UNSW Business School

Working Paper

Never Stand Still | Business School | Economics

UNSW Business School Research Paper No. 2015 ECON 7

Intimidation: Linking Negotiation and Conflict

Sambuddha Ghosh
Gabriele Gratton
Caixia Shen

Intimidation: Linking Negotiation and Conflict

Sambuddha Ghosh, Gabriele Gratton, and Caixia Shen*

Abstract

Challenger demands a resource from Defender. In each period, Challenger chooses whether to attack; if attacked, Defender chooses whether to concede the resource forever. Each player might be committed to fighting until victory. Before conflict begins, Defender can make finitely many offers; conflict begins if Challenger rejects all offers. In equilibrium, all offers except the last are unacceptable. Negotiation cannot eliminate conflict because a larger offer makes conflict increasingly attractive for Challenger. If negotiation fails, prolonged conflict can happen in equilibrium, even when uncertainty is vanishingly small. We provide comparative statics regarding the probability and length of conflict.

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

---

*Ghosh: Department of Economics, Boston University, 270 Bay State Road, Boston, MA 02215 USA, sghosh@bu.edu. Gratton: School of Economics, UNSW Business School, 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 Gaurav Aryal, Laurent Bouton, Alessandra Casella, Micael Castanheira, 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. 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. Ghosh is indebted to Costas Cavounidis and Juan Ortner for extended discussions.

1 Introduction

Violent conflicts are long, costly, and often impossible to avoid. The latter fact is particularly puzzling: 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] One solution to this puzzle is to argue that political leaders are irrational. For one, this view sits uneasily with the idea that by and large leaders are rational; for another, it fails to generate any useful predictions and policy prescriptions.

We build a model of conflict to show that this can last several periods even if no party benefits from it. Our model is built on the concept of *intimidation*: repeatedly inflicting losses on one's opponent to scare him that further losses will be incurred unless a concession is made. For example, terrorists carry out attacks to raise the specter of further attacks until and unless their demands are met; similarly, workers go on strike to raise the fear of further strikes.[2] We model this as a game of two-sided incomplete information between Challenger and Defender. Our results are driven by *reputation*—with an arbitrarily small probability either party to the dispute is a tough type, who does not experience the disutility of conflict. Slight irrationality is magnified by equilibrium play into a significant force that protracts conflict.

We show that this same idea of intimidation also explains why negotiations fail: even if Defender can make offers to Challenger before conflict begins, equilibrium offers are always rejected with positive probability. The basic intuition is that, short of conceding everything, Defender cannot make offers that Challenger cannot refuse: a better offer affords Challenger a better chance to signal her toughness and trigger conflict on terms that are to her advantage. Moreover, our model also explains the phenomenon of *brinkmanship* in negotiations: conflict is averted only at the last minute, if at all, even when there are multiple chances of making an offer that would avert conflict. In particular, this means that the Coase conjecture fails in our setting, as discussed later.

[1] In a seminal paper, Fearon (1995) poses this as a "central puzzle" that rationalist explanations fail to solve.

[2] The analogy between strikes and wars dates back at least to Waltz (1979, p. 114).

Part I of this paper explains protracted conflicts. We model conflict as a repeated game with two-sided incomplete information. A resource that yields flow utility is currently in Defender's possession. There are infinitely many periods. In each period until the end of conflict, Challenger and Defender play a two-stage extensive-form game. In the first stage, Challenger decides whether or not to attack. In the second stage, Defender decides whether or not to concede the resource to Challenger's demand. Each player is privately informed about his or her type—*tough* or *normal*. A tough Challenger is able and committed to attack costlessly until Defender concedes; a normal Challenger incurs a cost in each period she attacks and maximizes the discounted sum of future utilities. A tough Defender never concedes as he does not incur the cost of conflict; a normal Defender suffers a loss from each attack and maximizes the discounted sum of future utilities. We find conditions under which long conflicts can occur even when irrationality becomes vanishingly small, i.e. both players are normal with probability arbitrarily close to 1.

In Part II we enrich the model with pre-conflict negotiations: Defender can make a sequence of offers in the form of a fraction of the contested resource. If Challenger accepts an offer, she immediately gets the promised share of the resource, Defender enjoys the rest, and conflict is averted. Conflict begins if Challenger rejects all offers. One might imagine that conflict is not avoided because of Defender's fear of revealing his type. We show that this intuition is incomplete because even if offers do not reveal Defender's type, the opportunity to make an offer is a double-edged sword for him. On one hand, offers that have a higher probability of being accepted increase the utility of Defender. Indeed, if beliefs were held fixed both before and after the offer, Defender could completely avoid conflict with the normal Challenger by offering slightly more than her expected value of entering conflict. On the other hand, if such an offer were made and accepted with certainty in equilibrium, Challenger would prefer to reject it to signal that she is tough, forcing Defender to concede as soon as conflict began. This detrimental effect of more generous offers rules out the possibility that negotiations succeed for sure. In equilibrium, Defender's optimal offer is accepted with probability strictly between 0 and 1. Thus, negotiations can mitigate but not avert conflict.

When Defender has multiple chances to make offers, in equilibrium he makes only unacceptable offers (in the sense that no Challenger would ever accept them) until his very last chance, when he makes exactly the same offer he makes when he has a single chance. Thus, long negotiations resolve in *brinkmanship*: the parties make no progress towards a peaceful solution up until the last opportunity before conflict begins.

The policy message is that although negotiations might fail to avoid conflict, these are not a waste of time. Both Challenger and Defender benefit from the opportunity to negotiate, as does any third party interested in mitigating conflict: if negotiations succeed, then conflict is avoided; if they fail, conflict is deemed to be shorter. However, when a deadline for making offers exists, multiple rounds of negotiations do not help to avert conflict any more than a single round.

In Part III we use our model to make a variety of comparative predictions regarding the likelihood and length of the conflict. We divide them into two groups. Suppose that a conflict is in progress, i.e. Challenger has attacked at least once and Defender has not conceded yet. The first set of predictions pertain to what happens next. A longer conflict is to be expected if either Challenger or Defender value the resource more, the costs of fighting are smaller, or players are more patient. Most importantly, only these parameters affect the probability that the conflict continues.

Our second set of predictions answer two questions: (i) what is the probability of a first attack? (ii) if Challenger attacks, what is the probability that the Defender concedes before any further attack? To answer these questions, we introduce a notion of what it means for one player to be more committed than the other. In equilibrium, when Challenger is more committed, she attacks for sure. Conversely, when Defender is at least as committed as Challenger, then Challenger mixes between conceding and attacking. In Section 10 we show how these predictions help explain some stylized facts about terrorist conflicts and workers' strikes.

Related Literature

Our main contribution is to identify the interaction between negotiations and conflict. The underlying model is one of reputation *à la* Kreps and Wilson (1982) and Milgrom and Roberts (1982)—a player with private information about his type mimics a commitment type so as to *guarantee* himself high payoffs.[3] Formally, we model post-negotiations conflict as a *costly* war of attrition with incomplete information.[4] We believe our framework to be particularly well-suited to the study of long and costly conflicts. To appreciate why, let us first compare our model to the closely related ones in Ponsatì and Sàkovics (1995) and Abreu and Gul (2000). Both study related continuous-time models of the war of attrition with two-sided incomplete information;[5] the latter shows that as the time interval between offers goes to zero all equilibria of the discrete-time game converge to the unique equilibrium in continuous time.

Our model differs in ways that make it more natural for our motivating examples. First, a key feature of these models is that concession can be delayed for a very brief time, making 'each' delay virtually costless. This hardly catches the case of terrorist conflicts where the decision not to concede to the terrorists might trigger a new, potentially devastating attack. The sequential structure of our model with significant costs captures this aspect. Second, time discounting is the only cost of waiting in these model, while ours accom-

[3]A commitment type is, quite naturally, wedded to a particular strategy, such as never conceding.

[4]The genesis of the extensive literature on the war of attrition is Maynard-Smith (1974). Hendricks *et al.* (1988) study a *complete-information* war of attrition. Kornhauser *et al.* (1989) apply reputational techniques to a war of attrition in discrete time, with committed types conceding with a constant probability in each period; the main goal is a set of conditions under which the less patient player concedes immediately. In Ordover and Rubinstein (1986) two players fight a war of attrition with a given *finite* deadline. Only one player privately knows who will 'win' if the game doesn't end by the deadline. Either the uninformed player concedes at the start or players remain indifferent between conceding and persisting. Hörner and Sahuguet (2010) model a war of attrition with two-sided asymmetric information about the *valuation* of the good to each player. The distinguishing feature is that both players can signal their types by sinking costly bids. Conflict never goes beyond two periods.

[5]The latter studies a bargaining model where players are, or pretend to be, irrationally committed to mutually incompatible demands.

modates an explicit cost of fighting, which is necessary to force concessions in our model because the resource is initially held by one party, as is often the case in economic and political applications. In doing so, we permit flow costs and benefits. Since delaying a concession carries the risk of further losses, the normal Defender strictly prefers to concede if he believes Challenger to be tough with sufficient probability, no matter what he expects a normal Challenger to do next. A similar statement holds for the normal Challenger. In this sense, our model of conflict is closest to a model of entry deterrence with two-sided incomplete information that Kreps and Wilson (1982) touch on, without deriving the equilibrium rigorously.[6]

These models identify sufficient conditions under which conflict resolves almost immediately as irrationality disappears.[7] Thus, comparative statics on the length and probability of conflict are highly sensitive to the prior probability of irrationality. Our model allows us to identify conditions that permit prolonged conflict in equilibrium even when the prior irrationality vanishes. These conditions are useful in understanding extended conflict.

The literature on bargaining under incomplete information can be traced back to Chatterjee and Samuelson (1983). Fudenberg, Levine, and Tirole (1985) and Gul, Sonnenschein and Wilson (1986) present dynamic models where a monopolist makes repeated offers to a buyer with unknown valuation. The main result is a verification of the Coase conjecture: as offers become frequent, the monopolist immediately offers a price that is acceptable to the lowest valuation buyer.[8] Our brinkmanship result, however, says that in our setup increasing the number of offers does not change the probability that negotiations succeed. Inefficiency is independent of the number of potential offers.

More broadly, we share our interest in the length of conflicts with a num-

---

[6]In Chatterjee and Samuelson (1988) both a buyer and a seller have private information and alternately make offers. Although flow utilities are absent, the qualitative features are shared: the soft type mimics the hard type initially. Uniqueness does not hold in their richer setting.

[7]See Fearon (2013) for a brief review of why most bargaining models fail to match the observation that conflicts are usually long.

