Long Wars*

Sandeep Baliga
Kellogg School of Management, Northwestern University

Tomas Sjöström
Department of Economics, Rutgers University

October 2023

Abstract

We study whether the Coase conjecture holds in a model of bargaining during conflict due to Powell [18] and Fearon [7]. Two players, A and B, contest a divisible resource. At any time during the conflict, they can make a binding agreement to share the resource. The conflict continues until they make an agreement or one side collapses. Player B privately knows whether he is a strong or a weak type, with a greater probability of collapse if he is weak. The "lemons condition" says that player A would rather fight to the end than make a generous offer at the beginning of the conflict that both types of player B would accept. If this condition holds then the expected length of the conflict is bounded away from zero, even if negotiations are frictionless. Thus, the Coase conjecture does not hold. We study how the minimum length of conflict depends on the parameters, and the impact of third party intervention.

1 Introduction

Asymmetric information is an important component of many conflicts. For example, the ability of a rebel group to fight a long civil war depends on

*We thank Marco Battaglini, Adam Meirowitz and seminar participants at Cornell, Utah and Vanderbilt for useful comments. Any errors are our responsibility.

its financial resources and its support among the population. It is plausible that the rebel group has private information about these variables (Walter [24]).[1] Such information asymmetries can lead to war (Fearon [6]). However, Wagner [23] points out that "real war", as defined by Clausewitz [3], allows negotiations to continue after the war has started.[2] How long will it take to negotiate an agreement which terminates the war? In a buyer-seller model with one-sided incomplete information and frictionless bargaining, the *Coase conjecture* says that an agreement will be reached very quickly (Gul, Sonnenschein and Wilson [12]). If a similar result holds for bargaining during conflict, why do many civil wars last a long time despite many opportunities to negotiate a settlement? How can we explain prolonged (inefficient) delay before an agreement is reached?

To answer these important questions, we investigate a dynamic conflict bargaining game due to Powell [18] and Fearon [7].[3] Powell [18] assumed two players engage in sequential battles. The uninformed player A makes repeated proposals between battles, and the war ends when a proposal is accepted or one player has collapsed. If a proposal is accepted, then it represents a binding agreement to share the contested resource. If a player collapses then the opponent takes all of the resource. A collapse can only occur during battle. Battles have fixed length and cannot be terminated once started. Powell [18] showed that if player B privately knows his probability of collapse (the case of interdependent values), and the prior probability that player B is weak is large enough, then the probability of having at least one battle does not go to zero as bargaining *between* battles becomes frictionless.[4] Fearon [7] pointed out that this happens because the battle is the only screening device. Since battles are indivisible, player A retains some commitment ability even if the time between proposals goes to zero. However, a lack of commitment ability is a key assumption behind the Coase conjec-

[1] Even if the rebel group has hard (verifiable) information about its strength, it may prefer to keep it secret. For example, if it reveals the sources of its financial support, the government could block these sources (Walter [24]).

[2] Clausewitz [3], p. 605: "..war is simply a continuation of political intercourse, with the addition of other means. We deliberately use the phrase 'with the addition of other means' because we...want to make clear that war in itself does not suspend political intercourse or change it into something completely different."

[3] Other models of negotiations during wartime have been provided by Wittman [25], Filson and Werner [10], Slantchev [20], Smith and Stam [21] and Heifetz and Segev [13].

[4] Powell [18] also considered the case of private values (player B privately knows his own cost of fighting), but we do not consider this case.

ture. Fearon [7] suggested that for the commitment ability to vanish, the length of each battle should shrink to zero. The model then becomes similar to well-known buyer-seller bargaining models, with the exception that each player may collapse during the negotiations. Each bargaining period corresponds to one battle. Player B is a strong or a weak type, with a greater probability of collapse in any period if he is weak. Player B's true type is his private information. Player A may also collapse in any period, but his probability of collapse is common knowledge. The question is: will an agreement be reached very quickly if each bargaining period (the interval between proposals) is very short?

Suppose player A has all the bargaining power, and can make a take-it-or-leave-it offer before each battle. If player A could commit to a sequence of offers then the standard logic of optimal screening would apply. If player B is very likely to be strong, then player A would make an initial pooling offer which is high enough that both types accept. This would give player B a large informational rent in the unlikely event that he is weak. If player B is more likely to be weak, then pooling is not optimal. Instead, player A would initially make a low offer; if player B rejects then he must fight a long time before getting a better offer. This screening would optimally reduce the informational rent for the weak type, who is averse to fighting and therefore accepts the low offer. However, if player A cannot commit to a sequence of offers then this optimal screening mechanism is not credible. Indeed, if player B rejects the initial low offer then player A knows that player B is strong, so player A has an incentive to make a high offer which terminates the conflict as soon as possible. But since player B will anticipate this, player A must also increase the weak type's information rent by a more generous initial offer. This commitment problem is more serious if negotiations are frictionless in the sense that each period is very short. The Coase conjecture states that, as the period length shrinks to zero, it becomes impossible to extract any information rent from the weak type, and an offer will be accepted by both types almost immediately. If the Coase conjecture holds, then the conflict will be short, as long as there are no impediments to bargaining.

