Bargaining and Conflict with Incomplete Information

Santiago Sanchez-Pages*

October 2009

Prepared for the
Oxford Handbook of the Economics of Peace and Conflict

Abstract

This paper studies bargaining and conflict under incomplete information, provides an overview and a critical account of the literature on the topic and contributes with original research. We first revise models of mechanism design and sequential bargaining that take confrontation as final. Conflict and inefficiencies are to be expected in these models whenever parties have optimistic prospects on the outcome of the all-out conflict. After examining the causes and reasons for this optimism, we move to the analysis of the recent literature that considers the existence of limited confrontations that allow bargaining to resume. In the presence of private information, these limited conflicts convey information and thus become a bargaining instrument. The paper closes with a discussion on the related empirical literature, the challenges that it faces and some potential avenues for further research.

JEL codes: C78, D74, D82.

Keywords: Bargaining, Conflict, Incomplete information, Power, Optimism, Hicks paradox, Uneven contenders paradox.

*School of Economics, University of Edinburgh, 31 Buccleuch Place, EH8 9JT Edinbugh, United Kingdom. E-mail: ssanchez@staffmail.ed.ac.uk

"Vito: I swore that I would never break the peace.
Michael: But won't they take that as a sign of weakness?
Vito: It is a sign of weakness."

The Godfather (1972).

1 Introduction

Conflict and bargaining are inextricably linked. In many situations, a failure in striking an agreement precipitates a confrontation between the parties involved. States engage in wars over territories, legal disputes end up in trial, unions engage in strikes, firms fight price wars against their competitors, couples argue on the distribution of chores. All these conflicts entail large losses in money, time, output and equipment. But then, why is conflict so pervasive? Why are agents unable to realize the existence of mutually beneficial agreements?

This paradox puzzled economist John Hicks, who attributed costly delays in labour negotiations to mistakes and irrational behavior (Hicks, 1932). Roughly at the same time, in the midst of the social turmoil of the Great Depression, sociologist E. T. Hiller argued that strikes occurred because parties could only estimate they comparative strength by actually engaging in conflict (Hiller, 1933). With that, Hiller was advancing the explanation to the *Hicks paradox* that we will be examining in detail in this chapter: Incomplete information. Given that strengths and power can be difficult to observe and measure, it is likely that parties hold private information over them. Moreover, parties have clear incentives to misrepresent their strength in order to obtain an advantage in the bargaining table. In that case, even if a mutually beneficial agreement actually exists, disagreement is to be expected, and thus conflict.[1]

The purpose of this chapter is also to challenge some of the commonly accepted ideas on conflict and bargaining within Economics. Economists' views on this issue are still profoundly influenced by Nash's (1950) conception of the bargaining problem, where conflict is reduced to a disagreement payoff that parties are supposed to receive when agreement fails and bargaining shuts down completely. That approach can be useful in many contexts, like bilateral trade; in that case, when buyer and seller disagree on the price for the good, they can walk away and look for alternative trading partners. But in many other situations, like the examples mentioned above, it is wrong to think of conflict and peace as two distinct phenomena. As Schelling (1960) put it in his seminal study on conflict, "most conflict situations are essentially bargaining situations."

The differences in the conducts of real conflicts falsify the conception of peaceful diplomacy and conflict as mutually exclusive alternatives. Many disputes do not entail any hostilities, like for instance the Second Moroccan Crisis in 1911 or 80-90% of contract negotiations (Kennan and Wilson, 1989). In many other occasions, hostilities precede agreement. As a matter of fact, about 65% of interstate wars end with a negotiated settlement (Leventoglu and Slantchev,

[1] Commitment problems constitute another powerful explanation for conflict. See Fearon (1995) for an illuminating discussion on this issue.

2007). Only a few disputes become absolute conflicts in which parties engage in a fight to the finish. Models of bargaining should be able to account for all these different patterns by incorporating conflict as part of the negotiation process.

In this chapter, we will provide an overview and a critical account of the literature on bargaining and conflict under incomplete information. We will first revise the basic economic approach to the topic that takes confrontation as final. That implies that conflict and inefficiencies are to be expected whenever parties have optimistic prospects on the outcome of the all-out conflict. We examine the causes and reasons for this optimism. We then move to the analysis of the recent literature that has challenged this view. These new contributions consider the existence of limited confrontations that allow bargaining to resume. In the presence of private information on the balance of strengths, that means that conflict can convey information and potentially become a bargaining instrument. We will contribute also with original research on this area. The chapter closes with a discussion on the related empirical literature, the challenges that it faces and some potential avenues for further research.

2 Disagreement as an all-out conflict

The most basic setting concerns two risk neutral agents denoted by $i \in \{1, 2\}$ who attempt to resolve how to divide one dollar. For the time being, we will assume that an all-out conflict ensues disagreement between these two agents. Each of them has a strength s_i within the set of possible strengths is $S_i = [\underline{s}_i, \overline{s}_i] \subset \mathbb{R}_+$ and with $S \equiv S_1 \text{x} S_2$. These strengths could be determined through agents' investments previous to bargaining, either by building arms, mobilizing supporters or by hiring quality lawyers, but here we will take them as given.

If a settlement is not reached then both parties resort to confrontation and fight an all-out conflict. After this conflict, no further interaction between them is possible. Moreover, this conflict is costly. A fraction $1-\theta$ of the dollar is lost if it occurs, where $0 < \theta < 1$. This accounts for the fact that war may render parts of the contested territory unproductive, legal services need to be paid and firms output is reduced by strikes.

The outcome of the all-out conflict is determined by the relative strengths of the parties, that is, it is determined by the balance of power. Given a pair of strengths $(s_1, s_2) \in S$, the probability that agent 1 wins the all-out conflict against agent 2 is $p(s_1, s_2) \in [0, 1]$. Agent 2 wins the conflict with probability $1 - p(s_1, s_2)$. We assume that agent i's winning probability is strictly increasing in his own strength and strictly decreasing in the strength of his opponent. Hence, stronger agents potentially benefit more from engaging in conflict. One possible formalization of this probability can be the class of Contest Success Functions (CSFs) axiomatized by Skaperdas (1996), and in particular the *difference functional form* (Hirshleifer, 1991):

$$p(s_1, s_2) = \frac{e^{\gamma s_1}}{e^{\gamma s_1} + e^{\gamma s_2}} = \frac{e^{\gamma (s_1 - s_2)}}{e^{\gamma (s_1 - s_2)} + 1}, \tag{1}$$

where the decisiveness parameter $\gamma \in [0, \infty)$ measures how noisy the outcome of the all-out conflict is: when $\gamma = 0$, it just becomes a fair lottery whereas when $\gamma \to \infty$ it is won with certainty by the strongest agent.

Note that under full information, a peaceful agreement should always be possible. If strengths are common knowledge, parties can easily compute their expected payoff from conflict. In line with Nash (1950), they are the *disagreement payoffs* of our bargaining problem.

$$d = \{d_1, d_2\} = \{\theta p(s_1, s_2), \theta(1 - p(s_1, s_2))\}.$$

Then any division of the dollar that gives to each agent more than d_i constitutes a potential peaceful agreement.

The presence incomplete information constitutes a fundamental problem here. When strengths are private information to the parties, a range of potential peaceful agreements may not exist. Following the pioneer work of Harsanyi (1967-68), let us identify strengths with "types" and assume that each agent knows her own type but not necessarily their opponent's. The distribution of agent i type is $f_i(s_i)$ with full support in $[\underline{s}_i, \overline{s}_i]$, that is, with $f_i(s_i) \in [0, 1]$ for all types $s_i \in S_i$ and zero otherwise. These types are selected independently by nature at the beginning of the game.

Given these beliefs about the strength of the opponent and their actual types (s_1, s_2), agents can derive their expected payoff from the all-out conflict:

$$\begin{aligned} E_1(d_1) &= \theta E_1(p) = \theta \int_{\underline{s}_2}^{\overline{s}_2} p(s_1, s_2) df_2, \\ E_2(d_2) &= \theta(1 - E_2(p)) = \theta(1 - \int_{\underline{s}_1}^{\overline{s}_1} p(s_1, s_2) df_1). \end{aligned}$$

Under these circumstances, agreement is *impossible* if beliefs are such that $E_1(d_1) + E_2(d_2) > 1$, that after some rewriting becomes

$$\int_{\underline{s}_2}^{\overline{s}_2} p(s_1, s_2) df_2 - \int_{\underline{s}_1}^{\overline{s}_1} p(s_1, s_2) df_1 > \frac{1 - \theta}{\theta} \equiv R. \tag{2}$$

We refer to R as the *Loss ratio*. A peaceful agreement becomes impossible if the difference between the two agents' expectations about agent 1 probability of winning the conflict is too big. If both agents are very optimistic about their own winning probability it is more likely that no settlement will satisfy both of them at the same time. On the other hand, note that as the Loss ratio increases, that is, as the all-out conflict becomes costlier, an agreement becomes more likely to be possible.

The main question at this stage is thus whether it is possible to solve the conflict of interests between the parties and to avoid a costly conflict, at least partially. This is the purpose of the next section.

3 The mechanism design approach

Mechanism design is a very powerful tool in the study of situations in which there exists a conflict of interest among agents. Its goal is ambitious: To find social contracts, peace-keeping institutions or arbitration bodies whose rules are able to reduce, and even eliminate, the incidence of conflict.

This represents the *normative* side of mechanism design. However, this approach has also a *positive* side. It can be used also to derive very general results of bargaining games. Mechanism design is a "game-free" approach, that is, it does not posit any assumption on the specific bargaining process to be followed. Because of that, this approach offers a class of results that any equilibrium of any unmeditated bargaining game must satisfy (Ausubel et al., 2002).

The generality of the mechanism design approach stems from the *revelation principle* (Myerson, 1979). It states that we do not need to consider the potentially very complex strategies that parties can follow. We can just restrict our attention to mechanisms that ask participants to announce their types. And within this class of mechanisms, called direct revelation mechanisms, we should look for the ones that are (Bayesian) incentive compatible, that is, those that make players reveal (and thus behave according to) their actual type.

The application of the tools of mechanism design to the study of conflict was pioneered by Banks (1990). We will now follow him in the formalization of the problem. A mechanism is a pair of functions $< \varkappa, \pi >: S \to [0,1]\text{x}[0,1]$ determining a division of our initial dollar and a probability of the all-out conflict occurring. For each pair of reported types (s_1', s_2'), the function $\varkappa(s_1', s_2')$ is the share of the dollar received by Player 1 (Player 2 receives $1 - \varkappa(s_1', s_2')$). With probability $\pi(s_1', s_2')$ the conflict occurs, and then each player receives d_i.