[8]Cho (1990) extends the above results to two-sided incomplete information.

ber of papers in the political science literature, for example Powell (2004) and Fearon (2013), who study conflict as a problem of bargaining with private information (about the likelihood of winning the conflict at each period). Powell (1996) presents an alternating offer game where players can impose a settlement at a cost players have private information about; when parties become too pessimistic they take the outside option of conflict. In Brito and Intriligator (1985), and the closely related Sobel and Takahashi (1983), the *uninformed* party uses active conflict to screen among various types of informed parties, who wish to appear stronger than they really are in order to secure a better bargain. Sanchez-Pages (2009a) shows the converse—the *informed* party uses 'limited conflict' to convey credible information to an uninformed party about the eventual outcome of rejecting agreements and triggering 'absolute conflict'; this makes the latter more amenable to agreements even if he were optimistic enough initially to render agreements infeasible. Similarly, in Heifetz and Segev (2005) a party delays making an acceptable offer to credibly signal its true stand, and escalation makes resolution more attractive. See Baliga and Sjöström (2013) and Sanchez-Pages (2009b) for a highly integrated view of the above literature and more.

Lapan and Sandler (1988) model terrorism as a repeated game between players who are irrational with some probability. In their model, absent a concession, the probability of being a commitment type jumps up to an arbitrary and exogenously given quantity. Hodler and Rohner (2012) make this endogenous, but they have only two periods; this 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 of attrition 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.

An immediate implication of our results is that Defender wants to be perceived as tough. Schelling (1956, 1960, 1966) and Crawford (1982) 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. Yet, the advantage of showing commitment should not be overstated: once conflict begins, the expected payoff for (normal) Defender is independent of his probability of being tough. To see this, notice that the entire advantage of being perceived as tough comes from the fact that this induces 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 our model.

Part I

Conflict

Our model of conflict presents what happens if negotiation fails.

2 Setup

There are infinitely many periods $t \in \mathbb{N}$ and two players—Challenger (she) and Defender (he). Until conflict ends, a two-stage sequential game is played in each period:

Stage 1. Challenger chooses whether to fight (attack) or concede;

Stage 2. Defender chooses whether to fight or concede.

A contested resource is initially in Defender's possession. If Defender has not conceded yet, he enjoys a rent $d > 0$ from the resource in the current period; if he concedes, Challenger enjoys a rent $c > 0$ in the current period and in each subsequent period. In particular, periods after concession do not involve any actions, only flow rents, as concession prevents all future attacks.

Each player can be of two types—*tough* or *normal*. Challenger is tough with probability $\mu_0 > 0$; Defender is tough with probability $\pi_0 > 0$. Normal Challenger pays a cost $A > 0$ in each period she attacks; normal Defender suffers a loss $L > 0$ from each attack. Notice that Defender suffers the loss even if he concedes the resource immediately after the attack. Normal Challenger and Defender maximize the sum of future utilities discounted by factors $\delta^C > 0$ and $\delta^D > 0$ respectively.

Tough Challenger and Defender have no costs associated with attacks. Tough Challenger attacks until Defender concedes. Tough Defender never concedes.

We make a parametric assumption to restrict attention to interesting cases. The first part of Assumption 1 says that the loss from one attack in the next period exceeds the flow utility of the contested resource in the current period. The second part says that the cost of attacking is strictly less than the discounted value of getting the contested resource forever, starting from the current period. It immediately follows from this that Challenger strictly prefers to attack if she knows that one attack will result in Defender conceding immediately.

Assumption 1. *$\delta^D L > d$; $A < c\left(1 - \delta^C\right)^{-1}$.*

Note that if Assumption 1 fails, then either Challenger never attacks or Defender never concedes.

Our solution concept is perfect Bayesian equilibrium (henceforth *equilibrium*), since subgame-perfection does not have bite in such models of incomplete information.

2.1 Strategies and Beliefs

At each period t, in stage 1 the state of the game is the vector (μ_{t-1}, π_{t-1}), where μ_{t-1} is Defender's belief that Challenger is tough, and π_{t-1} is Challenger's beliefs that Defender is tough. In stage 2, the state vector is (μ_t, π_{t-1}), where Defender's belief about Challenger's type has been updated from μ_{t-1} to μ_t in light of Challenger's action at stage 1.

A (behavior) strategy[9] for Challenger is a sequence of mappings $\sigma_t^C : [0,1]^2 \to [0,1]$, $t \in \mathbb{N}$, where $\sigma_t^C(\mu_{t-1}, \pi_{t-1})$ is the probability that the normal type of Challenger concedes (does not attack) in period t as a function of the public beliefs. A strategy for Defender is a sequence of mappings $\sigma_t^D : [0,1]^2 \to [0,1]$, one for each $t \in \mathbb{N}$, where $\sigma_t^D(\mu_t, \pi_{t-1})$ is the probability with which Defender concedes in period t when the public beliefs are μ_t and π_{t-1}.[10]

Since tough players never concede, the *average conditional* probabilities of concession by Challenger and Defender respectively are obtained by multiplying the probability of the normal type by the probability that the (respective) normal type concedes:

$$\begin{aligned} \bar{\sigma}_t^C(\mu_{t-1}, \pi_{t-1}) &= (1-\mu_{t-1})\,\sigma_t^C(\mu_{t-1}, \pi_{t-1})\,; \\ \bar{\sigma}_t^D(\mu_t, \pi_{t-1}) &= (1-\pi_{t-1})\,\sigma_t^D(\mu_t, \pi_{t-1})\,. \end{aligned} \tag{1}$$

Obviously, $\mu_t = 0$ at any history where Challenger concedes. If Challenger has not conceded until period t, the updated belief μ_t that Challenger is tough is derived by Bayes' rule from μ_{t-1} and σ_t^C:

$$\mu_t = \frac{\mu_{t-1}}{1-\bar{\sigma}_t^C(\mu_{t-1}, \pi_{t-1})}. \tag{2}$$

[9] Writing the strategies as a function of public beliefs is without loss of generality. To see this point, notice that normally a strategy is defined on histories. However, in our setting, as soon as one party concedes, the game is over. Therefore, our formulation is the general one, where a strategy at t is a function of t and all possible histories after which the game is still being played.

[10] Note that Challenger's action in period t depends on the beliefs at the end of period $t-1$, as is standard. In contrast, Defender observes Challenger's move at t, updates her belief about Challenger's type to μ_t, and only then chooses an action.

Similarly, if Defender has refused to concede throughout we have

$$\pi_t = \frac{\pi_{t-1}}{1 - \bar{\sigma}_t^D(\mu_t, \pi_{t-1})}, \tag{3}$$

with $\pi_t = 0$ if Defender concedes at t.

3 Threshold Beliefs

An instance of our model is given by prior beliefs (μ_0 and π_0), costs (A and L), flows of utility (c and d) and discount factors ($\delta^C$and δ^D). In this section we derive some conditions that all equilibria must satisfy. This greatly simplifies our model as an instance of it will be given by only the prior beliefs and two *threshold beliefs*, $\bar{\mu}$ and $\bar{\pi}$.

Clearly, each player concedes if he or she believes the other to be tough with probability 1. Since payoffs are continuous in beliefs, if Defender believes Challenger to be tough with sufficiently high probability, then Defender will concede immediately, even if he knows that normal Challenger will never attack again. Similarly, if Challenger believes Defender to be tough with sufficiently high probability, then Challenger will concede immediately, even if she knows that normal Defender will concede with certainty in response to a further attack. The following lemma shows that optimal strategies have this threshold form.

Lemma 1 (Threshold beliefs)**.** *Let $\bar{\mu}$ and $\bar{\pi}$ be given by*

$$\begin{aligned} \bar{\mu} &:= \frac{1}{\delta^D}\left[1 + \left(1 - \delta^D\right)\frac{L}{d}\right]^{-1}, \text{ and} && (4) \\ \bar{\pi} &:= 1 - \left(1 - \delta^C\right)\frac{A}{c}. && (5) \end{aligned}$$

In any equilibrium,

(i) *if normal Defender concedes after a single attack, then normal Challenger strictly prefers to attack if $\pi_t < \bar{\pi}$, is just indifferent at $\bar{\pi}$, and strictly prefers not to attack otherwise;*

(ii) *if normal Challenger concedes after Defender does not concede once, then Defender strictly prefers not to concede if $\mu_t < \bar{\mu}$, is just indifferent at $\bar{\mu}$, and strictly prefers to concede otherwise.*

Proof. In Appendix. □

We shall use the fact that $\bar{\mu}$ and $\bar{\pi}$ are not functions of priors μ_0 and π_0. These thresholds depend in a very natural way on the discount factors of the players. To illustrate, $\bar{\pi} \to 1$ as Challenger becomes more patient ($\delta^C \to 1$), i.e. Challenger concedes only if Defender is almost certainly tough. Similarly, a very patient Defender concedes only if Challenger is almost certainly tough.[11]

For simplicity, we make the following genericity assumption, which rules out a negligible set (of zero Lebesgue measure) of priors but gives us uniqueness.

Assumption 2. *The quantities $\ln \bar{\pi} / \ln \pi_0$ and $\ln \bar{\mu} / \ln \mu_0$ are not integers.*

4 Equilibrium in the Game of Conflict

Our conflict game has a unique equilibrium. Depending on parameter values it is one of two *mutually exclusive* types. Both types have similar qualitative features, analogous to those in continuous-time wars of attrition. Propositions 1 and 2 stated below and proved in Appendix B.4 fully characterize the unique equilibrium.

First, solely from parameters we can compute a finite number of periods during which normal players fight with strictly positive probability. Beliefs during this phase are always below their respective thresholds. Thereafter, beliefs cross the thresholds and normal players concede.[12] This maximum duration of conflict between normal players is given by the *conflict order*.

[11] As $\delta^C, \delta^D \to 1$, i.e. as the interval between periods becomes smaller, our model gets closer to the continuous-time models in Ponsatì and Sàkovics (1996) and Abreu and Gul (2000). In the limit, if $\delta^C / \delta^D \to 0$, then (normal) Challenger never attacks; if $\delta^D / \delta^C \to 0$, Challenger attacks with probability 1 in period 1, and (normal) Defender concedes immediately.

[12] Conflict continues forever if and only if both players are tough.

Definition 1. A conflict is of order n if n is the largest non-negative integer such that $\mu_0 < \bar{\mu}^n$ and $\pi_0 < \bar{\pi}^n$.

Note that the order is 0 if either $\bar{\mu} < \mu_0 < 1$ or $\bar{\pi} < \pi_0 < 1$; otherwise it is non-zero. We assume that the conflict order is at least 1.[13]

Second, we determine which type of equilibria prevails. This depends on the *relative commitment* of Challenger and Defender.