Gul, Sonnenschein and Wilson [12] and Gul and Sonnenschein [11] showed that the Coase conjecture holds in stationary equilibrium of a buyer-seller model where the monopolistic seller makes a sequence of price offers. The seller cannot credibly sustain a high price, because early rejections signal that the consumer's willingness to pay is low. In the conflict model, rejections are also informative, but so is the very fact that player B does not collapse. This suggest that the commitment problem may be even worse than in the buyer-seller model (where there is never any "buyer collapse"). The possibility that player B might collapse in the future strengthens player A's incentive to fight rather than negotiate, but this must be balanced against the risk that player A himself may collapse.

In Gul, Sonnenschein and Wilson [12] and Gul and Sonnenschein [11], values are private because the buyer's valuation of the good does not directly impact the seller's payoff. In the Powell-Fearon model, values are interdependent because the probability that player B will collapse directly impacts player A's expected value of delaying an agreement. Deneckere and Liang [4] show that the Coase conjecture does not hold in a buyer-seller model with interdependent values. However, in their model nobody collapses during negotiations, and the private information concerns the value of the good that is traded. In the conflict model, collapse is possible and it is the probability of collapse that is subject to private information. What happens on the battlefield is informative about player B's type, quite apart from what happens at the bargaining table. This distinguishes war bargaining from economic bargaining. As argued by Wagner [23], the possibility of collapse opens the door to faster learning, which would seem to also open the door for the Coase conjecture.

We show this door is still closed in the seminal Powell-Fearon conflict model with no indivisible battles of fixed length. We find that if screening is optimal for player A, which we refer to as the "lemons condition", then the war cannot end quickly when player B is strong. Thus, wars caused by one-sided incomplete information can be long, even if bargaining is frictionless and binding peace agreements are feasible. The intuition is simple. In the Powell-Fearon model, unlike the buyer-seller model, each player has the "outside option" of fighting to the end, hoping or believing that the opponent will collapse. This has two implications. First, player B's strong type would rather fight than accept a low offer. Second, if player B is sufficiently likely to be weak, i.e., if the lemons condition holds, then player A would rather fight than make a high offer. These two implications prevent the Coasean logic from taking hold. A lot of fighting is required to dissipate the weak type's rents from bluffing, before player A becomes willing to make a high offer that the strong type is willing to accept. As in the classic Akerlof lemons problem studied by Deneckere and Liang [4], when there is sufficient risk of giving a lemon (the weak type) too much informational rent, the outcome is far from Pareto efficient. We also argue that third party intervention to help player B can backfire by making the lemons problem worse thus prolonging the conflict. Meaningful interventions should be drastic, and attempt to make weak and strong types sufficiently similar to eliminate the incentive for screening.

Rather than specifying any particular bargaining protocol, we consider outcomes that are incentive-compatible (IC) and individually rational (IR). If the lemons condition holds, then IC and IR imply a strictly positive lower bound on the expected duration of war. This has implications for the sequential equilibrium of any bargaining protocol, since these IC and IR conditions must necessarily be satisfied. For example, suppose player A makes a sequence of take-it-or-leave-it offers to player B. If player B accepts then the game ends with a binding agreement. If player B rejects, then they fight for Δ units of time, where Δ is the length of a "battle". During the battle some player may collapse; if not then the conflict continues with player A making another proposal before the next battle, etc. This game has a sequential equilibrium by Theorem 6.1 of Fudenberg and Levine [9]. Since any sequential equilibrium must satisfy IC and IR, the expected duration of war cannot be less than the lower bound we derive. Thus, if the lemons condition holds, then long wars must occur in any sequential equilibrium, even if the time between successive offers is arbitrarily short. This refutes the Coase conjecture.

In independent work, Fearon and Jin [8] also have revisited the Powell-Fearon model. They identify a refined PBE where wars can last a long time. Our analysis is simpler: instead of constructing an equilibrium, we establish a lower bound of the length of conflict (and thus an upper bound on welfare) that holds for any equilibrium (and any bargaining protocol). Unlike us, Fearon and Jin [8] also study the case where binding agreements are not feasible (i.e., player A can rescind offers).

2 Dynamic Conflict without Negotiations

Two players, A and B, contest a divisible resource (perhaps a disputed territory). In this section, we derive expected payoffs under the assumption that the conflict must be settled by fighting. The winner takes all; agreeing to share the resource peacefully is impossible.

Time is continuous and the discount rate is $r > 0$. Fighting starts at time $t = 0$. Player A collapses with constant hazard rate $\alpha > 0$. Player B privately knows his own type: weak with hazard rate β_W, or strong with hazard rate β_S, where $\beta_W > \beta_S \geq 0$. The prior probability that player B is weak is q_0, where $0 < q_0 < 1$. Everything except player B's true type is common knowledge.[5] During the conflict, each player $i \in \{A, B\}$ gets a flow payoff of $\pi_i > 0$. When a player collapses, the fighting stops and the conflict is over. The other player wins and controls all of the resource forever. After the conflict has ended the resource yields a flow benefit of 1 forever. Assume $\pi_A + \pi_B < 1$, so fighting is inefficient: it reduces the total flow benefit by $1 - \pi_A - \pi_B > 0$.