Define now as

$$\begin{aligned} U_1(s_1' \quad | \quad s_1) &= \int_{\underline{s}_2}^{\overline{s}_2} [(1-\pi(s_1', s_2))\varkappa(s_1', s_2) f_2(s_2) ds_2 + \pi(s_1', s_2) d_1] df_2 \\ &= \int_{\underline{s}_2}^{\overline{s}_2} (1-\pi(s_1', s_2))\varkappa(s_1', s_2) df_2 + \theta \int_{\underline{s}_2}^{\overline{s}_2} \pi(s_1', s_2) p(s_1, s_2) df_2, \end{aligned}$$

the expected utility of player 1 from announcing type s_1', and

$$U_1(s_2' \mid s_2) = \int_{\underline{s}_1}^{\overline{s}_1} (1-\pi(s_1, s_2'))(1-\varkappa(s_1, s_2')) df_1 + \theta \int_{\underline{s}_1}^{\overline{s}_1} \pi(s_1, s_2')(1-p(s_1, s_2)) df_1,$$

the analogous for player 2 when he announces type s_2'.

The mechanism $< \varkappa, \pi >$ is *incentive compatible* if for both types it is the case that

$$U_i(s_i \mid s_i) \geq U_i(s_i' \mid s_i) \text{ for all } s_i \in S_i. \tag{3}$$

The second type of restrictions we will impose on our mechanism is that they must be *individually rational,* that is, they should give in expectation to participants at least what they can obtain by triggering the all-out conflict. This

was implicitly assumed in our initial analysis in the previous Section. Formally, the *participation constraints* imply that

$$U_i(s_i \mid s_i) \geq E_i(d_i) \text{ for all } s_i \in S_i. \tag{4}$$

The fact that the all-out conflict is costly ensures that, contrary to Myerson and Satterthwaite (1983), a peaceful agreement is always ex-post efficient. That is why we can evaluate the quality of a mechanism by how likely it is to avoid confrontation. In particular, we focus on mechanisms that are *peaceful* (Bester and Warneryd, 2006), that is, mechanisms such that

$$\pi(s_1, s_2) = 0 \text{ for all } (s_1, s_2) \in S.$$

Peaceful mechanisms are both ex-post and ex-ante efficient. The problem is that to impose participation constraints implies that when no mechanism can be peaceful if there exists a realization of types such that agreement is impossible, i.e. condition (2) holds, .

Proposition 1 *There exists a threshold $\widehat{R} > 0$ on the Loss ratio with*

$$\widehat{R} \equiv \int_{\underline{s}_2}^{\overline{s}_2} p(\overline{s}_1, s_2) df_2 - \int_{\underline{s}_1}^{\overline{s}_1} p(s_1, \overline{s}_2) df_1, \tag{5}$$

such that a peaceful mechanism exists if and only if $R < \widehat{R}$.

The expression for $\widehat{R}$ comes naturally from (2): Given that p is strictly increasing in s_1 and strictly decreasing in s_2, it is clear to see that for any $R \leq \widehat{R}$ a peaceful mechanism does not exist[2].

When $R > \widehat{R}$ our mechanism must admit the occurrence of conflict for some realizations of types. This is because if the all-out conflict is not too destructive then it may become too attractive for stronger agents. Peace cannot prevail when both highest types are drawn. This result was first stated by Spier (1994) in the context of pretrial negotiations (where the all-out conflict takes the form of going to court), and later by Fey and Ramsay (2009) in the context of international relations. It fits with the idea that it was the Mutual Assured Destruction what prevented the Cold War from escalating into a nuclear conflict; such war would had been so destructive, i.e. the Loss ratio was close to infinity, that a range of peaceful mechanisms surely existed. Let us reiterate that this result is very general: the impossibility of peaceful mechanism applies to any game in which the cost of an all-out conflict is small enough.

Note that the threshold $\widehat{R}$ potentially depends upon the technology of conflict and the distribution of types. For example, suppose that $S_1 \equiv S_2 \equiv [0, 1]$ and that $p(s_1, s_2)$ takes the form in (1). Then it is easy to show that,

$$\widehat{R} = \frac{1}{\gamma}(\ln(e^{\gamma} + 1) + \ln\left(e^{-\gamma} + 1\right) - 2\ln 2),$$

[2] That $\widehat{R} > 0$ can be established just by seeing that $\widehat{R} > P(\overline{s}_1, \overline{s}_2) - P(\overline{s}_1, \overline{s}_2) = 0$.

that in turn is increasing in γ. In this case, as the decisiveness of the conflict technology increases, the existence of a peaceful mechanism is less likely. The intuition for this is that as γ the stronger types win the conflict with higher probability and that make this option more attractive.

On the other hand, Bester and Warneryd (2006) explored what distributions of strengths make conflict more likely. This question has been central in Political science for a long time and has produced all possible answers: The preponderance-of-power school argues that when one of the parties is weak, it is more likely to accept a negotiated settlement, whereas the balance-of-power school argues that that has the opposite effect because the stronger party will demand more and those greater demands are more likely to be rejected[3]. The results in Bester and Warneryd (2006) show that if the distribution of strengths of one or both agents shift to the right according to the first order stochastic dominance criterion, then $\widehat{R}$ decreases, making the existence of a peaceful mechanism more likely. This implies that if one agent becomes relatively more powerful than the other, peace is more likely to occur. This at first glance supports the school that defends the preponderance of power as a guarantor of peace. However, this conclusion would be wrong. Note that this result holds regardless of whether the party that becomes more powerful was weaker or stronger than her opponent. Then, the likelihood of conflict does not seem to depend on the balance of strengths in a meaningful way (Wittman, 1979).

These authors also show that if the distributions of skills become riskier according to the second order stochastic dominance criterion, that makes $\widehat{R}$ increase and conflict more likely. This relates with the idea that incomplete information difficults agreement. When parties hold more precise information about their opponents the potential for a peaceful settlement increases.

The next step is to characterize the mechanism. Recently, Frey and Ramsay (2009) generalized the analysis of Banks (1990) along this line. The latter author showed that when only of the parties is uninformed about the balance of strengths then the mechanism must be monotonic, that is, the stronger the informed player the more she receives in case of a peaceful settlement. This generalizes to the case of two-sided incomplete information:

Proposition 2 *Suppose that $p(s_1, s_2)$ is continuously differentiable in s_1 and s_2. Then*

(i) $U_i(s_i') = U_i(s_i) = E_i(d_i \mid \overline{s}_i)$ for any $s_i', s_i \in S_i$ if a peaceful mechanism exists,

(ii) $U_i(s_i') > U_i(s_i)$ for $s_i' > s_i$ if a peaceful mechanism does not exist.

[3] Claude (1962) belong to the former school, whereas Organski and Kugler (1980) and Blainey (1973) belong to the latter. The balance of power school received empirical support in Bueno de Mesquita et al. (1997).

The intuition for this result is very simple. If a peaceful mechanism did offer different equilibrium payoffs to different types then the worst off types would have an incentive to mimic the better off types. That is profitable because they do not incur in any risk of conflict by doing that given that the mechanism is peaceful. Moreover, the mechanism awards agents the expected payoff from the all-out conflict for the strongest types. Notice that this is feasible precisely because a peaceful mechanism exists. When this is not the case, stronger types must receive higher expected payoffs but this must be at the cost of a higher probability of conflict. Type-contingent payoffs are necessary to elicit truth-telling: Stronger players suffer a higher probability of conflict because that reduces the incentive of weaker agents to exaggerate their report on their own type. Better settlements on the other hand compensate stronger types for this by giving them higher rents. This is called the *risk-return trade-off*.

Finally, let us clarify that the mechanism design approach does make assumptions. As we will see next, it typically does assume a very basic structure. In the case of bargaining and conflict, this structure follows the general setting laid out in the previous section: One or both parties have private information about their strength and a failure to strike a bargain provokes an all-out conflict that is probabilistic and costly and that precludes any further interaction. Very different mechanism could be obtained if, as we will see in the next sections, conflict is not a game-ending move.

4 Optimism as a cause of conflict

Let us now reflect on the analysis performed so far. Optimism emerges as a valid explanation to the Hicks paradox. Excessive optimism can explain the paradox without resorting to bounded rationality because it is capable of generating incompatible demands. As we have seen in the previous section, and unlike in Myerson and Satterthwaite (1983), breakdown of negotiations can occur despite the existence of gains from trade being common knowledge.

Optimism has become a widespread explanation to the ubiquity of war. In his now classic book, *The Causes of War* (1973) military historian Geoffrey Blainey concluded that the main reason for armed conflicts throughout history has been the high hopes that countries had in the eve of war, hopes that made countries believe that they could achieve more through conflict than through diplomacy. Because of that, it was impossible to bring one or both of them to the bargaining table. Since then, mutual optimism has been widely cited as a source of interstate war (Wittman, 1979). In the previous Section we have seen that one-sided incomplete information is able to generate this result. So optimism does not even need to be *mutual* to preclude agreements.

One important question, specially when looking at particular applications, is where does this optimism come from. Following the influential study by Fearon (1995), let us discuss three main answers.

The first one follows the standard argument of economists: parties just hold different information. Otherwise they should have exactly the same estimates

about winning the conflict, that is, the should have *common priors* (Harsanyi, 1967-68). Superior information can stem from a better knowledge of military capabilities, resilience of union members in case of a protracted strike or exclusive evidence that litigants may hold about the case. Incomplete information comes thus from a difference between observable and non-observable capabilities.

The second set of explanations does not discard that agents hold potentially the same information about the balance of strengths. It assumes instead that they reach different estimates because they process information differently. One reason for this comes from relaxing the assumption of unitary actors: Military assessors, factions within political parties or lawyers may distort decision making within organizations. As Blainey (1973) pointed out, political leaders tend to be surrounded by people who feed them with mostly positive information rather than with realistic estimates. Another reason is that information acquisition and processing is costly. Decision makers face such a volume of information that they make mistakes when they compute estimates. All these reasons contribute to generate what Schelling (1960) called an "imperfect process of decision."

