Causes of War

Sandeep Baliga, Northwestern University*
Tomas Sjöström, Rutgers University†

March 19, 2023

Abstract

Even though war is a Pareto dominated outcome, attempts to negotiate a peaceful settlement may be unsuccessful. First, negotiations can be impeded by *opportunistic bargaining tactics.* Attempts to create a *fait accompli* can trigger a war that neither side had intended. Opportunistic bargaining tactics also include misrepresentation of private information. With asymmetric (private) information, incentive compatibility and individual rationality conditions can be inconsistent with peace. Second, war can be unavoidable if peace agreements are *unenforceable*, i.e., vulnerable to opportunistic behavior ex post. With incomplete contracting ability, a security-seeker may attack out of fear that the opponent will strike first (preemptive war) or is becoming too powerful (preventive war).

1 Introduction

Disputes can be settled peacefully or by fighting, i.e., by war. But war destroys assets and creates suffering:

> "Wars are usually Pareto inferior outcomes of a conflict in that both parties would be better off if the expected loser compensated the expected victor by means of a transfer of resources without actually going to war" (Brito and Intriligator, 1985, p. 945).

*Email: baliga@kellogg.northwestern.edu
†Email: tsjostrom@economics.rutgers.edu

Therefore, factors such as the balance of power, the distribution of resources and contested territories are, by themselves, insufficient to explain war. We must also explain why the parties do not solve their disputes peacefully. The *rationalist approach to war* (Fearon, 1995) emphasizes that strategic behavior can impede negotiations, and peace agreements are incomplete contracts which are difficult to enforce. These fundamental obstacles to Pareto optimality – opportunistic behavior and the inability to write complete binding contracts – are familiar to economists, at least since the work of Oliver Williamson (1985). The formal theory of conflict was greatly advanced by the work of James Fearon and Robert Powell. The organization of our survey was influenced in particular by Fearon (1995) and Powell (1999).

Schelling (1960) emphasized that bargaining involves both common and opposing interests: both sides can benefit if the available (social) surplus is large, but each wants a bigger share at the other's expense. Typically, there is a trade-off, and the optimal bargaining strategy involves a calculated risk of war. A player may engage in brinkmanship, taking forceful actions in hopes that the opponent will concede rather than risk a confrontation. For example, a country may send soldiers to a contested territory in order to create a *fait accompli.* A credible commitment not to withdraw from the territory can give the country a strong bargaining position (a first-mover advantage).[1] But if both sides use such tactics, the outcome may be an unintended war, like a car crash in a game of chicken.[2] They cannot negotiate an agreement to ban such tactics, because these tactics are an integral part of the bargaining process itself. This is illustrated by Schelling's haggling metaphor:

> "If each party knows the other's true reservation price, the object is to be first with a firm offer. Complete responsibility rests with the other, who can take it or leave it as he chooses (and who chooses to take it). Bargaining is all over; the commitment (that

[1] After World War II, the Soviet Union occupied Eastern Europe. Its commitment to defend the territory was credible because it could not have retreated without a huge loss of prestige (Schelling, 1960).

[2] In the 1750's, the French and British governments took aggressive actions (sending expeditions, building forts, etc.) in the disputed Ohio River Valley. According to Higonnet (1968), these actions were intended to lay claim to the disputed territory, not to start a war. But British and French forces clashed, and neither side found it possible to back down. There was war, although "neither the French nor the English government wanted to come to blows" (Higonnet, 1968, p. 60).

> is, the first offer) wins. Interpose some communication difficulty. They must bargain by letter; the invocation becomes effective when signed but cannot be known to the other until its arrival. Now when one person writes such a letter the other may already have signed his own or may yet do so before the letter of the first arrives. There is then no sale; both are bound to incompatible positions" (Schelling, 1960, p. 26).

In Section 3 we consider a bargaining procedure based on Crawford (1982), Ellingsen and Miettinen (2008) and Baliga and Sjöström (2020). There is no private information, but the players are unsure about whether commitment attempts will be successful. After eliminating dominated strategies, the procedure can be represented by a 2×2 Hawk-Dove game with strategic substitutes, where Hawk represents a strategic move (i.e., a commitment attempt). If the cost of attempting a commitment is sufficiently small, then there is a unique equilibrium, and war occurs with strictly positive probability. Section 4.1 shows that the same is true if commitment costs can be large but are privately known.

Authors such as Thucydides, Hobbes and Rousseau argued that wars are caused by *fear*:

> "It is quite true that it would be much better for all men to remain always at peace. But so long as there is no security for this, everyone, having no guarantee that he can avoid war, is anxious to begin it at the moment which suits his own interest and so forestall a neighbor, who would not fail to forestall the attack in turn at any moment favorable to himself, so that many wars, even offensive wars, are rather in the nature of unjust precautions for the protection of the assailant's own possessions than a device for seizing those of others" (Rousseau, quoted by Jervis, 1976, p. 63).

If there is a *first-strike (offensive) advantage*, i.e., if the technology of war favors the offense, then a security-seeker, who prefers the status quo to war, may attack preemptively.[3] The decision to attack is caused by fear and distrust, like hunting rabbit in a game of stag hunt. In Section 4.2

[3] World War I is often cited as an example:

we consider a Hawk-Dove game where Hawk represents a decision to start a war. Actions are strategic complements due to the first-strike advantage, and binding peace agreements are infeasible. If it is common knowledge that both players are security-seekers, then there exists a peaceful equilibrium where both players choose Dove. But if the players are uncertain of each other's type, then security-seekers with a sufficiently small cost of choosing Hawk will choose Hawk in self-defense. This creates a "Hobbesian trap", a cascade of security-seekers with higher and higher costs who all choose Hawk. Binding peace treaties can prevent the Hobbesian trap. If peace treaties cannot be enforced, then in certain situations, the trap can be prevented by costly signaling (Kydd, 1997) or by peace talks using cheap messages (Baliga and Sjöström, 2004).

Chatterjee and Samuelson (1983) and Myerson and Satterthwaite (1983) showed how opportunistic behavior in the form of *misrepresentation of private information* can prevent traders from finding a mutually acceptable price. By analogy, diplomats may find it impossible to identify a mutually acceptable compromise. This may lead to war, even if binding agreements are feasible:

> "Indeed one can almost suggest that war is usually the outcome of a diplomatic crisis which cannot be solved because both sides have conflicting estimates of their bargaining power... In peace time the relations between two diplomats are like relations between two merchants. While the merchants trade in copper or transistors, the diplomats' transactions involve boundaries, spheres of influence, commercial concessions and a variety of other issues which they have in common. A foreign minister or diplomat is a merchant who bargains on behalf of his country. He is both buyer and seller, though he buys and sells privileges and obligations rather than commodities. The treaties he signs are simply

---

> "But in July and August, 1914, the primary motivation for the precipitous decisions to mobilize and to launch attacks was the fear of each power that by waiting it would enable the enemy to strike a decisive blow first" (Iklé, 2005, p. 8).

However, Fischer (1967) argued that Germany had aggressive aims of territorial and commercial expansion in both world wars. The ensuing debate over Germany's war aims illustrates the difficulty in discerning true motives.

> more courteous versions of commercial contracts. The difficulty in diplomacy, as in commerce, is to find an acceptable price for the transaction... In diplomacy each nation has the rough equivalent of a selling price – a price which it accepts when it sells a concession – and the equivalent of a buying price. Sometimes these prices are so far apart that a transaction vital to both nations cannot be completed peacefully; they cannot agree on the price of the transaction. The history of diplomacy is full of such crises" (Blainey, 1988, p. 113-114).

In Section 5, we consider bargaining with asymmetric (private) information.[4] We distinguish information about military capabilities (e.g., the quality of weapons) from information about preferences (e.g., the cost of fighting). By definition, military capabilities determine the likely outcome of a war, while preferences determine the evaluation of different outcomes.

In Section 5.1 there is asymmetric information about preferences. The ultimatum game (Fearon, 1995) illustrates Brito and Intriligator's (1985) insight that war is consistent with optimal screening. Because of a built-in first-mover advantage, the proposer has all the bargaining power: he is committed to fight if his proposal is rejected. If he does not know the responder's cost of fighting, then he must trade off the risk of being overly generous if the responder's cost is high against the risk of war if the cost is low. The equilibrium offer may be such that if the responder's cost is low, then the responder would rather fight than appease the proposer.[5] On the other hand,

[4] We follow convention in economics by deriving beliefs from a common prior, representing disagreement as the result of private signals. Smith and Stam (2004) consider disagreement caused by non-common prior beliefs.

[5] As in Section 3, the commitment that makes the ultimatum credible may come from a strategic move, such as the occupation of a contested territory. Such occupations, followed by attempts to negotiate, precede many wars:

> "The plunderers usually believed that they could snatch territory without provoking a war or that, if war erupted, the campaign would be swift and victorious. No war was apparently expected by Louis XIV when in 1701 he quietly began to poach territory from Spain, by Frederick the Great when in 1740 he invaded Austrian Silesia, and by Joseph of Austria when in 1778 his white-coated troops marched into the south of Bavaria. Those annexations however were followed by strenuous fighting" (Blainey, 1988, p. 70).

In the ultimatum game, the first-mover advantage is exogenously given. If there is competition to become the first-mover, the risk of war is compounded. They will risk with asymmetric information only about preferences, and with no opportunities for strategic moves or first-mover advantages, there exist bargaining games that guarantee a peaceful outcome (Fey and Ramsay, 2011). With an equilibrium allocation of resources proportional to military strength, there is no need for binding peace treaties, since each player prefers his share to war.[6]

Blainey (1988) argued that there will be war if there is disagreement about the balance of power, and each side thinks it can gain more by fighting than by negotiating. In Section 5.2 we consider asymmetric information about military capabilities. In view of the agents' private information and outside option (war), Incentive Compatibility (IC) and Individual Rationality (IR) conditions must be satisfied. If fighting is not too expensive, then the IC and IR conditions imply a strictly positive probability of war (Bester and Wärneryd, 2006, Fey and Ramsay, 2009, 2011). No bargaining game can guarantee peace, even if unrestricted transfers can be made and a peace treaty, once signed, would be completely binding.