Define

$$V_{BS}^* \equiv \frac{r\pi_B + \alpha}{r\left(\alpha + \beta_S + r\right)}, \tag{1}$$

$$V_{BW}^* \equiv \frac{r\pi_B + \alpha}{r\left(\alpha + \beta_W + r\right)} \tag{2}$$

and for any $q \in [0, 1]$,

$$V_A^*(q) \equiv q\frac{r\pi_A + \beta_W}{r\left(\alpha + \beta_W + r\right)} + (1 - q)\frac{r\pi_A + \beta_S}{r\left(\alpha + \beta_S + r\right)}. \tag{3}$$

It is easy to verify that, at the start of the conflict, the expected payoffs for player B's strong and weak types, and player A, respectively, are V_{BS}^*, V_{BW}^* and $V_A^*(q_0)$. The longer they fight, the more pessimistic player A becomes about his chances of winning, so his continuation payoff falls. If they have been fighting for $t > 0$ units of time, without collapse, then player A's continuation payoff is $V_A^*(q_t)$, where q_t is the probability that player B is weak conditional of having fought for t units of time. By Bayes rule,

$$q_t = \frac{q_0 e^{-\beta_W t}}{\left(1 - q_0\right) e^{-\beta_S t} + q_0 e^{-\beta_W t}} = \frac{q_0}{\left(1 - q_0\right) e^{(\beta_W - \beta_S)t} + q_0}. \tag{4}$$

Since $\beta_W > \beta_S$, q_t is decreasing in t and $V_A^*(q)$ is decreasing in q. Player B does not learn anything about player A, however, so his continuation payoff is fixed at either V_{BS}^* or V_{BW}^*.

During fighting, flow surplus of $1 - \pi_A - \pi_B > 0$ is lost. This loss stops when either player A collapses (which happens at rate α) or player B collapses (which happens at rate $\beta \in \{\beta_S, \beta_W\}$). A third party seeking to maximize

[5] Adding additional private information about player B would make bargaining even more inefficient. We show bargaining can be inefficient with just one dimension of private information.

total welfare should minimize the expected cost of fighting by speeding up the hazard rates of collapse. To increase a particular player's expected payoff, the third party should reduce that player's rate of collapse and/or increase the opponent's rate of collapse. For given hazard rates, policies that increase flow benefits during conflict increase both total and individual welfare.

With player A as the "government" and player B as an "insurgent group", the Powell-Fearon model captures civil war. During the conflict, the two players each control some part of the country, obtaining flow payoffs π_A and π_B, respectively. Ending the conflict eliminates the cost of fighting, so the total flow payoff increases to $1 > \pi_A + \pi_B$. With player A as country A and player B as country B, the Powell-Fearon model captures asymmetric interstate war. The two players fight over a territory which belongs to country B. Country A is larger and has offensive capability; country B is smaller and can only defend itself. During the conflict, the flow benefits from the territory are dissipated by the cost of conflict; the two players each control some share with flow benefits π_A and π_B, respectively. If one side collapses then the other side takes control of all of the disputed territory.

The nature and technology of war determine the probability of collapse. A weak government with inadequate policing in a poor country, where terrain makes it easy for insurgents to hide, might mean a low α and β. The government has enough resources to survive and is unlikely to face decisive defeat by a small group of rebels. The guerrilla might "swim in the people as the fish swims in the sea" (Mao [15]) if the population is exploited by the government and covertly supports the insurgency. An insurgent group or defending country facing a would-be hegemon might have financial or military support from a third party. Or they might have their own sources of revenue from natural resources or illicit trade. In these cases, conflict in the absence of negotiations might resemble a war of attrition with low hazard rates. On the other hand, the attacker might have an advantage depending on the technology of offense and defence. Siege cannons destroyed fort walls so wars without negotiations would be short. Napoleon's army moved quickly through enemy land, capturing territory with little resistance. Blitzkrieg tactics played a similar role at the start of World War II.[6] In these cases, α and β might be big and wars quick and decisive.

[6] See McNeill [16] for these and other examples.

3 Binding Agreements and Failure of the Coase Conjecture

Now we introduce the possibility of making a binding agreement to share the contested resource peacefully. The conflict can end either with a collapse as in the previous section, where the winner takes all, or with an agreement. Formally, an *agreement* is simply a real number x, representing player B's share of the resource. Let $X \subseteq [0, 1]$ represent the set of feasible agreements. It is convenient to assume X is a finite set. The set of non-negative integers is denoted $N \equiv \{0, 1, 2, ..., n, ...\}$. An agreement can be made at any time $t = n\Delta$, where $n \in N$ and $\Delta > 0$ is the fixed length of each "bargaining period". If the agreement $x \in X$ is made at time t then player B will control the share x in perpetuity, which is worth x/r at time t. Player A's share $1 - x$ is worth $(1 - x)/r$.