The third type of explanations states that parties somehow neglect information. Blainey (1973) believed that nationalism and patriotism often made leaders "evade reality" and get carried away by *animal spirits.* For instance, it seems that racial stereotypes where at the root of the Russian optimism on the eve of the Russian-Japanese war of 1904 and on MacArthur overconfidence before the Korean War. On the other hand, Goemans (2000) argued that weak autocrats have often used war as a gamble to ensure their political survival despite obvious and evident unfavorable odds. This is because, as democratic leaders, they are likely to lose power in case of a defeat. But contrary to elected politicians, being overthrown may entail severe domestic punishments. This can lead them to neglect "hard" information. Finally, Johnson (2004) uses recent advances in evolutionary psychology to argue that overconfidence in conflict was an adaptive trait in our evolutionary past. Optimistic assessments of the own strength enhance performance and the probability of winning confrontation and that explain their pervasiveness.

However, note that although these three explanations can account for differences in parties' estimates, they do not preclude that agents act rationally. Contrary to what Fearon (1995) argues, it may be perfectly rational for parties to have different estimates even when they have the same information. The non-unitary actor or the political survival hypotheses can deliver that[4]. But, to a certain extent, the causes of optimism are not ultimately important. It may be that parties holds optimistic beliefs because there is an innate human tendency towards "positive illusions" and self-deception. But that does not preclude that, given those beliefs, parties can still make rational choices and play their best response to whatever their opponents do. An investigation on the likely causes of optimism can shed light on how parties factor in the new information that they learn through the bargaining process into their decisions.

[4]The non-unitary actor explanation for optimism, although outside the scope of this chapter, deserves extensive consideration.

The next question is under which circumstances should we expect optimism. The causes of optimism can also tell us when it is more likely to emerge. According to the non-unitary actors hypothesis, excessive demands will take place when parties's decision process is not transparent and the internal debate is not open. That has led Johnson (2004) to conjecture that optimism is more likely to occur in dictatorial regimes than in democratic ones, because in the latter overconfident leaders can be better counteracted. On the other hand, Blainey (1973) argued that optimism is more likely when power is distributed equally. Wagner (1994) built up on this by making the distinction between observable and unobservable capabilities. When the balance of the former is more equal, the role of the latter should be more important and hence the more likely is that parties will feel optimistic. If the balance of observable capabilities is very unequal, then the importance of unobservable characteristics should be much smaller and hence optimism must be less likely to precipitate war. However, this hypothesis is problematic: There exists many contexts in which one can observe clearly small and weak agents fighting against much larger and powerful ones; cases range from David against Goliath to the Vietcong against the US. This phenomenon was first noted by Clausewitz (1832) and is commonly referred to as the *Uneven contenders paradox* (Sanchez-Pages, 2003, and Jackson and Morelli, 2007). Hence, any model that explores the relation between conflict and incomplete information should be able to offer an explanation to this puzzle. We will elaborate more on this later on.[5]

The reader may be perplexed at this point. It seems that each time incomplete information render parties' demands incompatible confrontation will ensue. Conflict is however known to be a destructive option and it is therefore common knowledge that there must exist some Pareto superior agreement. This was precisely what puzzled Hicks in the first place and what had lead some authors to look at the optimism hypothesis with distrust. By observing a mutual willingness to fight, parties should be able to infer that optimism is prevalent and that should lead them to revise their demands, up to a point in which one of them prefers to settle (Fey and Ramsay, 2007). This is just the logical extension of Aumann's (1976) idea that rational players should never "agree to disagree" when their rationality is common knowledge. Whereas this argument is not valid in the context of conflict, it allows us to understand better the pervasiveness of conflict despite the apparent paradox.

Firstly, confrontation, although mutual in essence (it occurs only if at least two parties fight), can be imposed by just one party. So even if one of the parties prefers to settle, the other will continue fighting if by doing so he thinks that he can obtain a better deal. But more importantly, parties always have an incentive to display a willingness to fight because they have an incentive to misrepresent their private information. As we saw in the previous section, a peaceful mechanism must be flat because otherwise players have an incentive to bluff and exaggerate their strength. The possibility of conflict occurring was the

[5] Wrangham (1999) explain this paradox in terms of the evolutionary adaptiveness of military incompetence. He argues that agents are more likely to become military or political leaders if they have positive illusions and tend to deceive themselves more often.

only way in which truth-telling could be elicited. If negotiation fails then parties must commit to fight (and suffer the costs from it). Otherwise their demands cannot be credible. No cheap-talk type of communication, no declaration of intentions can help to avert confrontation by itself. When optimism precipitates conflict, it is very likely that one party will regret it because the fact that hostilities began proves that her expectations were wrong.

Learning seems then the only way out from the gloomy prospect of an all-out conflict. The next section describes models of sequential bargaining in which parties can learn form disagreement or can try to signal credibly their private information through the process of negotiation. After that, we will describe a second and, until recently, neglected source of information: conflict itself.

5 Learning by bargaining

In retrospect, mechanism design may be regarded as "too general", because it can not incorporate important features of specific applications. One could argue that models that make explicit and detailed assumptions about the bargaining protocol are richer and perhaps more "realistic". But at the same time they may not be robust or their equilibria may be difficult to find, something that ultimately reduces their potential applicability. Banks (1990) himself advocated a "two-step" approach: Mechanism design should be used first to establish general results and then additional predictions should be obtained from using a game form that captures the main features of the application of interest.

But there exists a second and important motivation for the analysis of sequential models of bargaining: Parties can learn about each other through the process of negotiation itself, a possibility that was advanced by Schelling (1960) in Economics and Pillar (1983) in International Relations. Parties can update their prior information by observing the offers that are on the table or from the rejection of the offers they made to their opponents. Hence, agreements must incorporate the information that players learn through the process of bargaining. Notice that mechanism design cannot tackle this issue appropriately. It assumes no structure in the process of bargaining so it cannot yield specific predictions about how this learning takes place. Sequential bargaining is much better suited to analyze this issue in particular.

Sequential bargaining models with incomplete information flourished in the 80s. In them, negotiations occurs through a dynamic process of bilateral negotiations according to a pre-specified protocol and under a particular informational structure. Two main protocols are mainly considered: one-sided offers and alternating offers. On the other hand, incomplete information could be one-sided or two-sided. This spawns four possible bargaining structures. In this Section we will consider two of them in detail: When offers are made only by the informed party and when they are made only by the uninformed party. We will adapt these standard bargaining models to the analysis of conflict. In the last part of the Section we will also discuss other protocols.

In order to simplify the exposition, we will consider models with only two

periods. Hence $t = 1, 2$. Players are impatient and discount the future at a common factor $\delta \in [0, 1]$. In each period only one offer can be on the table. An offer $x_t \in [0, 1]$ specifies the share of the cake that Player 1 will receive if the offer made at period t is accepted . His payoff is then $\delta^t x_t$. Player 2 in that case receives $\delta^t(1 - x_t)$. If none of the two offers is accepted then the all-out conflict takes place and the game ends.

For the sake of exposition we will collapse the distribution of types into a two type support, so $S_i = \{\underline{s}_i, \overline{s}_i\}$. When one of the strengths is common knowledge, Player 1's probability of victory in the all-out conflict can thus be reduced to $p \in \{p_L, p_H\}$ with $1 > p_H > p_L > 0$. So instead of specifying priors over the strength of the informed party, we will just specify them over p at the beginning of the game: the uninformed party believes it to be p_H with probability μ_o and p_L with probability $1 - \mu_o$.

A history $h(t)$ of the game specifies at t the offers made so far and whether they were accepted or rejected. Beliefs consist then of a probability distribution $\mu(\cdot \mid h(t))$ over the set of types. Hence, a strategy for the player making the offers is a map from the set of histories and types into the set of actions $\{x_1, x_2\}$. Similarly, a strategy for the player that receives the offers is a map from the set of histories into $\{Accept, Reject\}$.

Perfect Bayesian Equilibrium is the standard equilibrium concept employed in this type of games. It consists of a strategy profile and a system of beliefs that are consistent with each other via Bayes' rule and such that all players are playing best responses to each other given those beliefs. This concept requires that we specify even off-the-equilibrium-path beliefs since Bayes' rule imposes no restriction on them. This leads to a multiplicity of equilibria that pervades most of the bargaining literature (Ausubel et al., 2002)

5.1 The uninformed party makes all the offers

The following follows the analysis in Fudenberg and Tirole (1983), extended by Sobel and Takahashi (1983) to a continuum of types and an infinite horizon. The dynamics of negotiation in these models take the form of *screening:* The uninformed party makes offers that are accepted first by players with a lower outside option. Learning occurs through negotiations: a rejection can potentially signal that the informed party has a high value from resorting to the all-out conflict. In turn, that prompts the uninformed party to concede and raise their initial offer. Therefore, three situations may arise: Agreement is reached in the first round, after and impasse or never. The final outcome depends on the incentives of the uninformed player, Player 2 in our case, to engage in screening.

The strategies for Player 1 are the probabilities of accepting the offer at each period. Let us denote them as $\{q_{L1}(x_1), q_{L2}(x_2)\}$ and $\{q_{H1}(x_1), q_{H2}(x_2)\}$ for types L and H respectively. No offer will be greater than θp_H. Hence we know that $q_{H1}(x_1) = 1$ for $x_1 \geq \delta\theta p_H$ and 0 otherwise. Hence, we only need to care about $q_{L1}(x_1)$.

Rejection of the offer conveys information about the type of Player 1. Given the initial beliefs about Player 1 being strong, Player 2's beliefs in the second

stage are just

$$\mu(H \mid q_{L1}) = \frac{\mu_o}{\mu_o + (1-\mu_o)(1-q_{L1})}, \tag{6}$$

which is greater than μ_o. At the second stage, type i will accept the offer x_2 if and only if $x_2 \geq \theta p_i$. Offers by Player 2 will then only take two values, θp_L and θp_H. It will be the latter if and only if

$$\mu(H \mid q_{L1}) \geq \frac{p_H - p_L}{R + p_H - p_L},$$

where recall that R is the loss ratio. Suppose that this the case. Then, both types of Player 1 will accept the first period offer if and only if $x_1 \geq \widetilde{x}_1 = \delta\theta p_H$. Hence $\widetilde{x}_1$ is a *pooling offer* that is accepted by both types immediately. If we specify that in case of an off-the-equilibrium path rejection, Player 2 believes she is facing a strong opponent then this pooling offer constitutes a candidate for a sequential equilibrium.

Now assume that Player 2 offers $x_2 = \theta p_L$. That can happen if and only if

$$q_{L1} \leq q \equiv 1 - \frac{\mu_o}{1-\mu_o}\frac{R}{p_H - p_L} = 1 - \mu_o \frac{R}{\widehat{R}},$$

where the last inequality just comes from rewriting (5) in the current setup. Note that if the L type is randomizing it must be that $x_1 = \delta\theta p_L$. Hence, a second candidate for equilibrium arises is a *separating* one, in which Player 2 offers $\delta\theta p_L$ in period 1 that the H type rejects and the L type accepts with probability q (since that leaves indifferent Player 2 between his two possible offers at $t = 2$). The following results summarizes this findings and fully characterizes the conditions under which each of these candidates constitutes an equilibrium.

