ESCALATION AND DELAY IN PROTRACTED INTERNATIONAL CONFLICTS*

Aviad Heifetz† Ella Segev‡

This version: June 2003

Abstract

Why do escalations in protracted international conflicts sometimes hasten the pace of negotiations? And why is it sometimes the case that the resulting terms of agreement were deemed unacceptable to one or both sides before the escalation? We analyze these issues in a game-theoretic setting with asymmetric information, in which the delay a party exercises before it makes an acceptable offer is served to signal credibly its true stand, of which the other side is initially uncertain. Escalation makes both sides more eager to settle than before, as an agreement would end the increased level of hostilities. We analyze how this effect may loosen the incentives to exercise long delays in the course of bargaining, and hence shorten the time to agreement. However, it turns out that the larger is the overall increase in violence implied by escalation, the higher are also the chances that its initiator will eventually regret its own decision to escalate. These insights emerge both with one-sided and two-sided asymmetric information.

*We benefited from insightful comments of Anat Admati, Steve Brams, Bruce Bueno de Mesquita, Raymond Cohen, Jim Fearon, Gilat Levy, Zeev Maoz, Motty Perry, Nicholas Porteiro and Robert Powell.

†The School of Economics, Tel Aviv University, heifetz@post.tau.ac.il. Support from the Hammer fund for Economic Cooperation, Tel Aviv University and Prix FSR, CORE, Universite Catholique de Louvain are gratefully acknowledged.

‡The Faculty of Management, Tel Aviv University, segeve@post.tau.ac.il

1 Introduction

One of the puzzles of protracted international conflicts is that breakthrough and agreement sometimes follow soon after military escalation takes place. Moreover, it is sometimes even the case that when the conciliatory terms of the eventual agreement were proposed by one of the belligerents before the onset of escalation, they were rejected with confusion or mistrust by the opponent.

For example, the 1973 war between Israel and Egypt was not only intensive by itself, but led also to an accelerated arms-race in subsequent years. Nevertheless, the war was shortly followed by a series of disengagement agreements and finally a peace treaty in 1979. In this treaty Israel conceded the entire Sinai peninsula in exchange for peace. These very terms were proposed by Egypt already in 1971,[1] but Israel rejected them with mistrust.[2]

Similarly, the 1987-1990 Palestinian *Intifada* (upheaval) was soon followed by the Israeli-Palestinian 1993 Oslo Accords. These accords should have led, in a gradual process, to the establishment of a Palestinian state – an intolerable idea to most of the Israeli public before the *intifada.* However, Israel's initiative in 2000-2001 to *hasten* the negotiation process and terminate the conflict right away – suggesting a far-reaching territorial compromise and a *de-facto* division of Jerusalem[3] – was rejected with deep mistrust by the Palestinians,[4] and led to the outburst of a second, fierce *Intifada.*

In Ireland, the 1998 "Good Friday" agreement granted self-ruling to Northern Ireland in exchange for an appeal to decommission the paramilitary groups, but without a verification procedure for such a

[1] In an official document presented to the UN mediator Jarring on February 15, 1971 (Riad 1981, p. 188)

[2] In an interview to *The Times* on March 13, 1971 (cited in Cohen 1990, p. 158), the Israeli prime minister Golda Meir affirmed that president Sadat was the first Egyptian leader to say that he was "prepared to make peace." "At least, he said it. But does he mean it?"

[3] Shavit (2001).

[4] Agha and Malley (2001) report that Arafat's main aim in the 2000 Camp David summit was to evade coercion, since he viewed the summit as a coordinated Israeli-American trap, and therefore did not believe that the Israelis actually intend to abide by their own proposals.

decommission. However, when the IRA announced a unilateral cease-fire and a willingness to negotiate already in August 1994, it was encountered by confusion on the part of the British government, who demanded a unilateral complete decommission of the IRA, as a prior condition to negotiations. The cease-fire ended in February 1996 with an IRA bomb in London and the continuation of violence.[5]

What is the internal logic of such tragic chains of events? What is the "merit" of mutually-painful escalations in the course of protracted international conflicts? And why are proposals for mutual concessions confronted with mistrust when they are not preceded by long-enough suffering?

In this article we show how these phenomena are consistent with the idea that *the mere passage of time* in a protracted conflict may reveal substantive information to one or both parties about the reservation stand of the opponent, i.e. that *tacit bargaining* takes place while the conflict continues on at a given intensity. Moreover, we show that the *speed* in which this information gets revealed may be *increased* with the onset of escalation. To this end, we generalize a bargaining model in continuous time, in which delay per se serves in signaling.[6]

The equilibrium behavior in this model has the following property. Suppose that one of the parties hasn't yet come up with a serious, acceptable offer up to a given point in time. This very fact reveals that this side ascribes a lesser merit to settling (at some given terms) than in case this side had valued reconciliation more, and would therefore had strive to settle sooner. The delay that precedes the eventual offer thus becomes a credible signal regarding the actual reservation stand of the proposer, and hence also about the acceptable terms for reconciliation. When there is a mismatch between the delay and the content of the offer, it constitutes an off-equilibrium behavior, which is bound to be rejected by the opponent.

When the conflict escalates to a higher level of hostilities, the stakes at the negotiation table are higher as well, because an agreement will save the belligerents from the increased level of harm they inflict on one another. We identify the conditions under which these higher stakes will indeed shorten the delay to

[5] Mallie and McKittrick (2001, p.175-202).

[6] Admati and Perry (1987), Cramton (1992), Wang (2000).

agreement. Under these conditions, when one of the sides foresees large gains in compromising, it is less tempted than before to delay the settlement just in order to ameliorate the terms of agreement. Thus, when this side actually sees only mild gains to reconciling, a shorter delay already portrays to the opponent a behavior that wouldn't have been followed by a more reconciliatory type. Since this reasoning obtains for every such pair of distinct types, escalation will enable each type to come forward with the acceptable offer sooner than without escalation.[7]

This poses a dilemma to the other side. By escalating the conflict, it can hasten the pace of the tacit bargaining and hopefully end the conflict sooner. Also, escalation can influence the terms of the eventual agreement – the aggressor will eventually be better off if it is able to inflict a large flow of damages upon its opponent at relatively little cost to itself. On the other hand, the aggressor will have to endure the cost of escalation as long as the conflict goes on.[8] If the opponent sees rather little gains to agreement, the aggressor might eventually find itself yielding to agreement terms it would not have accepted at the outset, in order to end the harm it suffers in the more intensive conflict. In the worst case, it might turn out that the opponent prefers to fight indefinitely rather than settle even at the escalated level of the conflict. At that stage, however, it might be impossible to de-escalate and resume the more tolerable, ex ante level of hostilities.

We characterize the conditions under which a potential aggressor will prefer the gamble embodied by escalation over the status quo level of the conflict. With further specification of assumptions we are also able to demonstrate that the larger the extent of the escalation, the higher is also the chance that the aggressor will regret it ex post.

Our article belongs to a large body of literature which analyzes the interplay between negotiation and

[7]Except for the type who forsees the largest benefits, and makes its offer immediately in both cases.

[8]In real life the mutual damage inflicted by escalation may vary with time, in ways which are difficult to predict. In this article we abstract from this extra dimension of uncertainty, in order to isolate the interplay between the extent of the damages – whatever these turn to be – and the issue of timing in bargaining with asymmetric information about fundamental stands.

the use of force in the international arena. As in the spatial model of crisis bargaining (Morgan (1994), Morrow (1986)), our model allows for a large spectrum of bargaining outcomes.[9] As in Downs and Rocke (1987), it involves tacit bargaining, in the sense that actions – here the length of time before an offer is made– rather than words constitute the medium by which information is exchanged. The model provides a rationalist explanation to violence in the sense of Fearon (1995). According to his categorization, there are two kinds of such explanations: Those based on private information about relative capabilities or resolve, and explanations based on commitment problems in which one or more sides would have an incentive to renege upon the agreement terms. Our explanation belongs to the first kind, as do many other works which incorporate asymmetry of information into International Relations, in order to generate predictions concerning the role of information in determining crisis bargaining outcomes (e.g. Powel 1987, Morrow 1989,1992, Bueno de Mesquita & Lalman 1989, Bueno de Mesquita et al 1997, Fearon 1994, 1997, Banks 1990). These models predict the potential paths leading to either war or some non-violent resolution in an extensive-form game with finitely many stages. However, in most of these essays war is an outside option, and not a part of the negotiation process. Only a failure to resolve the dispute can lead to war, and there is no possibility of reaching an agreement after war has erupted – there is only a winning side and a losing one.

Several recent models have incorporated the use of force into the bargaining process. In the model of Wagner (2000), the outcome of a limited war may narrow the gap between the belligerents' initially-misaligned beliefs about their chances in a hypothetical all-out war. In particular, the defeated side may update its belief, and consequently yield to a deal it refused at the outset. In the articles of Filson and Werner (2002), Powell (2001), Ponsati (2002), and Smith and Stam (2001), a war or a conflict is a series of discrete costly events called battles, and negotiation can only take place before or after such an event. These models are the closest to ours, and we therefore turn to elaborate the differences between them and

[9]However, unlike in the spatial model, which predicts the probability of various bargaining outcomes or war, our model stipulates a definite outcome for each combination of reservation stands.

the model we offer here.

First, in our model the conflict is a continuos phenomenon rather than a series of discrete events. This seems a natural modeling choice in the context of protracted international conflicts, in which fighting per se may be sporadic, but the economic burden and the reduction in the quality of life due to the conflict are endured continuously. Typically, negotiation is conducted in parallel and not necessarily in interludes of the conflict.

Second, in the above models information is transmitted via the content of explicit proposals made in the negotiation phase, and then by the outcome of battles. As noted above, our analysis focuses on tacit rather than explicit bargaining, and in particular on the use of delay for information transmission.

Third, in all of the above essays war erupts when one side rejects the other's offer. In our model the decision to escalate is a *calculated risk* taken by one of the parties at some point in time during the protracted conflict, in the hope of influencing favorably the time to agreement and possibly the terms of agreement. In particular, it does not have to be preceded by a turned-down offer.