Definition 2. In a conflict of order n, Challenger is *more committed* if $\bar{\mu}^{n+1} < \mu_0 < \bar{\mu}^n$ and $\pi_0 < \bar{\pi}^{n+1}$; Defender is *at least as committed* if $\mu_0 < \bar{\mu}^n$ and $\bar{\pi}^{n+1} < \pi_0 < \bar{\pi}^n$.

Proposition 1. *If Challenger is more committed, normal players play the following strategies in the unique equilibrium:*
(i) in period 1*, Challenger attacks with probability* 1 *and Defender concedes with probability*

$$1 - \frac{\pi_0}{\bar{\pi}^n}\frac{1-\bar{\pi}^n}{1-\pi_0}; \tag{6}$$

(ii) subsequently, Challenger and Defender concede with probabilities $\frac{1-\bar{\mu}}{1-\mu_t}$ *and* $\frac{1-\bar{\pi}}{1-\pi_t}$ *respectively until period* $n+1$*, Stage* 1*, and concede with probability* 1 *thereafter .*

Proposition 2. *If Defender is at least as committed, normal players play the following strategies in the unique equilibrium*
(i) in period 1*, Challenger concedes with probability*

$$1 - \frac{\mu_0}{\bar{\mu}^n}\frac{1-\bar{\mu}^n}{1-\mu_0}; \tag{7}$$

(ii) subsequently, Challenger and Defender concede with probabilities $\frac{1-\bar{\mu}}{1-\mu_t}$ *and* $\frac{1-\bar{\pi}}{1-\pi_t}$ *respectively until period* n*, and concede with probability* 1 *thereafter.*

The key difference between the two cases is that when Challenger is more committed, she attacks with probability 1 at the start, whereas she strictly mixes when Defender is at least as committed.

[13] Otherwise the game ends immediately as one prior is above its threshold belief.

The following results follow immediately from the probabilities of concession above.

Remark 1. If Challenger is more committed than Defender, then normal Challenger's expected payoff is

$$u^C := \left(1 - \frac{\pi_0}{\bar{\pi}^n}\right) \frac{c}{1 - \delta^C} - A \tag{8}$$

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

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

Remark 2. If Defender is at least as committed as Challenger, then normal Challenger's expected payoff is 0 and normal Defender's expected payoff is

$$u^D := \left(1 - \frac{\mu_0}{\bar{\mu}^n}\right) \frac{d}{1 - \delta^D} - \frac{\mu_0}{\bar{\mu}^n} L. \tag{9}$$

The unconditional probability of an attack in period $t \in [1, n+1]$ is given by

$$\Pr\left(\text{attack at } t\right) = \frac{\mu_0}{\bar{\mu}^n} \left(\bar{\mu}\bar{\pi}\right)^{t-1}.$$

4.1 Deriving the Unique Equilibrium

We now present a few lemmas that identify necessary conditions for equilibrium and thereby pin down a unique one in the game of conflict. Uniqueness permits comparative statics and facilitates the study of pre-conflict negotiations without grappling with equilibrium selection.

4.1.1 No Tough Play

We first ask if a normal player ever mimics the tough type.

Lemma 2. *In any equilibrium, normal types of both players concede with strictly positive probability in all periods, except possibly Challenger in period 1.*

Proof. In Appendix. □

We provide a simple intuition. If Challenger does not concede for several periods, this would force an immediate concession by Defender because Defender knows that (i) he will concede with positive probability at some future time t and that (ii) Challenger will not concede before then. The second fact means that at time t Defender's value of conflict is not larger than the value of conceding (otherwise he would not concede). The first fact implies that Defender should concede now with probability 1, as protracting conflict until t will only add costs for him. Should we conclude then that Defender concedes immediately because Challenger would not do so in the future? Of course not—if Defender deviates, Challenger would believe he is tough with probability 1 and concede immediately. Thus, in equilibrium both Challenger and Defender must always concede with positive probability (except possibly Challenger in period 1).

4.1.2 Mixing Throughout

Suppose that Challenger is indifferent between conceding at t and $t+1$. The former gives zero utility, whereas the latter requires Challenger to pay the cost of attack but also gives her some chance of getting everything starting from time t, in case Defender concedes at t. Hence indifference requires

$$0 = \bar{\sigma}^D \left(\mu_t, \pi_{t-1}\right) \frac{c}{1-\delta^C} - A$$

which implies $\bar{\sigma}^D \left(\mu_t, \pi_{t-1}\right) = 1 - \bar{\pi}$. Similarly, suppose Defender is indifferent between conceding at t and $t+1$. The former gives zero further utility, whereas the latter requires Defender to pay the cost of an attack if Challenger attacks again at $t+1$, but also gives Defender a rent at t and some chance of getting the same rent in all future periods, in case Challenger concedes at $t+1$. Hence indifference requires

$$0 = d + \bar{\sigma}^C \left(\mu_t, \pi_t\right) \frac{\delta^D}{1-\delta^D} d - \left[1 - \bar{\sigma}^C \left(\mu_t, \pi_t\right)\right] \delta^D L$$

which gives $\bar{\sigma}^C(\mu_t, \pi_t) = 1 - \bar{\mu}$. These indifference conditions lead to the lemma below.

Lemma 3. *If normal Challenger is indifferent between conceding at times t and $t+1$, then normal Defender's equilibrium concession probability and public beliefs about Defender's type are*

$$\tilde{\sigma}^D(\pi_{t-1}) := \frac{1-\bar{\pi}}{1-\pi_{t-1}} \text{ and } \pi_t = \frac{\pi_{t-1}}{\bar{\pi}} \tag{10}$$

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

$$\tilde{\sigma}^C(\mu_t) := \frac{1-\bar{\mu}}{1-\mu_t} \text{ and } \mu_{t+1} = \frac{\mu_t}{\bar{\mu}}. \tag{11}$$

Remark 3. 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 $t = 1$.

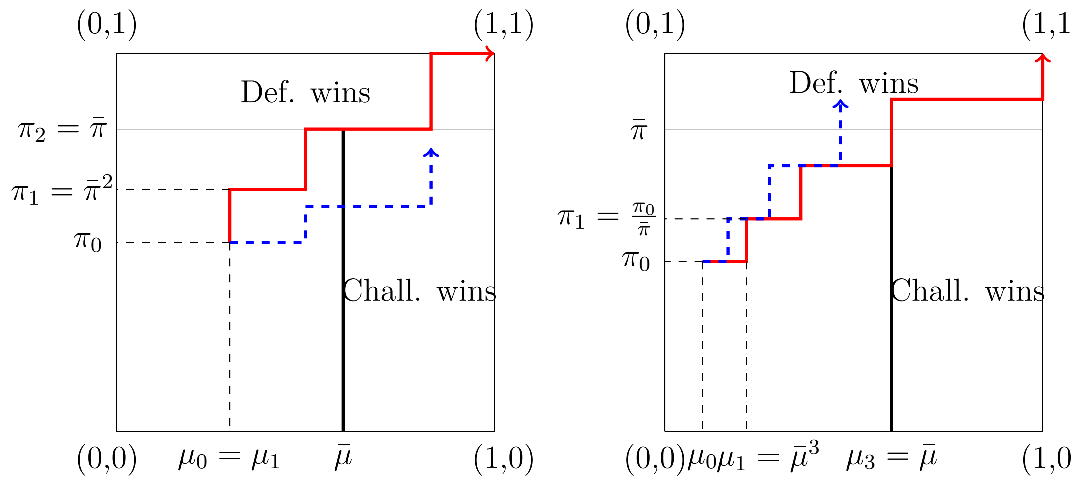
Combining Lemmas 2 and 3, in equilibrium both players concede with the probabilities in Lemma 3 above, except possibly at $t = 1$. Beliefs evolve according to Lemma 3, except possibly at $t = 1$ and until they hit $\bar{\mu}$ or $\bar{\pi}$.

4.1.3 Crossing the Thresholds

If players mix between conceding and fighting, public beliefs that they are tough increase with every act of fighting. Conflict ends for sure only when a threshold is crossed. Lemma 4 says that when an equilibrium belief crosses its threshold, the other belief is exactly equal to its own threshold.

Lemma 4. *In equilibrium (i) $\pi_t < \bar{\pi}$ and $\mu_{t+1} < \bar{\mu}$ implies $\pi_{t+1} \leq \bar{\pi}$; (ii) $\mu_t < \bar{\mu}$ and $\pi_t < \bar{\pi}$ implies $\mu_{t+1} \leq \bar{\mu}$.*

Proof. In Appendix. □



(a) Challenger is more committed, $n = 2$ (b) Defender is more committed, $n = 3$

Figure 1: How Beliefs Evolve

The intuition is simple and has much in common with the classic logic of reputation. Suppose that case (i) is true. By Lemma 1, if Defender knows that fighting once will move the game into his *region of victory* (Challenger's belief about Defender is above its threshold); then he will fight with probability 1, thereby violating Lemma 2. In other words, this affords Defender too much opportunity to modify the beliefs of Challenger to his own advantage. The implication is that when normal types finish conceding, one reputation is exactly at its threshold, i.e. either $\pi_n = \bar{\pi}$ or $\mu_n = \bar{\mu}$.

4.1.4 The First Period

What do our previous results imply about period 1's probability of attack? By Lemma 3, from period 2 onwards beliefs must grow by a factor $\bar{\mu}^{-1}$ and $\bar{\pi}^{-1}$, respectively. The solid line in Figure 1a depicts the equilibrium evolution of beliefs in a conflict of order 2 with Challenger more committed than Defender. The dashed line represents the evolution of beliefs if it is common knowledge that both Defender and Challenger play the strategies in lemma 3 from period 1 onwards. In this case, $\pi_1 < \bar{\pi}$ and $\mu_2 > \bar{\mu}$, violating Lemma 4. In equilibrium, Defender must concede with sufficiently large probability in period 1 so as to *level the playing field* with Challenger and guarantee $\pi_2 = \bar{\pi}$. Since Defender is conceding with a higher probability than what would make Challenger indifferent, in period 1 Challenger strictly prefers to attack.

Remark 4. If Challenger is more committed than Defender, in period 1 Challenger attacks with probability 1 and Defender concedes with probability $1 - \pi_0/\bar{\pi}^n > 1 - \bar{\pi}$.

Figure 1b depicts the equilibrium evolution of beliefs in a conflict of order 3 with Defender at least as committed. Now Challenger must concede with sufficiently high probability in period 1 so as to level the playing field with Defender and guarantee $\mu_n = \bar{\mu}$.

Remark 5. If Defender is at least as committed as Challenger, in period 1 Challenger attacks with probability $\mu_0/\bar{\mu}^n$.