Proposition 3 *In the two-period game in which the uninformed party makes all offers, for low values of the loss ratio, i.e.* $R \leq \frac{\widehat{R}}{\mu_o}$*, there exists a threshold discount rate*

$$\delta_{\mu_o} \equiv \frac{\mu_o}{\mu_o + \frac{\widehat{R}}{R}\frac{\frac{\widehat{R}}{R} - \mu_o}{1 + \frac{\widehat{R}}{R} - \mu_o}},$$

such that if $\delta \geq \delta_{\mu_o}$ *the separating profile of offers is the unique equilibrium. Otherwise, the pooling offer constitutes the unique equilibrium.*

The pooling profile implies that a settlement is reached immediately and conflict does not occur. This is what we already obtained in Section 3: Conflict is averted if the cost of the all-out conflict is high enough. Otherwise, conflict occurs because Player 2 prefers to take a calculated risk: He discriminates between the two types of Player 1 by making a low offer in period one and a higher one in period 2, but not high enough to appease the strong type so conflict occurs with probability μ_o. This risk is necessary for this profile to be

an equilibrium. If Player 2 were to appease the H type by making a high offer in the second period, the L type could mimic him by rejecting x_1. In addition, one can easily show that δ_{μ_o} is increasing in μ_o, so peace becomes more likely the more likely Player 2 believes Player 1 to be of the strong type.

This also shows that the feasibility of an agreement, in the sense of $R \leq \widehat{R}$ is a necessary but not a sufficient condition for a settlement to occur. Player 2 can engage in separation when the cost of conflict is moderate, i.e. $R \in (\widehat{R}, \frac{\widehat{R}}{\mu_o})$. This is because he can obtain a better deal by being deliberately tough, even though that entails the risk of an all-out conflict with the strong type.

In the context of bilateral trade, Sobel and Takahashi (1983) showed that with infinitely many periods and a continuum of types, the uninformed seller uses an ascending profile of offers, that gradually appease buyers with higher reservation values. Low valuation buyers wait to get a better price and use delay as a signal of their valuation. As noted by Kennan and Wilson (1993), this provides some foundations to the observation that unions make lower demands on wages as strikes progress (the decreasing union "resistance curve" described by Hicks, 1932), or to the fact that initially demanding countries concede more to their rivals as hostilities last longer (Slantchev, 2004). This result rests on the fact that delay is more costly for high valuation buyers or more profitable firms. As the time between offer becomes smaller, or alternatively, as the discount factor approaches one, delay is less effective in screening types and the "Coase conjecture" applies (Coase, 1972): agreement is immediate. Ultimately, this implies that disagreement can be sustained only if there exists an exogenously imposed lapse between offers. However, the Coase conjecture does not hold in models of bargaining under the shadow of conflict because time itself does not help parties to screen their opponents.

5.2 The informed party makes all the offers.

When the uninformed agents makes all offers, he learns from disagreement; the informed party can, to a certain extent, signal credibly his type by rejecting the first period offer. This enabled the uninformed agent to screen the type of their opponent. On the other hand, when the informed party is the one making the offers, information is conveyed through offers themselves. Then *signaling* occurs. Next we will follow a modified version of the model in Sanchez-Pages (2009) to describe this bargaining problem.

Let us first show that a two-period separation profile, in which the weak type settles immediately and the strong type settles at $t = 2$ after making a non serious offer at $t = 1$ cannot be an equilibrium. The reason is straightforward. In that case, Player 2 knows he is facing the H type in period 2 so it must be that $x_2 \geq 1 - \theta(1 - p_H)$. For this equilibrium to exist it must be that $x_1 \geq \delta(1 - \theta(1 - p_H))$, because otherwise the L type could be better off by mimicking the behavior of H. But if that is the case then the H type himself has an incentive to mimic L and settle immediately. This implies that the informed party cannot use time to credibly signal its type. Sanchez-Pages (2009) argues

that this is a very general property of this type of bargaining models.

Hence, only two types of profiles can constitute an equilibrium. Both types should make the same offer at $t = 1$ (the two-type assumption precludes the construction of a fully revealing schedule of offers), or the L type settles at $t = 1$ and the strong one makes a series of non serious offers that precipitates conflict. Two elements are critical here. First, how the second period unravels. If the second period is reached because both types made the same offer at $t = 1$, posterior beliefs coincide with priors and time does not transmit any information. Given that, the minimal acceptable offer to Player 2 is simply

$$\widetilde{x}_2 = 1 - \theta(1 - \mu_o p_H - (1 - \mu_o)p_L).$$

For it to be a candidate for an equilibrium, the H type must prefer to make it than to make an non-serious offer, i.e. $x_2 = 0$, that would precipitate conflict. That is,

$$\widetilde{x}_2 \geq \theta p_H \;\Leftrightarrow\; R \geq \widehat{R}.$$

For some other range of parameters, separation can occur at $t = 2$. Type L can reveal his true type by making an offer $x_2 = \underline{x} = 1 - \theta(1 - p_L)$.But it must be such that H type would never make it, that is, $\underline{x} < \theta p_H$. This implies that separation can be sustained only when

$$R < p_H - p_L = \frac{\widehat{R}}{1 - \mu_o}.$$

Under these two scenarios, the expected payoff for Player 2 from rejecting a pooling offer at $t = 1$ is just $1 - \widetilde{x}_2$. So immediate settlement can occur only if $x_1 \leq 1 - \delta + \delta\widetilde{x}_2$. Whether this scenario arises or not will depend on the discount factor.

The second element to be considered are beliefs. Recall that the PBE concept does not impose any restrictions on beliefs off-the-equilibrium path. This gives rise to a multiplicity of equilibria. These beliefs are critical and the equilibrium outcomes are quite sensitive to their specification. For our purposes consider the following set of "reasonable" beliefs:

$$\mu(H \mid x_1 \leq 1 - \delta + \delta\underline{x}) = \mu_1 \in [0, \mu_o]. \tag{7}$$

These beliefs are "reasonable" because they assume that if Player 2 observes any serious offer at $t = 1$, that is, an offer not smaller than his absolute minimal acceptable offer at that time, he does not believe that he is more likely to be facing the strong type than at $t = 0$. We will require that these beliefs will be maintained in $t = 2$.

Now we are in the position of characterizing a set of PBE of this game

Proposition 4 *In the two-period game in which the informed party makes all offers and the uninformed player holds beliefs (7), for low enough values of the loss ratio, i.e.* $R \leq \frac{\widehat{R}}{1-\mu_0}$, *there exists a threshold discount rate*

$$\delta_{\mu_1} \equiv \frac{1+R}{1+R+\frac{1-\mu_1}{1-\mu_0}\widehat{R}},$$

such that if $\delta \geq \delta_{\mu_1}$ *the L types offers* $x_1 = 1-\delta+\delta\underline{x}$ *and the H type precipitates conflict in the PBE of the game. Otherwise, both types offer* $x_1 = 1-\delta+\delta\widetilde{x}_2$ *and the uninformed agent accepts.*

Note that it is again the case that the feasibility of agreement is not enough to ensure a peaceful settlement. That can be the case only if $\mu_1 = \mu_o$. When the conflict loss is small and the discount factor is high enough so the H type cannot be satisfied with any acceptable agreement for Player 2, then conflict happens with probability μ_0. The reason for this, again, lies in the fact that the strong type cannot use time to separate himself from the weak type. When conflict and delay are not too costly, conflict becomes inevitable. This is aggravated by beliefs: The more likely Player 2 thinks he is facing a weak rival after observing a serious offer the smaller the sets of parameters under which immediate agreement occurs. In that sense, these beliefs become self-fulfilling: only weak types make an offer in the first period, whereas the strong types resort to conflict.

The fact that the passage of time cannot convey information is an unsatisfactory feature of this model. That cannot explain why sometimes, after a long period of hostilities, one party accepts an agreement with terms identical to those proposed before confrontation started. The negotiations between Israel and Egypt during the 70s or between the IRA and the British government in the 90s found a breakthrough only after a protracted confrontation. The fact that the agreements finally reached almost mirrored the initial proposals suggests that parties needed time to reveal substantive information about their stands.

This is the precisely starting point of the article by Heifetz and Segev (2005), who extend the classic analysis on signaling in bargaining by Admati and Perry (1987) to the study of escalation and conflict. In their model incomplete information is about the reservation stand of one of the parties. The informed agent can choose the timing of his offer so that self-imposed delay credibly signals his reservation stand. The uninformed one has also the opportunity to become the aggressor and impose an escalation of the conflict, which implies imposing costs to both parties. In the equilibrium of the game, agreement follows after the first offer, although low valuation types wait longer. The effect of escalation in the model is that it makes agreement to be reached sooner, because parties find delay more costly, and that it may improve the terms of the agreement. However, this option does not necessarily benefit the aggressor. The authors characterize conditions under which he can regret ex-post his decision to escalate.

5.3 Other bargaining protocols

In an influential article, Powell (1996) explored an infinite horizon bargaining model with two-sided incomplete information and alternating offers. In that model, parties can exchange offers or impose a settlement (an all-out conflict) when they become too pessimistic about the outcome of the negotiations. The costs of imposing this settlement are private information to the parties. In fact, incomplete information on the costs of conflict rather than on the balance of strengths is the typical modelling choice of political scientists (Brito and Intriligator, 1985, Fearon, 1995). Whereas this feature is, to a large extent, just a matter of taste, exploring the similarities between these two formulations is an avenue for further research[6]. Powell's model displays two distinguished features that will give us the opportunity to discuss two important aspects of negotiations under the shadow of conflict.