Another aspect in which our model differs from those mentioned above is the source of uncertainty. Asymmetric information in most of the above models stems from either an uncertainty about fighting capabilities (which may differ from commonly known indices of power), or about the costs of fighting. We focus here on a third dimension - asymmetric information about the opponents' reservation values regarding some issue in dispute. Such asymmetry of information is reasonable in protracted conflicts, while fighting capabilities and cost of fighting are less likely to remain private knowledge after the fighting goes on for a long while.[10]

[10]For example, in the Egyptian-Israeli 1979 peace treaty, Israel returned to Egypt the Sinai Peninsula even though Israel had the superior military position at the end of the 1973 war (its troops crossed the Suez canal and were only 101 Kilometers from Cairo, while the Egyptian 3rd Army in Sinai was under Israeli siege (Hertzog 1975)). Moreover, Israel's chances in an absolute war (in which its alleged nuclear superiority would be central) were not undermined by the war. Therefore, it is probably not a change in the assessments of fighting capabilities *per se* that have led to the breakthrough in the negotiations.

Finally, the above models do not address, at least not explicitly, the length of the time it takes the belligerents to reconcile. To address the issue of the speed of information revelation in tacit bargaining, and the influence of escalation on the time it takes to reach an agreement, we therefore turn to a model of bargaining in continuous time.

In the following section we present our basic model. A potential Aggressor with a known reservation stand is uncertain about the reservation stand of its rival, the Proposer, who should choose how long to wait before it comes up with an acceptable offer to resolve the dispute.

In section 3 we generalize this setting into a completely symmetric model, in which both sides can choose whether or not to escalate at the outset, while they are each uncertain about the reservation stand of the other. At equilibrium, a group of types of each side choose not to escalate, while the remaining types escalate. This initial decision therefore conveys some information to the other side. While enduring the resulting level of violence, both sides engage in a "war of attrition," each waiting for the other to approach the negotiation table. The party who foresees larger gains to agreement does so first, and the time it took it do so reveals its stand. The other party then waits further until it comes up with a serious offer which terminates the conflict. This full-fledged model delivers similar insights to those of the basic model. We conclude in section 4.

2 The Basic Model

Our basic model generalizes and adds an escalation stage to a bargaining model of Admati and Perry (1987) and Cramton (1992). We start by describing the bargaining model, in which one of the parties signals its stand by the delay it exercises. Next, we proceed with the modeling of escalation and the analysis of its implications.

2.1 Signaling Resolve by Delay

Two parties, A and B, are the sides to a protracted conflict. There is a one-dimensional set of potential resolutions to the dispute. These resolutions will be parametrized by real numbers S. We assume that the opposing sides have conflicting interests: If $S_1 < S_2$ are two potential resolutions, then A prefers S_1 to S_2, while B prefers S_2 to S_1.

Denote by Q the reservation stand of A – the resolution for which side A is indifferent between the continuation of the conflict and settling at the terms Q. In the basic model, we assume that this stand Q is common knowledge.[11] Similarly, denote by P the reservation stand of B – the resolution for which side B is indifferent between the continuation of the conflict and settling at P. Initially, side A does not know the stand P of B, which is private information of B and a priori distributed according to a probability distribution G with support $[\underline{P}, \overline{P}]$.

We scale the family of potential resolutions S such that $\overline{P} - S$ is the increase in the flow of welfare for side B with reservation stand $\overline{P}$ once the conflict is resolved at the terms S. In other words, the numbering S of potential real-world settlements corresponds to the subjective benefits the settlements accrue to B, and not necessarily to any objective measurement.[12]

Next, we assume that for every $P \in [\underline{P}, \overline{P}]$ there is a function $u_P : R \to R$ such that $u_P(S)$ is the increase in the flow of welfare for B when it has a reservation stand P once the conflict is resolved at the terms S. From the above description we have that u_P is a decreasing function of S (because B prefers settlements parametrized by lower S for any of its reservation stands $P \in [\underline{P}, \overline{P}]$), and that $u_P(P) = 0$ (the resolution P neither increases nor decreases the welfare of B with reservation stand P relative to the on-going conflict). To keep the model tractable, we proceed with the assumption that $u_P(S) = P - S$ for

[11] This assumption will be relaxed in the next section.

[12] For example, if S_{80}, S_{90}, S_{100} represent potential resolutions of the Israeli-Palestinian conflict in which the Palestinians establish an indepedent state on 80%, 90%, 100% of the West Bank and Gaza Strip, it need not be the case that $S_{100} - S_{90} = S_{90} - S_{80}$.

every $P \in [\underline{P}, \overline{P}]$.[13]

Similarly, we assume that $v(S-Q)$ is the welfare-flow increase to side A once the conflict is resolved at the terms S, where $v : R \to R$ is increasing, twice continuously differentiable and concave with $v(0) = 0$.[14]

Future welfare flows are discounted at the rate $r > 0$, and we define the *payoff* of each party (from a given path of welfare flows over time) to be *the increase of its present-value of welfare relative to the present-value of indefinite continuation of the on-going conflict.*

Thus, if no agreement is ever reached and the conflict continues at its present level, the payoffs for the sides are zero. If an agreement is signed in time t at the terms S, the payoff for A is

$$V(S,t) = \int_0^t 0re^{-r\tau}d\tau + \int_t^\infty v(S-Q)\,re^{-r\tau}d\tau = e^{-rt}v(S-Q) \tag{2.1}$$

while B's payoff is

$$U_P(S,t) = \int_0^t 0re^{-r\tau}d\tau + \int_t^\infty (P-S)re^{-r\tau}d\tau = e^{-rt}(P-S) \tag{2.2}$$

For both sides the increase in welfare occurs only from time t and on, and from present-day perspective a welfare increase at the future instant τ is discounted by $re^{-r\tau}$.

Without loss of generality we can now assume the normalization $Q = 0$.[15]

The sides alternate in making offers for an agreement, with the minimal elapse t_0 between the offers. We denote by $\delta = e^{-rt_0}$ the discount associated with this minimal elapse.

The first side to offer is B, which we therefore call the Proposer. It has to decide when to make its offer S. If the offer is made at time t and A accepts the offer, the payoffs of the sides are given by (2.1) and (2.2)

[13] This specific structure relates the welfare-flow increase functions u_{P} of the different types $P \in [\underline{P}, \overline{P}]$. It *does not* involve in any way, however, an inter-personal utility comparison between A and B.

[14] This generalizes Cramton (1992), who carries out the analysis with the linear function $v(S-Q) = S-Q$

[15] If this were not the case in the first place, we could apply everywhere the change of variables $\overline{S} = S-Q$, $\overline{P} = P-Q$, $\overline{Q} = Q-Q$, and continue to work with $\overline{S}, \overline{P}, \overline{Q}$ throughout instead of S, P, Q, respectively.

above. If A turns down the offer, A can make a counter-offer $\widetilde{S}$ to B after the minimal delay t_0. A can delay the counter-offer as long as it likes (possibly indefinitely) beyond t_0. If A made its counter-offer at time $\widetilde{t}$ and it was accepted by B , the payoffs are as above with S replaced by $\widetilde{S}$ and t by $\widetilde{t}$. The alternating offers scheme repeats itself.

We look for a separating equilibrium, in which different types $P > Q = 0$ exercise different delays, so that the equilibrium delay time $\beta(P)$ of each type $P > 0$ reveals its identity to A, while if there are also types $P \leq 0$ (for whom there exist no mutually-beneficial resolution S to end the conflict), they exercise an infinite delay and never come up with an offer.[16] Thus, if and when the delay is over, the remaining game essentially becomes one of complete information. In the limit when $\delta \to 1$, i.e. when the minimal time between consecutive offers, t_0, goes to zero,[17] the Proposer's subgame-perfect equilibrium offer will tend to the Nash bargaining solution

$$S(P) \equiv \arg\max_{S} (P - S)\, v(S) \tag{2.3}$$

and A will accept it immediately (Rubinstein 1982). In particular, at equilibrium the Proposer believes that if it ever makes a less generous offer $S < S(P)$ following the delay $\beta(P)$, then A will reject the offer.[18]

Proposition 1 *The eventual settlement at equilibrium $S(P)$ is increasing with P.*

Proof. In the Appendix. ■

[16] Under some conditions there may also exist pooling or partially-pooling equilibria for types $P > 0$ (see the discussion in Admati and Perry 1987). However, such equilibria do not survive forward-induction refinements (see e.g. Banks 1991, chapter 2).

[17] The assumption $\delta \to 1$ does *not* mean that the proponents tend to be infinitely patient. The restriction to the limit case $\delta \to 1$ is only in order to simplify the exposition and to avoid cumbersome computations. The qualitative results remain intact also when $\delta < 1$.

[18] A could then simply delay its counter-offer indefinitely in order to deter the Proposer from such a deviation, but Cramton (1992) specifies also a "milder" off-equilibrium reaction which still achieves this deterring effect.

To compute the equilibrium delay $\beta(P)$, denote by $P(\Delta)$ the inverse function, which specifies the reservation stand of the Proposer who delays its offer by Δ at equilibrium. The Proposer's problem is therefore to find the delay Δ which balances best between the losses from the delay Δ and the terms of settlement $S(P(\Delta))$ which will be accepted after that delay. Precisely, the Proposer with reservation stand P chooses Δ so as to maximize

$$U_P(\Delta) = e^{-r\Delta}(P - S(P(\Delta))) \tag{2.4}$$

Proposition 2 The equilibrium delay time. *On the separating equilibrium path, when $\delta \to 1$ the delay time of a Proposer with reservation stand $P > Q = 0$ tends to*

$$\beta(P) = \frac{1}{r}\int_P^{\overline{P}} \frac{S'\left(\tilde{P}\right)}{\tilde{P} - S\left(\tilde{P}\right)} d\tilde{P} \tag{2.5}$$

A Proposer of type $P \leq 0$ will delay indefinitely, i.e. $\beta(P) = \infty$.

Proof. In the appendix. ■

Remarkably, the equilibrium delay depends on the highest potential reservation stand $\overline{P}$ of B, *but on no other feature of the probability distribution G.*

For example, with a linear welfare-flow increase function $v(S) = S$ for A and with $\overline{P} = 1$, we have $S(P) = \frac{P}{2}$ and therefore

$$\beta(P) = \begin{cases} -\frac{1}{r}\ln P & \text{For } P \in (0,1] \\ \infty & \text{For } P < 0 \end{cases} \tag{2.6}$$

This delay is logarithmic in the Proposer's type. The type who has the most to gain from an agreement, $P = 1$, will propose immediately, while other types will wait longer the less they have to gain. Ultimately, the types $P \leq 0$, with whom there is no wedge for a mutually beneficial agreement, will never come up

with an offer.

