d_H denotes the Hausdorff distance in Euclidean space. Now from Lemma 2 we get

$$\lim_{\Delta \downarrow 0} \underline{P}^D(\Delta, \bar{\alpha}) = P^D(0, \bar{\alpha}) = \lim_{\Delta \downarrow 0} \bar{P}^D(\Delta, \bar{\alpha}), \tag{29}$$

which by (26) implies that the size of the steps (the difference between the lower and upper bounds in (24)) vanishes: $\lim_{k\to\infty}\ \underline{x}\left(\Delta^k, \bar{\alpha}^k\right) = P^D(0, \bar{\alpha}) = \lim_{k\to\infty}\ \bar{x}\left(\Delta^k, \bar{\alpha}^k\right)$. Hence there exists $K_2 \geq K_1$ such that for all $k \geq K_2$ we have

$$d_H\left(P^D(0, \bar{\alpha}), \left[\underline{x}\left(\Delta^k, \bar{\alpha}^k\right), \bar{x}\left(\Delta^k, \bar{\alpha}^k\right)\right]\right) < \eta/2.$$

The triangle inequality for Euclidean spaces implies that $d_H\left(x^k, P^D(0, \bar{\alpha})\right) < \eta$ for all $k \geq K_2$. In other words, the limit of $\left(x^k, \bar{\alpha}^k\right)$ as k tends to infinity is $\left(P^D(0, \bar{\alpha}), \bar{\alpha}\right)$, which lies in $F(0)$ by (26). Therefore, F is upper hemicontinuous at $\Delta = 0$.

Step 2. For any sequence $\left(\Delta^k\right)_{k\geq 1}$ decreasing to 0, fix $(x, \bar{\alpha}) \in F(0)$, which implies $x = P^D(0, \bar{\alpha})$. Then pick any sequence $\left(x^k, \bar{\alpha}^k\right)_k$ such that each $\bar{\alpha}^k = \bar{\alpha}$, and $\left(x^k, \bar{\alpha}\right) \in F\left(\Delta^k\right)$. By (29) and Lemma 2 the bounds in (24) and the expression in (23) reduce to $P^D(0, \bar{\alpha})$ as $\Delta^k \to 0$. Therefore, (26) implies that we have $x^k \to P^D(0, \bar{\alpha}) = x$; hence F is lower hemicontinuous at $\Delta = 0$.

Since F is both upper and lower hemicontinuous at $\Delta = 0$, it is continuous at $\Delta = 0$.

The objective function $U^D_\Delta(x, \bar{\alpha})$ is jointly continuous in the variables $(x, \bar{\alpha})$ and the parameter Δ. Then the maximum theorem immediately implies that the set of optimal solutions is also upper hemicontinuous at $\Delta = 0$. In other words, as Δ goes to zero all optimal solutions of the constrained maximization problem $\mathcal{P}_\Delta$ approach the unique solution of the problem $\mathcal{P}_0$, the limiting problem we solve in the body of the paper.

Thus there exists $\bar{\Delta} > 0$ such that for any $\Delta < \bar{\Delta}$ the optimal offer x_Δ will be accepted with probability $\bar{\alpha}_\Delta(x_\Delta) > 0$. By Remarks 4 and 5, we know that an offer $x < u^C\left(\pi^C, \pi^D\right)$ cannot be accepted with positive probability, while an offer $x > u^C\left(1, \pi^D\right)$ cannot be rejected with positive probability. Thus an optimal offer x_Δ lies in $\left[u^C\left(\pi^C, \pi^D\right), u^C\left(1, \pi^D\right)\right]$. Furthermore, once again by upper hemicontinuity of the solution set, normal Challenger mixes in response to any optimal offer with probabilities close to $\bar{\alpha}^*/(1-\pi^C) \in (0, 1)$, and thus bounded away from 1. □