First, he assumes that the party that makes the offer can also impose a settlement. In contrast, the models described above implied an exogenously imposed deadline to negotiations. That may fit situations in which a third party (an international organization, an arbitration body) is imposing a negotiation protocol, or parties agree on using a particular one. But that begs the question of why such process was agreed upon. But, more in general, assuming that the unilateral use of confrontation is always an option for parties makes perfect sense in the context of conflict. Powell (1996) shows that this assumption reduces considerably the potential complexity of his model. In the end, the equilibrium outcome is unique and involves one party making a take-it-or-leave-it offer that his opponent can accept or reject by imposing a settlement. This contrasts with the results in bargaining models of bilateral trade (Fudenberg et al., 1987) or in pretrial negotiations (Spier, 1992) in which adding an outside option creates a multiplicity of equilibria.[7]

It is easy to see that this property generalizes to the protocols considered above. When the uninformed party makes all the offers, the separating equilibrium vanishes because the strong type does not need to wait until the second period to precipitate conflict. Hence, the game collapses into one with a take-it-or leave offer: Depending on the parameters, the uninformed party either makes a pooling offer or one that only the weak types accepts. Sanchez-Pages (2009) shows that the same happens in the game in which the informed party makes all the offers. The strong type does not need to make two non serious offers to trigger conflict; he does it immediately unless a pooling offer can leave him better-off. Then agreement or disagreement must be immediate. A further implication of this result is that it is possible to obtain fundamental properties of bargaining situations in the shadow of conflict by looking just at the single-

[6] Fey and Ramsay (2009) compare the two under the mechanism design approach and obtain significantly different results.

[7] In a recent article, Leventoglu and Tarar (2008) show that this simplicity disappears if the model is extended by considering that war can happen at any point of the game and not only after offers are made. They obtain qualitatively different equilibria, including one in which war does not happen and several offers are exchanged, a feature that is commonly not present in models of bargaining under the shadow of conflict.

offer case. The papers by Brito and Intriligator (1985) and Fearon (1995) in international conflict or by Bebchuk (1984) and Schweizer (1989) on pretrial negotiations can thus offer important insights despite the apparent simplicity of their single-offer setup.

A second important feature of Powell (1996) is that there exists a previous status-quo that parties aim to revise. This assumption is very common in the international relations literature because in reality states tend to bargain over pre-existing issues. This status-quo defines who is the dissatisfied agent at any point in time and ultimately determines who makes the take-it-or-leave-it offer and the likelihood of conflict. But the existence of a status-quo allocation can have deeper implications on the bargaining process. The fact that a party triggers a crisis can convey information. Let us go back to the mechanism design approach: Banks (1990) briefly considered the idea of the informed party having a status-quo payoff. Resorting to a mechanism, or to a bargaining protocol, to solve a dispute means that that party expects to get from doing so at least what he obtains from the status-quo. And that in turn should affect the prior beliefs of the uninformed party in case negotiations commence. This corresponds to the "selection bias" observed in the political science literature (Morrow, 1989). Given that, as we have seen in the models discussed above, stronger parties enjoy greater benefits from bargaining, conditional on a crisis occurring, posterior beliefs of the uninformed party should put more weight on stronger types.

6 Learning by fighting

6.1 Limited conflict as part of the bargaining process

So far we have approached conflict purely as a fight to the finish that takes place once parties fail to reach an agreement completely. This approach however is very limited. Evidence shows that parties continue negotiating after a clash starts and that most wars end with a negotiated settlement (Pillar 1983, Kecskemeti, 1958). Assuming that conflict is just a game-ending move, a final outcome that precludes any further interaction, limits the analysis to the initiation of conflict and leaves out the study of how and why conflicts end. But, most importantly, it neglects another source of learning under incomplete information: conflict itself. When conflict is final no further learning can occur after it starts. But if we admit that bargaining can continue after conflict takes place then, we are allowing parties to learn new information by fighting.

The critical element here is to recognize that parties can choose the scope of the conflicts they fight. An all-out conflict is a possibility, but not the only one. Countries, political parties, married couples, unions and litigants engage in limited confrontations that do not preclude the continuation of negotiations. Skirmishes and battles, family arguments, holdouts or discovery procedures constitute limited forms of confrontation in war, marriages, labour negotiations and legal disputes respectively. After these clashes, negotiations over territories, chores, wages and compensations can resume. This distinction between all-out

and limited conflicts was first noticed by Carl von Clausewitz in his seminal book *On War* (1832). He coined the term "Absolute war" to refer to duels in which parties tried to disarm each other completely so they could impose their will on their rival without opposition. And he employed the term "Real war" to refer to those restrained confrontations (Schelling, 1966) that do not preclude a later agreement because parties "employ no more strength than is absolutely necessary" and that are fought to alter the perceptions and beliefs of the enemy.

This distinction is vital when incomplete information is present because "real" or limited conflicts reveal information about the power and resolve of the parties involved. It was the great German sociologist Georg Simmel who seventy years after Clausewitz, took further these ideas and in his work *The Sociology of Conflict* (1904, p. 501) pointed out the following: "the most effective prerequisite for preventing struggle, the exact knowledge of the comparative strength of the two parties, is very often attainable only by the actual fighting out of the conflict". Paradoxically, the occurrence of conflict is at the same time a necessary conclusion and a solution to the problem of optimism. Again seventy years later, Blainey (1973) pursued this provocative idea and asserted that "war itself provides the stinging ice of reality", because it helps to dissolve optimism and conflicting expectations about the outcome of war.

The idea of two different levels of conflict had been considered in the Economics literature by Cramton and Tracy (1992), who presented a model in which unions could choose the intensity of the labour dispute by opting between strikes and holdouts. However, it was Wagner (2000) who first put together all these ideas within the framework of bargaining with incomplete information. If the outcome of limited conflicts is related to the outcome of absolute ones via strengths, he said, then fighting will change parties beliefs about the outcome of the potential all-out conflict. In consequence, disputants will be able to modify their demands during the course of the negotiation according to the events in the battlefield. In this sense, conflict is part of the bargaining process.

Limited confrontations create a type of information substantially different from the one transmitted through offers and rejections. The latter type of information is "soft" because it is highly manipulable and that in turn allows beliefs to generate multiple equilibria. The battlefield, on the other hand, provides noisy but non-manipulable information: Two types can behave in exactly the same manner but the events on the battlefield will discriminate between them. For instance, when the uninformed party receives a pooling offer, his posterior beliefs should remain unchanged. But if rejection triggers a battle, its outcome will induce him to revise his beliefs. Hence, models that incorporate both offers and limited confrontations allow parties to face conflicting bits of information; a rejected offer may convey strength but a defeat in a skirmish signals weakness.

The seminal work by Wagner spawned a second generation of bargaining models in which conflict was no longer necessarily final and in which parties used the information generated in the battlefield. In most of these models, limited conflicts or *battles* are a by-product of disagreement because they only take place after the rejection of an offer.

Filson and Werner (2002) explored a model of screening very similar to the

one considered above with the addition that a battle takes place after rejection of the first offer. The probability of each agent winning that conflict depends upon the type of the informed player. Hence, after a rejection, the uninformed party learns information also through the outcome of the limited conflict. A victory makes the uninformed party more demanding next period whereas a defeat reduces his demands. The equilibrium profile of the game consists of an increasing schedule of offers and a battle. Hence, both conflict and a peaceful schedule of offers coexist along the equilibrium path. Apart from that, this equilibrium displays qualitatively the same features as the one in the model without the battle: an increase in the prior probability that Player 2 attaches to the strong type makes agreement more likely.

On the other hand, Smith and Stam (2003) developed an infinite horizon model in which information transmission only occurs in the battlefield. They proposed a random-walk model of conflict in which multiple battles are possible. These skirmishes enable parties to capture "forts" -landmarks, resources- such that they can totally defeat their opponent by capturing all of them. The authors then focus on the process of fighting and how limited confrontations shape agreements and the duration of hostilities. This formulation provides a simple characterization of optimal behavior, simply based on the number of forts hold at a particular time. This is because the authors assume that both parties have complete information about their probabilities of prevailing in each of this small conflicts but their estimates are heterogenous. That is, they hold non-common priors or, in Aumann's (1976) words, they "agree to disagree" on the balance of strengths. That simplifies the problem generated by "soft" information and ultimately makes the model very mechanistic. But, most importantly, this common knowledge of non-common prior constitutes a very problematic assumption and it is ultimately incompatible with full rationality. However, it helps the authors to generate a very important observation: parties keep triggering battles as long as the continuation value of doing so exceeds the one from settling, *independently of whether optimism is present or not.* Realistic offers are postponed until both parties find that the amount they expect to extract from their opponent by fighting an additional battle is offset by its costs[8]. We will elaborate more on this next.

6.2 Limited conflict as a bargaining tool

Assuming that limited conflicts are a by-product of disagreement is certainly a step forward. But it fails to asses fully the role of conflict in negotiations. In the models just mentioned, the impact of confrontation is limited to the computation of the offers that make the opponent indifferent between acceptance

[8] Slantchev (2003) extended Smith and Stam (2003) by assuming common priors. He employed an alternating-offers protocol so the informed party can screen her opponent and the uninformed signals her type through non-serious offers. Hence, information transmission occurs both in the battlefield and in the bargaining table. The author shows that when players are sufficiently patient the uninformed party engages in costly separation of types in the equilibrium of the game.

and rejection. Leventoglu and Tarar (2008) argue that this formulation is unsatisfactory because it still builds in the *risk-return trade-off* that characterize the models in which conflict is final. The uninformed party engages in screening depending on the risk of losing the limited conflict, that in turn depends upon the degree of optimism. These models can thus account for the simultaneous occurrence of conflict and diplomacy and for the occurrence of final conflicts, but do not generate equilibrium paths along which offers are peacefully made, simply because conflict is assumed to take place after an offer is rejected.

However, there is a second and more important reason to consider limited conflicts as an option in itself. When limited conflict is just a by-product of disagreement a convergence in the estimates of the balance of strengths can be enough to ensure agreement. This was certainly the view held by Blainey (1973). He claimed that "wars usually cease when the fighting nations agree on their relative strength". If that is the case, a limited conflict cannot lead to inefficiencies that were not present when it was not available. However, as Clausewitz (1832) and Wittman (1979) later pointed out, if parties can use conflict to screen their opponents or alter their expectations about the balance of power, the incentives to use confrontation as a bargaining tool may remain even when optimism has vanished because parties can use confrontation to extract better terms from the opponent. It is the fact that, paraphrasing Clausewitz, conflict is the continuation of diplomacy with the addition of other means, what ultimately explains the pervasiveness of conflict in negotiations.

Next, we will develop two simple models that extend the ones considered in the previous Section by assuming that the party that makes the offers can also trigger a limited conflict. Depending on the information structure, this limited confrontation becomes either a screening device or a credible signal. We then use the result obtained to shed light on some of the issues discussed so far.