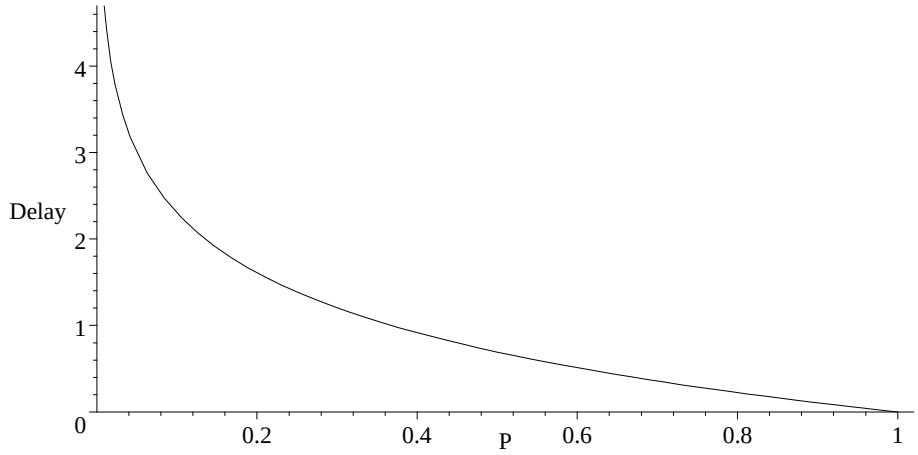

Fig 1: The delay time $\beta(P)$ multiplied by r

In case $\underline{P} > Q = 0$ there is common knowledge between the parties that there exist mutually beneficial agreements among them.[19] In such a case no type of B will delay indefinitely, and the graph above should then be truncated to the domain $[\underline{P}, 1]$.

2.2 Escalation

We now add to this bargaining model the possibility of escalation by side A, which we therefore call the Aggressor. Just before the onset of bargaining, A can choose whether or not to escalate. The Aggressor bases this choice on its belief G about the plausibility of the various possible reservation stands of B, before this stand is revealed.

[19] This is sometimes called "the gap case" (see e.g. Fudenberg and Tirole 1991, chapter 10). For this case there exist other bargaining protocols in which the delay vanishes as the two sides become infinitely patient (i.e. $r \to 1$). To re-iterate, in this work we have nothing to contribute as to why the sides to a protracted conflict fail to engage in the most efficient bargaining mechanism one could possibly design. Rather, we take the delay embodied in their behavior as a *point of departure*, in order to investigate (in the the following subsections) how this delay is sensitive to the onset of escalation.

Escalation implies that side A incurs a subjective welfare-flow decrease C as long as the dispute is not resolved with an agreement. Similarly, escalation implies that side B suffers a subjective welfare-flow decrease D as long as the conflict is not over. Since C and D are subjective measures, they need not be directly related to physical damages. For example, it may very well be the case that $C < D$ even though the physical damages[20] suffered by A due to the escalation are much higher than those suffered by B.

Suppose then that A chose to escalate. If B makes a proposal of terms S at time t and A accepts it, S is implemented immediately and A's payoff is (recall the normalization $Q = 0$)

$$V^E(S,t) = \int_0^t (-C)re^{-r\tau}d\tau + \int_t^\infty v(S)re^{-r\tau}d\tau = e^{-rt}(v(S)+C) - C \tag{2.8}$$

while the payoff of B when it has reservation stand P is

$$U_P^E(S,t) = \int_0^t (-D)re^{-r\tau}d\tau + \int_t^\infty (P-S)re^{-r\tau}d\tau = e^{-rt}(P-S+D) - D \tag{2.9}$$

Otherwise, if A turns down the offer, it continues to inflict the subjective damage D upon B while bearing the subjective cost C. After a minimal delay t_0, A gets a chance to choose when and what counter-offer $\widetilde{S}$ to make to B, and so on, as before.

2.2.1 The Implications of Escalation

The first important effect escalation has on the process of bargaining pertains to the timing of resolution. We claim that escalation, in some situations we identify below, may very well shorten the time to agreement.

Once the conflict has escalated, side A would prefer any resolution S with which $v(S) > -C$, i.e. $S > v^{-1}(-C)$, rather than continuing to endure the subjective flow of damages C. Similarly, side B with an original reservation stand P would now prefer any settlement with $P - S > -D$ rather than continuing to sustain the subjective flow of damages D. Thus, an agreement S is mutually beneficial if and only if

[20] Economic hardship, casualties etc.

$P + D > S > v^{-1}(-C)$. Denote by $S_E(P)$ the Nash Bargaining solution between the Aggressor and a Proposer of type P under complete information after the conflict has escalated, i.e. with the threat points $-C$ and $-D$ for the Aggressor and the Proposer, respectively:

$$S_E(P) = \arg\max_S (P - S + D)(v(S) + C) \tag{2.10}$$

Here again, $S_E(P)$ is the subgame-perfect equilibrium offer (at the limit when $\delta \to 1$) that P will make when its stand eventually gets revealed (Rubinstein 1982).

We again analyze the separating equilibrium path in which there is a one-to-one correspondence between the Proposer's types (who exercise a finite delay) and the extent of the delay. Denote by $\beta_E(P)$ the delay time of type P at equilibrium with escalation and by $P_E(\Delta)$ the inverse function, i.e. the type that waits Δ at equilibrium with escalation. At equilibrium, once the delay is over the Proposer will make the offer $S_E(P)$, and the Aggressor will accept it immediately.

Proposition 3 Equilibrium delay given escalation. *On the separating equilibrium path, as $\delta \to 1$ the delay of a Proposer with a reservation stand $P > v^{-1}(-C) - D$ is given by*

$$\beta_E(P) = \frac{1}{r}\int_P^{\overline{P}} \frac{S'_E\left(\tilde{P}\right)}{\tilde{P} + D - S_E\left(\tilde{P}\right)} d\tilde{P} \tag{2.11}$$

A Proposer of type $P \leq v^{-1}(-C) - D$ will delay indefinitely, i.e. $\beta_E(P) = \infty$.

Proof. *In the appendix* ■

Returning to the example $v(S) = S$ and $\overline{P} = 1$, we get that $S_E(P) = \frac{1}{2}(P + D - C)$ and

$$\beta_E(P) = \begin{cases} -\frac{1}{r}\ln\left(\frac{P+C+D}{1+C+D}\right) & \text{For } P \in (-C - D, 1] \\ \infty & \text{For } P \leq -(C + D) \end{cases} \tag{2.12}$$

The resulting delay time is, once again, logarithmic in the difference between the sides' reservation stands, but this time normalized to the new, larger range of B's types who will come up with an acceptable offer in a finite time. While the type $P = 1$ is still the one to make an immediate offer, it is now the case that any type with original stand $P > -(C + D)$ will eventually make an acceptable offer. As explained above, this is because the agreement will put an end to the suffering of the damage flows C, D, and thus any terms between $-C$ and $P + D$ (and in particular the terms $S_E(P) = \frac{1}{2}(P + D - C)$) are preferred by both sides to the on-going level of escalation.

The important insight of this example is that once escalation has occurred, the delay time (2.12) of all the Proposer types $P \in (-C - D, 1)$ will be shorter than the delay time without escalation (2.6). *With the increased level of hostilities, it takes less time for the Proposer to signal credibly its true stand to the other party.*

To grasp the intuition for this result, consider any two types $P' > P$. At equilibrium, the type P, who gains less than P' from each given settlement S, signals its position by waiting longer than P', but eventually offers to settle at better terms for itself than the terms that P' offers earlier. Given the eventual Rubinstein (1982) subgame-perfect offers once the positions are revealed, how much longer should P wait in order to differentiate itself credibly from P'? In other words, what should be the extra waiting time $\beta(P) - \beta(P')$ so that P' would not be tempted to imitate P's waiting time and offer? The crucial observation is that *this necessary extra waiting-time decreases with the subjective damage levels* C, D. The harsher is the on-going violence in the eyes of the belligerents, the less tempting it becomes for P' to postpone its offer further just in order to improve the eventual terms of agreement. Consequently, the smaller also becomes the gap in waiting time $\beta(P) - \beta(P')$ which P has to exercise, in order to differentiate its own behavior from one which P' might like to imitate.

Figure 2 depicts the delay time with and without escalation for the case $C + D = \frac{1}{2}$ and $r = 1$

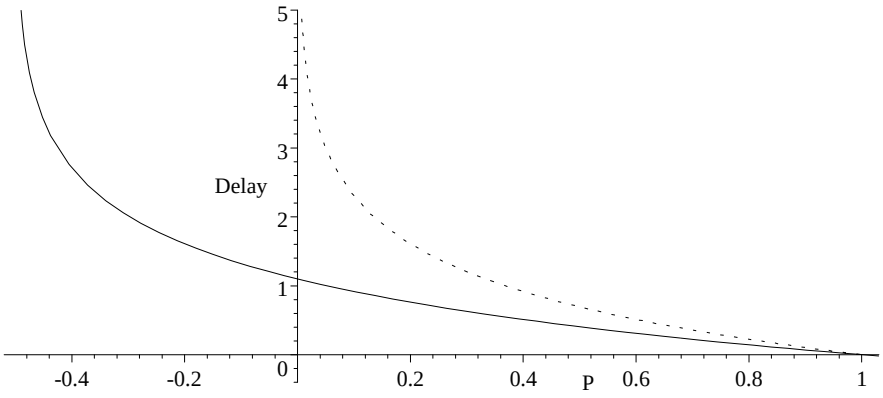

Fig 2: Delay time with escalation (solid line) **and without** (dashed line)

In the following proposition we identify sufficient conditions for this insight to obtain.

Proposition 4 *For every $P > v^{-1}(-C) - D$ the delay time under escalation $\beta_E(P)$ is shorter than the delay time $\beta(P)$ without escalation if*

$$\frac{P + D - S_E(P)}{S'_E(P)} > \frac{P - S(P)}{S'(P)}$$

which obtains when either:

(*i*) *The Arrow-Pratt coefficient of absolute risk aversion of the Aggressor, $-\frac{v''}{v'}$, is decreasing in P and for each reservation stand $P \in (v^{-1}(-C) - D, \bar{P})$ of side B, $S_E(P) \leq S(P)$*

or

(*ii*) *The Arrow-Pratt coefficient of absolute risk aversion of the Aggressor, $-\frac{v''}{v'}$, is constant and for each reservation stand $P \in (v^{-1}(-C) - D, \bar{P})$ of side B, $S_E(P) < S(P) + D$*

or

(*iii*) *The Arrow-Pratt coefficient of absolute risk aversion of the Aggressor,* $-\frac{v''}{v'}$*, is increasing in* P *and for each reservation stand* $P \in (v^{-1}(-C) - D, \bar{P})$ *of side* B, $S(P) < S_E(P) < S(P) + D$.