5 Two Limits

5.1 Vanishingly Small Uncertainty

We explore the limit of our model as uncertainty about players' types becomes vanishes small. Our main interest is in determining conditions under which conflict can continue even when both parties are very likely to be normal. From the previous sections, a *conflict* (game) is completely characterized by the prior probabilities of being tough and the thresholds, $(\mu_0, \pi_0, \bar{\mu}, \bar{\pi})$, which is in $(0,1)^4$. Fixing $(\bar{\mu}, \bar{\pi}) \in (0,1)^2$, we can define a sequence of conflict games $\{\Gamma_k\}$ by varying the prior $(\tilde{\mu}_k, \tilde{\pi}_k)$; let the order of Γ_k be denoted by $n(k)$. We consider sequences that converge to complete information, i.e. $(\tilde{\mu}_k, \tilde{\pi}_k) \to (0,0)$. Obviously, $n(k) \to \infty$ along such sequences.

For such a sequence of conflicts converging towards rationality, the following definition ensures that no player is infinitely more committed (see Definition 2) than the other.

Definition 3. A sequence of conflicts $\{\Gamma_k\}$ is *equipoised* if

$$\lim_{k\to\infty} \frac{\tilde{\mu}_k/\bar{\mu}^{n(k)}}{\tilde{\mu}_k/\bar{\mu}^{n(k)} + \tilde{\pi}_k/\bar{\pi}^{n(k)}} \in (0,1)\,.$$

Define the ratio $Z_k := \{\tilde{\pi}_k/\tilde{\mu}_k\} / \{\bar{\mu}/\bar{\pi}\}^{n(k)}$. Henceforth we consider sequences along which the limit exists in the extended reals. The condition above may equivalently be stated as requiring that Z_k converges to a strictly positive real number, rather than limiting to 0 or infinity. Results below establish that equipoised sequences feature a non-vanishing probability of prolonged conflict even as we limit to complete information.

Proposition 3 considers a sequence of conflicts such that Defender is at least as committed as Challenger.

Proposition 3. *If the sequence $\{\Gamma_k\}$ of conflict games is equipoised and the Defender is at least as committed for all $k \in \mathbb{N}$, $\liminf_k \Pr(\text{attack at } t) > 0$ for all $t \in \mathbb{N}$.*

Proof. In Appendix. □

As an example, consider a symmetric conflict of order n with $\bar{\mu} = \bar{\pi}$ and equal priors. As uncertainty vanishes, the order of conflict grows. Along the natural sequence $\tilde{\mu}_k = \tilde{\pi}_k$, conflict happens at any period t with probability at least $\bar{\mu}\,(\bar{\mu}\bar{\pi})^{t-1}$. Notice that the probability of conflict diminishes to zero with t; we are interested above in how it behaves as k goes to infinity.

We now consider a sequence of conflicts such that Challenger is more committed than Defender.

Proposition 4. *If the sequence $\{\Gamma_k\}$ of conflict games is equipoised and Challenger is more committed for all $k \in \mathbb{N}$, $\liminf_k \Pr(\text{attack at } t) > 0$ for all $t \in \mathbb{N}$.*

Proof. In Appendix. □

It is easy to see that when the equipoise condition fails, i.e. the limit is either 0 or 1, one of the two players must concede with probability approaching unity. If we are in the case where Challenger is more committed, then this means that the first attack by Challenger (which happens with certainty) leads to almost certain concession. If we are in the case where Defender is more committed, this means that Challenger will trigger conflict with a vanishing small probability.

Abreu and Gul (2000) show a result which might appear to be in contradiction with ours. They consider a model which is identical to ours in the limit as the interval between periods goes to 0 or, equivalently, $\delta^C, \delta^D \to 1$. Notice that as $\delta^C, \delta^D \to 1$, then $\bar{\mu}, \bar{\pi} \to 1$. Their result is that in any sequence of such limit-conflicts with

$$\lim_{k \to \infty} \frac{\tilde{\pi}_k}{\tilde{\mu}_k + \tilde{\pi}_k} \in (0, 1)$$

and approaches complete information, the limit probability of conflict after the first instant is 0. They call such sequences *balanced*.

In the limit as $\delta^C, \delta^D \to 1$, balance and equipoise coincide. Nonetheless, for any strictly positive interval between periods, all equipoised sequences are *not* balanced unless $\bar{\mu} = \bar{\pi}$. Thus, studying our discrete model before taking the continuous-time limit allows us to exploit the thresholds $\bar{\mu}$ and $\bar{\pi}$ and understand under which conditions conflict protracts.

Notice that if a sequence is *not* balanced it may or may not be equipoised. Consider a sequence which is not balanced because the Defender is much more likely to be rational: $\tilde{\pi}_k / \tilde{\mu}_k \to 0$. If the thresholds are such that $\bar{\mu} < \bar{\pi}$, then Z_k goes to zero, because both factors go to zero (recall that $n(k)$ becomes infinitely large). On the other hand if $\bar{\mu} > \bar{\pi}$, the two factors in Z_k go to zero and infinity respectively. Then we have to look at the *rate* at which $\tilde{\pi}_k / \tilde{\mu}_k$ is going to zero. It seems intuitive that if it goes to zero *fast enough*, conflict will be brief (with Defender very likely to concede after the first attack); equipoise identifies precisely what is *not* fast enough.[14] The conclusion of the above is that our equipoise condition refines the balance condition in Abreu and Gul, and is necessary and sufficient for a non-vanishing probability of extended

[14]For example, it may be checked that with $\bar{\mu} = 1/2$ and $\bar{\pi} = 1/3$, conflict vanishes along the sequence $(\tilde{\mu}_k, \tilde{\pi}_k) = (1/k, 1/k^2)$ for $k \geq 2$, whereas it survives along the sequence $(\tilde{\mu}_k, \tilde{\pi}_k) = (1/k, 1/k^{3/2})$ for $k \geq 2$.

conflict even with vanishingly small uncertainty.

5.2 Continuous-time Limit

For analytical simplicity, we study the limit of our discrete game of conflict, i.e. we allow the time interval between periods to vanish. Time is non-negative and Δ is the interval between periods. Normal players discount time at the rate r such that $\delta = e^{-r\Delta}$. When we refer to 'limit equilibrium' it should be understood as the limit of the unique equilibria in discrete-time games as the time between successive periods of the war vanishes. This essentially allows us to differentiate.

Lemma 5. *In the limit equilibrium, when Challenger is more committed than Defender, then Defender concedes at time $t = 0$ with total probability*

$$P^D = 1 - \pi_0 \mu_0^{-\sigma^D/\sigma^C} > 0, \text{ where } \sigma^D \equiv \frac{A}{c/r}, \; \sigma^C \equiv \frac{L-d}{d/r}.$$

Proof. In Appendix. □

Part II

Negotiation

6 A Model of Negotiation

Our model predicts protracted conflict. Since conflict is wasteful, the normal Defender should prefer to sacrifice a positive fraction of the resources to avoid conflict. This begs a number of questions. Can negotiations avoid conflict? Who benefits from the possibility of negotiations? To answer these questions we extend the basic model of conflict to allow for negotiations before conflict.

There is a given deadline $T \geq 0$ and K points of time $t_1, \ldots, t_K$ in $[0, T)$ at which Defender can offer a fraction of the resource to Challenger. After each offer, Challenger decides whether to accept or reject an offer immediately thereafter.[15] If all offers are rejected, the conflict game presented in Part I is played out starting at time T. In other words, Challenger can first initiate conflict at time T. This captures the real-life gap between breakdown of negotiations and the initiation of hostilities.[16] Our model makes a very natural distinction between the negotiation and conflict phases—the former allows compromises, but the latter is all-or-nothing because the tough type of each player demands everything.

At $t = 0$, Challenger privately observes her type, while Defender does not know either type. Defender will privately learn his own type at $t = T$ if he is attacked. For example, in the context of terrorism, by facing attacks the government learns whether the public is willing and able to brave it. Since Defender does not possess private information, negotiation failure does not arise from his fear of revealing her type by making an offer. In Section 9 we discuss some alternative models, such as one where the privately informed agent makes offers.

Conditional on being tough, Defender's payoff is calculated as for the normal type but with $L = 0$ (although a tough Defender could for the purposes of Section 2 be thought of as a Defender with loss strictly less than d/δ^D, we simplify calculations by assuming that the tough type has a zero cost of fighting).

We focus on the case in which Challenger is (relatively) more committed than Defender. In this case, Challenger gets a strictly positive utility from conflict at the given priors; therefore, if no offer is made, at $t = T$ Challenger strictly prefers to attack. If all offers are rejected, Defender believes Challenger to be tough with probability $\mu_T \geq \mu_0$. So Challenger will initiate conflict and Defender will concede with mass probability P^D given by Lemma 5, with μ_0 replaced by μ_T.

[15]'Immediately' signifies that Challenger can start enjoying the resources from the time of the offer if she accepts it; this is merely to simplify calculations by eliminating discounting between the time of the offer being made and decided on.

[16]At the national and international levels T could be a significant length of time. We see no difficulty in allowing T to be random, but no significant gain in insight either.

7 Why Negotiation Fails ($K = 1$)

To illustrate the failure of negotiations in a transparent way, we begin with the simplest scenario: Defender can make a single offer at time $t = 0$. Defender's problem is to choose the optimal offer. Defender's utility from making an offer x will depend on the strategy of Challenger. If offer x is accepted by the normal type with probability $\alpha(x)$,[17] the total probability of acceptance is $\beta(x) = (1 - \mu_0)\,\alpha(x)$. Defender's total utility, as measured from date 0, as a function of his offer x is then given by

$$V_0^D\left(x, \beta(x); \mu_0, \pi_0\right) = \beta(x)(1-x)\frac{d}{r} + \\ + (1-\beta(x))\left\{\pi_0 \frac{d}{r} + (1-\pi_0)\left(1 - e^{-rT}\right)\frac{d}{r}\right\}.$$

The first term is clearly the payoff from the portion of the resource that stays with Defender; the second represents the expected payoff in case the offer is rejected: if Defender is tough, his continuation payoff (from time 0) will be d/r; if Defender is normal, we know that his continuation payoff will be the same as if he were to concede immediately at time T, i.e. $\left(1 - e^{-rT}\right) d/r$.

Challenger's decision depends on the comparison between the offer x and the alternative of conflict. The value of the offer x to Challenger is xc/r; the value of conflict is, of course, $P^D ce^{-rT}/r$, where P^D has been defined earlier and is increasing in the updated belief $\mu_T = \mu_0/\left(1 - \beta(x)\right)$.