If the balance of power changes over time, then the inability to enforce contracts can cause war, even if there are no first-mover advantages and no private information (Fearon, 1995, 1998). For example, suppose the government must appease a rebel group in order to avoid a civil war. If the government is expected to be stronger in the future (an exogenous power shift), then it has a credibility problem. To appease the rebels, it must promise them a large permanent share of the country's wealth. However, if the rebels think the government will renege on the deal once it has grown stronger, then appeasement is impossible and war unavoidable.

If no exogenous power shift is expected, but wealth transfers make the rebels stronger because "wealth is power", then the rebels have a credibility problem. If they cannot commit to staying peaceful, then any attempt at appeasing them will only increase the threat to the government. If the rebels cannot make any commitment at all, then war may be unavoidable (Jackson and Morelli, 2007, Beviá and Corchón, 2010). However, if the rebels can

war both to get the first-mover advantage, as in Section 3, and to optimally exploit the advantage if they get it, as in Section 5.1.

[6] We assume war is a costly lottery where the winner takes all of the available resource, and utility functions are concave. Concavity implies that it is better to get a share p of the (divisible) resource than to pay the cost of war and win all of the resource with probability p.

make short-term commitments, and power is a continuous function of wealth, then the rebels are appeased by a sequence of transfers, and war is avoided (Fearon, 1996, Schwarz and Sonin, 2008). We discuss power shifts and their concomitant commitment problems in Section 6.

If the disputed resource is *indivisible* and the ability to make transfers is limited, then war may be unavoidable. For example, Hassner (2003) argued that disputes over "sacred spaces" admit no compromise, because the only acceptable outcome for either side is complete victory. There is no mutually acceptable middle ground.[7] In many cases, however, compromise alternatives exist. As Fearon (1995) argued, compromises can be generated by creating linkages across issues, and settlements may involve transfers of both land and other resources.[8] In this survey, we downplay the role of indivisibilities.

The survey is organized as follows. Section 2 introduces basic notation and definitions. Section 3 shows that bargaining with two-sided commitment opportunities can be represented by a Hawk-Dove game with strategic substitutes. Hawk represents a strategic move, and war is an unintended consequence of brinkmanship. Section 4 introduces asymmetric information and finds conditions for uniqueness of equilibrium in a class of 2×2 games. Section 4.1 applies these results to the bargaining game of Section 3. In Section 4.2, Hawk represents a decision to attack, and a first-strike advantage causes the game to have strategic complements. Preemptive war is a deliberate decision caused by fear (the Hobbesian trap). Section 5 considers the mechanism design approach. Section 5.1 shows how optimal screening implies a strictly positive probability of war when there is asymmetric information about preferences. Section 5.2 shows how IC and IR conditions imply a strictly positive probability of war when there is asymmetric information about military strength. In Section 6, preventive war is caused by imperfect contracting ability. Section 6.1 considers exogenous shifts in the balance of power, and Section 6.2 considers endogenous shifts induced by resource transfers. Section 7 concludes with a brief discussion of omitted topics.

[7]In theory, an indivisible object can be allocated by a coin toss. But Powell (2006) pointed out that it may be impossible to make a binding commitment to respect the outcome, so a contracting problem is fundamental here as well.

[8]For example, in 238 BC, Rome threatened Carthage with war. Rather than fight, Carthage agreed to hand over Sardinia plus 1,200 talents. Centuries later, a weakened Rome would itself habitually pay barbarians in order to avoid war. For more examples, see Beviá and Corchón (2010).

2 Baseline model

There are two players, A and B, and a perfectly divisible resource (e.g., a contested territory), normalized to size one. Let x_i denote player i's share of the resource, $i \in \{A, B\}$. The allocation $x = (x_A, x_B) \in \mathbf{R}^2$ is *feasible* if and only if $x_A \geq 0$, $x_B \geq 0$ and $x_A + x_B \leq 1$. Let X denote the set of all feasible allocations. There is a *status quo* allocation $\omega \in X$ such that $\omega_A + \omega_B = 1$. We refer to ω_i as player i's initial (status quo) endowment. If the allocation $x \in X$ is implemented then player $i \in \{A, B\}$ gets utility $u_i(x_i)$ from his share of the resource. For each $i \in \{A, B\}$, the function u_i is continuous, concave and strictly increasing on $[0, 1]$, and we normalize so that $u_i(0) = 0$ and $u_i(1) = 1$.

In standard bargaining theory, if the players cannot agree on an allocation of resources then there is a "disagreement outcome". Here, disagreement means war. Thus, the set of feasible (deterministic) outcomes is $X^* \equiv X \cup \{war\}$.[9] The set of feasible randomized outcomes is denoted $\Delta(X^*)$. War is a "costly lottery": each player $i \in \{A, B\}$ pays a cost $\phi_i > 0$ and the winner takes all of the contested resource. Suppose player $i \in \{A, B\}$ wins the war with probability $p_i > 0$. War is always decisive, so $p_A + p_B = 1$. Player i's payoff from war is

$$p_i u_i(1) + (1 - p_i) u_i(0) - \phi_i = p_i - \phi_i$$

using the normalizations. Since war is a costly lottery and utility functions are concave, war is a strictly Pareto dominated outcome: there exists $x \in X$ such that $u_i(x_i) > p_i - \phi_i$ for each $i \in \{A, B\}$. Indeed, giving player $i \in \{A, B\}$ the share $x_i = p_i$ strictly Pareto dominates a war which player i wins with probability p_i.

3 Bargaining and commitment with complete information

A strategic commitment is a "voluntary but irreversible sacrifice of freedom of choice" (Schelling, 1960, p. 22). In an *ultimatum bargaining game*, only the proposer (the "first-mover") can make a commitment. The commitment

[9] In the baseline model, the outcome "war" is unambiguous: it does not matter who attacks whom, or how much of the resource each controls. This will be generalized later.

represents a *fait accompli*. Rejecting the ultimatum means war, so the responder faces a stark choice between war and appeasement. With complete information, there is a unique subgame perfect Nash equilibrium where the responder gets the smallest share he (weakly) prefers to war.

The *Nash demand game* is a model of two-sided commitment (Nash, 1953). Each player $i \in \{A, B\}$ simultaneously demands a share $x_i \geq 0$ of the contested resource. If $x_A + x_B \leq 1$ then each gets what he demanded. But if $x_A + x_B > 1$ then they have made incompatible commitments, so there is war. The Nash demand game has many Nash equilibria.[10]

In Crawford's (1982) modified Nash demand game, each player decides whether to attempt a commitment, and if so, how much to demand. A commitment attempt is a *strategic move* which, in the simplest version, is *successful* with probability q and *unsuccessful* with probability $1 - q$. Ellingsen and Miettinen (2008) found that if there is a small but strictly positive cost of attempting a commitment, then both players will try to commit, and they make the maximum demand an uncommitted opponent would accept.[11] If both attempts succeed then there is war, so the equilibrium probability of war is q^2. War is not a deliberate choice, but rather an unintended "accident" caused by brinkmanship, i.e., by strategic moves meant to convince the opponent to back down.[12]

3.1 A two-stage bargaining game

The following game is based on Crawford (1982), Ellingsen and Miettinen (2008) and Baliga and Sjöström (2020). In stage 1, the players simultaneously

[10] If each player prefers war to getting nothing, then there exists an inefficient Nash equilibrium where each player demands all of the resource, and the equilibrium outcome is war. There are also peaceful equilibria. Due to space constraints, we focus on models that do not have any peaceful equilibrium. Nevertheless, models that have both peaceful and non-peaceful equilibria can be very insightful (e.g., Slantchev, 2003).

[11] In some cases, a purely verbal (non-costly) message could turn into a credible commitment: backing down might involve a massive loss of reputation, or the message might become a policy that subordinates will follow in a crisis. However, in many situations a credible commitment requires a costly action, such as occupying part of a contested territory. For a classic discussion, see Schelling (1960, 1966).

[12] Accidental wars can also be caused by mistakes, misunderstandings, and technological problems such as malfunctioning early warning systems. For example, a country may fire its missiles in the mistaken belief that a war has already started. See Acemoğlu and Wolitzky in this Volume.

decide whether or not to challenge the status quo. A challenge by player $i \in \{A, B\}$ consists of a demand x_i such that $\omega_i < x_i \leq 1$, and an attempt to make a commitment. Making the challenge costs $c_i > 0$, and the commitment attempt succeeds with probability q, where $0 < q < 1$.[13] If neither player makes a successful commitment, then the status quo allocation ω remains in place. If both players make successful commitments then there is war, because the commitments are irrevocable and the demands are incompatible ($x_A + x_B > \omega_A + \omega_B$). Player $i \in \{A, B\}$ wins the war with probability p_i.

Suppose only one player, say player i, makes a successful commitment. Then we go to stage 2, where player $j \neq i$ (the "second-mover") accepts or rejects player i's demand. Rejection means war, which player i wins with probability p_i. Player j's expected payoff from rejection is $p_j - \phi_j$, since the winner takes all and $u_j(1) = 1$. Assume for simplicity that $0 < p_j - \phi_j < u_j(\omega_j)$ for each $j \in \{A, B\}$, so player j prefers the status quo to war but would rather fight than get nothing. Sequential rationality implies that player j accepts the demand x_i if and only if $x_i \leq x_i^*$, where $x_i^* \in (\omega_i, 1)$ satisfies

$$u_j(1 - x_i^*) = p_j - \phi_j. \tag{1}$$

In any subgame perfect equilibrium, player $i \in \{A, B\}$ will either do nothing in stage 1, or challenge the status quo and demand x_i^* such that Equation (1) holds.[14] Referring to those two actions as Dove and Hawk, respectively, we can represent the two-stage bargaining game by a 2×2 payoff matrix: the Hawk-Dove game. Player i chooses a row, player j a column, and only player i's payoff is indicated:

$$\begin{array}{lcc} & \text{Hawk} & \text{Dove} \\ \text{Hawk} & v_i^{HH} - c_i & v_i^{HD} - c_i \\ \text{Dove} & v_i^{DH} & u_i(\omega_i) \end{array} \tag{2}$$

[13]For example, player i could attempt to cross the status quo demarcation and take control of additional territory, corresponding to the demand $x_i > \omega_i$. If the incursion succeeds, then retreating would imply an intolerable loss of face.