Proof. In the appendix. ■

In the example above with $v(S) = S$ we are in case (*ii*): The Arrow-Pratt coefficient of absolute risk aversion of the Aggressor is constant, and $S_E(P) = \frac{P+D-C}{2}$ is indeed strictly smaller than $S(P)+D = \frac{P}{2}+D$.

2.2.2 The Aggressor's Gamble

From the point of view of the aggressor, escalation has several implications. The first and most obvious effect is its cost, which has to be endured until an agreement is reached. In case the Proposer's reservation stand is so low that it prefers to endure the increased level of violence indefinitely rather than settle at any acceptable terms, the Aggressor too will have to endure the cost indefinitely. On the other hand, escalation may change the time to agreement with all or with some of the Proposer types with whom the Aggressor would have reached an agreement without escalation. Escalation can induce some Proposer types to come up with an acceptable offer, even though without escalation they preferred the status quo level of the conflict to any feasible settlement. Escalation can also change the agreement terms with any given type of the Proposer, and the new terms could be either worse or better for the Aggressor. We should recall, however, that even if escalation shortens the time to agreement and improves the terms of agreement for the Aggressor, it might still not compensate for its cost.

Taking all these considerations into account and computing the expected payoff according to the Aggressor's beliefs about its opponent stand, the Aggressor has to choose whether to escalate or not. If the

Aggressor chooses to escalate, its expected payoff is

$$U_A^E = \int_{\underline{P}}^{\max\left(\underline{P},\, v^{-1}(-C)-D\right)} (-C)g(P)dP + \int_{\max(\underline{P},\, v^{-1}(-C)-D)}^{\bar{P}} \left((-C)\left(1-e^{-r\beta_{\mathsf{E}}(P)}\right) + v(S_E(P))e^{-r\beta_{\mathsf{E}}(P)}\right) g(P)dP \tag{2.13}$$

where g is the density of G. The first term describes the Aggressor's expected payoff in the event that the opponent has a low reservation stand, in the range $P \leq v^{-1}(-C) - D$, in which case the Aggressor endures the escalation cost indefinitely. The second term describes its expected payoff when confronting an opponent with reservation stand $P > v^{-1}(-C) - D$, with whom an agreement is eventually reached after the equilibrium delay $\beta_E(P)$.

If, in contrast, the Aggressor chooses not to escalate, its expected payoff is

$$U_A^{NE} = \int_0^{\bar{P}} v(S(P))e^{-r\beta(P)}g(P)dP \tag{2.14}$$

In order to maximize its expected payoff, the Aggressor will choose to escalate if and only if

$$U_A^E > U_A^{NE} \tag{2.15}$$

It is easy to see then that the answer to the question "when will the Aggressor prefer to escalate?" depends on its prior belief G on the distribution of the Proposer's reservation stands P, and on the subjective cost C it will have to endure following the escalation. We can only know for sure that if $\beta_E(P) \geq \beta(P)$ for every P (i.e. escalation causes a longer delay) and also $S_E(P) \leq S(P)$ for every P (escalation results in worse terms of agreement for the Aggressor) then the Aggressor will choose not to escalate. In all other cases it will have to weigh the positive and negative effects of escalation.

However, ex post, after finding out side B's stand, side A might very well regret its decision to escalate. This happens in the following cases:

1. The dispute continues indefinitely at its escalated level, i.e. side B's reservation stand is so low ($P < v^{-1}(-C) - D$) that it prefers to endure the increased level of violence indefinitely rather than settle. This happens with probability $G(v^{-1}(-C) - D)$, which is decreasing as C and D increase (for a given

belief G). When the escalation subjective damages C, D are very large, almost all of the Proposer types will prefer – sooner or later – to settle with the Aggressor.

2. The Aggressor eventually yields to an offer $S \in (v^{-1}(-C), 0)$ it would not have accepted without enduring the cost of escalation. Without escalation, an agreement S is acceptable by the Aggressor only if $S > 0$, while with escalation an agreement is acceptable by the Aggressor already when $S > v^{-1}(-C)$. Thus, when P is such that $S_E(P) \in (v^{-1}(-C), 0)$, the Aggressor yields to an offer it would not have accepted at the outset.

3. The Aggressor yields to an offer $S > 0$ it would initially prefer to the status quo, but the damage it has to endure until agreement is reached overshadows any improvement in the final outcome (if there is at all such an improvement, which may stem from a shorter delay $\beta_E(P) < \beta(P)$, better terms $S_E(P) > S(P)$, or both):

$$C\left(1 - e^{-r\beta_{\mathsf{E}}(P)}\right) > v(S_E(P))e^{-r\beta_{\mathsf{E}}(P)} - v(S(P))e^{-r\beta(P)} \tag{2.16}$$

With further specifications we can analyze how the overall probability of regret (i.e. the overall probability of cases 1-3) is related to the damage levels C, D:

Example 1 *Suppose that $v(S) = S$, $[\underline{P}, \overline{P}] = [-1, 1]$, and G is the uniform distribution on $[\underline{P}, \overline{P}]$. Then the probability that the Aggressor will eventually regret the escalation – carried out by its own initiative – is increasing in the sum of the subjective escalation damages $C + D$, and decreasing in their difference $D - C$.*

Proof. In the Appendix ■

In this example escalation is thus a two-edged sword. The larger is the sum of subjective damages it entails, the higher is also the probability that its initiator will eventually regret it, even if a priori it conceived it to be a worthwhile gamble.[21]

[21] Our model is related but different than that of Cramon and Tracy (1992), which dealt with strikes in a model where

3 A Symmetric Two-Sided Model

In this section we extend the above results to a symmetric, two-sided model. Both sides will now be uncertain about the other's reservation stand, and both would be able to induce escalation at the outset. Our purpose here is to show how the qualitative results above do not depend on the asymmetric roles of Aggressor and Proposer, by presenting an equilibrium with similar properties in the symmetric setting.

The model will be identical to the one above, with the following changes:

1. We will assume that A's reservation stand, which we now denote by a, is not known to side B, and B believes that a is distributed according to the probability distribution F. We stick with the assumption that A is uncertain regarding B's reservation stand, which we now denote by b. A believes that b is distributed according to the probability distribution G.

2. Before the bargaining starts, each side $i = A, B$, knowing its own reservation stand but not that of the other, can decide whether or not to initiate an escalation, with the subjective flow of damages C_i to itself and D_i to its rival. If both sides escalate, i endures a combined subjective damage flow

working contracts have a finite duration. In their formulation, the union (side A) is uncertain regarding the maximal wage the firm (side B) can afford to pay. The union can credibly commit to one out of two possible levels of sanctions – holdout or strike, in case its initial wage demand is declined by the firm. Thus, their model is more complicated than ours in the sense that A has to choose simultaneously two variables – the wage demand and the type of threat. Also, unlike in our case, the delay the firm can exercise until it makes its counter-offer is bounded by the duration of the contract. (Another technical difference is that unlike in our model, the damage to the firm during the sanctions varies with the actual reservation wage it can pay to the workers.) The qualitative results of Cramton and Tracy are nevertheless similar to ours: The equilibrium delay is decreasing in a measure akin to $C + D$ in our model. Moreover, when the initial wage of the workers is low – and therefore they foresee a large expected wage increase– they choose to strike, but when their initial wage is above a certain threshold they prefer a holdout. This is analogous to the equilibrium structure we explore in the next section, in which the types who foresee large potential gains from agreement are those who prefer to escalate.

E_i, satisfying

$$\max(C_i, D_j) < E_i \leq C_i + D_j$$

where j is the other party.[22]

3. Observing each other's decision whether or not to escalate, each of the opponents should then decide by when to approach the negotiation table if the other party will not yet have done so earlier. Once one side is seated at the negotiation table, the other one has to decide when to approach it as well and make the inaugurating offer. If this offer is accepted, it is implemented immediately. Otherwise the alternating offers scheme repeats itself as in the previous section.

We are able to demonstrate the existence of such an equilibrium for a fully symmetric case, in which $v(S) = S$, $C_A = C_B \equiv C$, $D_A = D_B \equiv D$, $E_A = E_B = C + D$, $G\tilde{}U[-1, 1]$ and $F\tilde{}U[0, 2]$. For a range of combinations of C and D, we construct an equilibrium in which the types who foresee large expected improvements from settlement choose to escalate, in order to hasten the process and extract sooner the gains from agreement. The remaining types, who foresee small or zero expected improvements from settlement, prefer to avoid escalation, in order to circumvent or decrease the suffering in the conflict, which they believe might last for long or even forever.

At first this equilibrium structure might seem surprising, because the types who foresee only small benefits to agreement – those who fight longer in the equilibrium analyzed in the previous section – are the ones who refrain from escalation here. However, notice that within the interval of the types who escalate, as well as within the interval of types who do not escalate, it is still the case that the types who expect to gain little from agreement do fight longer at equilibrium. The decision whether to escalate is

[22] In principle, we could have allowed for several consecutive stages in which escalation can be initiated or reciprocated. This, however, would create a preliminary stage of mutual signaling about resolve by the escalation decisions themselves, while the focus of the current paper is on *delay* as a signaling device. This, of course, is not to say that escalation is not useful or relevant as a signaling measure. Rather, our aim here is to disentangle the effect of signaling by delay from other signaling procedures.

of a different nature, because it is *dichotomous*. Consequently, the incentives of a given type to escalate depend crucially on what other types are expected to escalate at equilibrium, and hence with which pool of types this given type will prefer to associate itself. The equilibrium structure we propose is the "natural" one, because when the types in some interval choose to escalate, the gains from *escalation* are larger for the more reconciliatory types.[23,24].