Rather than considering a specific bargaining protocol, we consider a *revelation mechanism*. Player B reports his type $\beta \in \{\beta_W, \beta_S\}$ to a mediator before the conflict starts. The mechanism specifies, for each reported $\beta \in \{\beta_W, \beta_S\}$, all $n \in N$ and all $x \in X$, the probability $p(x|n\Delta, \beta)$ that the mediator will choose agreement x at time $t = n\Delta$ if the fighting is still going on at that time. The conflict ends when either an agreement is chosen by the mediator, or a player collapses, whichever happens first. Since no player can be forced to make peace against his will, we impose *Individual Rationality (IR)*: if $p(x|n\Delta, \beta) > 0$ then player A and player B's type β must (weakly) prefer agreement x to their outside options (reservation payoffs) at time $n\Delta$. The outside option is to fight until a collapse. Player B's reservation payoff is always either V^*_{BW} or V^*_{BS}, depending on his type (see Section 2). Player A's reservation payoff depends on his beliefs about player B's type, which in general changes over time. However, to prove our main result (that a short conflict is impossible), it is enough to consider player A's reservation payoff at time 0, which is $V^*_A(q_0)$ (see Section 2). Since any agreement must give player A at least what he could get by fighting until a collapse, at time $t = 0$ his expected payoff must be at least $V^*_A(q_0)$. *Incentive Compatibility (IC)* says that player B prefers to report his type truthfully before the conflict starts. In particular, the weak type should not gain by *feigning strength*, i.e., by falsely reporting that he is strong.

Given a revelation mechanism, let $\lambda(t)$ denote the probability that the conflict lasts for more than t units of time when player B is strong. The probability that the strong type makes an agreement without having to fight at all is $1 - \lambda(0)$. If $\lambda(t) = 0$, then the conflict is sure to end by time t. The conflict can end either with a collapse or an agreement. Without an agreement, the probability that player A and strong player B can fight until time t without either side collapsing is $e^{-(\alpha+\beta_S)t}$. Accordingly, $\lambda(t) = \rho(t)e^{-(\alpha+\beta_S)t}$, where $\rho(t)$ is the probability that, conditional on player B being strong and no player collapsing before time t, there is no agreement by time t. Conditional on player B being strong and no player collapsing before time t, the probability that the conflict lasts at most t units of time (due to an agreement made no later than time t) is $1 - \rho(t)$.

Lemma 1 *If a revelation mechanism satisfies IR, then (a) at time $t = 0$, player A's expected payoff must be at least $V_A^*(q_0)$; (b) for any $t \geq 0$, if $p(x|t, \beta_S) > 0$ then $x \geq rV_{BS}^*$; and (c) for any $t \geq 0$, the weak type's expected payoff from feigning strength is at least*

$$R(t) \equiv V_{BW}^* + e^{-(\alpha+\beta_W)t}(1 - \rho(t))e^{-rt}\left(V_{BS}^* - V_{BW}^*\right). \tag{5}$$

Proof. (a) Before the conflict starts, player A's outside option (fighting until a collapse) is worth $V_A^*(q_0)$.

(b) If player B is strong then at any time $t \geq 0$, his outside option is worth V_{BS}^*. Since an agreement $x \in X$ is worth x/r, IR requires $x \geq rV_{BS}^*$.

(c) Suppose player B is weak. The first term on the right hand side of Equation (5) is V_{BW}^*, the value of fighting until a collapse. The second term is the product of the following three components. First, $e^{-(\alpha+\beta_W)t}$ is the probability that both players fight without collapse for t units of time. Second, $1 - \rho(t)$ is the probability that, conditional on fighting without collapse for t units of time, player B's strong type does reach an agreement by time t. IR for the strong type implies that any such agreement will give player B at least the share rV_{BS}^*. Third, $e^{-rt}\left(V_{BS}^* - V_{BW}^*\right)$ is the weak type's gain from agreeing, at time t, to get the share rV_{BS}^* forever, instead of fighting until a collapse.

$R(t)$ is the weak type's expected payoff in the following hypothetical scenario: he falsely reports β_S to the mediator, then he obediently accepts any recommended agreement until time t, but if the conflict is still ongoing at time t then he fights to the end. Feigning strength will give him *at least* $R(t)$, because any agreement after time t would (by the strong type's IR) be better than continued fighting. ■

We will focus on the case where

$$V_{BS}^{*}+V_{A}^{*}(q_{0})>1/r. \tag{6}$$

We will refer to Condition (6) as the *lemons condition*. It implies that the probability of agreement at $t=0$ must be strictly less than 1. To see this, suppose for each $\beta\in\{\beta_W,\beta_S\}$ there is agreement at time 0 with probability one: $\sum_{x\in X}p(x|0,\beta)=1$. IR implies that player B's strong type must get the payoff V_{BS}^{*}, i.e., his share must be at least rV_{BS}^{*}. IC implies that the weak type must also get at least this much. Thus, player A can at most get the share $1-rV_{BS}^{*}$, which is worth $\left(1-rV_{BS}^{*}\right)/r$. But Condition (6) says that this is less than player A's reservation payoff $V_{A}^{*}(q_{0})$, contradicting IR.

The same argument shows why the expected length of the conflict must be bounded away from zero. For player B's strong type must get at least the share rV_{BS}^{*} in any agreement. If an agreement is reached very quickly, then no significant screening can occur, and IC implies that the weak type's expected payoff is almost V_{BS}^{*}. Then player A gets at most approximately $\left(1-rV_{BS}^{*}\right)/r$. Condition (6) implies that this violates IR for player A. This argument avoids the issue of beliefs, since it only involves player A's reservation payoff at $t=0$ (before he learns anything about player B's type), and each player can guarantee his reservation payoff (by fighting until the end) regardless of the opponent's behavior. We now make the argument more formal.