[14]Suppose player i's commitment is successful. If player $j \neq i$ also successfully commits, then player i gets $p_i - \phi_i$. But there is a strictly positive probability that player j does *not* become committed (since $q < 1$). Then demanding $x_i < x_i^*$ is strictly worse than demanding x_i^* because player j will accept x_i^* at stage 2. Demanding $x_i > x_i^*$ is also strictly worse, because player j will reject so player i gets $p_i - \phi_i < x_i^*$, where the inequality follows from concavity of u_j and Equation (1).

where[15]

$$v_i^{HH} \equiv q(1-q)u_i(x_i^*) + q(1-q)(p_i - \phi_i) + q^2(p_i - \phi_i) + (1-q)^2 u_i(\omega_i),$$

$$v_i^{DH} \equiv q(p_i - \phi_i) + (1-q)u_i(\omega_i),$$

and

$$v_i^{HD} \equiv q u_i(x_i^*) + (1-q)u_i(\omega_i).$$

Actions are strategic substitutes because $u_i(x_i^*) > u_i(\omega_i)$ implies $v_i^{HD} - u_i(\omega_i) > v_i^{HH} - v_i^{DH} > 0$. For intermediate levels of c_A and c_B, it is a game of chicken with two pure Nash equilibria, (Hawk,Dove) and (Dove,Hawk). But if c_A and c_B are small enough, then (Hawk,Hawk) is the unique Nash equilibrium, and the probability of war is $q^2 > 0$. War could be avoided if the two players could get together "before stage 1" and agree not to attempt any commitment. This is assumed to be impossible since, following Schelling (1960), *fait accompli* tactics are inherently a part of the bargaining process itself.

3.1.1 Large first-mover advantages and strategic complements

Actions will be strategic complements if a successful strategic move, such as an occupation of contested territory, brings a significant military advantage. Suppose that if only player i makes a successful commitment and player j rejects the ultimatum, then player i wins the war with probability $p_i + \theta$, where p_i is his inherent military strength and $\theta \geq 0$ the military advantage from the occupation. Assuming $0 < p_j - \theta - \phi_j < u_j(\omega_j)$, in stage 2 player j accepts the demand x_i if and only if $x_i \leq x_i^{**}$, where $x_i^{**} \in (\omega_i, 1)$ satisfies

$$u_j(1 - x_i^{**}) = p_j - \theta - \phi_j. \tag{3}$$

The entries in the payoff matrix of the Hawk-Dove game are modified as follows:

$$v_i^{HH} \equiv q(1-q)u_i(x_i^{**}) + q(1-q)(p_i - \theta - \phi_i) + q^2(p_i - \phi_i) + (1-q)^2 u_i(\omega_i),$$

[15] To derive v_i^{HH}, suppose both players try to commit. With probability $q(1-q)$, only player i succeeds, and his demand x_i^* is accepted. With probability $q(1-q)$, only player j succeeds, and player i gets $u_i(1-x_j^*) = p_i - \phi_i$. With probability q^2 both succeed and there is war, and with probability $(1-q)^2$ neither succeeds and the status quo remains in place. We derive v_i^{DH} and v_i^{HD} in a similar way.

$$v_i^{DH} \equiv q(p_i - \theta - \phi_i) + (1-q)u_i(\omega_i),$$

and

$$v_i^{HD} \equiv qu_i(x_i^{**}) + (1-q)u_i(\omega_i).$$

The game has strategic substitutes if $u_i(x_i^{**}) - u_i(\omega_i) > \theta$ for each $i \in \{A, B\}$, but strategic complements if the inequality is reversed. A large θ tends to make actions strategic complements (Baliga and Sjöström, 2020). For a range of commitment costs, it is a stag hunt game with two pure Nash equilibria, (Dove,Dove) and (Hawk,Hawk).[16]

4 Coordination with incomplete information

Chicken and stag hunt games are often used as metaphors for different kinds of conflicts (Jervis, 1978, Oye, 1986). In the stag hunt metaphor, aggression is triggered by fear and distrust, so a show of toughness can lead to escalation and war. In chicken, aggression is triggered by a *lack* of fear, so showing toughness can make the opponent back down. With complete information, chicken and stag hunt games have multiple Nash equilibria. This raises difficult issues of equilibrium selection and coordination.

A number of articles have modeled conflict as a coordination game with incomplete information (see Ramsay, 2017, for a recent survey). Some consider a global games framework, where each player observes a signal (his type) which is correlated with the true state of the world. Acharya and Ramsay (2013) found conditions such that all types take the aggressive action (Hawk) in the unique equilibrium. Chassang and Padro i Miquel (2010) identified a risk-dominance condition under which peace cannot be sustained in an infinite horizon interaction. Here we will follow Baliga and Sjöström (2004) and assume types represent preferences and are uncorrelated.

Consider a Hawk-Dove game where player i's cost of choosing Hawk, c_i, is his private information, his *type*. The type is *soft* (unverifiable) information. The cost-types c_A and c_B are independently drawn from a distribution with a continuously differentiable c.d.f. F and support $[\underline{c}, \overline{c}]$, where $\underline{c} \geq 0$. For simplicity, the players are symmetric *ex ante* (i.e., identical in all respects

[16] For a given θ, large costs of war make actions strategic substitutes, because x_i^{**} is increasing in ϕ_j, by Equation (3). Intuitively, if a car crash is sufficiently costly then it is a game of chicken.

except their cost-types), so we drop all subscripts except on the types. Endowments are $\omega_A = \omega_B = 0.5$. Player i chooses a row, player j a column, and only player i's payoff is indicated:

$$\begin{array}{ccc} & \text{Hawk} & \text{Dove} \\ \text{Hawk} & v^{HH} - c_i & v^{HD} - c_i \\ \text{Dove} & v^{DH} & u(0.5) \end{array} \tag{4}$$

In Bayesian equilibrium, for each $j \in \{A, B\}$ there must be a cutoff point $z_j \in [\underline{c}, \overline{c}]$ such that player j chooses Hawk if $c_j < z_j$ and Dove if $c_j > z_j$.[17] Thus, player j chooses Hawk with probability $F(z_j)$. Define

$$\Gamma(z_j) \equiv F(z_j)(v^{HH} - v^{DH}) + (1 - F(z_j))(v^{HD} - u(0.5)).$$

Each player's cutoff must maximize his ex ante expected payoff. Player i's best response to z_j is the cutoff point $z_i = BR(z_j)$, where

$$BR(z_j) \equiv \begin{cases} \underline{c} & \text{if} \quad \Gamma(z_j) < \underline{c} \\ \Gamma(z_j) & \text{if} \quad \underline{c} \leq \Gamma(z_j) \leq \overline{c} \\ \overline{c} & \text{if} \quad \Gamma(z_j) > \overline{c} \end{cases}$$

If there is sufficient uncertainty about types, specifically if

$$F'(c) < \frac{1}{|v^{HH} - v^{DH} - v^{HD} + u(0.5)|} \text{ for all } c \in [\underline{c}, \overline{c}], \tag{5}$$

then $|\Gamma'(z_j)| < 1$ for all $z_j \in [\underline{c}, \overline{c}]$. This condition guarantees that there is a unique equilibrium, which must be symmetric.[18] The equilibrium cutoff point satisfies $z^* = BR(z^*)$.

A type is a *dominant strategy hawk* (resp. *dominant strategy dove*) if Hawk (resp. Dove) is a strictly dominant strategy for this type. If Condition (5) holds, then the support of F contains at least one kind of dominant strategy type, and the unique equilibrium can be obtained by the iterated elimination of (interim) dominated strategies.

We consider two applications of these results. Section 4.1 returns to the bargaining game of Section 3, and Section 4.2 considers the Hobbesian trap.

[17] We assume every type chooses an optimal action. Strictly speaking, Bayesian Nash equilibrium only requires this for almost every type. This has no substantive implications.

[18] With a uniform distribution, Condition (5) is equivalent to

$$\overline{c} - \underline{c} > \left| v^{HH} - v^{DH} - v^{HD} + u(0.5) \right|.$$

That is, the support must be sufficiently large.

4.1 Bargaining and uniqueness

Consider again the bargaining game of Section 3.1. Hawk represents a strategic move, i.e., a challenge to the status quo, and actions are strategic substitutes. If both players choose Hawk then there is war with probability q^2. Assuming the players are symmetric except for the c_i, in particular $\omega_A = \omega_B = p_A = p_B = 0.5$, we drop all subscripts from the matrix in Equation (2) (except for c_i) and obtain the matrix in Equation (4). It can be verified that $v^{DH} < v^{HH}$ and $v^{DH} < v^{HD}$ since $u(x^*) > u(0.5)$. If the costs c_A and c_B are private information, independently drawn from a distribution F which satisfies Condition (5), then there is a unique equilibrium.[19]

For example, suppose F is uniform on $[\underline{c}, \bar{c}]$. If $\underline{c} < v^{HD} - u(0.5)$ and $\bar{c} > v^{HH} - v^{DH}$, then the support of F contains both kinds of dominant strategy types and Condition (5) holds. The equilibrium cutoff point z^* must satisfy $\underline{c} < z^* < \bar{c}$. Solving the equation $z^* = \Gamma(z^*)$, we find the equilibrium probability that a player will choose Hawk:

$$F(z^*) = \frac{q\,(u(x^*) - u(0.5)) - \underline{c}}{\bar{c} - \underline{c} + q^2(u(x^*) - u(0.5))}. \tag{6}$$

With a unique equilibrium, it is straightforward to do comparative statics. Suppose there is an increase in the cost of fighting, ϕ. Equation (1) implies $u(1 - x^*) = 0.5 - \phi$, so x^* increases. That is, the second-mover makes larger concessions to avoid war. This makes it more profitable to challenge the status quo, so z^* increases.[20] An increase in the cost of war therefore makes war more likely.

4.2 The Hobbesian trap

Now suppose Hawk represents a deliberate decision to start a war. Accordingly, there is war if at least one player chooses Hawk. For simplicity, the players are again *ex ante* symmetric. If both players choose Dove then they

[19] For simplicity, only c_A and c_B are private information. With additional private information about the cost of fighting, for example, then the risk of war would be compounded: there is war if both are successfully committed, or only one player is committed but his ultimatum is rejected as in Section 5.1.