[23]To see this, suppose that the bunch of types of A who decided not to escalate is $[\underline{a},\overline{a}]$, and consider an interval of types $[\underline{b},\overline{b}]$ of B. If the types $[\underline{b},\overline{b}]$ are those who do not escalate, then for $a \in [\underline{a},\overline{a}]$ and $b \in [\underline{b},\overline{b}]$ the discounting due to the delay ((3.9) below) is $\frac{b-a}{\overline{b}-\underline{a}}$, the terms of the agreement ((3.10) below) will tend to be $\frac{b+a}{2}$ as $\delta \to 1$, and so the overall payoff to type b will tend to

$$U^{B}_{NE}(b;a) = \left(\frac{b-a}{\overline{b}-\underline{a}}\right)\left(b - \frac{b+a}{2}\right)$$

If, in contrast, the types $[\underline{b},\overline{b}]$ are those who do escalate, then for $a \in [\underline{a},\overline{a}]$ and $b \in [\underline{b},\overline{b}]$ the discounting due to the delay is $\frac{(b+C)-(a-D)}{(\overline{b}+C)-(\underline{a}-D)}$, so the flow of escalation cost C endured by b prior to the settlement is discounted to $C\left(1-\frac{(b+C)-(a-D)}{(\overline{b}+C)-(\underline{a}-D)}\right)$, while the eventual terms of agreement will tend to be $\frac{(b+C)+(a-D)}{2}$ as $\delta \to 1$. Thus, the overall payoff to type b will tend to

$$U^{B}_{E}(b;a) = -C\left(1-\frac{(b+C)-(a-D)}{(\overline{b}+C)-(\underline{a}-D)}\right) + \left(\frac{(b+C)-(a-D)}{(\overline{b}+C)-(\underline{a}-D)}\right)\left(\frac{(b+C)+(a-D)}{2}\right)$$

It is now straightforward to verify that the gains to escalation

$$U^{B}_{E}(b;a) - U^{B}_{NE}(b;a)$$

are *increasing in* b for every given $a \in [\underline{a},\overline{a}]$.

Now, also for the complementary set of types $[\underline{\alpha},\overline{\alpha}]$ of A who do escalate, the same computations and conclusions apply for $[\underline{a},\overline{a}] = [\underline{\alpha}-C,\overline{\alpha}-C]$, and by substituting $[\underline{b}+D,\overline{b}+D]$ for $[\underline{b},\overline{b}]$.

Thus we conclude that the gains to escalation are increasing in b for every possible type a of A. Therefore, also the expected gains to escalation (across A's types) are increasing in b. This means that the incentives of B to escalate are *larger* the *more reconciliatory* is its type (i.e., the larger b is). Yet, it is still the case that if the types $[\underline{b},\overline{b}]$ all chose to escalate (as well as in case they all chose to refrain from escalation), the more reconcilatory types among them (with larger b) will make an acceptable offer *sooner* and therefore end up fighting less at equilibrium.

[24]This observation does not exclude, by itself, the possibility of an equilibrium with the reverse structure, in which a bunch of types with large expected gains from settlement do not escalate (in analogy with the "signaling reversal" of Orzach and Tauman 1996). Yet, we did not succeed in constructing such a "reversed" equilibrium, both for the above parameters, as

Specifically, we construct an equilibrium characterized by a threshold $x > \frac{1}{2}$, such that the types of side A in the interval $[0, 1-x)$ escalate, those in $(1-x, 2]$ do not, and the type $a = 1-x$ escalates with probability $\frac{1}{2}$. Similarly the types of B in $(x, 1]$ escalate, those in $[-1, x)$ do not, and the type $b = x$ escalates with probability $\frac{1}{2}$.

Our intended equilibrium structure implies that, given the observed decisions about escalation, both sides will know that the *effective* reservation stands of A, B come from some intervals of the form $[\underline{a}, \bar{a}]$, $[\underline{b}, \bar{b}]$, respectively. These intervals are determined by the decisions regarding escalation and the resulting flows of damages. More precisely, these will be

	B escalated	B did not escalate
A escalated	$[\underline{a}, \bar{a}] = [-(C+D), 1-x-(C+D)]$ $[\underline{b}, \bar{b}] = [x+(C+D), 1+(C+D)]$	$[\underline{a}, \bar{a}] = [-C, 1-x-C]$ $[\underline{b}, \bar{b}] = [-1+D, x+D]$
A did not escalate	$[\underline{a}, \bar{a}] = [1-x-D, 2-D]$ $[\underline{b}, \bar{b}] = [x+C, 1+C]$	$(\underline{a}, \bar{a}] = [1-x, 2]$ $[\underline{b}, \bar{b}] = [-1, x]$

(3.1)

As in the previous section, we first turn to describe the timing behavior given the decisions to escalate or not, and then solve backwards to find the equilibrium choices regarding escalation.

3.1 War-of-Attrition Bargaining

Given any prior combination of decisions by the rivals to escalate or not, we will adopt the equilibrium behavior suggested by Wang (2000) for the bargaining stage. Given a pair of intervals $[\underline{a}, \bar{a}]$, $[\underline{b}, \bar{b}]$ of reservation stands, Wang (2000) analyses a host of equilibria for the bargaining stage, one of which is symmetric.

The equilibrium starts with a "war of attrition," in which each side waits for the other to approach the

well as for a case in which there is a gap between the parties' reservation stands, e.g. $F\tilde{}U[-3, -1]$, $G\tilde{}U[1, 3]$.

negotiation table first. This is because, at equilibrium, the second to approach is the one who will make the acceptable offer, and at the eventual Rubinstein (1982) subgame-perfect equilibrium the offeror gets a larger share of the gains from agreement.

At equilibrium, the types $\underline{a}$ and $\bar{b}$ approach the table without delay. In the sequel, the types of each side with larger conceivable gains from agreement approach the table before the types who foresee smaller gains. The timing of $a \in [\underline{a}, \bar{a}]$ and $b \in [\underline{b}, \bar{b}]$ is

$$t_A(a) = -\frac{1}{r} \ln \frac{((\bar{b} + \underline{a}) - 2a)}{(\bar{b} - \underline{a})} \tag{3.2}$$

$$t_B(b) = -\frac{1}{r} \ln \frac{(2b - (\bar{b} + \underline{a}))}{(\bar{b} - \underline{a})} \tag{3.3}$$

with the reverse functions

$$A(t) = \frac{1}{2}(1 - e^{-rt})(\bar{b} - \underline{a}) + \underline{a} \tag{3.4}$$

$$B(t) = \frac{1}{2}(1 + e^{-rt})(\bar{b} - \underline{a}) + \underline{a} \tag{3.5}$$

This staggering implies that for every given period of time by which nobody is yet at the table, each side understands that the type of the other comes from a smaller interval of types with whom the wedge of potential agreements is smaller. This induces the next-in-line types to approach the table, because by then they know that further delay, and the chance it brings with it to be the second to approach, would not compensate for the losses due to further delay.

Once one side has approached, its position becomes common knowledge by the delay it exercised. The other side then delays further its approach, as a function of its position, by

$$t_A(a)|_{B=b} = -\frac{1}{r} \ln \frac{b - a}{2b - (\bar{b} + \underline{a})} \tag{3.6}$$

$$t_B(b)|_{A=a} = -\frac{1}{r} \ln \frac{b - a}{(\bar{b} + \underline{a}) - 2a} \tag{3.7}$$

Thus, the overall time until both sides are at the table is given by

$$t(a, b) = \begin{cases} -\frac{1}{r} \ln \frac{(b-a)}{(\bar{b}-\underline{a})} & b > a \\ \infty & b \leq a \end{cases} \tag{3.8}$$

(Remarkably, this total delay time results also in all the other, non-symmetric equilibria with a similar structure that Wang (2000) introduces.) Hence, the discounting due to the delay is

$$\exp\left(-rt(a,b)\right) = \frac{b-a}{\overline{b}-\underline{a}} \tag{3.9}$$

Thus, when both sides are at the table, the delays they exercised have already revealed at equilibrium their effective positions. At this point Wang (2000) assumes, by definition, that the Rubinstein (1982) offer is then implemented. This tends to be

$$V(a,b) = \frac{a+b}{2} \tag{3.10}$$

as $\delta \to 1$. However, the more elaborate equilibrium structure in Cramton (1992) can be applied here, to show how the second to approach may be induced to make this offer at equilibrium by its beliefs about its rival's reaction in case of deviation.[25]

The Wang (2000) equilibrium implicitly assumes, that if there is a finite time by which a party should have approached[26] but it had failed to do so, no agreement is ever implemented. We will adopt this convention here. More precisely, we will assume that, at equilibrium, the preliminary decisions about escalation make the parties believe that the ranges of stands are given by table (3.1) above. Thus, any excessive delays, that could have corresponded to stands that protrude these intervals, would result in the lack of any settlement.

3.2 Escalation

Given the above behavior at the bargaining stage, we now look for conditions under which we will have a unique equilibrium characterized by the threshold x, as detailed at the beginning of this section. In such

[25] In fact, the setting and the equilibrium of Wang (2000) is a simplification of that of Cramton (1992), who assumes that also the first to approach makes an offer. That asumption makes the equilibrium structure more intricate.

[26] This is the case when $\overline{a} < \underline{b}$

an equilibrium, each type should carry out the escalatory behavior meant for it – escalate (E) or not escalate (NE). First of all, this equilibrium should have the property that

$$\text{The type } a = 1 - x \text{ of } A \text{ is just indifferent between escalating or not.} \qquad \text{(Indifference)}$$

By the symmetry of the construction, this condition will be mathematically equivalent to the condition that the type $b = x$ is indifferent between the two options. This equality constraint will pin down the candidate $x(C, D)$ for the threshold as a function of C, D.

The other incentive compatibility constraints say that

$$\text{Types } a \in [0, 1 - x) \text{ prefer to escalate rather than not} \qquad \text{(IC1)}$$

and

$$\text{Types } a \in (1 - x, 2] \text{ prefer not to escalate rather than to escalate} \qquad \text{(IC2)}$$

Analogous conditions should obtain for side B, but here again the mathematical conditions will be identical to those for (IC1) and (IC2), due to the symmetry of the equilibrium.

Notice, first, that if a type $a \in [0, 1 - x)$ ever decides not to escalate (contrary to the equilibrium behavior), B will believe it comes, in fact, from the interval $[1 - x, 2]$. Hence, a should better imitate the delay time of one of the types in $[1 - x, 2]$, as otherwise no agreement will ever be reached, by assumption.

Lemma 1 *If $a \in [0, 1 - x)$ were not to escalate, it would fare best imitating the delay time of $1 - x$ given that side B does adhere to its own equilibrium strategy.*

Proof. In the appendix ■

Similarly, if $a \in (1 - x, 2]$ contemplates to disregard the equilibrium recommendation and escalate, it should better plan what type in $[0, 1 - x]$ to imitate regarding the delay time in the bargaining stage

Lemma 2 *If $a \in (1-x, 2]$ were to escalate, it would fare best imitating the delay time of $1-x$ given that side B adheres to its own equilibrium strategy.*

Proof. In the appendix ■

These two lemmata pin down the best off-equilibrium behavior given a deviation at the escalation stage. The incentive-compatibility constraints are thus well defined. They consist of two inequalities which should obtain at equilibrium, and relate C, D, and $x(C, D)$.