6.2.1 Conflict as a screening device

Let us come back to the model of screening that we discussed in the previous Section: We will now assume that the uninformed party has two available actions at $t = 1$: He can make an offer x_1 or trigger the battle, in which case the game proceeds to $t = 2$.. The battle is a confrontation of limited scope that, contrary to the all-out conflict, does not entail the end of the game. We abstract from a specific interpretation of the battle; nature simply announces a winner and the second period is reached. Then, standard bargaining (offers) can take place both before and after hostilities have begun.

The outcome of this battle is related to the outcome of the all-out conflict via players' strengths. For simplicity, we will assume that the conflict technology is identical across the two types of confrontations. But that is not necessary. All that is needed is a correlation between the two, even a negative one[9]. Thus we will assume that Player 1's probability of winning the battle is p_L or p_H depending on his type.

[9] In military strategy it is often argued that the likelihood of success in guerrilla warfare is inversely related to the likelihood of victory in an open battlefield.

After the battle, two new information sets emerge depending on its outcome from Player 1's perspective: *Victory* (V) or *Defeat* (D). Therefore, the outcome of the battle conveys information about Player 1's type. Player 2, the one making the offers, will revise his beliefs taking this into account.

A history of the game at $t = 2$ can now consist of a rejected offer or the outcome of the battle. In the first case, Player 2's beliefs are identical to the ones in (6). In case a battle takes place beliefs are:

$$\begin{aligned} \mu(H \quad | \quad V) \equiv \mu^+ &= \frac{\mu_o p_H}{\mu_o p_H + (1-\mu_o)p_L}, \\ \mu(H \quad | \quad D) \equiv \mu^- &= \frac{\mu_o(1-p_H)}{\mu_o(1-p_H) + (1-\mu_o)(1-p_L)}. \end{aligned} \tag{8}$$

A straightforward comparison with (6) shows that a victory (defeat) of Player 1 in the battle makes Player 2 believe he is more (less) likely to be facing a strong opponent. Hence, a pooling offer may emerge under one or the two outcomes. But given that $\mu^+ > \mu^-$, a pooling offer will take place in both scenarios if and only if.

$$\mu^- \geq \frac{p_H - p_L}{R + p_H - p_L} \Leftrightarrow R \geq \frac{\widehat{R}}{\mu_o} \frac{1-p_L}{1-p_H}.$$

When the Loss ratio is sufficiently high, Player 2 prefers not to risk conflict by making an offer that will be rejected by the strong type. Otherwise, a separating offer can occur under V or in both scenarios. This is intuitive: Under V, Player 2 becomes more pessimistic about his prospects in case of an all-out conflict and is ready to make better offers. On the other hand, under D, he becomes more optimistic and thus more demanding, so conflict may become inevitable.

On the other hand, the case in which offers under both scenarios are separating occurs if and only if

$$\mu^+ \leq \frac{p_H - p_L}{R + p_H - p_L} \Leftrightarrow R \leq \frac{\widehat{R}}{\mu_o} \frac{p_L}{p_H},$$

and pooling under V and separation under D occurs in the remaining case.

Alternatively, Player 2 can just make an offer in period one. That will generate strategy profiles identical to the ones described in the model without the battle. Player 2 will choose the option that maximizes his payoff. Note then that he model still gives room to a peaceful exchanges of offers. It is straightforward to show that if the battle generates pooling or separation under both outcomes, triggering it is a dominated action. It can only benefit Player 2 in the intermediate case, when it helps him to screen his opponent. The next result characterizes conditions under which that constitutes an equilibrium. It does it in terms of the discount rate, that now has a natural interpretation: The cost of the battle. So the battle can only be part of an equilibrium profile if its cost is not too high.

Proposition 5 *In the game in which the uninformed party makes all the offers and the battle is available, triggering the battle*

(*i*) *Dominates a pooling offer if and only if*

$$\delta \geq \delta'_{\mu_o} \equiv \frac{1}{1+\widehat{R}(1-p_L)-R\mu_o(1-p_H)}.$$

(*ii*) *Dominates a separating profile of offers if an only if*

$$\delta \geq \delta''_{\mu_o} \equiv \frac{1+R}{1+R+\frac{\mu_o}{1-\mu_o}\frac{Rp_H-p_L\frac{\widehat{R}}{\mu_o}}{1-\mu_o\frac{R}{\widehat{R}}}}.$$

Moreover, the threshold δ'_{μ_o} (δ''_{μ_o}) *is increasing (decreasing) in* μ_o.

These two thresholds move in opposite ways as the probability that Player 2 attaches to Player 1 being strong increases. When Player 2 becomes more pessimistic, the battle becomes less attractive compared to the option of pooling, because an all-out conflict against a strong opponent becomes more likely. On the other hand, the battle becomes more attractive compared to a separating profile of offers because Player 2 can use it to screen his opponent. It allows him to avoid an all-out conflict in case Player 1 wins. Note that this latter effect occurs in sharp contrast with Filson and Werner (2002), where the battle was just a by-product of rejected offers and where an increase in the prior probability that Player 2 attaches to the strong type always made immediate agreement more likely.

But the most important result is that the limited conflict displays a double-edged effect. On the one hand, the battle has a potential efficiency enhancing effect. When the loss ratio is moderately low, i.e. $R \in [\frac{\widehat{R}}{\mu_o}\frac{p_L}{p_H}, \frac{\widehat{R}}{\mu_o}]$ and players are sufficiently patient, Player 2 triggers the battle. In that case, the all out conflict only occurs with probability $\mu_o(1-p_H)$ because if Player 1 wins, Player 2 becomes less optimistic about facing a weak opponent and that prompts him to make a high offer. Recall that in that area of the parameter space, Player 2 used a separating profile of offers when the battle was not available and then the all-out conflict occurred whenever Player 1 was of the strong type, i.e. with probability μ_o. Hence, the use of battling as a screening device has a potentially positive effect on efficiency.

However, the opposite is also true. Recall from Proposition 3 that a sufficient condition for immediate agreement was $R > \frac{\widehat{R}}{\mu_o}$. This is no longer the case. If Player 2 is sufficiently patient, i.e $\delta \geq \delta'_{\mu_o}$, he will use the battle to sort out types. He precipitates the limited conflict as a way to gain new information on his opponent and extract from him more surplus in case he manages to defeat him. This has a negative impact on efficiency. The possibility of using conflict as a screening device introduces delay when agreement was previously immediate. Contrary to the assertion by Schelling (1960), limited conflict does not need to increase the risk of an all-out one: the final effect depends on the outcome of such confrontation and on the incentives of parties to use it a bargaining tool.

6.2.2 Conflict as a credible signal

Let us now add to the model in which the informed party makes all the offers the option of fighting a battle in the first period. Its outcome conveys "hard" information on the type of Player 1 and will make Player 2 revise his beliefs. In that case, the informed party will use the limited conflict as a way to signal his strength. Again, the analysis of this signaling model is a modified version of the one in Sanchez-Pages (2009).

The informed party has now two available actions at $t = 1$: He can make an offer x_1 or trigger the battle, in which case the game proceeds to $t = 2$. In that period, Player 1 makes the second and final offer x_2. Rejection of the second offer triggers the all-out conflict.

We now explore the conditions under which the profile in which both types trigger the battle at $t = 1$ constitute a PBE of this game. It is straightforward to show that again it cannot be the case that one type settles first and the second type settles at the second period. Both types will either make an offer or trigger the battle. In the latter case, after observing its outcome, beliefs are identical to the ones in (8). This defines two possible pooling offers as a function of the result of the skirmish

$$\widetilde{x}_2^V = 1 - \theta(1 - \mu^+ p_H - (1 - \mu^+)p_L),$$

$$\widetilde{x}_2^D = 1 - \theta(1 - \mu^- p_H - (1 - \mu^-)p_L),$$

where clearly $\widetilde{x}_2^V < \widetilde{x}_2^D$: After a victory, Player 1 needs to offer less to a now more pessimistic Player 2 in order to obtain his acceptance. Against this, the H type has the option of making a non serious offer in the second period and trigger the all-out conflict. It is clear that if the pooling offer leaves him worse off compared to that option in case of victory, i.e. if

$$R \leq \widehat{R}\frac{1 - \mu^+}{1 - \mu_o},$$

conflict will ensue under both outcomes. By the same token, if under defeat, the strong type of Player 1 still prefers to make the offer $\widetilde{x}_2^D$, that is

$$R \geq \widehat{R}\frac{1 - \mu^-}{1 - \mu_o},$$

pooling under both outcomes will prevail. In the intermediate case, pooling occurs only under V.

Finally, consider the same set of beliefs (7) that we employed in the case without the battle. Then it is straightforward to see that if parameters are such that separation occurs under both outcomes, no type has any incentive to trigger the battle. The information conveyed by the battle is not enough to alter Player 2's belief significantly. Then, depending on its type, Player 1 is better off by settling immediately or triggering the all-out conflict with two non-serious offers. Things are different when pooling ensues under one or both

outcomes. In that case, Player 1 can gain extra concessions by fighting. Again, the discount factor plays a crucial role in the characterization of this case.

Proposition 6 *In the game in which the informed party makes all offers, the battle is available and the uninformed agent holds beliefs (7), there exists a PBE of the game in which both types fight the battle*

(*i*) *For relatively low values of the loss ratio, i.e.* $R \in [\widehat{R}\frac{1-\mu^+}{1-\mu_0}, \widehat{R}\frac{1-\mu^-}{1-\mu_0}]$, *when*

$$\delta \geq \delta' \equiv \frac{1+R}{1+R+p_L\mu^+\frac{\widehat{R}}{1-\mu_0}}.$$

(*ii*) *For higher values of the loss ratio, i.e.* $R \geq \widehat{R}\frac{1-\mu^-}{1-\mu_0}$, *when*

$$\delta \geq \delta'' \equiv \frac{1+R}{1+R+(p_L\mu^+ + (1-p_L)\mu^-)\frac{\widehat{R}}{1-\mu_0}}.$$

Moreover, both δ' *and* δ'' *are decreasing in* μ_0.

Contrary to the screening case, an increase in the probability that Player 2 attaches to Player 1 being strong has an unambiguous effect: It makes the battle more likely. This is because when μ_0 increases the pooling offer increases as well. Player 2 is less optimistic and accepts lower offers. That makes the strong type of Player 1 to resort to the battle and not to the all-out conflict.

Again, the battle can open the door to agreement in situations in which it was not possible without it. When the battle is not too costly, the profile in which both players trigger the battle dominates separation through offers. The all-out conflict occurs only if the strong type loses the battle and therefore inefficiencies are reduced. But, again, it may be the case that the informed party triggers the battle as a way to obtain a concession when immediate agreement was feasible when it was not available.