[20] The derivative of the right hand side of Equation (6) with respect to x^* has the same sign as $\underline{c}q + \bar{c} - \underline{c}$. This expression is strictly positive because $\bar{c} > v^{HH} - v^{DH}$ and $\underline{c} < v^{HD} - u(0.5)$. Thus, z^* is increasing in x^*.

coexist peacefully and each gets $u(0.5)$. If only one player chooses Hawk then he wins the war with probability $0.5+\theta$, where θ represents a first-strike advantage, $0<\theta<0.5$. His expected payoff is therefore $v^{HD}=0.5+\theta-\phi$, using the normalizations, while the opponent gets $v^{DH}=0.5-\theta-\phi$, where $\phi>0$ is the cost of war. If both choose Hawk then the first-strike advantage goes to whoever can mobilize and launch an attack faster. Assuming both have the same chance of being the fastest, each expects

$$v^{HH}=\frac{1}{2}v^{HD}+\frac{1}{2}v^{DH}=0.5-\phi<u(0.5), \tag{7}$$

since u is concave. Condition (5) is now

$$F'(c)<\frac{1}{|\phi+u(0.5)-0.5|} \text{ for all } c\in[\underline{c},\overline{c}]. \tag{8}$$

The game has strategic complements, since

$$v^{HH}-v^{DH}=\theta>\theta-\phi-(u(0.5)-0.5)=v^{HD}-u(0.5)$$

Peace, i.e., (Dove,Dove), maximizes the social surplus.

Player i is a dominant strategy hawk if

$$c_i<0.5+\theta-\phi-u(0.5)$$

and a security-seeker if

$$0.5+\theta-\phi-u(0.5)\leq c_i\leq\theta.$$

Security-seekers want to match the opponent's choice. They would rather not fight, but if they think the opponent plans to attack then they will try to deprive him of the first-strike advantage. Assume

$$\underline{c}<0.5+\theta-\phi-u(0.5)<\bar{c}<\theta \tag{9}$$

so the support of F contains security-seekers and dominant strategy hawks, but no dominant strategy doves. Clearly, it is a Bayesian equilibrium for all types to choose Hawk. It must be the only equilibrium if F satisfies Condition (8), and it can be reached via the iterated elimination of dominated strategies. In the first round of iteration, the dominant strategy hawks eliminate Dove. This sets off a contagion, a "Hobbesian trap", where security-seekers with higher and higher cost-types eliminate Dove out of fear that the opponent will choose Hawk. There are no dominant strategy doves to stop the contagion from reaching all types in the support of F. Thus, if Conditions (8) and (9) hold then there is war with probability one.

4.2.1 Binding agreements

If the players can sign a binding peace treaty before playing the Hawk-Dove game, then peace can be achieved. Suppose Conditions (8) and (9) hold, and consider the following two-stage game. In stage 1, each player simultaneously chooses whether or not to sign the treaty. These decisions are publicly observed. In stage 2, they play the Hawk-Dove game with one modification: if both have signed the treaty, then both must choose Dove. If at least one player did not sign, then each player is free to choose either Hawk or Dove. Thus, signing is a commitment to choose Dove conditional on the opponent also signing. There is a perfect Bayesian equilibrium where all types sign, backed up by the threat that if some player refuses to sign, both players will choose Hawk.[21] Since all types prefer (Dove,Dove) to (Hawk,Hawk), all types prefer to sign.

4.2.2 Signaling

If binding treaties are not feasible, peace may still be possible if the players can take unilateral actions that signal peaceful intentions (Kydd, 1997). Suppose Conditions (8) and (9) hold, and consider the following two-stage game. In stage 1, the two players simultaneously choose to be armed or unarmed. These decisions are publicly observed and, for simplicity, have no direct cost. In stage 2, they play the Hawk-Dove game with one modification: a player who is unarmed cannot choose Hawk. Thus, to be unarmed is a unilateral commitment to choose Dove. In a symmetric perfect Bayesian equilibrium, there must be a cutoff point z^* such that player $i \in \{A, B\}$ is unarmed if $c_i > z^*$ and armed if $c_i < z^*$. Dominant strategy hawks will always be armed and choose Hawk, so $z^* \geq v^{HD} - u(0.5)$. Even security-seekers will choose Hawk if both players are armed. For peace to be possible, the cutoff must satisfy $z^* < \bar{c}$ so that sufficiently high cost-types are unarmed.

Security-seekers choose Dove against unarmed opponents. Thus, an unarmed security-seeker will live in peace and get $u(0.5)$ if the opponent is also a security-seeker, but he will be attacked and get $v^{DH} = 0.5 - \theta - \phi$ if the opponent is a dominant strategy hawk. For this gamble to be worthwhile,

[21] The threat is credible because there are no dominant strategy doves in the support of F.

dominant strategy hawks must be sufficiently rare. Specifically, if

$$F\left(0.5+\theta-\phi-u(0.5)\right)<\frac{\theta}{u(0.5)+\phi+\theta-0.5} \tag{10}$$

then there exists an equilibrium where security-seekers with sufficiently high cost-types gamble for peace: the cutoff point is $z^{*}<\bar{c}$.[22] Thus, if Conditions (8), (9) and (10) hold, then there exists a signaling equilibrium with a peaceful outcome for pairs of security-seekers such that at least one has a cost-type above z^{*}.

4.2.3 Cheap-talk

Suppose it is impossible to make any kind of commitment, but the players can engage in peace talks. Since talk is cheap, a dominant strategy hawk can pretend to be peace-loving by sending a dovish message, and then he attacks anyway. Yet peace talks can be meaningful as long as dominant strategy hawks are sufficiently rare (Baliga and Sjöström, 2004).[23] Assume Conditions (8) and (9) hold, and consider the following two-stage game. In stage 1, each player simultaneously sends either a dovish or a hawkish message to the opponent. In stage 2, they play the Hawk-Dove game. If messages are uninformative ("babbling"), then Conditions (8) and (9) guarantee that there is war with probability one. For peace talks to be meaningful, messages must be informative. This is possible because different types have different

[22] A player who is unarmed gets expected payoff

$$F\left(v^{HD}-u(0.5)\right)v^{DH}+\left(1-F\left(v^{HD}-u(0.5)\right)\right)u(0.5). \tag{11}$$

An armed security–seeker with cost-type c_i gets

$$F\left(z^{*}\right)\left(v^{HH}-c_i\right)+\left(1-F\left(z^{*}\right)\right)u(0.5). \tag{12}$$

The cutoff type $c_i=z^{*}$ must be indifferent between being armed and unarmed, which requires

$$F\left(z^{*}\right)\left(u(0.5)-v^{HH}+z^{*}\right)-F\left(v^{HD}-u(0.5)\right)\left(u(0.5)-v^{DH}\right)=0. \tag{13}$$

The left hand side of Equation (13) is strictly negative for $z^{*}=v^{HD}-u(0.5)$, and strictly positive for $z^{*}=\bar{c}$ if Condition (10) holds. By continuity, Equation (13) is satisfied for some z^{*} such that $v^{HD}-u(0.5)<z^{*}<\bar{c}$.

[23] Baliga and Sjöström (2004) assumed $v^{HH}=\underline{c}=0$. Kang (2022) shows that their results are valid without this restriction.

objectives: dominant strategy hawks only want to increase the probability that the opponent chooses Dove, but security-seekers also want to know the opponent's action in order to match it. For some parameter values, there exist informative cheap-talk equilibria. Security-seekers with the highest cost-types send the dovish message, and if they also receive a dovish message then they choose Dove in stage 2. Dominant strategy hawks (the lowest cost-types) also send the dovish message, but they always choose Hawk in stage 2. The hawkish message is sent by intermediate types – security-seekers with a relatively low cost of choosing Hawk – who want to match the opponent's action. Security-seekers with high cost-types are willing to gamble for peace (i.e., trust the dovish peace message) if the prior probability that the opponent is a dominant strategy hawk is small enough. Thus, cheap-talk can create peace among types that have a high cost of choosing Hawk. If such types do not exist, then peace talks are meaningless, as illustrated by Acharya and Ramsay's (2013) negative results on cheap-talk in a global games model.

5 Mechanism design

We now return to the baseline model of Section 2, without first-strike advantages but with asymmetric (private) information. Does there exist a bargaining game such that the two players always negotiate a peaceful outcome? If some bargaining game generates a positive probability of war, does there exist another bargaining game which is better for both ? To answer these questions, we use theory of mechanism design. A mechanism, or game form, specifies the "rules of the game": what actions the players can take and how the outcome depends on these actions. The set of feasible outcomes is $\Delta(X^*)$, i.e., the set of probability distributions over $X^* \equiv X \cup \{war\}$. A game form becomes a "game" if we add assumptions about payoffs. A player's type specifies his privately known characteristics. As in Section 4, the type is unverifiable. A pair of types, one for each player, is a state of the world. The game starts by drawing a state according to a commonly known distribution. A *state-contingent outcome* is a function from the set of possible states to $\Delta(X^*)$, i.e., an assignment of a feasible outcome to each pair of types.

The mechanism cannot have built-in knowledge about the true state, since types are private information. The outcome can depend on the true state only indirectly, via the players' strategies. Each player's strategy specifies, for each of his possible types, what actions he will take. The players strategies therefore generate a state-contingent outcome. This state-contingent outcome must satisfy incentive compatibility (IC) and (interim) individual rationality (IR) conditions. The IC conditions say that each type's expected payoff must be at least as large as what he could get by mimicking some other type. The IR conditions say that each type's expected payoff must be at least as large as what he could get by unilaterally declaring war instead of negotiating. These are *necessary* conditions which must be satisfied even if binding commitments are feasible.[24] The setup allows for the existence of an impartial mediator who receives messages from the two players and then recommends an outcome. The IC and IR conditions would still be necessary, assuming the mediator does not know the true state and cannot prevent the players from declaring war.

In Section 5.1, there is private information about preferences, specifically, the cost of going to war. War may occur with ultimatum bargaining, but other mechanisms exist that guarantee peace. In Section 5.2, military capabilities are private information. If the cost of war is not too high, then no mechanism can guarantee peace.