Using Equations (1) and (3), Condition (6) can be rewritten as

$$q_{0}>r\frac{\left(\alpha+\beta_{W}+r\right)\left(1-\left(\pi_{B}+\pi_{A}\right)\right)}{\left(\beta_{W}-\beta_{S}\right)\left(\alpha+r\left(1-\pi_{A}\right)\right)}. \tag{7}$$

Notice that this inequality is automatically satisfied if r is small. The right hand side of Condition (7) is less than 1 if and only if

$$\left(\beta_{W}-\beta_{S}\right)\left(\alpha+r\left(1-\pi_{A}\right)\right)>r\left(\alpha+\beta_{W}+r\right)\left(1-\left(\pi_{B}+\pi_{A}\right)\right).$$

Define

$$t^{*}\equiv\frac{1}{\left(\alpha+r+\beta_{W}\right)}\ln\frac{\left(r\pi_{B}+\alpha\right)\left(\beta_{W}-\beta_{S}\right)q_{0}}{r\left(1-\left(\pi_{A}+\pi_{B}\right)\right)\left[\left(\alpha+\beta_{W}+r\right)-q_{0}\left(\beta_{W}-\beta_{S}\right)\right]} \tag{8}$$

and

$$\lambda^{*}(t)\equiv e^{-\left(\alpha+\beta_{S}\right)t}-r\frac{\left(1-\left(\pi_{A}+\pi_{B}\right)\right)\left[\left(\alpha+\beta_{W}+r\right)-q_{0}\left(\beta_{W}-\beta_{S}\right)\right]}{\left(r\pi_{B}+\alpha\right)\left(\beta_{W}-\beta_{S}\right)q_{0}}e^{\left(r+\beta_{W}-\beta_{S}\right)t} \tag{9}$$

Condition (7) implies $t^* > 0$. Also, $\lambda^*(t^*) = 0$ and $0 < \lambda^*(t) < 1$ for all t such that $0 < t < t^*$. Recall that $\lambda(t)$ denotes the probability that player B's strong type will fight for more than t units of time. Importantly, neither t^* nor $\lambda^*(t)$ depend on Δ.

Theorem 2 *If a revelation mechanism satisfies IR and IC, then $\lambda(t) \geq \lambda^*(t)$ for any $t \geq 0$.*

Proof. For a given revelation mechanism, let v_{BS}, v_{BW} and v_A denote the expected payoffs for player B's strong and weak types, and player A, respectively, before the conflict starts. Since the resource gives a flow payoff of 1,

$$v_A + (1 - q_0)\, v_{BS} + q_0 v_{BW} \leq 1/r. \tag{10}$$

If IR and IC hold then Lemma 1 implies $v_A \geq V_A^*(q_0)$, $v_{BS} \geq V_{BS}^*$ and $v_{BW} \geq R(t)$ for all $t \geq 0$. Substituting these inequalities into (10) yields

$$V_A^*(q_0) + (1 - q_0)\, V_{BS}^* + q_0 V_{BW}^* + q_0 e^{-(\alpha + r + \beta_W)t}(1 - \rho(t))\,(V_{BS}^* - V_{BW}^*) \leq 1/r. \tag{11}$$

Substituting from Equations (1), (2) and (3) into (11) and rearranging, we get

$$\lambda(t) = \rho(t) e^{-(\alpha + \beta_S)t} \geq \lambda^*(t).$$

■

If the lemons condition holds then $t^* > 0$, and $\lambda^*(t) > 0$ when $t < t^*$, so Theorem 2 implies that the expected length of the conflict is strictly positive. This is inefficient, since $\pi_A + \pi_B < 1$.

The intuition behind Theorem 2 is as follows. The strong type would rather settle the dispute by fighting (an option not available in the buyer-seller model) than accept a share smaller than rV_{BS}^*. If the lemons condition holds, then at the beginning of the conflict, player A would rather fight than take the share $1 - rV_{BS}^*$. For such a deal to become attractive to player A, he has to become sufficiently pessimistic about his chances of winning on the battlefield, and this requires significant fighting. To put it differently, since the payoffs for player B's strong type and player A must be at least V_{BS}^* and $V_A^*(q_0)$, respectively, Condition (6) implies that the weak type must get a significantly lower payoff than V_{BS}^*. In order to dissipate the weak type's surplus, the strong type must fight a long time before reaching an agreement.

By looking at revelation mechanisms, we avoid specifying a particular bargaining protocol. By the arguments familiar from mechanism design theory, as long as agreements are voluntary and the weak type can mimic the strong,[7] any Bayesian-Nash equilibrium for any protocol corresponds to a revelation mechanism that satisfies IR and IC. For example, suppose player A has all the bargaining power. The interval between $t = n\Delta$ and $t' = (n+1)\Delta$ corresponds to a "battle" during which no agreement can be made. For each $n \in N$, before the nth battle begins player A makes a proposal $x \in X$ that player B must accept or reject. Since X is finite, a sequential equilibrium exists by Theorem 6.1 of Fudenberg and Levine [9]. Theorem 2 implies that the probability that player B's strong type must fight for at least t units of time is at least $\lambda^*(t)$, regardless of Δ (the interval between offers). Thus, if the lemons condition holds, then the expected length of the conflict is bounded away from zero as $\Delta \to 0$. There must be significant fighting in any sequential equilibrium even if bargaining is "frictionless", so the Coase conjecture does not hold.