Unfortunately, the three conditions cannot be solved explicitly so as to isolate $x(C, D)$ or to isolate the restrictions on C as a function of D, because they involve high-degree polynomials. Nevertheless, the equilibrium threshold $x(C, D)$ can be solved numerically given particular values of (C, D). We have done so for a wide range of values of (C, D), and verified that the incentive compatibility constraints (IC1) and (IC2) obtain. For example, it turns out that for $C = 0.1$ and $D = 0.5$, the unique equilibrium of the above structure is with the threshold $x = 0.832\,53$.

Since all three constraints vary smoothly with C, D, existence and uniqueness will typically be preserved at some open neighborhood of parameters for which the equilibrium exists and is unique.

Theorem 1 *There is an open range of parameter values (C, D) for which there is a unique threshold equilibrium of the above structure.*

Proof. In the Appendix ■

Here again the parties will sometimes regret ex post their decision to escalate. Most notably, when both sides escalate, the eventual terms of agreement are the same as when they do not escalate (due to the symmetry) but until agreement is reached they suffer the cost of escalation and its damages. For example, the types $a = 1 - x$ and $b = x$, if they escalate, will settle at terms that tend to $\frac{1}{2}$ as $\delta \to 1$, and their

payoffs will tend to be

$$U_{a=1-x} = U_{b=x} = -(C+D)(1-e^{-rt(1-x-C-D,x+C+D)}) + \left(\frac{1}{2} - (1-x)\right) e^{-rt(1-x-C-D,x+C+D)}$$
$$= -(C+D)\left(1 - \frac{2x-1+2(C+D)}{1+2(C+D)}\right) + \left(\frac{1}{2} - (1-x)\right)\left(\frac{2x-1+2(C+D)}{1+2(C+D)}\right)$$
$$= -(1-x)\left(1 - \frac{2(1-x)}{1+2(C+D)}\right) + x - \frac{1}{2}$$

which is *decreasing* in the overall level of the sum $C+D$.

4 Conclusion

We use a continuous-time bargaining model to demonstrate the implications of a decision to escalate (by one or both sides) during a protracted international conflict. In this bargaining model the parties signal their reservation stand by the delay they exercise before approaching the negotiation table and making a serious offer. As a state waits longer, it signals that it values a peace agreement (at some given terms) less than another possible type of this state which would have made an offer earlier . The delay thus becomes a credible signal of the state's stand, and hence also of the reconciliation terms on which it would insist.

The flow of damages that follows escalation makes both sides more eager to settle than before. The stakes at the negotiation table are higher when one or both sides escalate, because an agreement will also save the belligerents from the increased level of harm they inflict on one another. We identify the conditions in which these higher stakes will shorten the delay to agreement.

Under these conditions escalation loosens the incentives of the sides to sustain further the burden of the conflict, when they try convey to each other their true positions in a credible way. Types who see less possible gains from a settlement, can use shorter delays to credibly signal their resolve, since those types who foresee large gains in compromising will not be willing to bear the increased costs of escalation to the same extent they were willing before the escalation, and thus those more reconciliatory types will make an offer much sooner.

We show that as the overall level of damage caused by escalation increases, the less tempting it becomes to postpone one's offer further just in order to improve the eventual terms of agreement. Consequently, the smaller also becomes the additional waiting time, which has to be exercised in order to differentiate one's own behavior from the behavior which others might like to imitate. Since this reasoning obtains for every pair of distinct types, escalation will enable each type to come forward with the acceptable offer sooner than without escalation.

However, the decision to escalate bears with it a considerable amount of risk. In an asymmetric model in which only one side (the Aggressor) can make the decision to escalate we show that ex post the Aggressor might very well regret its decision to escalate.

On the one hand, escalation may shorten the time to an agreement (with those proposers willing to settle), and may improve the terms of such an agreement (if the Aggressor is able to inflict a large flow of damages upon its opponent at relatively little cost to itself). Escalation can also induce some of the opponent types to come up with an acceptable offer, even though without escalation they preferred the status quo level of the conflict to any feasible agreement.

On the other hand, all of that might not compensate for the increased damage done while the conflict is still going on. If the opponent sees rather little gains to agreement, the Aggressor might eventually find itself yielding to agreement terms it would not have accepted at the outset, in order to end the harm it suffers in the more intensive conflict. In the worst case, it might turn out that the opponent prefers to fight indefinitely rather than settle even at the escalated level of the conflict.

The Aggressor bases its decision whether or not to escalate on its belief about the plausibility of the various possible reservation stands of its opponent before this stand is revealed, and on the subjective cost it will have to endure following the escalation. We characterize the conditions under which a potential Aggressor will prefer the gamble embodied by escalation to the status quo level of the conflict. We also show (with further specification of assumptions) that the probability that the Aggressor will indeed regret

ex post its decision to escalate increases with the overall level of damages entailed by escalation.

Finally we show that in a symmetric model in which both sides have the ability to escalate, our major insights remain intact. In this model, after observing each other's decision whether or not to escalate, each of the opponents should decide when to approach the negotiation table if the other party will not yet have done so earlier. Once one side is seated at the negotiation table, the other one delays its offer further in order to signal credibly its resolve, and then makes an offer which is accepted and implemented immediately.

Here again, if one or both sides have decided to escalate, the overall time it will take them to reach an agreement will be shorter than in case they do not escalate. We also show that the parties will sometimes regret ex post their decision to escalate and might even yield to agreement terms they would not have accepted at the outset.

We believe that the insights and the causal links highlighted in this model do provide a relevant perspective on the interplay between violence and diplomacy.

5 Appendix

Proof of Proposition 1. The eventual equilibrium offer $S(P)$ satisfies the first order condition

$$(P - S(P))\, v'(S(P)) - v(S(P)) = 0$$

Differentiating with respect to P yields

$$(1 - S'(P))\, v'(S(P)) + (P - S(P))\, v''(S(P))S'(P) - v'(S(P))S'(P) = 0$$

or

$$S'(P) = \frac{1}{2 - \frac{v''(S(P))}{v'(S(P))}\,(P - S(P))}$$

which is positive, since $\frac{v''(S(P))}{v'(S(P))}$ is negative by the concavity of ν.[27] ¥

Proof of Proposition 2: The Proposer B has to choose the delay Δ which maximizes its payoff

$$U_P(S(P(\Delta)),\Delta) = e^{-r\Delta}(P - S(P(\Delta)))$$

We take the derivative with respect to Δ and find the first order condition

$$\frac{d}{d\Delta}U_P(S(P(\Delta)),\Delta) = -re^{-r\Delta}\left(P - S(P(\Delta))\right) - e^{-r\Delta}S'(P(\Delta))\frac{dP(\Delta)}{d\Delta} = 0$$

At equilibrium the maximum is achieved at $\Delta = \beta(P)$, i.e. $P(\Delta) = P(\beta(P)) = P$. We get

$$\frac{dP(\Delta)}{d\Delta}|_{\Delta=\beta(P)} = -r\frac{P - S(P)}{S'(P)}$$

or equivalently

$$\frac{d\beta(P)}{dP} = -\frac{1}{r}\frac{S'(P)}{(P - S(P))}$$

with the initial condition that $\beta(\bar{P}) = 0$, i.e. that the type with the most to gain from any given agreement will offer immediately. This condition obtains at equilibrium if A believes that an offer made at $t = 0$, either on or off the equilibrium path, is made by $\bar{P}$. With such a belief of A, the equilibrium offer of type $\bar{P}$ will indeed be made at $t = 0$ (because then any offer after a positive delay will be inferior to making the same offer without delay).[28]

Integrating then gives

$$\beta(P) = \int_P^{\bar{P}} \left(\frac{1}{r}\frac{S'(\tilde{P})}{\left(\tilde{P} - S(\tilde{P})\right)}\right) d\tilde{P}$$

[27] In fact, we see here that $S'(P)$ would be positive even if v were convex (side A would then be risk-loving) but not too convex, as long as $\frac{v''(S(P))}{v'(S(P))}(P - S(P)) < 2$.

[28] Here again, any other equilibrium structure with $\beta(\overline{P}) > 0$ does not survive forward-induction equilibrium refinements (Banks 1991).

Finally we need to check that $\beta(P)$, which was derived from the local analysis of the first order conditions, is indeed the globally optimal delay for type P. This follows by a standard argument[29] from the fact that $S'(P) > 0$ (proposition 1) and that the preferences of side B over settlement-delay pairs (S,t) satisfy the single-crossing property (i.e. any indifference curve of type P crosses at most once any indifference curve of a different type P'). The latter property obtains, since the slope of B's indifference curves at (S,t) is given by

$$-\frac{\frac{\partial U_{\mathrm{P}}(S,t)}{\partial S}}{\frac{\partial U_{\mathrm{P}}(S,t)}{\partial t}} = -\frac{-e^{-rt}}{-re^{-rt}(P-S)} = -\frac{1}{r(P-S)}$$

which is monotonic in P. ¥

Proof of Proposition 3: The proof is completely analogous to that of proposition 2, and so we only sketch the computations. Under escalation, an agreement S is mutually beneficial if and only if $P + D > S > v^{-1}(-C)$. Thus, if $P \leq v^{-1}(-C) - D$ there are no possible gains from an agreement and the sides will continue to endure the damages forever.