5.1 Asymmetric information about preferences

Following Fearon (1995), consider an ultimatum game where player A is the proposer. Player A has no private information. Player B's privately known type is his own cost of war, ϕ_B. Assume $\phi_B \in \{\phi_B^L, \phi_B^H\}$, where $0 < \phi_B^L < \phi_B^H < p_B$. Type ϕ_B gets expected payoff $p_B - \phi_B$ from war. The probability of winning, p_B, is the same for both types. This is a private values environment, in the sense that for any outcome in $\Delta(X^*)$, player A's expected payoff does not depend on player B's type.

Player A thinks $\phi_B = \phi_B^H$ with probability h, where $0 < h < 1$. Assume $u_i(x_i) = x_i$ for each $i \in \{A, B\}$. If player A offers the share $x_B = p_B - \phi_B^L$ to his opponent, then both types accept, so the outcome is $(1 - x_B, x_B) \in X$ in each state. This state-contingent outcome is incentive-compatible, individually rational and maximizes the social surplus (since there is no war in any state). Player A's payoff is $1 - (p_B - \phi_B^L) = p_A + \phi_B^L$. However, this may not

[24] The interim IR conditions would not be necessary if the players could negotiate a binding agreement before knowing their own types, but this is assumed to be impossible.

be player A's optimal offer, since type ϕ_B^H gets an *information rent*, a strictly higher expected payoff than from war: $x_B > p_B - \phi_B^H$. If player A offers the smaller share $x_B = p_B - \phi_B^H$ then only type ϕ_B^H will accept, so player A's expected payoff is

$$h\left(1 - (p_B - \phi_B^H)\right) + (1-h)(p_A - \phi_A) = p_A + h\phi_B^H - (1-h)\phi_A,$$

where $\phi_A > 0$ is player A's cost of war.

Player A prefers to offer $p_B - \phi_B^H$ if

$$h\phi_B^H - (1-h)\phi_A > \phi_B^L. \tag{14}$$

If this inequality holds, then in equilibrium there is war with probability $1 - h > 0$. Player A has the first-mover advantage, and therefore all the bargaining power, and he prefers to eliminate his opponent's information rent even at the cost of a reduced social surplus.[25]

5.1.1 Optimal screening

Even if there is war in equilibrium, ultimatum bargaining is an efficient mechanism in the following sense: no state-contingent outcome which satisfies the IC and IR conditions can give each player a higher expected payoff.[26] Specifically, there is *optimal screening*: the proposer's expected payoff is maximized, given the responder's private information and outside option (war). Consider any state-contingent outcome. The expected payoff for type ϕ_B^K is

$$u_B^K \equiv (1 - y_B^K)x_B^K + y_B^K(p_B - \phi_B^K),$$

where $K \in \{L, H\}$, y_B^K is the probability of war for type ϕ_B^K, and x_B^K is his expected share if there is no war. Since type ϕ_B^K can behave just like type ϕ_B^M, the two IC constraints are:

$$u_B^K \geq (1 - y_B^M)x_B^M + y_B^M(p_B - \phi_B^K) \tag{15}$$

[25] In the ultimatum game, neither the status quo nor the balance of power plays any role in determining whether there is war (since neither ω_i nor p_i appear in Condition (14)). In Powell's (1999) alternating offers bargaining game, war is more likely if there is a "mismatch" between the status quo and the balance of power.

[26] In the terminology of Mas-Colell, Whinston and Green (1995), it is "constrained efficient".

for $K \in \{L, H\}$ and $M \neq K$. The two IR constraints are:

$$u_B^K \geq p_B - \phi_B^K \tag{16}$$

for $K \in \{L, H\}$.[27] Player A's expected payoff is

$$h\left((1 - y_B^H)(1 - x_B^H) + y_B^H(p_A - \phi_A)\right) + (1-h)\left((1 - y_B^L)(1 - x_B^L) + y_B^L(p_A - \phi_A)\right). \tag{17}$$

Choosing $(x_B^L, x_B^H, y_B^L, y_B^H)$ to maximize player A's expected payoff subject to the IC and IR constraints reproduces the equilibrium outcome of the ultimatum game. Thus, ultimatum bargaining is an efficient mechanism, even if war reduces the social surplus.

5.1.2 Surplus-maximizing mechanisms

However, there exist bargaining games such that there is never war in equilibrium, and each player always prefers the equilibrium outcome to war (Fey and Ramsay, 2011). This is true even if each player has private information about his own preferences, e.g., the cost of going to war or the utility of the contested resource. It remains true if preferences have a common value component, and each player has private information correlated with the true value.[28] What is required for peace is only that the balance of power, represented by (p_A, p_B), is the same in every state of the world, and therefore common knowledge. For example, it is a Bayesian Nash equilibrium of the Nash demand game for each type of player $i \in \{A, B\}$ to demand $x_i = p_i$. The equilibrium outcome is $(p_A, p_B) \in X$, which both players strictly prefer to war in every state. The Nash demand game has other equilibria as well (although sharing in proportion to military power may be a focal point). But other mechanisms exist such that in every equilibrium, the outcome is $(p_A, p_B) \in X$ with probability one. For example, a mediator (the "mechanism designer") could propose the allocation $(x_A, x_B) = (p_A, p_B)$, and the players sequentially either accept this allocation or declare war. In perfect Bayesian equilibrium, both players accept with probability one. There is no

[27] If $y_B^K = 1$ then the IR constraint holds trivially for type ϕ_B^K. If $y_B^K < 1$ then the IR constraint is equivalent to $x_B^K \geq p_B - \phi_B^K$.

[28] For example, the contested resource could be an oil field, and each player observes a private signal of the amount of oil it contains.

"bad equilibrium".[29] Thus, with asymmetric information only about preferences, there exist bargaining games that guarantee peace.

5.2 Asymmetric information about military power

Now suppose each player's private information is his military strength. To be specific, suppose the two types t_A and t_B are independently drawn from a uniform distribution with support $[0, 1]$. Only player i knows t_i. If there is war, then player $i \in \{A, B\}$ wins with probability $p_i(t_A, t_B)$. Assume $p_A(t_A, t_B) = \mu + \alpha t_A - \beta t_B$ where μ, α and β are non-negative constants. Since war is decisive,

$$p_B(t_A, t_B) = 1 - p_A(t_A, t_B) = 1 - \mu - \alpha t_A + \beta t_B.$$

To guarantee that p_A and p_B are between 0 and 1, assume $\beta \le \mu \le 1 - \alpha$. Since a player's expected payoff from war depends on his opponent's type, values are interdependent. Note that the bigger are α and β, the bigger is the uncertainty about the balance of power. The baseline model of Section 2 corresponds to the case $\alpha = \beta = 0$. For simplicity, we again assume $u_i(x_i) = x_i$ for each i.

Theorem 1 *For any state-contingent outcome which satisfies the IC and (interim) IR conditions, the (ex ante) probability of war is at least*

$$\frac{1}{2} - \frac{\phi_A + \phi_B}{\alpha + \beta}.$$

Proof. The expected payoff for player $i \in \{A, B\}$ in state (t_A, t_B) is

$$(1 - y(t_A, t_B))x_i(t_A, t_B) + y(t_A, t_B)(p_i(t_A, t_B) - \phi_i),$$

where $y(t_A, t_B)$ denotes the probability of war, and $x_i(t_A, t_B)$ is the expected share player i gets if there is no war. Let $V_i(t_i)$ denote the interim expected

[29] This mechanism does not have built-in knowledge of the state of the world. It is however "parametric" in the sense that it depends directly on the parameters (p_A, p_B), which are constant across states. But as long as p_A and p_B are common knowledge among the two players, more complicated *non-parametric* mechanisms will uniquely implement the constant state-contingent outcome $(x_A, x_B) = (p_A, p_B)$. There is no need for the mechanism designer to know (p_A, p_B).

payoff and $y_i(t_i)$ the interim probability of war for type t_i. By a standard argument, V_i is non-decreasing, and

$$V_A(1) - V_A(0) = \alpha \int_0^1 y_A(t_A) dt_A = \alpha E\{y\}, \tag{18}$$

where $E\{y\}$ is the ex ante probability of war. Similarly,

$$V_B(1) - V_B(0) = \beta E\{y\}. \tag{19}$$

Adding Equations (18) and (19) we get

$$(\alpha + \beta) E\{y\} = V_A(1) + V_B(1) - V_A(0) - V_B(0). \tag{20}$$

Since V_i is non-decreasing and there is one unit of the resource available,

$$V_A(0) + V_B(0) \leq 1. \tag{21}$$

If type $t_A = 1$ declares war without knowing the opponent's type, his expected payoff is

$$\int_0^1 p_A(1, t_B) dt_B - \phi_A = \mu + \alpha - \beta \int_0^1 t_B dt_B - \phi_A = \mu + \alpha - \frac{1}{2}\beta - \phi_A.$$

Therefore, the following IR condition must hold:

$$V_A(1) \geq \mu + \alpha - \frac{1}{2}\beta - \phi_A. \tag{22}$$

Similarly,

$$V_B(1) \geq 1 - \mu - \frac{1}{2}\alpha + \beta - \phi_B. \tag{23}$$

Substituting from Conditions (21), (22) and (23) into Equation (20) we obtain

$$E\{y\} \geq \frac{1}{2} - \frac{\phi_A + \phi_B}{\alpha + \beta}.$$

■

Thus, if the costs of war are small then the probability of war cannot be much less than one half. More generally, war must happen with strictly positive probability if $\phi_A + \phi_B < (\alpha + \beta)/2$. As shown by Bester and Wärneryd (2006) and Fey and Ramsay (2009, 2011), if the players disagree about the balance of power and war is not too costly, then no mechanism can guarantee peace. The problem is worse, the more uncertain is the balance of power, i.e., the bigger are α and β.