Thus, the opportunity to make an offer is a double-edged sword for Defender, with one advantageous and one detrimental side. The advantageous effect is clear—offers that have a higher probability of being accepted increase the utility of Defender. Indeed if beliefs were held fixed both before and after the offers, Defender could completely avoid conflict with the normal Challenger by offering (slightly above) her expected value of entering conflict. But this means that Defender would concede at T if the initial offer were to be rejected.

[17] We can simply assume that tough Challenger doesn't accept anything short of the whole resource, but we can get this property for free because, as we shall see, Defender will never make an offer that would be large enough to be accepted by tough Challenger.

This is the detrimental effect of making better offers—because they are more likely to be accepted, bigger offers when rejected move beliefs further against Defender in the conflict phase. It is not *a priori* clear if the negative effect outweighs the positive effect. The following proposition says that (i) the advantageous effect always dominates at the low end, while (ii) the detrimental effect always dominates at the high end, resulting in the normal Challenger strictly mixing between accepting and rejecting the equilibrium offer.

Proposition 5. *The optimal offer has an acceptance probability* $\beta^* \in (0, 1 - \mu_0)$ *independent of* $T \geq 0$.

Proof. In Appendix. □

As it becomes more likely that the offer will be accepted, the endogenous outside option interacts with the offer (via the type) to produce a 'cat-chasing-its-tail effect': the value of rejecting the offer rises fast relative to the gain from accepting it and manages to stay ahead of the offer itself. Thus, negotiations can mitigate but not preempt conflict. Surprisingly enough, this holds regardless of the (absolute or relative) probabilities of being committed, as long as Challenger is relatively more committed. Informal discourse often suggests that we should surrender when the battle for reputations is skewed; our model shows that the undesirability of a bigger μ_0 is not enough that Defender should concede everything

The next proposition answers a very natural question that comes up often in discussions of conflict: should we negotiate? More formally, we compare two possible models—the first model is the one presented in this section; whereas the second affords no opportunity of making an offer at time 0 (so that nothing happens until Challenger attacks at time T and trigger conflict). It shows that negotiation benefits Defender, which is unsurprising since Defender can always make a zero offer. More importantly, *both* types of Challenger get a higher utility in the first model than in the second. Since $\beta^* > 0$, Challenger gets the same utility from accepting and rejecting the offer. Rejecting raises the probability of Challenger being committed, raises the probability of mass acceptance by Defender at $t = T$, and therefore gives the Challenger a higher utility. This is formally stated below. The extra opportunity to signal is thus valuable for both types of Challengers—it gives the tough type an extra opportunity to demonstrate toughness and raise her payoff; the normal type gets an extra opportunity to mimic the tough type without paying for an attack. Thus negotiation is Pareto optimal (for any fixed T).

Proposition 6. *Let u^{*Cn}, u^{*Ct}, u^{*D} denote the utilities at time 0 of the normal Challenger, the tough Challenger, and the Defender in the unique equilibrium of the game in which there is no chance to make an offer. Let $u^{*Cn}(\beta^*)$, $u^{*Ct}(\beta^*)$, $u^{*D}(\beta^*)$ denote the equilibrium utilities in the game in which the only offer is accepted with probability β^*. Then*

$$u^{*Cn}(\beta^*) > u^{*Cn},\ u^{*Ct}(\beta^*) > u^{*Ct},\ u^{*D}(\beta^*) > u^{*D}.$$

Proof. See discussion above the proposition. □

8 Brinkmanship ($K > 1$)

We now turn to the more general case with multiple rounds of offers in the interval $[0, T)$. It seems plausible that Defender might be able to progressively 'screen', i.e. take out a portion of normal types at each round, lowering the probability of conflict, which could conceivably go to 0 as the number of rounds increases. While our single-offer model is *a priori* restrictive, we now show that this is not in fact the case: our results are robust to an arbitrary number of offers in the time interval $[0, T]$. All offers except the one in the last period will be unacceptable; if negotiation succeeds it will do so on the brink of conflict. Thus brinkmanship emerges quite naturally from the model, driven by the same forces that could trigger conflict.

Let the posterior probability of the tough type at the beginning of round k be μ_k. A strategy of Defender comprises finitely many functions x_k for $k = 0, 1, \ldots, K$, mapping from $[0, 1]$ to $[0, 1]$ such that the k^{th} offer is $x_k(\mu_k)$. Let $(\beta_k \mid k = 0, 1, \ldots, K)$ be the corresponding 'acceptance probability' mappings from $[0, 1]$ to $[0, 1]$; the offer made in round $k \in \{0, 1, \ldots, K\}$ is accepted with probability $\beta_k(\mu_k)$. The strategy of Challenger decides the probability with which an offer is accepted in each round; we omit the notation for this.

Proposition 7. *All equilibria must satisfy the following for any prior μ_0:*

$$\beta_0^* = \ldots = \beta_{K-1}^* = 0 < \beta_K^* = \beta^*.$$

Proof. In Appendix.

For intuition assume that $K = 2$. First, note that even if there is early acceptance with positive probability, Defender will always make an offer at the last step that has positive probability of being accepted; this follows from Proposition 5. If an early offer is accepted with positive probability, early rejection raises the probability that Challenger is tough; this in turn raises the share that Defender offers later. Second, all Challengers get the same utility in equilibrium since she must be indifferent between accepting early and late. This simply means that Defender buys out a bigger proportion of normal types at a higher cost than is optimal in the model with only a single offer to make. Recall that the utility of the defender depends only on what proportion of normal types enter the conflict and what it costs to buy out the proportion that leaves. Therefore Defender buys out too big a proportion at too high a cost. □

This *brinkmanship* result explains why negotiations in the shadow of conflict sometimes make no progress until the very last chance. An instructive anecdote comes from the negotiations of the Treaty of Porthsmouth in the Russo-Japanese War (1904–1905). Negotiations for the treaty, which subsequently gained President Theodore Roosevelt the Peace Nobel Prize, were held while Russia brought further troops to Manchuria, a move that would have given Russia an advantage in case of conflict. For the whole length of the negotiations, the Japanese delegation demanded the southern part of the island of Sakhalin *and* war reparations. Only upon the arrival of further four Russian divisions, at what was conceivably their last chance to negotiate, did the Japanese drop their claim for reparations and avert conflict.[18]

[18]See Stanton (2010).

9 Extensions

This section discusses possible extensions of our model. We shall assume that we are in the 'Challenger wins' case, i.e. absent a change in beliefs Challenger strictly prefers to attack.

Extension 1. **One-sided private information with two-sided offers.** Our main model allows Defender to make offers, but does not allow Challenger to make demands. What happens if we allow rounds *à la* Stahl and Rubinstein? Just for the sake of maintaining consistency with the timing within rounds of the conflict game, Challenger first makes a demand, to which the Defender responds; if the response is a rejection, Defender makes a counteroffer. An acceptance as before ends the game with the accepted split of resources.

Results could depend on how the tough type behaves during negotiations. A reasonable interpretation of irrationality within the bounds of our conflict game is that the tough type never demands anything less than the whole pie. Assume that there is a single round. By demanding less than everything, the normal Challenger immediately reveals her rationality. Since our model is skewed towards Defender in terms of the initial allocation of resources, this would simply lead to Defender never offering anything to Challenger.[19] It is now easy to see that a multi-round version of this would be the same as the model with one-sided offers and multiple rounds.

Extension 2. **Two-sided private information with one-sided offer.** Suppose that Defender makes a single offer, but with knowledge of own type. Unlike the previous cases, the offer will move beliefs about his type since the tough type never makes any offer.[20] Since any positive offer reveals normalcy, the only offer that would be accepted by Challenger is the discounted value of the entire pie. Defender's problem is to then compare his utility when he

[19] It can be formally proved that in the limiting equilibrium of the conflict game the party who reveals rationality effectively ends up accepting the offer of the other player.

[20] This is a consequence of our assumption that there is no cost of fighting for the tough type.

makes this offer and the utility when he offers zero.

Part III

Applications and Conclusion

10 Comparative Statics and Applications

We provide some comparative statics regarding the length of conflict, conditional on (i) a first attack, and (ii) a first non-concession by Defender. This is well-defined thanks to our proof of uniqueness.

We first show that the maximum length of a conflict depends on the absolute likelihood that Challenger and Defender are tough. That is, on the conflict order n. If Challenger is more likely to be tough, then there is still an attack with positive probability in period $n+1$. Otherwise, attacks must end with period n. Thus, our model predicts that conflicts between players believed to be tough last for fewer periods.

Corollary 1. *Unless both players are tough, the maximum length of a conflict is determined by the conflict order n. If Challenger is more committed, there is never an attack after period $n+1$. If Defender is at least as committed, there is never an attack after period n.*

Remark. Recall that as uncertainty vanishes, the order of conflict and its maximum length tend to infinity.

Conditional on the conflict continuing after period 1, the probability that it lasts until $t \leq n$ is independent of which player is more committed or how much the players are likely to be tough. Between period 2 and n, the survival probability of the conflict depends only on the threshold values $\bar{\mu}$ and $\bar{\pi}$, which do not depend on the priors μ_0 and π_0. Thus, two sets of conflicts, one with a large and one with small order n, are empirically indistinguishable from period 2 until period $n+1$. Indeed, in each period $t \leq n$, the probability of a further attack is fixed at $\bar{\mu}$.

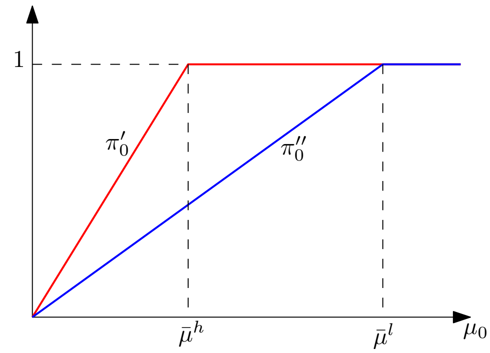


Figure 2: Probability of First Attack

Corollary 2. *Conditional on there being a conflict at the end of period* 1*, the probability of an attack in period* $t \in [2, n]$ *is*

$$\bar{\mu}^{t-1}\bar{\pi}^{t-2}.$$

Remark. In each period $t \in [2, n]$, if the conflict has not yet ended, Challenger attacks with constant probability $\bar{\mu}$.

The following comparative statics follows from the previous corollary and the definition of $\bar{\mu}$ and $\bar{\pi}$.

Corollary 3. *Conditional on there being a conflict at time* $t > 1$*, the probability of an attack in period* $t' > t$ *is increasing in* c*,* d*,* δ^C *and* δ^D*, and decreasing in* A *and* L*.*

We now turn to the question of when a conflict is more likely to begin, i.e. there is a first attack. Figure 2 depicts the total probability of first attack as a function of μ_0 for two different values of π_0, $\pi_0' \in (\bar{\pi}^{h+1}, \bar{\pi}^h)$ and $\pi_0'' \in (\bar{\pi}^{l+1}, \bar{\pi}^l)$. In both cases, the probability of first attack is strictly increasing for low values of μ_0 and equals 1 for higher values.