For comparison, consider the classic Akerlof lemons model. A buyer (player A) buys a car from a seller (player B) who knows its quality: lemon or peach. The buyer thinks the car is a lemon with probability q_0. A peach is worth rV_{BS}^* to the seller, a lemon is worth less. A peach is worth $1 - rV_{BS}^*$ to the buyer, but a lemon is only worth $1 - rV_{BW}^*$, so values are interdependent. When the lemons condition is violated, the buyer is willing to pay a price rV_{BS}^* that the peach owner is willing to accept, because

$$q_0 \left(1 - rV_{BW}^*\right) + \left(1 - q_0\right) \left(1 - rV_{BS}^*\right) \geq rV_{BS}^*.$$

However, if the lemons condition holds, then the price the buyer is willing to pay is unacceptable to the peach owner, so no fully efficient equilibrium is possible. In this case, it follows from Deneckere and Liang [4] that the Coase conjecture fails in the buyer-seller model. We have shown that it also fails in the Powell-Fearon conflict model, which does not follow directly from Deneckere and Liang [4]. There is no risk of collapse in the buyer-seller model; buyers only drop out when they accept an offer. In the conflict model, it is not only rejections that reveal information about the responder's type,

[7] The weak type's probability of collapse will always be β_W, not β_S. However, *until a collapse occurs*, the weak type can "feign strength" and behave just like the strong type. If he does, then as long as the conflict is ongoing, he will have the same probability distribution over agreements as the strong type.

but also the fact that there has not yet been a collapse. Indeed, this is the key distinction made in the seminal contribution of Wagner [23]. Since belief updating causes the commitment problem which drives the Coase conjecture, learning could be faster, exacerbating the commitment problem and making the Coase conjecture more likely to hold in the conflict model. However, Theorem 2 reveals that the commitment problem is overcome by the fact that player A's outside option makes it credible for him not to be too generous too soon.

Independently, Fearon and Jin [8] found that the Coase conjecture is violated in a refined PBE of the Powell-Fearon model, where player A increases his offers over time as in Gul, Sonnenschein, and Wilson [12]. Our result applies to any sequential equilibrium (in fact, to any Bayesian-Nash equilibrium) for any bargaining protocol where agreement is voluntary. In the next section we consider comparative statics and discuss some of the trade-offs that arise from third party intervention.

4 Comparative Statics

Third party interventions can change the nature and technology of conflict. One possibility is the third party has "no dog in the fight" and seeks only to maximize welfare by shortening the conflict. Another that it seeks to help player B but does not know player B's type. Finally, the conflict between player A and B might be a proxy war between player A and the third party. The third party might seek to use the proxy war to damage player A. Specifically, we are interested in how third party interventions change the parties' ability to solve the conflict by negotiations rather than by fighting.

In general, the effect of third party interventions will depend on the details of the bargaining protocol, and on which sequential equilibrium is being played. To obtain some intuition regarding the forces at play, in this section we assume the lemons condition holds and explore the implications of third party policies on $\lambda^*(t)$, the minimum probability that the strong type will fight for at least t units of time, and t^*, the minimum time before the conflict is sure to be over.

If the lower bound of Theorem 2 is attained then the probability that the strong type makes an agreement before the fighting starts (at $t = 0$) is

$$1 - \lambda^*(0) = r\frac{(1 - (\pi_A + \pi_B))\left[(\alpha + \beta_W + r) - q_0\left(\beta_W - \beta_S\right)\right]}{(r\pi_B + \alpha)\left(\beta_W - \beta_S\right) q_0} < 1 \qquad (12)$$

where the inequality is due to the lemons condition. Conditional on player B being strong and no player collapsing before time $t > 0$, the probability of an agreement no later than time t is

$$1-\rho(t) = 1-e^{(\alpha+\beta_S)t}\lambda^*(t) = e^{(\alpha+r+\beta_W)t}\frac{(\alpha+\beta_W(1-q_0)+q_0\beta_S+r)\,r\,(1-(\pi_A+\pi_B))}{q_0\,(r\pi_B+\alpha)\,(\beta_W-\beta_S)} \tag{13}$$

This expression reveals that the hazard rate of agreement for the strong type is

$$\frac{-\rho'(t)}{\rho(t)} = \frac{(\alpha+r+\beta_W)}{e^{(\alpha+r+\beta_W)t}\frac{(\alpha+\beta_W(1-q_0)+q_0\beta_S+r)r(1-(\pi_A+\pi_B))}{q_0(r\pi_B+\alpha)(\beta_W-\beta_S)}-1}.$$

We consider several cases.

First, suppose only β_S decreases. For example, a third party gives air defense weapons to player B but these weapons are useful only if player B is strong and knows how to use them. Thus, the probability that strong player B collapses in a battle decreases, with a direct effect of prolonging conflict in the absence of agreement. The indirect effect is that the probability of agreeing without fighting, $1-\lambda^*(0)$, falls. This happens because rW^*_{BS}, the minimum share that is acceptable to the strong type, increases. To dissipate the weak type's rent from mimicking the strong, the strong type has to fight with a higher probability. Also, the hazard rate of agreement for the strong type, $-\rho'(t)/\rho(t)$. To put it differently, the lemons problem gets worse when the relative hazard rate $\beta_W-\beta_S$ increases. The direct and indirect effects suggest the policy may backfire.