A.4 Proof of Proposition 3

Proof. Preliminaries. The negotiation stage lasts for K rounds $k = -K+1, \ldots, -1, 0$. In round k, a decision history h^D_k at which Defender moves comprises a vector $\boldsymbol{x}_{k-1} = (x_{-K+1}, \ldots, x_{k-1})$ of rejected past offers; the decision history h^C_k where Challenger moves is captured by $\boldsymbol{x}_k$, which comprises a vector of rejected past offers and the offer x_k currently on the table. A strategy of Defender maps $\boldsymbol{x}_{k-1}$ into an offer in $[0, 1]$; a strategy of normal Challenger maps each $\boldsymbol{x}_k$ into a probability of acceptance in $[0, 1]$. The acceptance probabilities of normal Challenger are denoted by $\alpha_k(\boldsymbol{x}_k)$ while the total probabilities are $\bar{\alpha}_k(\boldsymbol{x}_k) = \left(1 - \pi^C_{k-1}\right)\alpha_k(\boldsymbol{x}_k)$, where π^i_{k-1} is the probability with which player $i \in \{C, D\}$ is publicly believed to be tough at the *end* of round $k-1$. Since an offer $x_k < 1$ is rejected by tough Challenger, beliefs about Challenger are determined by Bayes' rule and Challenger's equilibrium strategy. The belief that Defender is tough equals π^D throughout the negotiation phase, since Defender cannot signal what he does not know.

The optimal offer at every history is Markovian, i.e., it deterministically depends on current beliefs. To see this, notice that the optimal offer at 0 is Markovian and deterministic, depending only on π^C_{-1}. Hence, Challenger's expected payoff of rejecting an offer at -1 depends only on π^C_{-1}. In turn this implies that normal Challenger's optimal response to any offer at -1 depends only on the current offer x_{-1} and the beliefs $\left(\pi^C_{-1}, \pi^D\right)$; thus the optimal offer at -1 depends only on the beliefs. By induction it follows that the optimal offer in all rounds $k = -K+1, \ldots, -1, 0$ is Markovian.

Limit as $\Delta \downarrow 0$ and $K = 2$. First, consider the limiting model as Δ goes to zero and fix $K = 2$. Let $\mathbb{A}\left(\pi^C, \pi^D\right)$ be the unique solution in $\bar{\alpha}$ to the first-order condition (15); let $\mathbb{X}\left(\pi^C, \pi^D\right)$ be the corresponding value of x. From an inspection of (15) it follows that $\mathbb{A}$ is increasing in π^C (strictly increasing unless $\mathbb{A}\left(\pi^C, \pi^D\right) = 1$), and so are the posterior probability $\pi^C / \left(1 - \mathbb{A}\left(\pi^C, \pi^D\right)\right)$ and the offer $\mathbb{X}\left(\pi^C, \pi^D\right)$.

Take any candidate equilibrium of the 2-offer game with equilibrium acceptance probabilities $(\bar{\alpha}_{-1}, \bar{\alpha}_0)$ such that $\bar{\alpha}_{-1} > 0$. The total probability (summed over rounds and types) that conflict will not begin is then given by $p = \bar{\alpha}_{-1} + (1 - \bar{\alpha}_{-1})\, \bar{\alpha}_0$. By Bayes' rule, the public belief that Challenger is tough at the start of conflict is

$$\pi^C_0 = \frac{\pi^C_{-1}}{1 - \bar{\alpha}_0} = \frac{\pi^C}{(1 - \bar{\alpha}_{-1})\,(1 - \bar{\alpha}_0)} = \frac{\pi^C}{1 - p}.$$