Corollary 4. *Fix the likelihood* π_0 *that Defender is tough. The probability that Challenger begins to attack is increasing in the likelihood of Challenger being tough* μ_0*. It is* strictly *increasing if and only if Defender is at least as committed as Challenger.*

For Defender, an image of toughness can pay: if Defender is at least as committed as Challenger, then the probability of a conflict is strictly less than 1. In this case, the probability of a conflict is $\mu_0/\bar{\mu}^n$, where n is the largest natural number such that $\pi_0 \leq \bar{\pi}^n$. Thus, if π_0 increases, the probability of conflict decreases. Figure 2 depicts the total probability of first attack as a function of μ_0 for $n = h$ (blue line) and $n = l < h$ (red line). Increasing π_0 from π_0' to $\pi_0'' > \pi_0'$ moves the line representing the probability of first attack to the right.

Corollary 5. *Let Defender be at least as committed as Challenger. Then, the probability that Challenger begins to attack is decreasing in* π_0.

The advantage of being perceived as tough should not be overstated. Recall that after the first attack, the expected payoff for Defender is $-L$, independently of π_0. Indeed, in equilibrium, Defender is indifferent between conceding and resisting whenever he plays.

10.1 Terrorism

Terrorist conflicts can last from a few months to many decades. A common distinction in the literature is the one between religious and secular terrorism. Religious terrorist groups are believed to be fiercely committed to an ultimate goal to the exclusion of compromise. For example, Taheri (1987) notes that a key difference between Islamic and secular terrorism is that the first "is clearly conceived and committed as a form of Holy War which can only end when total victory has been achieved" (p. 7) (see also Hoffman, 1995). With this imterpretation, our model predicts that religious terrorist groups are more likely to start conflict, but that such conflicts will be of shorter duration when they start.

Figure 3 depicts the number of attacks by year-quarter for four major terrorist groups from 1970 to 2012.[21] All four terrorist groups are ultimately

[21] Our calculations on data from National Consortium for the Study of Terrorism and Responses to Terrorism (START), 2012, Global Terrorism Database [globalterrorismdb_1012dist]. Retrieved from http://www.start.umd.edu/gtd.

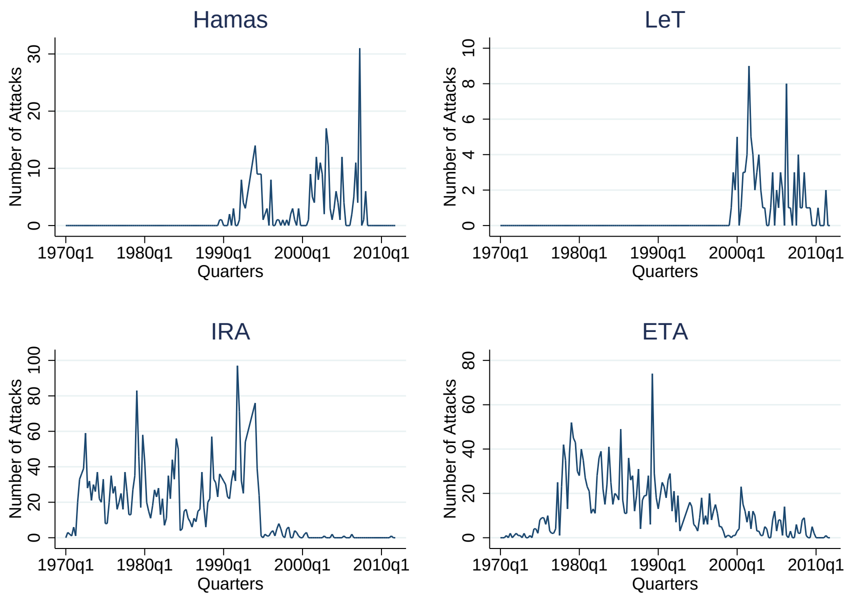


Figure 3: Number of Attacks by Quarter of Major Terrorist Groups

motivated by a separatist goal. Hamas and Lashkar-e-Taiba (LeT) are Islamic terrorist groups originating respectively in Palestine and Afghanistan (active in Pakistan and India). The (Provisional) Irish Republican Army (IRA) and Basque Fatherland and Freedom (ETA) are separatist military groups respectively in Ireland and the Basque Country in Spain and France. The pattern of conflicts is consitent with our model. Both religious groups, Hamas and LeT, show short burst of conflicts followed by periods of little or no activity. The time-series of Hamas's attacks, in particular, shows clearly three brief periods of conflict corresponding to the First Intifada (from Hamas foundation in 1988 to 1993), the Second Intifada (September 2001 to 2005) and the conflict culminated with the Gaza War (June 2006 to January 2009). In stark contrast, both secular groups, the IRA and ETA, show no such a pattern. The Troubles in Ireland and the UK which begun in 1969 lasted almost uninterrupted until the late 90s, when the Good Friday Agreements were signed (April 1998). Similarly, ETA's attacks continued uninterrupted from the assassination of of Admiral Luis Carrero Blanco (Franco's chosen successor) in 1973 to the *Spirit* *of Ermua* massive demonstrations against ETA in 1997. A second conflict started with a car bomb on January 21, 2000.

10.2 Workers' Unions and Macroeconomic Conditions

When workers go on strikes, they effectively incur a cost for themselves *and* impose a cost on their employers. The implicit threat to the employers is that the working conditions are so poor that the workers would lose their current salary rather than continuing to work. The committed type workers are therefore those who effectively have a sufficiently good outside option—typically, another employment. On the other hand, employers deny a concession to signal that they are unable to raise pay without jeopardizing the future of their firms. A committed employer is thus one who prefers losing human capital to raising pay.

The relevant stylized empirical facts are that strikes (i) are more likely in good economic times, and (ii) longer in bad economic times (Kennan and Wilson, 1990). Both facts agree with the predictions of our model.

In a typical business cycle, in good times, workers are more likely to have a good outside option—μ_0 is larger—and businesses are more likely to be able to raise pay—π_0 is smaller. Thus, our model predicts that strikes are more likely in good economic times (Challenger is relatively more likely to be committed) than in bad times. Furthermore, the cost of a strike is greater (at least for the businesses) in good economic times, i.e. L is greater in good economic times. Thus, our model predicts longer strikes in bad times than in good ones.

11 Conclusion

We provide a framework to make sensible predictions about both the onset and progression of wars of intimidation, where a challenger repeatedly attacks to convince the defender that more attacks will come if the latter doesn't concede. The model is built on small probabilities with which each player could be a type that faces no cost of fighting, and therefore persists with conflict until the opponent concedes. Our model identifies conditions for prolonged conflicts even as the uncertainty about the players' types is arbitrarily small. Our comparative static predictions find support in applications of interest in economics and political science.

Wars of intimidation begin when negotiations fail. Our model allows us to uncover a novel reason for why this happens: 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. On the other hand, 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. Our model ties together three key elements of conflict—brinkmanship, failed negotiations, and extended conflict—with the thread of intimidation and reputation.

References

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

[2] **Baliga, Sandeep and Tomas Sjöström (2013).** Bargaining and War: A Review of Some Formal Models, *The Korean Economic Review*, 29(2): 235–266.

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

[4] **Chatterjee Kalyan and William Samuelson (1983).** Bargaining under Incomplete Information, *Operations Research*, 31(5): 835–851

[5] **Chatterjee Kalyan and Larry Samuelson (1988).** Bargaining under two-sided incomplete information: The unrestricted offers case. *Operations Research*, 36(4): 605–618.

[6] **Cho, In-Koo (1990).** Uncertainty and Delay in Bargaining, *Review of Economic Studies*, 57: 575–596.

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

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

[9] **Fearon, James D. (2013).** Fighting Rather Than Bargaining, *mimeo*, Stanford University.

[10] **Fudenberg, Drew, David Levine and Jean Tirole (1985).** Infinite horizon models of bargaining with one-sided incomplete information. In: Roth, Al (ed.) *Game theoretic models of bargaining.* Cambridge: Cambridge University Press.

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

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

[13] **Hendricks, Ken, Andrew Weiss and Charles Wilson (1988).** The War of Attrition in Continuous Time with Complete Information. *International Economic Review*, 29(4), 663–680.

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

[15] **Hoffman, Bruce (1995).** "Holy Terror": The Implications of Terrorism Motivated by Religious Imperatives, *Studies in Conflict & Terrorism*, 18(4): 271–284.

[16] **Hörner, Johannes and Nicolas Sahuguet (2010).** A War of Attrition with Endogenous Effort Levels, *Economic Theory*, 47: 1–27.

[17] **Kennan, John and Robert Wilson (1990).** Can Strategic Bargaining Models Explain Collective Bargaining Data?, *The American Economic Review*, 80(2): 405–409.

[18] **Kornhauser, Lewis, Ariel Rubinstein and Wilson, Charles (1989).** Reputation and Patience in the'War of Attrition'. *Economica*, 15–24.

[19] **Kreps, David and Robert Wilson (1982).** Reputation and imperfect information. *Journal of Economic Theory*, 27, 253–79.

[20] **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.

[21] **Maynard-Smith, John (1974).** The Theory of Games and the Evolution of Animal Conflicts, *Journal of Theoretial Biology*, 47: 209–22.

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

[23] **Ordover, Janusz A. and Ariel Rubinstein (1986).** A Sequential Concession Game with Asymmetric Information, *The Quarterly Journal of Economics*, 101(4): 879–888.

[24] **Ponsatì Clara and Sákovics József (1995).** The war of attrition with incomplete information, *Mathematical Social Sciences*, 29(3), pages 239-254.

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

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

[27] **Sanchez-Pages, Santiago (2009a).** Conflict as a part of the bargaining process, *Economic Journal*, 119: 1189–1207.

[28] **Sanchez-Pages, Santiago (2009b).** Bargaining and Conflict with Incomplete Information, *working paper*.

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

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

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

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

[33] **Sobel, Joel and Ichiro Takahashi (1983).** A Multistage Model of Bargaining, *Review of Economic Studies*, 50: 411–426.

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

[35] **Taheri, Amir (1987).** *Holy Terror: The Inside Story of Islamic Terrorism*, Sphere Books, London.

[36] **Waltz, Kenneth N. (1979).** *Theory of International Politics*, Reading, MA: Addison-Wesley.