Third, suppose only β_W decreases. A third party trains player B in the use of advanced weapons but the training is necessary only if player B is weak. Thus, the probability that weak player B collapses in a battle decreases. The hazard rate of agreement for the strong type, $-\rho'(t)/\rho(t)$, may decrease as β_W decreases, because it takes longer to dissipate the weak type's rents. But other things being equal, the lemons problem is mitigated as $\beta_W-\beta_S$ decreases. These countervailing effects mean the total effect of a decrease in β_W on t^* is ambiguous:

$$\begin{aligned}\frac{dt^*}{d\beta_W} &= -\frac{1}{(\alpha+r+\beta_W)^2}\ln\frac{(r\pi_B+\alpha)(\beta_W-\beta_S)\,q_0}{r\,(1-(\pi_A+\pi_B))\,[(\alpha+\beta_W+r)-q_0\,(\beta_W-\beta_S)]}\\ &\quad+\frac{1}{(\alpha+r+\beta_W)}\frac{1}{(\beta_W-\beta_S)}\frac{[(\alpha+\beta_S+r)]}{[(\alpha+\beta_W+r)-q_0\,(\beta_W-\beta_S)]}.\end{aligned}$$

The first negative term captures the fact that a high hazard rate β_W facilitates rapid rent dissipation, making it easier to reach a fast agreement. The second positive term captures the fact that a high relative hazard rate $\beta_W - \beta_S$ means the net benefit to feigning strength is large, making it harder to agree quickly. The first term dominates when weak player B's rate of collapse β_W is high and strong player B's rate of collapse β_S is low. Then a reduction in β_W caused by third party intervention tends to delay agreement. However, if β_W falls further and approaches β_S, the second term dominates and time to agreement falls. In fact, if it falls enough, then the lemons condition becomes violated and fighting is no longer necessary.[8] Thus, the relationship between β_W and the length of conflict tends to be non-monotonic. When β_W is very high, weak player B collapses quickly and the conflict is short. When β_W is close to β_S, the lemons condition is violated and agreement is immediate. Intuitively, a third party who wants to maximize welfare should go "all in" and employ policies that help the weak type so much that types will be hard to distinguish on the battlefield. Half measures could backfire by prolonging the conflict.

Fourth, suppose α increases because a third party cuts off player A's supply of weapons and weapons parts. This has the direct effect of shortening the conflict simply because player A is more likely to collapse in the absence of agreement. But on the other hand, as types are interdependent, an increase in α affects player A's payoffs from fighting weak and strong opponents differently. As $\beta_W > \beta_S$, player A's payoff from fighting the weak type until a collapse decreases less rapidly than his payoff from fighting the strong type. This makes it easier to satisfy the lemons condition: the right-hand-side of (7) decreases in α. This means the effect of an increase in α on t^* is ambiguous:

$$\begin{aligned}\frac{dt^*}{d\alpha} &= \frac{-1}{(\alpha + r + \beta_W)^2} \ln \frac{(r\pi_B + \alpha)(\beta_W - \beta_S) q_0}{r(1 - (\pi_A + \pi_B))[(\alpha + \beta_W + r) - q_0(\beta_W - \beta_S)]} \\ &\quad + \frac{1}{(\alpha + r + \beta_W)} \frac{[(\beta_W(1 - q_0) + r(1 - \pi_B)) + q_0\beta_S]}{(r\pi_B + \alpha)[(\alpha + \beta_W + r) - q_0(\beta_W - \beta_S)]}.\end{aligned}$$

When α is low enough that the lemons condition is just barely satisfied, the second term dominates and an increase in α prolongs the conflict. When α increases, eventually the first term will dominates and the high probability that player A collapses or agreement is reached reduces the length of conflict.

[8] Mathematically, the two inequalities in Equations (7) and (12) are equivalent.

Again, it is half measures, where player A's capabilities are just slightly diminished, that may backfire by exacerbating the lemons problem. (This stands in contrast to the scenario without negotiations, where increasing α surely increases welfare.)

Fifth, suppose the third party puts restrictive trade embargoes and sanctions on players A and B, so π_A and π_B decrease. This has a direct effect on welfare by increasing the cost of conflict. But it becomes more costly for player A to screen different types of player B by fighting, so the lemons problem is mitigated and conflict ends more quickly, i.e., $\frac{d\lambda^*(t)}{d\pi_A} < 0$. Also, other things being equal, the weak type's marginal benefit from feigning strength decreases in π_B so $\frac{d\lambda^*(t)}{d\pi_B} < 0$. So, while the net effect on welfare is ambiguous, the effect on the expected length of the conflict is unambiguous. If the third party wants to decrease the duration of the conflict, strong sanctions are optimal.

The overall message is that third party interventions in favor of player B can backfire by making it harder for the two sides to come to an agreement. Specifically, policies that make player B even stronger when he is already strong – such as giving weapons that benefit only the strong type or sanctioning player A – tend to promote fighting rather than negotiating. Policies that make player B stronger when he is weak tend to mitigate the lemons problem, but half measures may backfire because the rate at which the weak type's rent is dissipated slows down. In contrast, a dramatic decrease in the weak type's probability of collapse can eliminate the lemons problem altogether, making immediate agreement possible and rent dissipation redundant.