A Proposer B of type $P > v^{-1}(-C) - D$ has to choose the delay Δ, which maximizes its payoff

$$U_P^E(S_E(P(\Delta)),\Delta) = \int_0^{\Delta}(-D)\,re^{-r\tau}d\tau + \int_{\Delta}^{\infty}(P - S_E(P_E(\Delta)))re^{-r\tau}d\tau = e^{-r\Delta}\left(P + D - S_E\left(P_E\left(\Delta\right)\right)\right) - D$$

The first order condition with respect to Δ is

$$\frac{d}{d\Delta}U_P^E(S_E(P(\Delta)),\Delta) = -re^{-r\Delta}(P + D - S_E(P(\Delta))) - e^{-r\Delta}S'_E(P(\Delta))\frac{dP_E(\Delta)}{d\Delta} = 0$$

At equilibrium the maximum is achieved at $\Delta = \beta_E(P)$ i.e. $P_E(\beta_E(P) = P$. We get

$$\frac{dP_E(\Delta)}{d\Delta}|_{\Delta=\beta_{\mathrm{E}}(P)} = -r\frac{P + D - S_E\left(P\right)}{S'_E\left(P\right)}$$

or

$$\frac{d\beta_E(P)}{dP} = -\frac{1}{r}\frac{S'_E\left(P\right)}{P + D - S_E\left(P\right)}$$

[29] See e.g. Fudenberg and Tirole 1992, p. 261-262, thm 7.3 and its corollary.

again with the initial condition $\beta_E(\overline{P}) = 0$. Integration then gives

$$\beta_E(P) = \int_P^{\bar{P}} \left(\frac{1}{r} \frac{S'_E\left(\tilde{P}\right)}{\tilde{P} + D - S_E\left(\tilde{P}\right)} \right) d\tilde{P}$$

¥

Proof of Proposition 4: A sufficient condition for having

$$\int_P^{\bar{P}} \left(\frac{1}{r} \frac{S'_E\left(\tilde{P}\right)}{\tilde{P} + D - S_E\left(\tilde{P}\right)} \right) d\tilde{P} = \beta_E(P) < \beta(P) = \int_P^{\bar{P}} \left(\frac{1}{r} \frac{S'(\tilde{P})}{\left(\tilde{P} - S(\tilde{P})\right)} \right) d\tilde{P}$$

is the inequality between the integrands

$$\frac{1}{r} \frac{S'_E(P)}{P + D - S_E(P)} < \frac{1}{r} \frac{S'(P)}{P - S(P)}$$

for $P > v^{-1}(-C) - D$, i.e.

$$\frac{P + D - S_E(P)}{S'_E(P)} > \frac{P - S(P)}{S'(P)} \tag{A1}$$

Now, from the proof of proposition 1 we know that

$$S'(P) = \frac{1}{2 - \frac{v''(S(P))}{v'(S(P))}(P - S(P))} \tag{A2}$$

and a similar computation yields

$$S'_E(P) = \frac{1}{2 - \frac{v''(S_E(P))}{v'(S_E(P))}(P + D - S_E(P))} \tag{A3}$$

Substituting (A2) and (A3) into (A1) yields

$$2(P + D - S_E(P)) - \frac{v''(S_E(P))}{v'(S_E(P))}(P + D - S_E(P))^2 > 2(P - S(P)) - \frac{v''(S(P))}{v'(S(P))}(P - S(P))^2 \tag{A4}$$

Now it is easy to see that if the Arrow-Pratt coefficient of absolute risk aversion, $-\frac{v''}{v'}$, is decreasing and $S_E(P) \leq S(P)$ then (A4) obtains. Also, if $-\frac{v''}{v'}$ is constant and $S_E(P) < S(P) + D$ then (A4) obtains as well. Finally, (A4) also holds if $-\frac{v''}{v'}$ is increasing and $S(P) < S_E(P) < S(P) + D$. ¥

Proof of properties of Example 1:

1. With probability $\frac{1-C-D}{2}$, P turns to be in the range $[-1, -C-D]$, in which case the dispute continues indefinitely at its escalated level.

2. With probability C, P turns to be in the range $(-C-D, C-D)$. In such a case an agreement is reached after a finite delay but the Aggressor yields to an offer it would not have accepted at the outset, because then
$$S_E(P) = \frac{(P+D)-C}{2} < 0$$

3. With probability $\frac{1}{2}\left(1 - \sqrt{(D-C)\left(1+\frac{1}{C+D}\right)} + D - C\right)$, P turns to be in the range $\left[C-D, 1-\sqrt{(D-C)\left(1+\frac{1}{C+D}\right)}\right)$. In this case the eventual terms of the agreement would be acceptable by the Aggressor in the first place, but the damage it has to endure until agreement is reached overshadows the difference in the final outcome:
$$\int_0^{-\frac{1}{r}\ln\left(\frac{P+C+D}{1+C+D}\right)} Cre^{-r\tau}d\tau > \int_{-\frac{1}{r}\ln\left(\frac{P+C+D}{1+C+D}\right)}^{\infty} \frac{(P+D)-C}{2} re^{-r\tau}d\tau - \int_{-\frac{1}{r}\ln P}^{\infty} \frac{P}{2} re^{-r\tau}d\tau$$
i.e.
$$C\left(1 - \frac{P+C+D}{1+C+D}\right) > \frac{(P+D)-C}{2}\left(\frac{P+C+D}{1+C+D}\right) - \frac{P^2}{2}$$

Overall, the Aggressor regrets the escalation ex post when P turns to be in the range $\left[-1, 1-\sqrt{(D-C)\left(1+\frac{1}{C+D}\right)}\right)$, i.e. with probability
$$1 - \frac{1}{2}\sqrt{(D-C)\left(1+\frac{1}{C+D}\right)}$$

This probability is *increasing* in the sum $C+D$, though, as could be expected, it decreases with the difference $D-C$ between the subjective damages B and A have to sustain due to the escalation. ¥

Proof of Lemma 1: Suppose that a type $a \in [0, 1-x)$ decides not to escalate (contrary to its equilibrium behavior). We know that if a chooses to wait less than what type $1-x$ would have waited, B will not negotiate with it at all (it follows from B's off the equilibrium path beliefs). Thus a should better imitate

the delay time of one of the types $k \in [1-x, 2]$ (actually $k \in [1-x, 1+C+D]$ since a would never choose to imitate a type k for which $k \in [1+C+D, 2]$ because then it would never reach an agreement with side B). If a sees that side B has escalated it believes that $b \in [x, 1]$ and that its own expected payoff is (the *effective* reservation stand of a is $k-D$ and of b is $b+C$):

$$U^{NE}_{a\in[0,1-x)}(k)|_{b\in[x,1]} = \left(\tfrac{2}{1-x}\right)\left(\int_x^{\max(k-C-D,x)}(-D)f(b)db\right.$$
$$\left.+\int_{\max(k-C-D,x)}^{1}\left(\left(\tfrac{(b+C)+(k-D)}{2}-a\right)\tfrac{(b+C)-(k-D)}{C+D+x}-D\left(1-\tfrac{(b+C)-(k-D)}{C+D+x}\right)\right)f(b)db\right)$$

This is a continuous non increasing function of k and its maximum in the interval $[1-x, 1+C+D]$ is achieved at $k = 1-x$. Thus in this case, if a saw B escalating, the best it can do is imitate $k = 1-x$, and then its expected payoff would be

$$U^{NE}_{a\in[0,1-x)}(1-x)|_{b\in[x,1]}$$
$$= \frac{1}{6}\frac{-2x^2+7x-2+3(C+D)(1-2a)+3a(1-3x)}{x+C+D}-\frac{1}{2}(D-C)$$

If on the other hand, a sees that side B did not escalate, a believes that $b \in [-1, x]$ and that its own expected payoff is (the *effective* reservation stand of a is k and of b is b):

$$U^{NE}_{a\in[0,1-x)}(k)|_{b\in[-1,x]} = \left(\frac{2}{1+x}\right)\left(\int_{-1}^{\min(k,x)}0f(b)db+\int_{\min(k,x)}^{x}\left(\frac{b+k}{2}-a\right)\frac{b-k}{2x-1}f(b)db\right)$$

This is also a continuous non increasing function of k and its maximum in the interval $[1-x, 1+C+D]$ is achieved at $k = 1-x$. Thus in this case also the best a can do is imitate $k = 1-x$, and then its expected payoff would be

$$U^{NE}_{a\in[0,1-x)}(1-x)|_{b\in[-1,x]} = \left(\frac{2}{1+x}\right)\left(-\frac{1}{2}\left(x-\frac{1}{2}\right)a-\frac{1}{6}x^2+\frac{5}{12}x-\frac{1}{6}\right)$$

Then we see that no matter what B has done, a chooses to imitate $k = 1-x$ and its expected payoff is

therefore:

$$
\begin{aligned}
&U^{NE}_{a\in[0,1-x)}(1-x)\\
&= U^{NE}_{a\in[0,1-x)}(1-x)|_{b\in[x,1]} \times \Pr(b\in[x,1]) + U^{NE}_{a\in[0,1-x)}(1-x)|_{b\in[-1,x]} \times \Pr(b\in[-1,x])\\
&= \frac{1}{12}(1-x)^2\frac{3a-2(1-x)}{C+D+x} + \frac{1}{4}(1-x)(1-(D-C)) - \frac{1}{4}a + \frac{1}{12}(2x-1)(2-x)
\end{aligned}
$$

¥

Proof of Lemma 2: Suppose that $a \in (1-x, 2]$ chooses to escalate (contrary to its equilibrium behavior). Then side B believes that $a \in [0, 1-x]$. If a escalates but waits longer than what $a = 1-x$ would have waited, B would never settle at equilibrium. Thus the best a can do is imitate also the delay of some $k \in [0, 1-x]$. Thus if a sees that B has escalated its expected payoff is (the *effective* reservation stand of a is $k-C-D$ and of b is $b+C+D$):

$$
\begin{aligned}
&U^{E}_{a\in(1-x,2]}(k)|_{b\in[x,1]} = \left(\tfrac{2}{1-x}\right) \times \left(\int_x^{\max(k-2(C+D),x)} (-(C+D)) f(b)db\right.\\
&\left. + \int_{\max(k-2(C+D),x)}^{1} \left(\left(\tfrac{b+k}{2} - a\right)\tfrac{(b+C+D)-(k-C-D)}{1+2(C+D)} - (C+D)\left(1 - \tfrac{(b+C+D)-(k-C-D)}{1+2(C+D)}\right)\right) f(b)db\right)
\end{aligned}
$$

This is a continuous non decreasing function of k and its maximum in the interval $[0, 1-x]$ is achieved when $k = 1-x$. Thus in this case, if a saw B escalating, the best it can do is imitate $k = 1-x$, and then its expected payoff would be

$$
U^{E}_{a\in(1-x,2]}(1-x)|_{b\in[x,1]} = -\frac{1}{6}\frac{2x^2+9xa-7x+6(C+D)(2a-x)+2-3a}{1+2C+2D}
$$

If, on the other hand, it sees that B did not escalate its expected payoff is (the *effective* reservation stand of a is $k-C$ and of b is $b+D$):

$$
\begin{aligned}
&U^{E}_{a\in(1-x,2]}(k)|_{b\in[-1,x]} = \left(\tfrac{2}{1+x}\right) \left(\int_{-1}^{\min(k-C-D,x)} (-C) f(b)db\right.\\
&\left. + \int_{\min(k-C-D,x)}^{x} \left(\left(\tfrac{b+D+k-C}{2} - a\right)\tfrac{b+D-(k-C)}{x+C+D} - C\left(1 - \tfrac{b+D-(k-C)}{x+C+D}\right)\right) f(b)db\right)
\end{aligned}
$$