Theorem 1 applies to any bargaining game, but we can illustrate by assuming the players simultaneously choose Hawk or Dove. There is war if at least one player chooses Hawk, otherwise the status quo remains in place. Suppose there exists a peaceful equilibrium where each player chooses Dove with probability one, so each type t_i gets $V_i(t_i) = \omega_i$. If type $t_A = 1$ chooses Hawk then his expected payoff is $\mu + \alpha - \frac{1}{2}\beta - \phi_A$, so he prefers Dove if and only if

$$\omega_A \geq \mu + \alpha - \frac{1}{2}\beta - \phi_A. \tag{24}$$

Similarly, type $t_B = 1$ prefers Dove if and only if

$$\omega_B \geq 1 - \mu + \beta - \frac{1}{2}\alpha - \phi_B. \tag{25}$$

Since $\omega_A + \omega_B = 1$, these two inequalities imply $\phi_A + \phi_B \geq (\alpha + \beta)/2$. Conversely, if $\phi_A + \phi_B \geq (\alpha + \beta)/2$ then there exists $(\omega_A, \omega_B) \in X$ such that Conditions (24) and (25) are satisfied. In this case, there exists a perfectly peaceful equilibrium.[30]

5.2.1 Mediation

To maximize the probability of peace when the balance of power is uncertain, it may be necessary to use some form of mediation (as opposed to direct face-to-face negotiations). As Fey and Ramsay (2009, 2011) point out, negotiations may reveal information about the true balance of power. This new information may give some player an incentive to attack. The interim IR conditions may not be sufficient to prevent this, since they presume each player knows only his own type. Hörner, Morelli and Squintani (2015) show how this problem can be solved by a mediator. The mediator's recommendation hides information contained in the players' messages to him, so that attack decisions cannot be made contingent on this information.

[30] If $\phi_A + \phi_B < (\alpha + \beta)/2$ then there is no peaceful equilibrium and the game may unravel, as in Akerlof (1970). Consider equilibrium with a cutoff $z_j^* < 1$ such that player j chooses Hawk if $t_j > z_j^*$. Since peace must be voluntary on both sides, player i knows that his own choice only matters if player j chooses Dove, i.e., if $t_j \leq z_j^*$. This encourages player i to choose Hawk, since a weak player j is easier to defeat. This will drive down the equilibrium cutoff points, like adverse selection in the market for lemons.

5.2.2 Hard information

Military strength depends on many variables: the quality of soldiers, weapons, leadership and planning, the ability to finance a war and to attract allies, the morale of the population, etc. Some of these variables may be hard (verifiable) information. For example, whether a country has nuclear capabilities could be verified by weapons inspections. However, hard information may be kept secret in equilibrium. For example, revealing information about weapons systems and plans could be militarily disadvantageous (Fearon, 1995). A rebel group's strength may depend on its financial resources, but if it reveals the sources of its support then the government can block these sources (Walter, 2009). A militarily weak country can be exploited by its enemies, but a strong country may be considered a threat to others. By hiding its true strength, it may be able to deter its enemies without provoking security-seekers (see Section 7.1).

6 Commitment problems and power shifts

In this section we consider dynamic models without private information or first-mover advantages. Contracts are incomplete, in the sense that long-term binding agreements are impossible. Thus, it is impossible to commit to staying peaceful in the future. We will consider Fearon's (1995, 1998, 2004) and Powell's (2004a) argument that the interaction of commitment problems with anticipated power shifts can trigger preventive war.[31] In Section 6.1, these power shifts are exogenous. In Section 6.2, power shifts are caused by resource transfers.

6.1 Power and time

There is an infinite number of periods, $t = 1, 2, 3...$ The common discount factor is $\delta < 1$. The contested resource is perfectly durable. The share controlled by player B at the end of period $t-1$ is denoted x_{Bt-1} and is referred to as the state of the world. The set of possible states is $[0, 1]$ and the initial state is $x_{B0} = \omega_B$. At the beginning of each period t, as long as there was no war in the past, player A either starts a war or demands

[31] A war meant to prevent the opponent from getting a first-strike advantage is "preemptive". A war meant to prevent the opponent from growing stronger is "preventive".

a new share x_{At}. If player B accepts, then $x_{Bt} = 1 - x_{At}$ becomes the new state; players A and B get utility x_{At} and $x_{Bt} = 1 - x_{At}$ in period t, and we move to period $t + 1$. Rejection means war, and the winner is determined instantaneously by a costly lottery. Let p_{it} denote player i's exogenously given probability of winning a war in period t (there is no first-strike advantage). War ends the game: the winner controls all of the resource forever, which is worth $1/(1 - \delta)$ since it generates one unit of utility per period.

Suppose $p_{B\tau+1} > p_{B\tau}$ for some period $\tau \geq 1$. Thus, a power shift in favor of player B is expected to occur between periods τ and $\tau + 1$ (the players have perfect foresight). Assume δ is close enough to 1 so that

$$\frac{\delta p_{B\tau+1} - p_{B\tau}}{1 - \delta} > \phi_A + \delta \phi_B, \tag{26}$$

where ϕ_i is player i's cost of war.

Suppose, in order to derive a contradiction, that there exists a subgame perfect equilibrium without war. Along the equilibrium path, player B's continuation payoff in period $\tau + 1$ cannot be smaller than his expected payoff from war in period $\tau + 1$, which is

$$\frac{p_{B\tau+1}}{1 - \delta} - \phi_B.$$

Since player B must get at least 0 in period τ, his continuation payoff in period τ must be at least

$$\delta \left(\frac{p_{B\tau+1}}{1 - \delta} - \phi_B \right).$$

Since the resource is worth $1/(1-\delta)$, player A's continuation payoff in period τ cannot exceed

$$\frac{1}{1 - \delta} - \delta \left(\frac{p_{B\tau+1}}{1 - \delta} - \phi_B \right). \tag{27}$$

But player A's expected payoff from war in period τ is

$$\frac{1 - p_{B\tau}}{1 - \delta} - \phi_A. \tag{28}$$

Condition (26) implies that Equation (28) exceeds Equation (27), a contradiction. The power shift leads to a reduction in bargaining power which, by Condition (26), is unacceptable to player A. Player B is unable to appease player A, because current transfers are limited by the size of the resource (so player B must get at least 0 in period τ), and future allocations are not contractible. Thus, war must occur.

6.1.1 Power shifts in intrastate conflicts

Power shifts are often discussed in the context of civil war (Walter, 1997, Wagner, 2000, Fearon, 2004).[32] Suppose the current government represents a majority ethnic group which is challenged by a minority rebel group. The government does not have enough resources currently available to buy off the rebels; the dispute therefore mainly concerns resources that will become available in the future. The rebels may be reluctant to accept even a very favorable offer from a temporarily weak government, because they fear that the government will renege on the deal once it has regained its strength. In a democratic country, the contracting problem might be partially solved by constitutional protections for minorities, e.g., supermajority rules, enforced by an independent justice system. However, constitutional protections can be overturned (Fearon, 1998). Authoritarian leaders may find it even more difficult to make credible commitments (Gehlbach, Sonin and Svolik, 2016), although some long-lived leaders have successfully cultivated a reputation for sharing the spoils of power with opposition groups (Azam, 1995).

If a settlement requires the rebel army to demobilize, then this in itself implies a power shift:

> "The key difference between interstate and civil war negotiations is that adversaries in a civil war cannot retain separate, independent armed forces if they agree to settle their differences" (Walter, 1997, p. 337).

After a demobilization, the rebels would be defenseless (Walter, 1997). To protect themselves against opportunistic behavior from the government, the rebels must stay armed. In effect, this may mean a political partition. However, it will be costly for each side to maintain its own military force (Garfinkel and Skaperdas, 1990). Moreover, the minority and majority populations may be intermingled and difficult to separate geographically, or one group may be based in an area with too few resources to be viable. Empirically, conflicts are most prevalent among partially separated ethnic groups with poorly defined boundaries (Lim, Metzler and Bar-Yam, 2007), perhaps because interspersed settlements create first-mover advantages (Posen, 1993).

[32] Walter (2009) argued that informational asymmetries as well as commitment problems are especially pervasive in intrastate conflicts. For example, it may be very hard to find reliable information about the strength of a rebel army, its support among the population and its financial resources.

Acemoğlu and Robinson (2000, 2001) looked at class conflict as a commitment problem. When the poor are relatively strong they have a good chance at staging a successful revolution. The rich may not have sufficient wealth to buy off the poor, and promises of future transfers are not credible because of anticipated power shifts. Powell (2006) considered the formal connections between Fearon's and Acemoğlu and Robinson's work, and the conditions under which current resources are sufficient for a negotiated settlement.

6.2 Power and resources

In this subsection, wealth (i.e., control of resources) is power.[33] If player B controls a share x_B of the contested resource, then the probability that he wins a war is $p_B = \lambda(x_B)$. Player A wins with probability $p_A = 1 - \lambda(x_B)$. Here $\lambda(x_B)$ is the *contest success function* (CSF) which captures how player B's military strength or power depends on his wealth.[34] The status quo is (ω_A, ω_B), so player B's initial power is $\lambda(\omega_B)$.

6.2.1 Self-enforcing allocations

Following Jackson and Morelli (2007) and Beviá and Corchón (2010), suppose the players cannot make any binding commitment to stay peaceful. At allocation $(x_A, x_B) \in X$, player A's payoff from war is $1 - \lambda(x_B) - \phi_A$ and his payoff from peace is $u_A(x_A)$, so player A prefers peace if

$$u_A(x_A) \geq 1 - \lambda(x_B) - \phi_A. \tag{29}$$

Similarly, at allocation (x_A, x_B) player B prefers peace if

$$u_B(x_B) \geq \lambda(x_B) - \phi_B. \tag{30}$$

The allocation (x_A, x_B) is *self-enforcing* if Conditions (29) and (30) both hold.

To prevent war, the two players must negotiate a self-enforcing outcome. Since each player can choose to fight rather than negotiate, the negotiated outcome (x_A, x_B) must also be *individual rational*: at the status quo, both sides must prefer (x_A, x_B) to war. At the status quo, player A gets $1 -$

[33] Cf. Section 3.1.1, where a player who occupies a contested territory gains a fighting advantage θ.

[34] On contest success functions, see Skaperdas (1996).

$\lambda(\omega_B) - \phi_A$ from war, and player B gets $\lambda(\omega_B) - \phi_B$. Thus, (x_A, x_B) satisfies the IR conditions if

$$u_A(x_A) \geq 1 - \lambda(\omega_B) - \phi_A. \tag{31}$$

and

$$u_B(x_B) \geq \lambda(\omega_B) - \phi_B. \tag{32}$$

However, there may not exist any self-enforcing allocation (x_A, x_B) that satisfies the IR conditions, so war may be inevitable. Suppose player B is a *revisionist*. That is, at the status quo, player B prefers war to peace:

$$u_B(\omega_B) < \lambda(\omega_B) - \phi_B.$$

To prevent war, player B must be appeased with a transfer of resources. However, if the transfer makes him stronger and he cannot commit to staying peaceful, then player A may prefer preventive war to appeasement.