5 Concluding Remarks

Leading historians and practitioners, such as Blainey [2] and Kissinger [14], have argued that uncertainty about the balance of power is a major cause of conflict. Blainey [2] argues that in most conflicts, each side starts out optimistic about the chances of winning on the battlefield. This makes it impossible to reach a diplomatic solution. But as the war grinds on, they become becomes more and more realistic about their chances of winning. Eventually, beliefs become sufficiently aligned to allow a diplomatic solution (c.f. Wagner [23]). At this time, the war ends. But after a period of peace, it is possible that beliefs will again start to diverge, triggering diplomatic crises and war. Such cycles are common in history:

> "In theory, of course, the balance of power should be quite calculable; in practice, it has proved extremely difficult to work out realistically. Even more complicated is harmonizing one's calculations with those of other states, which is the precondition for the operation of a balance of power. Consensus on the nature of the equilibrium is usually established by periodic conflict." (Kissinger, 1994, p. 63).

In game theory, it is well known that asymmetric information can explain inefficient outcomes, such as war. By convention, the "divergent beliefs" are assumed to be based on a common prior, but updated by private signals. However, if these beliefs converge quickly once war starts, then long wars would be a theoretical puzzle. We found that long wars *must* occur in the model proposed by Powell [18] and Fearon [7], because weak types can feign strength and beliefs do not converge quickly. This helps make uncertainty about the balance of power a plausible explanation for war.

References

[1] Ausubel Lawrence and Raymond Deneckere: "Reputation in Bargaining and Durable Goods Monopoly." *Econometrica*, Vol. 57, No. 3, May 1989, pp. 511-531

[2] Blainey, G. (1988) The Causes of War, The Free Press: New York; 3rd edition

[3] von Clausewitz, Carl (1832/1976): *On War*. Edited and translated by Michael Howard and Peter Paret. Princeton: Princeton University Press.

[4] Deneckere, Raymond J. and Meng-Yu Liang (2006), "Bargaining with Interdependent Values," *Econometrica*, Vol. 74, No. 5, pp. 1309-1364.

[5] Evans, Robert (1989): "Sequential Bargaining with Correlated Values," *Review of Economic Studies*, 56, 499-510.

[6] James Fearon (1995): "Rationalist Explanations for War," *International Organization*, 49, 3, Summer, pp. 379-414

[7] Fearon, James. (2013), “Fighting Rather than Bargaining,” mimeo, Stanford University.

[8] Fearon, James and Xin Jin (2021), “Armed Conflict Bargaining,” mimeo, Stanford University.

[9] Fudenberg, Drew, and David Levine (1983). "Subgame-perfect equilibria of finite-and infinite-horizon games." *Journal of Economic Theory* 31.2 (1983): 251-268.

[10] Filson, Darren and Suzanne Werner (2002): “A bargaining model of war and peace: anticipating the onset, duration, and outcome of war,” *American Journal of Political Science* 46(4):819–38.

[11] Faruk Gul and Hugo Sonnenschein (1988): “On Delay in Bargaining with One-sided Uncertainty,” *Econometrica*, 56(3): 601-611

[12] Gul, Faruk, Hugo Sonnenschein, and Robert Wilson, (1986), “Foundations of Dynamic Monopoly and the Coase Conjecture,” *Journal of Economic Theory*, vol.39, 155-190.

[13] Heifetz, A. and E. Segev, (2005), “Escalation and Delay in Protracted International Conflicts,” *Mathematical Social Sciences*, vol.49, 17-37.

[14] Kissinger, H. (1994), *Diplomacy*. New York: Simon and Schuster.

[15] Mao Tse-Tung (2000): *On Guerrilla Warfare*, translated by Samuel B Griffith. Champaign: University of Illinois Press.

[16] William McNeill (1982): *The Pursuit of Power*. Chicago: The University of Chicago Press.

[17] Andreu Mas-Colell, Michael Whinston and Jerry Green (1995): *Microeconomic Theory*. Oxford University Press: Oxford.

[18] Powell, Robert (2004), “Bargaining and Learning while Fighting,” *American Journal of Political Science*, vol.48, 344-361.

[19] Powell, R. (2012), "Commitment problems and shifting power as a cause of conflict”, in: M. Garfinkel and S. Skaperdas, eds.,*The Oxford Handbook of the Economics of Peace and Conflict*. New York: Oxford University Press.

[20] Slantchev, Branislav (2003) "The Principle of Convergence in Wartime Negotiations," *American Journal of Political Science*, 97(4): 621–32

[21] Smith, A. and A. Stam, (2004). "Bargaining and the Nature of War," *Journal of Conflict Resolution*, vol.48, 783-813.

[22] Daniel Vincent (1989): "Bargaining with Common Values," *Journal of Economic Theory*, 48:47-62.

[23] Wagner, R. Harrison (2000): "Bargaining and War." *American Journal of Political Science*, 44 (July):469–84

[24] Walter, Barbara (2009). "Bargaining Failures and Civil War." *Annual Review of Political Science* 12.1 (2009): 243-261.

[25] Wittman, Donald (1979) "How a War Ends: A Rational Model Approach" *Journal of Conflict Resolution*, Vol. 23, No. 4, pp. 743-763.