This result helps us to shed light on the *Uneven contenders paradox* that we outlined above. We argued that the fact that weak agents fight much stronger ones is at odds with the observation that optimism should be of little importance when the balance of observable strengths is very uneven. We conclude that this phenomenon takes places because weak parties use conflict as a way to alter the expectations of their opponents and to obtain more concessions from then. That can occur even if optimism is not present. That is the reason why these conflicts are rarely final: Guerilla warfare against a superpower or political demonstrations against a repressive regime do not aim to defeat the opponents in an absolute sense. They are just a form of tacit bargaining that weak parties use to obtain better terms via a change in their opponents' beliefs. Limited conflicts provides them with a way to signal privately known and unverifiable information about their strength.

6.3 Discussion and extensions

Let us now revisit the idea of optimism as a cause of conflict. There exists the view that at the end of many conflicts, specially military ones, expectations are not longer divergent and that is what necessarily leads to peace. The results above show that this observation is incomplete because it only looks at the starting and ending points of a conflict and not at how it is conducted. It is true that optimism is a necessary condition for conflict and that, therefore, confrontations can terminate only if parties' demands are compatible. But we have seen that the latter does not imply that a convergence in expectations ensures a settlement. This can be uncovered only if we abandon the idea that conflict is just a by-product of disagreement and assume that it is a tool that parties will choose to use as long as it is profitable.

One potential problem of the models discussed in this Section is that they employed a very restrictive protocol. Although they can offer important insights on the dynamics of conflict and negotiation, they are underscored by the limitations of their bargaining settings. Some other contributions have relaxed these assumptions. However, as we will see, a slightly more general analysis soon becomes very complex.

Powell (2004) extends Filson and Werner (2002) to an environment with an infinite horizon. The author departs from the modelling of incomplete information that we have considered so far. He adapts the bargaining model with inside options by Muthoo (1999) in which bargaining can collapse each period and the probability of breakdown is private information to one of the parties. Hence, the uninformed party updates his beliefs with rejections and with the fact that the uninformed party did not collapse. The difference with the initial model of screening we considered in Section 5.1 is that the passage of time does convey information. This comes at the cost of reducing the importance of limited fighting. Because a collapse did not happen, the uninformed party can only become more pessimistic after fighting. The actual outcome of the skirmish does not transmit any information. This shifts the emphasis of bargaining from the battlefield to the negotiation table. Because of that, equilibria display very similar properties to the ones on standard bilateral trade.

In a similar spirit to Smith and Stam (2003), Sanchez-Pages (2004) considers multiple battles but limits information transmission by assuming that the uninformed party is *unsophisticated* and does not make inference from offers he receives from the informed one. In this case, the proponent's strategy becomes an optimal stopping problem. His equilibrium strategy consists of two sequences of integers such that if the number of victories above the highest of the two, the informed player makes an acceptable offer, if it steps below the lowest one, triggers the all-out conflict, and keeps battling otherwise. The author shows that in this context, the stronger the informed player the more likely he is to use limited confrontations because he owns a better "persuasion device".

Finally, Ponsati (2006) studies the role of commitment. The author explores a Markov game in which bargaining states can give place to states in which one of the parties commits fully to a particular demand and transitions between

states are stochastic. This introduces frictions in the bargaining process but at the same time generates different types of equilibrium paths, including one with the appealing property of involving long confrontations and in which demands rejected initially are finally accepted.

7 Empirical evidence

Given that the literature on bargaining and conflict is relatively recent, only a few works have tried to examine data on real conflicts in the light of the models described above. As we have seen, these models tend to be complex and their estimation is difficult. There is for instance a fundamental problem with beliefs. They play a central role in the presence of incomplete information but they are unobservable, and it is not clear at all how could they be proxied. As argued by Morrow (1989), a possible way out of this problem is to develop models that allow us to derive testable empirical implications. Because such models already incorporate beliefs in them, the problem of misspecification can be ameliorated.

There exist two main avenues for empirical research in this context: On the one hand, the information provided by limited confrontations should affect the termination of conflicts and generate certain patterns of duration. On the other hand, there is the vital issue of conflict initiation; some disputes become full-fledged conflicts whereas others remain in the state of peaceful negotiations.

The empirical study of the duration and termination of conflicts can be carried out with the tools of duration analysis. This technique looks at the hazard rate of conflicts. The hazard rate is the probability of termination of an event (a war, a marriage, a strike) conditional on it having lasted a certain amount of time. When the hazard rate increases (decreases) in time, then the event is said to exhibit positive (negative) duration dependence.

The initial hypothesis is that if conflict is part of the bargaining process and it can be used as a negotiation instrument then this should have an impact on the duration patterns of real conflicts. For instance, the screening model of bargaining gives an indication of how these patterns should be. Since weaker types tend to settle first and offers increase with time, the a conflict should be more likely to end the more it lasts. This led Filson and Werner (2002) to conjecture that wars should display substantive positive duration dependence. There may be some other reasons for this. As pointed out by Sanchez-Pages (2009), one property of Bayesian updating is that the change in beliefs induced by limited conflicts should be decreasing in the number of skirmishes fought. Then, the expected concession obtained from fighting must be decreasing in time. When players are impatient, that implies that the use of conflict as a bargaining instrument must be a self-limiting phenomenon.

Filson and Werner (2002) in fact argued that the hazard rate of wars may be U-shaped. That could explain why previous studies on wars that only focused on monotonic hazard rates found no duration dependence (Bennett and Stam, 1996 and Vuchinich and Teachman, 1993)[10]. Sanchez-Pages (2009) proposes an-

[10] This contrasts with the negative duration dependence found in contract negotiations (Ken-

other explanation for the observed lack of duration dependence: these models did not distinguish between termination modes. In some of these conflicts incomplete information may have played no role. For instance, some colonial wars were treated by European powers as internal rebellions that only deserved to be crushed and therefore compromising was out of the question. The termination mode can gives us clues on this: If a conflict ended in a negotiated settlement then it was more likely to have a significant bargaining component. Of course, this is just a proxy, because a total collapse is a possibility in most conflicts. Hence, bargaining could also have been an important issue in some confrontations that ended in an absolute way. Still, Sanchez-Pages (2009) shows that colonial and imperial wars that ended in the complete annihilation of one of the parties displayed a flat hazard rate, whereas those that ended in a negotiated settlement display positive duration dependence.

However, this type of duration analysis assumes that there exists an underlying battling process that makes the passage of time to reveal information. But that does not need to be the case. Some confrontations may contain prolonged periods of inactivity whereas in others hostilities may be constant and intense. There exist two solutions to this problem. One is to consider specific case studies and analyze how participants' estimates changed over the course of the conflict. This is the route taken by Goemans (2000), who in an impressive study of the First World War, proves that the Germans designed the attack at Verdun to influence France's estimate of its own relative strength.

The second possibility is to disaggregate the data and look at the battlefield level. This approach is potentially very powerful but it is plagued with problems on the availability and accuracy of data. Ramsay (2008) uses data on battle days for a restricted set of interstate wars, to estimate the effect of the information transmitted through limited confrontations on the probability of war termination. The author finds that this effect is positive during the first 40 days of the conflict but vanishes after that, a finding that seems to run against the informative role of confrontations. On the other hand, Slantchev (2004) found that the events in the battlefield, that he proxied with the ratio of casualties between the contenders, overcome quickly the effect of observable capabilities on duration. He also obtains that terms of settlement worsen for war initiators as conflicts proceed, a result that is in line with the screening hypothesis.

If longer conflicts and more intense fighting induce a stronger convergence of beliefs, the resulting settlements should be more stable and the probability of further conflicts should be lower. This leads to some testable implications on the likelihood of new conflicts. Schnell and Gramm (1982) show that lagged strike experience reduces the propensity of unions to strike again. On the other hand, Box-Steffensmeier et al. (2003) obtained that peace is more fragile after a war that ended in a stalemate than after a war that ended in a decisive victory, and that this effect vanishes with time. That may explain why Sanchez-Pages (2009) obtains that the number of previous disputes between two contenders does not have an impact on the hazard rate of imperial and colonial wars.

nan and Wilson, 1989).

The second issue that the empirical research on bargaining and conflict faces is the problem of selection bias. Duration analysis cannot tell us anything about why certain disputes turn into conflicts whereas others do not. As argued by Morrow (1989), the set of conflicts observed cannot be treated just as a random sample. This selection bias is well known in the literature on legal disputes because only a small fraction of cases end up in trial (Wittman, 1988). As we observed above, stronger types are more likely to initiate disputes because they are the ones that have most to gain from that. This, of course, compounds with the problem of misspecification stemming from the unobservability of beliefs. Hence, the observed conflicts are a subset of cases with specific balances of strengths and distributions of beliefs. Bueno de Mesquita et al. (1997) developed a model that looked into dispute initiation and estimated a U-shaped effect of observed capabilities on the probability of conflict between two states. They obtain that a peaceful settlement is more likely when capabilities are similar. This may explain the findings by Slantchev (2004), who obtains that wars between more similar adversaries (in terms of observable capabilities) tend to last longer. If parties know this, then they should be more reluctant to enter into a conflict and more willing to settle their differences peacefully.

8 Conclusions and further research

The results and models described in this chapter offer a comprehensive explanation for the pervasiveness of confrontations in negotiations that puzzled Hicks so much. Once we acknowledge that conflict itself is part of the bargaining process and that it can become a screening or a signaling device, we need to restate his assertion: A mutually beneficial agreement always exists, but if at least one of the parties beliefs that he can improve his position in the bargaining table by fighting then peace cannot be attained. A settlement will only take place when all the parties involved find that the returns from standard diplomacy outweigh the returns of using conflict as a bargaining tool.

Limited conflicts play a role that other mechanisms could also potentially fulfill. The necessary conditions are that its outcome must be related in some way to the outcome of a potential final confrontation and that it must be public. It does not even need to be costly. As a matter of fact, it does not even need to be fought against the other bargaining party. For instance, the defeat of a colonial power against a colony will convey information to other dependencies also engaged in a self-determination fight. Or the conduct of a strike in one particular plant may transmit information on the profitability of the firm to workers in other plants, as in Gu and Kuhn (1998). This information could also be transmitted unilaterally, by, for instance, launching long range missiles into the sea, as North Korea has being doing in recent years. In that sense, the use of limited conflicts is related to the concept of "verifiable disclosure", in which sellers provide information to buyers that external institutions or buyers themselves can check in order to induce them to buy their products[11]. The main

[11] For a survey on this literature see Milgrom (2008).

difference between these two types of information is that conflict is typically noisy (as reflected by the parameter γ in the formulation (1)). It is a gamble that parties will sometimes find profitable to take.