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

A Omitted Proofs

A.1 Proofs of Section 5.1

Proof of Proposition 3

Proof. Defender is at least as committed as Challenger is: $\frac{\tilde{\pi}_k}{\bar{\pi}^{n(k)}} \in (\bar{\pi}, 1)$. Thus, $\liminf \frac{\tilde{\pi}_k}{\bar{\pi}^{n(k)}} \geq \bar{\pi}$. By Proposition 2, for any $k \in \mathbb{N}$, the equilibrium probability of attack at time $t \in \mathbb{N}$ is given by

$$\Pr(\text{attack at } t) = (\bar{\mu}\bar{\pi})^{t-1} \frac{\tilde{\mu}_k}{\bar{\mu}^{n(k)}}, \text{ for all } t \in [1, n(k)].$$

$$\begin{aligned} \therefore \liminf \Pr(\text{attack at } t) &= (\bar{\mu}\bar{\pi})^{t-1} \liminf \frac{\tilde{\mu}_k}{\bar{\mu}^{n(k)}} \\ &\geq (\bar{\mu}\bar{\pi})^{t-1} \lim \frac{\tilde{\mu}_k/\bar{\mu}^{n(k)}}{\tilde{\pi}_k/\bar{\pi}^{n(k)}} \liminf \frac{\tilde{\pi}_k}{\bar{\pi}^{n(k)}} > 0, \text{ for all } t \in \mathbb{N}, \end{aligned}$$

where the last passage follows from noticing that $\lim \frac{\tilde{\mu}_k/\bar{\mu}^{n(k)}}{\tilde{\pi}_k/\bar{\pi}^{n(k)}} > 0$ for all equipoised sequences. □

Proof of Proposition 4

Proof. Challenger is more committed: $\frac{\tilde{\mu}_k}{\bar{\mu}^{n(k)}} \in (\bar{\mu}, 1)$. Thus $\liminf \frac{\tilde{\mu}_k}{\bar{\mu}^{n(k)}} \geq \bar{\mu}$. By Proposition 1, for any $k \in \mathbb{N}$, the equilibrium probability of attack at time $t \in \mathbb{N}$ is given by

$$\Pr(\text{attack at } t) = \bar{\mu}^{t-1}\bar{\pi}^{t-2} \frac{\tilde{\pi}_k}{\bar{\pi}^{n(k)}}, \text{ for all } t \in [2, n(k)+1].$$

$$\begin{aligned}
\therefore \liminf \Pr(\text{attack at } t) &= \bar{\mu}^{t-1}\bar{\pi}^{t-2} \liminf \frac{\tilde{\pi}_k}{\bar{\pi}^{n(k)}} \\
&\geq \bar{\mu}^{t-1}\bar{\pi}^{t-2} \lim \frac{\tilde{\pi}_k/\bar{\pi}^{n(k)}}{\tilde{\mu}_k/\bar{\mu}^{n(k)}} \liminf \frac{\tilde{\mu}_k}{\bar{\mu}^{n(k)}} > 0, \text{ for all } t \geq 2,
\end{aligned}$$

where the last step follows from noticing that $\lim \frac{\tilde{\pi}_k/\bar{\pi}^{n(k)}}{\tilde{\mu}_k/\bar{\mu}^{n(k)}} > 0$ for all equipoised sequences. □

A.2 Proof of Lemma 5

Proof. By definition of conflict order, when Challenger is relatively more committed, then $n \approx \ln(\mu_0)/\ln(\bar{\mu})$.

By Proposition 1, upon observing an attack, Defender concedes with total probability

$$\frac{\pi_0}{\bar{\pi}^n} \approx \pi_0 \mu_0^{-\frac{\ln \bar{\pi}}{\ln \bar{\mu}}}.$$

Using Lemma 1,

$$\frac{\ln \bar{\pi}}{\ln \bar{\mu}} = \frac{\ln\left[1 - (1-\delta) A/c\right]}{\ln\left[\delta\left(1 + (1-\delta) L/d\right)\right]}.$$

Replacing δ with $e^{-r\Delta}$, differentiating with respect to Δ, and taking the limit as Δ goes to 0 we get

$$\lim_{\Delta \downarrow 0} \frac{\ln \bar{\pi}}{\ln \bar{\mu}} = \frac{Ar/c}{r(L-d)/c} = \frac{\sigma^D}{\sigma^C},$$

and therefore

$$P^D = 1 - \frac{\pi_0}{\bar{\pi}^n} = 1 - \pi_0 \mu_0^{-\frac{\sigma^D}{\sigma^C}}. \tag{12}$$

It is worth noticing that σ^D and σ^C are the constant hazard rate of concession of Defender and Challenger in the limit case when $\Delta \downarrow 0$ (Appendix B.5 fully derives the concession strategies). □

A.3 Proofs of Section 7

Proof of Proposition 5

Proof. Let $\alpha(x)$ be the probability of normal Challenger accepting and $\beta(x) := (1-\mu_0)\beta(x)$. If Challenger is indifferent,

$$x = e^{-rT}\left\{1-\pi_0\left(\frac{\mu_0}{1-\beta}\right)^{-\sigma^D/\sigma^C}\right\}. \tag{13}$$

Let $\gamma := 1-e^{-rT}$. Defender's problem is to make the optimal offer, subject to constraint (13):

$$\beta^* := \underset{\beta\in[0,1-\mu_0]}{\arg\max}\ V_0^D\left(x(\beta),\beta;\mu_0,\pi_0\right).$$

Therefore β^* solves

$$\max V(\beta) = \beta(1-x) + (1-\beta)\left\{\pi_0 + (1-\pi_0)\gamma\right\}, \tag{14}$$

with x as function of β from (13) above. Differentiating (14) and noting that everything other than β and x are parameters,

$$\begin{aligned}\frac{dV(\beta)}{d\beta} &= \gamma + \pi_0(1-\gamma)\left(\frac{\mu_0}{1-\beta}\right)^{-\sigma^D/\sigma^C}\left(1-\frac{\beta}{1-\beta}\frac{\sigma^D}{\sigma^C}\right) - \{\pi_0 + (1-\pi_0)\gamma\} \\ &= \pi_0(1-\gamma)\left\{\left(\frac{\mu_0}{1-\beta}\right)^{-\sigma^D/\sigma^C}\left(1-\frac{\beta}{1-\beta}\frac{\sigma^D}{\sigma^C}\right)-1\right\}. \end{aligned} \tag{15}$$

We first show that it is always optimal to make an offer large enough that Challenger will accept the offer with non-zero probability:

$$\left.\frac{dV(\beta)}{d\beta}\right|_{\beta=0} = \pi_0(1-\gamma)\,\mu_0^{-\sigma^D/\sigma^C} > 0, \tag{16}$$

since $\mu_0 < 1$. Now,

$$\left.\frac{dV(\beta)}{d\beta}\right|_{\beta=1-\mu_0} = -\pi_0(1-\gamma)\frac{1-\mu_0}{\mu_0}\frac{\sigma^D}{\sigma^C} < 0, \tag{17}$$

i.e. it is suboptimal to completely avoid conflict with the normal Challenger.

To establish uniqueness, note that the acceptance probability β^* corresponding to an optimal offer satisfies

$$1 - \frac{\beta^*}{1-\beta^*}\frac{\sigma^D}{\sigma^C} = \left(\frac{\mu_0}{1-\beta^*}\right)^{\sigma^D/\sigma^C}. \tag{18}$$

The left hand side is strictly decreasing in β^*; hence the acceptance probability is unique. Denote by B the unique function satisfying equation (18). It is easy to see from equation (13) that when the offer increases so does the acceptance probability (until it hits $1-\mu_0$ when $x = e^{-rT}(1-\pi_0)$). Therefore the optimal offer is also unique. □

A.4 Proofs of Section 8

Proof of Proposition 7

Proof. We prove this by induction on the number of rounds. Let $K = 2$ henceforth.

The optimal offer at each state and each history is (i) Markovian, i.e. it depends only on beliefs about the type of Challenger; and (ii) deterministic. If B is the function such that $\beta^* = B(\mu_0)$ is the unique solution to the equation (18), then it is clear that B is increasing in μ_0 (strictly increasing unless it has hit 1), and so is the posterior probability $\mu_0/\left[1-B(\mu_0)\right]$.

Take any candidate equilibrium of the 2-offer game with the equilibrium acceptance functions (β_0, β_1). Clearly, sequential rationality requires that $\beta_1(\mu_1) = B(\mu_1)$. If possible, let $\bar{\beta}_0 := B(\mu_0) > 0$. Then the updated belief is $\mu_1 = \mu_0/\left(1-\bar{\beta}_0\right) > \mu_0$. The probability of acceptance in the second round is $\bar{\beta}_1 = B(\mu_1)$. The total probability (summed over rounds and types) that conflict will not start is then given by $\bar{\beta} = \beta_0 + (1-\beta_0)\beta_1$. The posterior at the end of round 2 is

$$\mu_2 = \frac{\mu_1}{1-\beta_1} = \frac{\mu_0}{(1-\beta_0)(1-\beta_1)} = \frac{\mu_0}{1-\bar{\beta}}.$$

By the above and equation (13), the sequentially rational offer in round 2 is

$$\bar{x} = e^{-r\Delta_{1,2}} \left\{ 1 - \pi_0 \left(\frac{\mu_0}{1-\bar{\beta}} \right)^{-\sigma^D/\sigma^C} \right\} \tag{19}$$

where $\Delta_{1,2} \geq 0$ is the time interval between the first and the second offer.

Incentive compatibility of the normal Challenger then requires that acceptance in any round give the same utility to Challenger in terms of round 2 shares. This, together with the fact that $\left(\bar{x}, \bar{\beta}\right)$ satisfies (19), means that the utility of Defender is the utility of a game in which he takes out a mass $\bar{\beta}$ of Challengers (at a cost of $\bar{x}$ per unit mass); this is denoted by $V_0^D\left(\bar{x}, \bar{\beta}; \mu_0, \pi_0\right)$). Let $\left(x^*\left(B\left(\mu_0\right)\right), B\left(\mu_0\right)\right)$ be the optimal pair when we have a single round with the prior μ_0. Since $\bar{x} > x^*$ and $\bar{\beta} > \beta^*$, it follows that

$$V_0^D\left(\bar{x}, \bar{\beta}; \mu_0, \pi_0\right) < V_0^D\left(x^*, \beta^*; \mu_0, \pi_0\right).$$

If Defender deviates in round 1 from β_0 and chooses $\beta_0'\left(\mu_0\right) = 0$ instead, it would have been sequentially rational to make the optimal offer x^* in round 2 and have it accepted with probability β^*. From the above we see that if two offers are accepted with positive probability in any equilibrium β^*, it is better to deviate and offer 0 at the first instant, so that the first offer is rejected. Hence there is no such equilibrium: the only possibility is $\beta_1^* = \ldots = \beta_{K-1}^* = 0$. □

B For Online Publication

B.1 Proof of Lemma 1

Proof. Challenger's payoff from choosing $\sigma_{t+1}^C\left(\mu_{t+1}, \pi_t\right) = 0$, i.e. attacking for sure at $t+1$, is maximum if after an attack at $t+1$ the normal type of Defender concedes for sure, *and* if Challenger concedes at $t+2$ if Defender does not concede at $t+1$. Note that when Defender concedes in period $t+1$, Challenger gets a flow payoff of c from the resource starting with the current period. When Defender is the commitment type, he doesn't concede at $t+1$ and Challenger gets none of the resource because she concedes at $t+2$. Hence Challenger's *maximum* payoff from attacking at $t+1$ is

$$-(1-\delta^C)A+(1-\pi_t)c+\pi_t\cdot 0. \tag{20}$$

Expression (20) is zero at $\bar{\pi}$, and negative above it. This proves part (i).