For example, suppose $u_i(x_i) = x_i$ and $\phi_i = 0.1$, for each $i \in \{A, B\}$. The CSF is $\lambda(x_B) = 0.5 + \zeta x_B$, where ζ is a constant such that $0 < \zeta \leq 0.5$. Further, suppose $\omega_B < 0.4/(1 - \zeta)$, which makes player B a revisionist who must be appeased in order to avoid war. To appease player B, the new allocation (x_A, x_B) must satisfy Condition (30). This condition requires

$$x_B \geq \frac{0.4}{1 - \zeta}. \tag{33}$$

However, player A is only willing to appease player B if Condition (31) holds. This condition requires

$$x_B \leq 0.6 + \zeta\omega_B. \tag{34}$$

Thus, if (x_A, x_B) is self-enforcing and individually rational, then x_B must satisfy both Condition (33) and Condition (34). Such x_B exists if and only if

$$\zeta\omega_B \geq \frac{0.4}{1 - \zeta} - 0.6.$$

War can be avoided if and only if this inequality holds. If $\zeta = 0.4$ then war is unavoidable if $\omega_B < 1/6$. If $\omega_B = 1/6$ then war is unavoidable if $\zeta > 0.4$. Intuitively, player A prefers war to appeasement if the status quo greatly favors him (ω_B is small) and the balance of power is very sensitive to changes in the wealth distribution (ζ is big).

6.2.2 Short-term commitments: the continuous case

Fearon (1996) found that war can be avoided if player B can make short-term commitments not to attack, as long as power is a continuous function of wealth.[35] Consider again the infinite horizon game from Section 6.1, but now wealth is power. There are no *exogenous* power shifts. If the state at the end of period $t-1$ is x_{Bt-1}, and player B rejects the new offer x_{Bt}, then there is a war which player B wins with probability $\lambda(x_{Bt-1})$, where λ is a *continuous* CSF. Thus, the expected payoffs from war in period t are

$$\frac{1-\lambda(x_{Bt-1})}{1-\delta} - \phi_A \tag{35}$$

for player A, and

$$\frac{\lambda(x_{Bt-1})}{1-\delta} - \phi_B \tag{36}$$

for player B. For simplicity, assume

$$\frac{\lambda(0)}{1-\delta} > \phi_B \tag{37}$$

so player B's expected payoff from war is always strictly positive.

There exists a Markov perfect equilibrium without war. In every state x_{Bt-1}, player A's equilibrium offer x_{Bt} makes player B indifferent between accepting and rejecting. Along the equilibrium path, player B's continuation payoff always equals his payoff from rejection, given by Equation (36). Since player B accepts in equilibrium and the social surplus is $1/(1-\delta)$, player A's expected continuation payoff at state x_{Bt-1} is

$$\frac{1}{1-\delta} - \left(\frac{\lambda(x_{Bt-1})}{1-\delta} - \phi_B\right).$$

This exceeds the expression in Equation (35), since $\phi_A + \phi_B > 0$. Thus, in each period t, player A prefers peace to war. To explicitly construct player A's strategy, note that if player B accepts the period t offer x_{Bt} then his continuation payoff will be

$$x_{Bt} + \delta\left(\frac{\lambda(x_{Bt})}{1-\delta} - \phi_B\right).$$

[35] See also Powell (1996, 2002) and Schwarz and Sonin (2008). Powell (1993) and Jackson and Morelli (2009) show how wars may occur in the equilibria of infinite horizon guns-versus-butter models with limits on transfers.

Setting this equal to his payoff from war in period t yields the difference equation

$$(1-\delta)\,x_{Bt} + \delta\lambda(x_{Bt}) = \lambda(x_{Bt-1}) - (1-\delta)^2\phi_B \tag{38}$$

with initial condition $x_{B0} = \omega_B$. Condition (37) implies that for each $x_{Bt-1} \in [0,1]$ there is $x_{Bt} \in (0,1)$ such that Equation (38) is satisfied.[36]

For example, suppose the CSF is $\lambda(x_B) = 0.5 + \zeta x_B$, with $0 < \zeta \leq 0.5$. Condition (37) requires $(1-\delta)\phi_B < \lambda(0) = 0.5$. The explicit solution to the difference equation is

$$x_{Bt} = x_B^* + (\omega_B - x_B^*)\,b^t \tag{39}$$

for all $t \geq 1$, where

$$x_B^* = \frac{0.5 - (1-\delta)\,\phi_B}{1-\zeta}$$

and

$$b = \frac{\zeta}{1-\delta(1-\zeta)}.$$

Note that $0 < x_B^* < 1$ and $0 < b < 1$ since $\zeta \leq 0.5$ and $(1-\delta)\,\phi_B < 0.5$. Along the equilibrium path, player A's offers are given by Equation (39); they are all accepted, and player B's share converges to x_B^*. If $\omega_B > x_B^*$ then player A uses *salami tactics*: in each period t, player B is asked to give up a slice $x_{Bt-1} - x_{Bt} > 0$.

If $\omega_B < x_B^*$, then player A uses an appeasement strategy: in each period t, player B is offered an additional slice $x_{Bt} - x_{Bt-1} > 0$. Along the equilibrium path, player A buys one period of peace at a time. Player B is always indifferent between accepting and rejecting the offers. When he accepts he receives the additional slice. If he could attack right away, with the additional strength from the new slice, then he would get strictly more than his equilibrium continuation payoff. However, by accepting, player B has made a short-term commitment to stay peaceful until player A can offer another slice. Thus, there is no war in equilibrium.[37]

[36] If Condition (37) did not hold, there would still exist an equilibrium without war, but player B might accept an offer of $x_{Bt} = 0$.

[37] Technological constraints may be a substitute for the ability to commit. Suppose player B cannot make any short-term commitment, but it takes one period of time to convert new resources into military power. If the CSF is continuous, then there is a Markov perfect appeasement equilibrium just like the one described. In period t player B receives $x_{Bt} - x_{Bt-1} > 0$. He does not want to attack in period t because his power is still only $\lambda(x_{Bt-1})$. In period $t+1$ his strength increases to $\lambda(x_{Bt})$, but at that time player A will appease him with another slice.

6.2.3 Short-term commitments: the discontinuous case

If the CSF is discontinuous, then war may be unavoidable even if short-term commitments are possible (Fearon, 1996). Consider the same infinite horizon game, but now the CSF is

$$\lambda(x_B) = \begin{cases} \lambda_L & \text{if } x_B < \chi \\ \lambda_H & \text{if } x_B \geq \chi \end{cases}$$

for some χ such that

$$\omega_B < \chi < \lambda_L < \lambda_H.$$

Assume δ is close enough to 1 so that

$$\frac{\lambda_L - \chi}{1-\delta} > \phi_B \tag{40}$$

and

$$\frac{\delta\lambda_H - \lambda_L}{1-\delta} > \phi_A + \delta\phi_B. \tag{41}$$

Suppose, in order to derive a contradiction, that there exists a subgame perfect equilibrium without war. Let x_{Bt} denote player B's period t share in this equilibrium. Then $x_{Bt} \geq \chi$ in some period $t \geq 1$, for otherwise player B would get less than $\chi/(1-\delta)$, and then he would prefer war, by Condition (40). Since $\omega_B < \chi$ there is t such that $x_{Bt-1} < \chi \leq x_{Bt}$. Player B's expected payoff from war in period $t+1$ would be

$$\frac{\lambda(x_{Bt})}{1-\delta} - \phi_B = \frac{\lambda_H}{1-\delta} - \phi_B.$$

Player A's continuation payoff at period $t+1$ can therefore be at most

$$\frac{1}{1-\delta} - \left(\frac{\lambda_H}{1-\delta} - \phi_B\right) = \frac{1-\lambda_H}{1-\delta} + \phi_B \tag{42}$$

Thus, at the beginning of period t, player A's expected continuation payoff along the equilibrium path is at most

$$1 + \delta\left(\frac{1-\lambda_H}{1-\delta} + \phi_B\right). \tag{43}$$

Since $x_{Bt-1} < \chi$, player A's expected payoff from a war in period t is

$$\frac{1-\lambda_L}{1-\delta} - \phi_A. \tag{44}$$

Condition (41) implies that Equation (44) exceeds Equation (43), a contradiction.

Condition (40) implies that the status quo is unacceptable to player B. But appeasement would trigger a power shift which, by Condition (41), would be unacceptable to player A. The discontinuity means there is no mutually acceptable middle ground. Short-term commitments cannot solve this problem, so war is unavoidable.

6.2.4 Power, resources and deterrence

Consider a three-stage game, based on Gurantz and Hirsch (2017). Player A is the *status quo power* and player B the *challenger*. In stage 1, player B demands x_B. In stage 2, player A accepts or rejects the demand. If player A rejects, then there is war which player B wins with probability $\lambda(\omega_B)$, and the game ends. If player A accepts, then we move to stage 3. In stage 3, player B can either do nothing, in which case the final payoffs are $u_A(1-x_B)$ and $u_B(x_B)$, or start a war which he wins with probability $\lambda(x_B)$.

The game captures the logic of Section 6.2.1, as player B cannot commit to staying peaceful if player A accepts his demand. However, as in the ultimatum game, player B is committed to fight if player A rejects his demand.

Assume λ is continuous and strictly increasing. Player A is not a revisionist:

$$u_A(\omega_A) > 1 - \lambda(\omega_B) - \phi_A. \tag{45}$$

Player B's cost of war is either ϕ_B^L or ϕ_B^H, where

$$\lambda(\omega_B) - \phi_B^H < u_B(\omega_B) < \lambda(\omega_B) - \phi_B^L. \tag{46}$$

Thus, player B is a revisionist if and only if $\phi_B = \phi_B^L$. Finally, assume that if $\phi_B = \phi_B^L$ then no individually rational allocation is self-enforcing (as defined in Section 6.2.1).