This function is also a continuous non decreasing function of k and its maximum in the interval $[0, 1-x]$

is achieved when $k = 1 - x$. Thus in this case also the best a can do is imitate $k = 1 - x$, and then its expected payoff would be

$$U^E_{a\in(1-x,2]}(1-x)|_{b\in[-1,x]} = \tfrac{1}{6(x+C+D)(1+x)}((2-x)(2x-1)^2 + 12ax(1-x-C-D)$$

$$+3x((D-C)+(C+D)(1-x)) - 3a(1-2C-2D)$$

$$+(C+D)((C+D)(C+D-3)+3(D-C)-3) - 3a(C+D)^2) + \tfrac{1}{2(1+x)}x(D-C)$$

Therefore a will imitate $k = 1 - x$ and its expected payoff will be

$$U^E_{a\in(1-x,2]}(1-x) =$$

$$U^E_{a\in(1-x,2]}(1-x)|_{b\in[x,1]} \times \Pr(b \in [x,1]) + U^E_{a\in(1-x,2]}(1-x)|_{b\in[-1,x]} \times \Pr(b \in [-1,x]) =$$

$$\tfrac{1}{12}\tfrac{1}{(1+2S)(x+S)}(2 - 11x - 13x^3 + 21x^2 + 2x^4 + 2S^4 - 3S^2(3+2S) + S^3 - S + 3aS(1-2S)(1+S)$$

$$-3xaS(x+4+4S) - 6xS(1-x)(1-2x-2S) - 3a - 3ax(1-x)(3x-5)) + \tfrac{1}{4}(1+x)R$$

where $S = C + D$ and $R = D - C$ ¥

Proof of Theorem 1: We can state the three conditions on the connections between C,D and $x(C,D)$ in their implicit form (where $S = C + D$, $R = D - C$).

1. Indifference: The type $a = 1 - x$ of A is just indifferent between escalating or not.

$$U^E_{a=(1-x)} = U^{NE}_{a=(1-x)} \Leftrightarrow$$

$$0 = -6R(x+S)(1+2S) + (1-x)\left(2 - 7x^3 + 16x^2 - 10x\right) - 2S^4$$

$$+(-6x+11)S^3 + \left(7x - 4x^2 + 8\right)S^2 + \left(-2 + 16x - 13x^2 + 3x^3\right)S$$

2. IC1: Types $a \in [0, 1-x)$ prefer to escalate rather than not

$$
\begin{aligned}
U^{E}_{a\in[0,1-x)} > U^{NE}_{a\in[0,1-x)} = U^{NE}_{a\in[0,1-x)}(1-x) \Leftrightarrow \\
0 < -\frac{1}{12(x+S)}a^3 + \frac{1}{4}\left(1+\frac{1-x}{1+2S}\right)a^2 - \frac{1}{4}(1-x+S)\left(1+\frac{(1-x)^2}{(S+x)(1+2S)}\right)a \\
+\frac{1}{12}\frac{(1-x)(x^2-5x+2) - S^2(3-S) - S(1-6R) - 3x(S-2R)}{x+S} \\
-\frac{1}{6}x(1-x) + \frac{1}{12}\frac{1-x^3+6x(1-x)S}{1+2S}
\end{aligned}
$$

3. IC2: Types $a \in (1-x, 2]$ prefer not to escalate rather than escalate

$$
U^{NE}_{a\in(1-x,2]} > U^{E}_{a\in(1-x,2]} = U^{E}_{a\in(1-x,2]}(1-x) \Leftrightarrow
$$

$$
\begin{cases}
0 < \frac{1}{12(2x-1)}(x-a)^3 + \frac{1}{4}\frac{1-x}{x+S}a^2 + W & \text{for} \quad 1-x < a \le x \\
0 < \frac{1}{4}\frac{1-x}{x+S}a^2 + W & \text{for} \quad x < a \le x+S \\
0 < -\frac{1}{12(x+S)}a^3 + \frac{1}{4}\frac{1+S}{x+S}a^2 - \frac{1}{4}(x+S)a + \frac{1}{12}(x+S)^2 + W & \text{for} \quad x+S \le a \le 1+S \\
0 < -\frac{1}{4}\left(\frac{S^2+3xS-2S+x^2-1}{x+S} + R\right)a + \frac{1}{12}(x+S)^2 & \\
+\frac{1}{12}(1+S)\frac{2S^2+3xS-2S-1}{x+S} + \frac{1}{4}R(1+S) + W & \text{for} \quad 1+S \le a \le 2
\end{cases}
$$

Where

$$
W = W(a,x,S,R) = -\frac{1}{4}\frac{3x^2-9x-6xS-2S^2+3S+5}{1+2S}a - \frac{1}{12}\frac{S\left(S^2-3S-6\right)}{(x+S)}
$$

$$
-\frac{1}{12}\frac{1-12x^3-10x^3S-3xS+18xS^2-11x+21x^2+2x^4+18x^2S-12x^2S^2}{(1+2S)(x+S)} - \frac{1}{2}R
$$

Since all three constraints involve expressions which are smooth as a function of C, D, existence and uniqueness will typically be preserved at some open neighborhood of the parameters C, D for which the equilibrium exists and is unique. For example, for $C = 0.1$ and $D = 0.5$ (that is, $S = 0.6$ and $R = 0.4$) the indifference condition 1. above gives $x = 0.832\,53$ as the unique solution in the relevant range $[\frac{1}{2}, 1]$. At these values, the derivative of the indifference condition 1. with respect to x is different than zero. The implicit function theorem then guarantees that there is an open neighborhood $\mathcal{O}$ of $(C, D) = (0.1, 0.5)$ with a unique, smooth solution $x(C, D) \in [\frac{1}{2}, 1]$ of the indifference condition for $(C, D) \in \mathcal{O}$. Moreover,

since the strict inequalities IC1 and IC2 obtain for $(C, D) = (0.1, 0.5)$ and $x = 0.832\,53$, they will continue to obtain for (C, D) and $x(C, D)$ in some open subset of $\mathcal{O}$. ¥

6 References

Admati A. R. and M. Perry (1987), "Strategic Delay in Bargaining", *Review of Economic Studies* 54, pp. 345-364.

Agha, H. and R. Malley (2001), "Camp David: The Tragedy of Errors", *The New York Review of Books*, August 9, 2001 http://www.nybooks.com/articles/14380

Banks, J.S. (1990), "Equilibrium Behavior in Crisis Bargaining Games", *American Journal of Political Science* 34(3), pp. 599-614.

Banks, J.S. (1991), Signaling Games in Political Science*, Harwood Academic Publishers.*

Bueno de Mesquita, B. and D. Lalman (1989), War and Reason*, Yale University Press.*

Bueno de Mesquita, B., J.D. Morrow and E.R. Zorick (1997), "Capabilities, Perception and Escalation", *The American Political Science Review* 91(1), pp. 15-27

Cohen, R. (1990), Culture and Conflict in Egyptian-Israeli Relations – a Dialogue of the Deaf, *Indiana University Press.*

Cramton, P.C. (1992), "Strategic Delay in Bargaining with Two Sided Uncertainty", *Review of Economic Studies* 59, pp. 205-225.

Cramton, P.C. and J.S. Tracy (1992), "Strikes and Holdouts in Wage Bargaining: Theory and Data", *The American Economic Review* 82(1), pp. 100-121.

Downs, G.W. and D.M. Rocke (1987), "Tacit Bargaining and Arms Control," *World Politics* 39, pp. 297-226.

Fearon, J.D. (1994), "Domestic Political Audiences and the Escalation of International Disputes", *The American Political Science Review* 88, pp. 577-592.

Fearon, J.D. (1995), "Rationalist Explanations for War." *International Organization* 49, pp. 379-414.

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

Filson, D. and S. Werner (2002), "A Bargaining Model of War and Peace: Anticipating the Onset, Duration and Outcome of War" *American Journal of Political Science* 46, pp. 819-38.

Fudenberg, D. and J. Tirole (1991), Game Theory, *MIT Press*.

Hertzog, Chaim (1975), *The War of Atonement*, Little Brown & Company (re-issued in 1998 by Stackpole Books).

Mallie, E. and D. Mckittric (2001), *Endgame in Ireland*, Hodder & Stoughton General.

Morgan, .T. C. (1994), *Untying the Knot of War: A Bargaining Theory of International Crises*, University of Michigan Press.

Morrow, J.D. (1986), "A Spatial model of International Conflict", *The American political Science Review* 80(4), pp. 1131-1150

Morrow, J.D. (1989), "Capabilities, Uncertainty and Resolve: A Limited Information Model of Crisis Bargaining", *American Journal of Political Science* 33(4), pp. 941-972

Morrow, J.D. (1992), "Signaling Difficulties with Linkage in Crisis Bargaining.", *International Studies Quarterly* 36, pp. 153-172.

Orzach, R. and Y. Tauman (1996), "Signaling Reversal", *International Economic Review* 37, pp. 453-464.

Ponsati, C. (2002), "The ABC of War: Advantage, Bargaining and Confrontation.", mimeo, Universitat Autonoma de Barcelona.

Powell, R. (1987), "Crisis Bargaining, Escalation and MAD", *The American Political Science Review* 81, pp. 717-735

Powell, R. (2001), "Bargaining while Fighting", mimeo, University of California at Berkeley.

Riad, M. (1981), *The Struggle for Peace in the Middle East*, Quartet Books, London.

Rubinstein A. (1982), "Perfect Equilibrium in a Bargaining Model", *Econometrica* **50**, pp. 97-110.

Shavit, A. (2001), "End of a Journey", an interview with Shlomo Ben Ami in *Haaretz Magazine*, September 14, 2001,

http://www.haaretzdaily.com/hasen/pages/ShArt.jhtml?itemNo=74353&contrassID=2

&subContrassID=14&sbSubContrassID=0&listSrc=Y

Smith, A. and A. Stam (2001). "Bargaining through Conflict." Unpublished manuscript, Department of Political Science, Yale University.

Wagner, R.H. (2000), "Bargaining and War", *American Journal of Political Science* **44**(3), pp. 469-484

Wang, R. (2000), "Separating equilibria in a continuous-time bargaining model with two-sided uncertainty", *International Journal of Game Theory* **29**, pp. 229-240