In the analysis above, we have abstracted from the physical meaning of limited conflicts. They were mostly instrumental. This comes at the price of ignoring the non-informational gains and losses that limited confrontations can generate. Smith and Stam (2003) and Slantchev (2003) assume that parties have a finite quantity of forts because they implicitly assume that parties can sustain only a limited number of losses before collapsing. One possible avenue for further research could make explicit assumptions on the meaning of the battles. That includes considering the possibility that the probability of victory an additional battle changes with the previous events in the battlefield. However, without a specific application in mind it is difficult to see how this could be implemented. Sometimes an initial setback precipitates the collapse of the loser but others it increases his conflict effort, like the case of the Allies in World War II. This may explain the conflicting results obtained in the duration analysis of wars.

New bargaining models should allow us a direct computation of the hazard rate. An open question is whether the type of uncertainty that parties face may have an effect on the duration dependence to be found. If uncertainty is over costs (as in Powell, 2004) or over reservation stands (as in Heifetz and Segev, 2005) it is the passage of time and not battles what provides information, so there is an extra incentive to wait as a way to learn. In short, another further avenue of research is the development of bargaining models with directly testable implications on duration patterns.

References

[1] Admati, A.R. and Perry, M. (1987). "Strategic Delay in Bargaining," *Review of Economic Studies*, 54(3): 345-64.

[2] Aumann, R.J. (1976). "Agreeing to Disagree," *The Annals of Statistics* 4(6): 1236-39.

[3] Ausubel, L.M., Cramton, P.C. and Deneckere, R.J. (2002). "Bargaining with incomplete information," in (R. J. Aumann and S. Hart, eds.), *Handbook of Game Theory vol. 3*, Amsterdam: Elsevier Science.

[4] Banks, J.S. (1990). "Equilibrium behavior in crisis bargaining games," *American Journal of Political Science*, 34(3): 599-614.

[5] Bebchuck, L.A. (1984). "Litigation and Settlement under Imperfect Information," *Rand Journal of Economics*, 15(3): 404-15.

[6] Bennett, D.S. and Stam, A.C. (1996). "The duration of interstate wars, 1816-1985," *American Political Science Review*, 90(2): 239-257.

[7] Bester, H. and Wärneryd, K. (2006). "Conflict and the social contract," *Scandinavian Journal of Economics*, 108(2): 231-249.

[8] Blainey, G. 1973. *The Causes of War*. New York: Free Press.

[9] Box-Steffensmeier, J.M., Reiter, D. and Zorn, C.J.W. (2003). "Nonproportional Hazards and Event History Analysis in International Relations," *Journal of Conflict Resolution*, 47: 33-53.

[10] Brito D.L. and Intriligator, M.D. (1985). "Conflict, war, and redistribution," *American Political Science Review*, 79: 943-957.

[11] Bueno de Mesquita, B., Morrow, J. and Zorick, E. (1997). "Capabilities, Perception and Escalation," *American Political Science Review*, 91(1): 15-27.

[12] Claude, I.L. (1962). *Power and International Relations*. New York: Random House.

[13] Clausewitz, C.v. (1932)[1976]. *On war.* Princeton: Princeton University Press.

[14] Coase, R.H. (1972). "Durability and Monopoly," *Journal of Law and Economics*, 15(l): 143-49.

[15] Cramton, P.C. and Tracy, J.S. (1992) "Strikes and holdouts in wage bargaining: Theory and data," *The American Economic Review*, 82(1): 100–121.

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

[17] Fey, M. and Ramsay, K.W. (2007). "Mutual Optimism and War," *American Journal of Political Science,* 51(4): 738-754.

[18] Fey, M. and Ramsay, K.W. (2009). "Uncertainty and Incentives in Crisis Bargaining: Game-Free Analysis of International Conflict." Unpublished manuscript.

[19] Filson, D. and Werner, S. (2002). "A Bargaining Model of War and Peace," *American Journal of Political Science,* 46(4): 819–838.

[20] Fudenberg, D., Levine, D., and Tirole, J. (1987). "Incomplete Information Bargaining with Outside Opportunities," *Quarterly Journal of Econonomics*, 102: 37–50.

[21] Fudenberg, D. and Tirole, J. (1983). "Sequential Bargaining with Incomplete Information about Preferences," *Review of Economic Studies*, 50(2): 221-47.

[22] Goemans, H. E. (2000). *War and punishment: The causes of war termination and the First World War.* Princeton: Princeton University Press.

[23] Gu, W. and Kuhn, P. (1998). "A Theory of Holdouts in Wage Bargaining," *American Economic Review*, 88(3): 428-49.

[24] Harsanyi, J. (1967-68). "Games with Incomplete Information Played by Bayesian Players," *Management Science,* 14:159-182, 320-334, 486-502.

[25] Heifetz, A. and Segev, E. (2005). "Escalation and Delay in Protracted International Conflicts," *Mathematical Social Sciences*, 49: 17-37.

[26] Hicks, J. (1932). *The theory of wages.* NY: Macmillan.

[27] Hiller, E.T. (1933). *Principles of sociology.* New York and London: Harper & Brothers.

[28] Hirshleifer, J. (1991). "The Technology of Conflict as an Economic Activity," *The American Economic Review*, 81(2): 130-134.

[29] Jackson, M.O., and Morelli, M. (2007). "Political Bias and War," *American Economic Review*, 97(4): 1353–1373.

[30] Johnson, D. (2004). *Overconfidence and War: The Havoc and Glory of Positive Illusions.* Harvard: Harvard University Press.

[31] Kennan, J. and Wilson, R. (1989). "Strategic bargaining models and interpretation of strike data," *Journal of Applied Econometrics*, 4: S87-S130.

[32] Kennan, J. and Wilson, R. (1993). "Bargaining with private information," *Journal of Economic Literature*, 31: 45-104.

[33] Leventoglu, B. and Slantchev, B (2007). "The Armed Peace: A Punctuated Equilibrium Theory of War," *American Journal of Political Science*, 51(4): 755-771.

[34] Leventoglu, B. and Tarar, A. (2008). "Does Private Information Lead to Delay or War in Crisis Bargaining?," *International Studies Quarterly*, 52(3): 533-553.

[35] Milgrom, P. (2008). "What the Seller Won't Tell You: Persuasion and Disclosure in Markets," *Journal of Economic Perspectives*, 22(2): 115-131.

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

[37] Muthoo, A. (1999). *Bargaining Theory with Applications.* New York: Cambridge University Press.

[38] Myerson, R.B. (1979). "Incentive Compatibility and the Bargaining Problem," *Econometrica* 47(1): 61-73.

[39] Nash, J. F. (1950). "The bargaining problem," *Econometrica*, 18(2): 155-62.

[40] Myerson, R.B. and Satterthwaite, M.A. (1983). "Efficient Mechansims for Bilateral Trading," *Journal of Economic Theory*, 29(2): 265-281.

[41] Organski, A. F. K. and Kugler, J. (1980). *The War Ledger*. Chicago: University of Chicago Press.

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

[43] Ponsati, C. (2006). "The continuation of bargaining by other means: an elementary theory of war." Unpublished manuscript.

[44] Powell, R. (1996). "Bargaining in the Shadow of Power," *Games and Economic Behavior,* 15: 255–289.

[45] Powell, R. (2004). "Bargaining and Learning While Fighting," *American Journal of Political Science,* 48(2): 344–361.

[46] Ramsay, K.W. (2008). "Settling it on the Field: Battlefield Events and War Termination," *Journal of Conflict Resolution*, 52(6): 850-879.

[47] Sanchez-Pages, S. (2003). *On Conflict and Power*. Ph.D dissertation, Universitat Autonoma de Barcelona.

[48] Sanchez-Pages, S. (2004). "The use of conflict as a bargaining tool against unsophisticated opponents." Unpublished manuscript.

[49] Sanchez-Pages, S. (2009). "Conflict as a Part of the Bargaining Process," *Economic Journal*, 119(539): 1189-1207.

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

[51] Schelling, T.C. (1966). *Arms and Influence*. New Haven and London: Yale University Press.

[52] Schnell, J. F. and Gramm, C. L. (1982) "Learning by striking: Estimates of the teetotaler effect," *Journal of Labor Economics*, 5(2): 221-240.

[53] Schweizer, U. (1989). "Litigation and Settlement under Two-Sided Incomplete Information," *Review of Economic Studies*, 56(2): 163-177.

[54] Simmel, G. (1904). "The sociology of conflict I," *American Journal of Sociology*, 9(4): 490-525.

[55] Skaperdas, S. (1996). "Contest success functions," *Economic Theory,* 7: 283-290.

[56] Slantchev, B. (2003). "The Principle of Convergence in Wartime Negotiations," *American Political Science Review,* 97(4): 621–632.

[57] Slantchev, B. (2004). "How Initiators End Their Wars: The Duration of Warfare and the Terms of Peace," *American Journal of Political Science*, 48(4): 813-829.

[58] Smith, A. and Stam, A.C. (2004). "Bargaining and the Nature of War," *Journal of Conflict Resolution,* 48(6): 783–813.

[59] Sobel, J. and Takahashi, I. (1983). "A Multi-Stage Model of Bargaining," *Review of Economic Studies*, 50(3): 411-26.

[60] Spier, K. (1992). "The Dynamics of Pretrial Negotiation," *Review of Economic Studies*, 51(1): 93- 108.

[61] Spier, K. (1994). "Pretrial bargaining and the design of fee-shifting rules," *Rand Journal of Economics*, 25:197-214.

[62] Vuchinich, S. and Teachman, J. (1993). "Influences on the Duration of Wars, Strikes, Riots and Family Arguments," *Journal of Conflict Resolution*, 37:544-568.

[63] Wagner, R. H. (1994). "Peace, war and the balance of power," *American Political Science Review,* 88(3): 593-607.

[64] Wagner, R.H. (2000). "Bargaining and War," *American Journal of Political Science,* 44(3): 469–484.

[65] Wittman, D. (1979). "How a War Ends: A Rational Model Approach," *Journal of Conflict Resolution,* 23(4): 743–763.

[66] Wittman, D. (1988). "Dispute Resolution, Bargaining, and the Selection of Cases for Trial: A Study of the Generation of Biased and Unbiased Data," *The Journal of Legal Studies*, 17(2): 313-352.

[67] Wrangham, R. (1999). "Is Military Incompetence Adaptive?," *Evolution and Human Behavior*, 20: 3-17.