For part (ii) note that $\sigma^D_{t+1}(\mu_t, \pi_t) = 1$ implies that if Defender does not concede in period $t+1$ then Challenger must put probability 1 on the tough type, i.e. $\pi_t = 1$ and hence will find it optimal to concede at $t+2$ by step (i) above. Then the payoff of attacking is given by (20) and the payoff of conceding is 0; hence $\bar{\pi}$ is again the point of indifference.

Part (iii) follows from a similar argument to part (i). □

B.2 Proof of Lemma 2

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 2 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 0s.

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! $\square$

B.3 Proof of Lemma 4

Proof. Suppose not. Let $\pi_t < \bar{\pi}$, $\mu_{t+1} < \bar{\mu}$ but $\pi_{t+1} > \bar{\pi}$. Lemma 1 implies that normal Challenger will concede with probability 1 at time $t+2$ if Defender does not concede at $t+1$. So if Defender does not concede at time $t+1$ he gets a continuation payoff of d from $t+2$ onwards if Challenger is the normal type; since Challenger is normal with probability $1-\mu_{t+1}$, Defender's payoff from $t+1$ (the current period) onwards is

$$\left(1-\delta^D\right)(d-L) + \delta^D\left[\left(1-\mu_{t+1}\right)d + \mu_{t+1}\left(-L\left(1-\delta^D\right) + 0\cdot\delta^D\right)\right].$$

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

$$\left(1-\delta^D\right)d + \delta^D\left[\left(1-\mu_{t+1}\right)d - \mu_{t+1}L\left(1-\delta^D\right)\right] > 0. \tag{21}$$

Inequality (21) reduces to $\mu_{t+1} < \bar{\mu}$, which is true by assumption. Therefore Defender strictly prefers to fight at $t+1$, i.e. $\sigma_{t+1}^D\left(\mu_{t+1}, \pi_t\right) = 0$—which contradicts Lemma 2, implying that $\pi_t < \bar{\pi}$ and $\mu_{t+1} < \bar{\mu}$ cannot lead to $\pi_{t+1} > \bar{\pi}$.

Let $\mu_t < \bar{\mu}$ and $\pi_t < \bar{\pi}$, but $\mu_{t+1} > \bar{\mu}$. By a similar logic Challenger strictly prefers to fight at $t+1$ if

$$-\left(1-\delta^C\right)A + \left(1-\pi_t\right)c + \pi_t\cdot 0 > 0.$$

The net utility for Challenger to fight at period t is $-A$ and with probability $1-\pi_t$ Defender will concede and Challenger will get c forever. The expression above reduces to $\pi_t < \bar{\pi}$. So Challenger strictly prefers to fight at $t+1$, i.e. $\sigma^C_{t+1}(\mu_t, \pi_t) = 0$—which contradicts Lemma 2. $\square$

B.4 Propositions 1 and 2

We begin with a preliminary result: the next lemma shows that along the equilibrium path, provided no one concedes, both reputations grow according to equations (2) and (3) from period 2 onwards until a time t when either $\mu_t = \bar{\mu}$ or $\pi_t = \bar{\pi}$.

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

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

Proceed by contradiction. If $\sigma^D_t(\pi_t) \neq \tilde{\sigma}^D(\pi_t)$, by Lemma 3, 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 2. Second, she strictly prefers to fight. But then by Lemma 2 she is Challenger in period 0 and $t = 1 < 2$. $\square$

The lemma above is useful in proving Propositions 1 and 2, which apply, respectively, to the cases where Challenger is more committed than Defender and where Defender is at least as committed as Challenger.

Proof of Proposition 1

Proof. Existence. We first show that the stategies σ^* defined in Proposition 1 constitute an equilibrium. From Lemma 3 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 σ^* once the thresholds are crossed, this is consistent with Lemmas 1. Since Defender concedes with a *larger* probability than $\tilde{\sigma}^D$ in the first period, Lemma 3 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 $\bar{\mu}^n$.

Uniqueness. If $\mu_0 \geq \bar{\mu}$, then Lemma 1 implies that the above is the only equilibrium; similarly for the case $\pi_0 \geq \bar{\pi}$. Therefore let $(\mu_0, \pi_0) < (\bar{\mu}, \bar{\pi})$, so that $n \geq 1$. If normal types follow $\tilde{\sigma}^C, \tilde{\sigma}^D$ defined in equations (11) and (10) up to and including time n, there will be a jump since $\pi_0/\bar{\pi}^n > \bar{\pi}$; but jumps are ruled out by Lemma 4. By Lemmas 3 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 (μ_{m+1}, π_m) with $\mu_{m+1} > \bar{\mu}$ and $\pi_m < \bar{\pi}$, contradicting Lemma 4.

Last, since Challenger cannot concede with probability less than 0, we have that $\mu_{n+1} > \bar{\mu}$. Thus, by Lemma 4, Defender must concede in period 1 with probability exactly σ_t^{D*}. □

Proof of Proposition 2

Proof. Existence. As before, it can be checked that Lemmas 1 and 3 imply that the above is an equilibrium. In particular, σ_1^{C*} and Bayes' rule imply that the equilibrium belief about Challenger's type after non-concession at $t = 1$ is given by $\bar{\mu}^n$.

Uniqueness. If $\mu_0 \geq \bar{\mu}$, then Lemma 1 implies that the above is the only equilibrium; similarly for the case $\pi_0 \geq \bar{\pi}$. Therefore let $(\mu_0, \pi_0) < (\bar{\mu}, \bar{\pi})$, so that $n \geq 1$. If normal types follow $\tilde{\sigma}^C, \tilde{\sigma}^D$ defined in equations (11) and (10) up to and including time n, there will be a jump since $\pi_0/\bar{\pi}^n > \bar{\pi}$; but jumps are ruled out by Lemma 4. By Lemmas 3 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 $\mu_n < \bar{\mu}$. 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 $\pi_1 \geq \pi_0/\bar{\pi}$ and therefore $\pi_n > \bar{\pi}$ i.e. there exists $m \leq n$ such that belief profile is (μ_m, π_m) with $\mu_m < \bar{\mu}$ and $\pi_m > \bar{\pi}$, contradicting Lemma 4. 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 $\mu_1 > \mu_0/\bar{\mu}$, $\mu_2 > \mu_0/\bar{\mu}^2$, etc. Since Proposition 1 implies that Defender's reputation is growing as the same rate $1/\bar{\pi}$ it follows from the Definition 2 and $\bar{\mu}^{n+1} < \mu_0$ that $\mu_n > \bar{\mu}$, i.e. a jump occurs by time n. Therefore, $\sigma_1^C \leq \sigma_1^{C*}$ is the only possibility in equilibrium.

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

B.5 Unique Equilibrium of the Continuous-time Conflict Game

We can directly derive the unique equilibrium of the continuous-time game, which turns out to be the limit of equilibria in our discrete-time games. The following equilibrium properties follow from known results for continuous-time wars of attrition (see, for example, Abreu and Gul, 2000):

- until the normal type has conceded with probability 1, the normal type is indifferent between conceding at t and $t + \Delta$, except at $t = 0$;
  - at most one player can have a mass probability of concession at time 0;
  - each cumulative concession distribution is continuous for all $t > 0$;
  - normal types of both players must finish conceding at the same time.

We now derive the equilibrium concession rates.[22]

[22] Strictly speaking, the derivations are heuristic because we use a discrete-time approximation, but the formal derivation for continuous variables can be easily supplied by differentiating integrals that represent utilities. We prefer to adopt this approach because it lends itself to interpretation readily.

Challenger's Decision If Challenger is indifferent between conceding at t and $t+\Delta$ for$\Delta > 0$, the total cost of delaying concession equals the total gain from delay. The cost is $A\Delta$ (flow A for a length Δ) that is committed to at time t; the gain is the discounted value of the resource, which is earned only if the opponent surrenders in the interim:

$$A\Delta\left(1-\sigma^{D}\left(t\right)\Delta\right)=e^{-r\Delta}\sigma^{D}\left(t\right)\Delta\frac{c}{r},$$

which, canceling common factors and letting Δ go to zero, yields the following constant hazard rate of concession for D:

$$\sigma^{D}\left(t\right)=\sigma^{D}=\frac{Ar^{C}}{c}=\frac{A}{c/r}\ \text{ for } t>0. \tag{22}$$

Defender's Decision If the normal type of Defender can choose between conceding at t and $t+\Delta$, indifference requires:

$$L\Delta\left(1-\sigma^{C}\left(t\right)\Delta\right)-d\Delta=e^{-r^{D}\Delta}\sigma^{C}\left(t\right)\Delta\frac{d}{r^{D}}.$$

The term in brackets is the net cost of waiting. This includes the gross cost $L\Delta(1-\sigma^{C}(t)\Delta)$, stemming from loss at the rate L that incurred is only if Challenger does not concede in the interim; recall that Challenger concedes with probability $\sigma^{C}\left(t\right)\Delta$ over this interval. To derive the net cost, the sure benefit $d\Delta$ from delaying concession and enjoying the resource from t to $t+\Delta$ must be deducted from the gross cost.

Collecting higher order terms in Δ together,

$$\left(1+r\Delta+\mathcal{O}(\Delta)\right)\left(L\Delta-L\Delta\sigma^{C}\left(t\right))\Delta-d\Delta\right)=\sigma^{C}\left(t\right)\Delta\frac{d}{r}.$$

Simplifying and taking limits as $\Delta\to 0$,

$$\sigma^{C}\left(t\right)=\sigma^{C}=\frac{L-d}{d/r}\ \text{for } t>0. \tag{23}$$

Equations (22) and (23) show how to interpret the exponent of μ_0 in the expression (12) for the limiting probability that D concedes after the first attack.