Clearly, sequential rationality requires that $x_0 = \mathbb{X}\left(\pi^C_{-1}, \pi^D\right)$, where $\pi^C_{-1} = \pi^C / (1 - \bar{\alpha}_{-1}) > \pi^C$ (since $\bar{\alpha}_{-1} > 0$). Hence $x_0 > \mathbb{X}\left(\pi^C, \pi^D\right)$. It must also be the case that Challenger is indifferent over the offer x_{-1} at round -1 and the offer x_0 at round 0. Thus the cost, measured at round 0, to Defender of getting the initial offer accepted with probability $\bar{\alpha}_{-1}$ at round -1 equals x_0. It follows that if the interval of time between rounds -1 and 0 is of length $\Delta_{-1,0}$ we must have

$$x_{-1} e^{\Delta_{-1,0}} = x_0.$$

Therefore, the utility of Defender is the same as if he were to strike a deal with a mass p of Challengers at a cost of x_0 per unit mass. Denote this utility by $\hat{U}^D\left(x_0, p; \pi^C, \pi^D\right)$. Since $x_0 \geq \mathbb{X}\left(\pi^C, \pi^D\right)$ and $p > \mathbb{A}\left(\pi^C, \pi^D\right)$, it follows that

$$\hat{U}^D\left(x_0, p; \pi^C, \pi^D\right) < \hat{U}^D\left(\left(\mathbb{X}\left(\pi^C, \pi^D\right), \mathbb{A}\left(\pi^C, \pi^D\right)\right); \pi^C, \pi^D\right), \tag{30}$$

whatever the time interval between the rounds of negotiation.

If, instead, Defender deviates in round 1 and chooses an offer so small (say $x_{-1} = 0$) that Challenger surely refuses ($\bar{\alpha}_{-1} = 0$), beliefs will be unchanged; in the continuation game it will be sequentially rational to make the optimal offer $\mathbb{X}\left(\pi^C, \pi^D\right)$ at time 0 and have it accepted with probability $\mathbb{A}\left(\pi^C, \pi^D\right)$. By inequality (30) this is a strictly profitable deviation. This establishes the result for the limiting model with $K = 2$.

Small but positive Δ. We now prove our proposition by induction on the number of rounds. First, consider $K = 2$. We proceed by contradiction.

If the proposition is not true, there exists $\varepsilon \in \left(0, 1-\pi^C\right)$, a sequence $\Delta^n \downarrow 0$, and a sequence of equilibrium outcomes $\left(\bar{\alpha}_{-1}\left(\Delta^n, \pi^C\right), x^n\right)$ of the game with interval Δ^n between consecutive periods of conflict such that the probability of the first-round equilibrium offer being accepted is bounded away from zero along the sequence: $\bar{\alpha}_{-1}\left(\Delta^n, \pi^C\right) \geq \varepsilon$. This immediately implies that the corresponding posteriors satisfy the bound:

$$\pi^C_{-1}\left(\Delta^n, \pi^C\right) := \frac{\pi^C}{1-\bar{\alpha}_{-1}\left(\Delta^n, \pi^C\right)} \geq \frac{\pi^C}{1-\varepsilon}.$$

Let γ be the function mapping the prior to the posterior in a one-offer limiting game (as $\Delta \downarrow 0$). Let $\eta := \gamma\left(\pi^C/\left(1-\varepsilon\right)\right) - \gamma\left(\pi^C\right)$. It can be seen from the first-order condition (15) that γ is continuous in π^C. Because the solution to the Defender's problem in the one-offer game is upper hemicontinuous in Δ (see Step 2 in A.3.2) and γ is continuous and increasing in the prior, there exists $\bar{\Delta}_1 > 0$ such that

$$\pi^C_0\left(\Delta^n, \pi^C\right) \geq \gamma\left(\frac{\pi^C}{1-\varepsilon}\right) - \frac{\eta}{4} \; \forall \Delta^n \in \left(0, \bar{\Delta}_1\right).$$

Furthermore, there exists $\bar{\Delta}_2 \leq \bar{\Delta}_1$ such that

$$\mid \pi^C_0\left(\Delta^n, \pi^C\right) - \gamma\left(\pi^C\right) \mid \leq \frac{\eta}{4} \; \forall \Delta^n \in \left(0, \bar{\Delta}_2\right).$$

These last two inequalities immediately imply that $\pi^C_0\left(\Delta^n, \pi^C\right) > \gamma\left(\pi^C\right) + \eta/2$ for sufficiently small Δ^n. That is, for sufficiently small Δ^n, the posterior belief that Challenger is tough after two rounds of negotiation in which the first offer is accepted with at least ϵ probability is greater than and bounded away from the posterior if the first offer is surely rejected. We can rewrite this condition as

$$\frac{\pi^C}{1-p\left(\Delta^n\right)} \geq \frac{\pi^C}{1-\bar{\alpha}^*} + \frac{\eta}{2},$$

where $p\left(\Delta^n\right)$ is the total (over two periods) probability with which an offer is accepted in the n^{th} game along the sequence. This inequality is clearly violated if there is a subsequence along which we have $p\left(\Delta^n\right) \to p\left(0\right)$. Hence we must have $p\left(\Delta^n\right) > p\left(0\right) + \zeta$ for some $\zeta > 0$. Now note that, for each Δ^n, any feasible solution of the two-offer game is also a feasible solution of the one-offer game, and vice versa. Thus, we have constructed a sequence of equilibrium acceptance probabilities of one-offer games with Δ^n going to 0 such that the equilibrum probability of acceptance of the single offer is bounded away from $p(0)$ from above along the sequence.

By upper hemicontinuity of the set of optimal solutions (see Step 2 in A.3.2), this means we have a new optimal solution of the limiting (as $\Delta \downarrow 0$) problem, with a probability of acceptance that is at least $p(0) + \eta$. This contradicts our proof that the one-offer limiting game has a unique solution with probability of acceptance $p(0)$.

We have therefore proven that a two-offer game is essentially the same as the one-offer game. Therefore, adding one more offer ($K = 3$) is equivalent to adding the second offer to the one-offer game. It follows that for any $K > 1$, all offers but the last one are accepted with vanishingly small probability. □

A.5 The role of Assumption 2 as $\Delta \to 0$

We solve the conflict part of our model for any positive Δ subject to Assumption 2. Suppose that, contrary to Assumption 2, $\ln \pi_0^C(x) / \ln \bar{\pi}^C = m \in \mathbb{N} \backslash \{1\}$ while $\ln \pi_0^D(x) / \ln \bar{\pi}^D < m$ (the case $\ln \pi_0^D(x) / \ln \bar{\pi}^D \in \mathbb{N} \setminus \{1\}$ is symmetric). Lemmas 4 to 6 do not rely on Assumption 2. Therefore, in all equilibria, beliefs move according to (5) from period 2, stage 1 onward. The first part of the proof of Proposition 1, then guarantees that there exists an equilibrium as in Proposition 1: at $t = 1$, Challenger attacks with probability 1 and Defender concedes with probability $1 - \pi_0^D(x) / \left(\bar{\pi}^D\right)^m$; $\pi_{m-1}^D(x) = \bar{\pi}^D$ and $\pi_m^C(x) = \bar{\pi}^C$. Yet, there exist also other equilibria. In all other equilibria, at $t = 1$, Challenger attacks with probability 1 and Defender concedes with probability p:

$$p \in \left[1 - \frac{\pi_0^D(x)}{\left(\bar{\pi}^D\right)^m}, 1 - \frac{\pi_0^D(x)}{\left(\bar{\pi}^D\right)^{m-1}}\right].$$

Therefore, $\pi_m^C(x) = \bar{\pi}^C$ and $\pi_{m-1}^D(x) < \bar{\pi}^D \leq \pi_m^D(x)$. Lemmas 4 to 6 guarantee that these are all the possible equilibria. While this means there are multiple equilibria, it is easy to see from Lemma 1 that

$$\left(1 - \frac{\pi_0^D(x)}{\left(\bar{\pi}^D\right)^{m-1}}\right) - \left(1 - \frac{\pi_0^D(x)}{\left(\bar{\pi}^D\right)^m}\right) \downarrow 0 \text{ as } \Delta \downarrow 0.$$

Therefore in the limit as $\Delta \downarrow 0$, the equilibrium strategies are uniquely determined by those in Proposition 1, Parts 1(a), 2, and 3, and Challenger's continuation payoff from conflict converges to the one given by Remark 1.