Let us for the moment suppose there is complete information. Thus, player B's true cost of war is common knowledge. Then there exists a peaceful subgame perfect equilibrium if and only if there exists a self-enforcing and individually rational allocation. Thus, war is unavoidable if player B's cost of war is low ($\phi_B = \phi_B^L$). Suppose instead that his cost is high ($\phi_B = \phi_B^H$). In this case, the status quo (ω_A, ω_B) is strictly self-enforcing and strictly individually rational, from (45) and (46).[38] Let x_B^* be the largest x_B such

[38] That is, all relevant inequalities, as defined in Section 6.2.1, are strict.

that $(1 - x_B, x_B)$ is both self-enforcing and individually rational. In the unique subgame perfect equilibrium, player B demands x_B^*, player A accepts, and player B does nothing in stage 3. Thus, player B uses his first-mover advantage to increase his share to $x_B^* > \omega_B$, although it is commonly known that his cost of war is high, so he is not a revisionist. Player B is not deterred from doing this, because he knows that player A will accept x_B^*. This is the problem of credible deterrence (Schelling, 1960).

Gurantz and Hirsch (2017) argued that uncertainty about the challenger's preferences can facilitate credible deterrence.[39] Thus, suppose player B's true cost of war, $\phi_B \in \left\{\phi_B^L, \phi_B^H\right\}$, is his private information. The following is a separating perfect Bayesian equilibrium. Player A accepts any demand $x_B \leq \omega_B$. But if player B demands $x_B > \omega_B$ then player A thinks $\phi_B = \phi_B^L$ and rejects the demand. If player B's cost of war is low then he challenges the status quo by demanding $x_B > \omega_B$, so there is war. But if his cost is high then he demands ω_B, and there is no war.

Player A's threat to reject any demand $x_B > \omega_B$ is credible for the following reason. If player B demands $x_B > \omega_B$, then player A concludes that he is facing a revisionist, so $(1 - x_B, x_B)$ cannot be both individually rational and self-enforcing. Either x_B is so big that $u_B(x_B) \geq \lambda(x_B) - \phi_B^L$, in which case $u_A(1 - x_B) < 1 - \lambda(\omega_B) - \phi_A$, so player A prefers to reject. Or else $u_B(x_B) < \lambda(x_B) - \phi_B^L$, but then player A expects player B to attack in stage 3 when his power has increased to $\lambda(x_B)$, so player A prefers a preventive war in stage 2 when player B's power is only $\lambda(\omega_B) < \lambda(x_B)$. Thus, if player B's cost of war is high, then he is credibly deterred from challenging the status quo.

Gurantz and Hirsch's (2017) theory of credible deterrence combines incomplete information and power shifts. If a status quo power is expected to resist an initial challenge, then only a highly aggressive opponent will make the challenge. If the initial challenge is not resisted, then such a dangerous (perhaps "crazy") opponent will become even more dangerous in the future. This renders it credible to resist the initial challenge, even if not much is directly at stake, and even if it is *a priori* very unlikely that the opponent is

[39] Credible deterrence can also be facilitated by the status quo power's reputational concerns in repeated interactions. That is, he may resist an initial challenge in order to create a reputation for toughness. In a multi-player scenario, he may resist an initial challenge in order to send a message to other potential opponents (Kreps and Wilson, 1982, Milgrom and Roberts, 1982).

"crazy".[40]

7 Concluding remarks

This selective survey has omitted many important topics. In closing, we will mention a few.

7.1 Guns or butter

Wars have sometimes been blamed on *arms races*. A security-seeker who arms himself for self-protection makes others feel less secure, and an arms race leads to increased fear all around. This is the *security dilemma* or *spiral model* (Jervis, 1976). Kydd (1997) formalized the security dilemma as a game of incomplete information, and showed that refraining from arming can signal peaceful intentions and reduce the risk of war. This was discussed in Section 4.2.2. In other parts of the survey, we did not consider so-called guns-or-butter choices, i.e., the allocation of resources between military and civilian uses. For example, in Section 6 we did not allow a player to compensate for an anticipated power shift by a military build-up. Powell (2012) argued that this strong assumption may have important consequences.

Guns-or-butter decisions have been analyzed by, among others, Garfinkel (1990), Hirshleifer (1988), Neary (1997), Skaperdas (1992) and Skogh and Stuart (1982). A basic guns-or-butter model has complete information and a single period with two stages. In the first stage, the two players simultaneously choose to be armed or unarmed. These decisions are publicly observed. To be armed is costly, but it makes the player militarily stronger. In the second stage, they bargain over the available resources. In case of disagreement, there is war, and the probability of winning depends on military strength.

[40] According to Acheson (1958), the US can credibly threaten to use nuclear weapons if there is an attack on western Europe. By making such an attack in spite of the nuclear threat, the enemy would reveal his true type, and the US would have to conclude that a nuclear showdown is inevitable:

> "Here, in effect, he (our potential enemy) would be making the decision for us, by compelling evidence that he had determined to run all risks and force matters to a final showdown, including (if it had not already occurred) a nuclear attack upon us" (Acheson, 1958, p. 87).

In equilibrium, there is no war.[41] But each side has an incentive to arm, because military power is bargaining power (as emphasized by Schelling, 1966), and arms control agreements are not enforceable. Garfinkel and Skaperdas (1990) showed how war emerges as an equilibrium outcome in a two-period version of the basic model. If there is no war in period 1, then in period 2 there will be costly arming just as in the basic one-period model. If there is war in period 1, then the loser is utterly defeated and the winner can save the cost of arming in period 2.[42] If the expected savings are large compared to the cost of war, then there will be war in period 1.

Meirowitz et. al. (2019) considered a version of the basic model where the stage 1 decision to arm is unobserved and there is Nash demand bargaining in stage 2. Armed players take advantage of unarmed opponents by demanding large shares of the contested resource. But when two armed players meet, their demands are incompatible so there is war. For a given probability of arming, peace talks can reduce the probability of war. Armed players can still take advantage of unarmed opponents, but two armed players can reduce the risk of war by exchanging cheap-talk messages before making their demands. However, since this benefits armed players, peace talks increase the incentive to arm, and the probability of war actually increases when arming decisions are taken into account. Meirowitz et. al. (2019) showed that *mediated* peace talks can solve the problem. Cheap-talk messages are sent privately to a mediator who makes proposals that do not fully reveal these messages. This secrecy favors the unarmed, and therefore reduces the incentive to arm.

Baliga and Sjöström (2008, 2012a) considered a model where the decision to arm is unobserved, but hard (verifiable) information about armaments can be produced by arms inspections. If player A is a security-seeker, then he may attack to eliminate a perceived threat if inspections reveal that player

[41] If bargaining is impossible, and stage 2 simply consists of attack decisions, then there may not exist any peaceful equilibrium. Suppose there are three possible armaments levels: High, Middle and Low. A player is able to attack and defeat the opponent if and only if he has a High level and the opponent a Low level. If there is an equilibrium without war, then neither player will choose High, since Middle is cheaper and sufficient for deterrence. But then they are safe from attack even with the cheapest level, Low, so both must choose Low. However, the best response against an opponent who chooses Low may be to choose High and attack. In this case, there is no peaceful equilibrium. Jackson and Morelli (2009) show that a similar logic holds in an infinite horizon model.

[42] Determining which player should "disappear" is analogous to the problem of allocating an indivisible resource. As mentioned in Footnote 7, a coin flip would not be a credible solution.

B is armed. But if player A is greedy, then he may attack in order to grab player B's oil fields if inspections reveal that player B is unarmed. If player B is not sure about player A's type, then player B's best option may be strategic ambiguity: he refuses arms inspections whether he is armed or not. Strategic ambiguity can prevent war, because a greedy player A can be "deterred by doubt" by an unarmed player B. This reduces the incentive for player B to arm, which in turn makes a security-seeking player A feel more secure. Thus, uncertainty about military strength can be good for peace. Cheap-talk messages can be used to trigger inspections when they are especially valuable to player A. But to preserve incentive compatibility, a message which makes inspections more likely must increase the probability that player B arms.

7.2 Institutional explanations

In this survey we treated each side of the conflict as a *unitary actor*. In reality, there may be differences of opinion within each side, and political leaders take domestic politics into account when they decide to go to war. Fearon (1994) argued that democratically elected leaders who back down during an international crisis suffer large "audience costs". Therefore, compared to autocracies, democracies are less likely to back down. Bueno de Mesquita et. al. (1999) considered a guns-or-butter trade-off: a country's leader can allocate resources either to a war effort, or to boost the private consumption of a *winning coalition* whose support he needs to stay in power. In an autocracy, the winning coalition is small, so its *per capita* cost of an increased war effort is large. Thus, autocrats are unwilling to devote resources to fighting, and they are unlikely to lose power if a war is lost. Democratically elected leaders suffer more from foreign policy failures, so they avoid wars where the chance of losing is significant, but they spend more on winning the wars they do fight. On average, autocrats may be more or less hawkish than democratically elected leaders. However, since democracies fight harder, they do not like to fight each other. The model of Bueno de Mesquita et. al. (1999) is consistent with the *democratic peace* (empirically, democratic states rarely if ever fight each other), and the fact that, compared to autocracies, democracies are more likely to win wars.

If an autocrat receives a disproportionate share of the gains from war, while the population at large pays the cost, then he has a *political bias* in favor of war. A democratically elected leader may be less biased, and therefore less inclined to go to war. Jackson and Morelli (2007) showed that if commitments are impossible (as in Section 6.2.1) and the probability of winning a war is proportional to wealth, then two countries will coexist peacefully if both leaders are unbiased, but not if they are biased in favor of war.[43] Also, an unbiased leader wants to fight only when the odds are in his favor. Thus, this model is also consistent with the democratic peace and the fact that democracies are more likely to win the wars they do fight.

Baliga, Lucca and Sjöström (2011) considered a Hawk-Dove game with strategic complements. War may be triggered by the fear of being attacked, as discussed in Section 4.2. If the representative citizen is a fearful security-seeker, then a democratically elected leader inherits this fearfulness. In an autocracy, with a winning coalition consisting of hawkish individuals who will punish foreign policy failures, the leader is even more fearful of being attacked. But a dictator who thinks he can stay in power whatever happens will be less inclined to choose Hawk out of fear. This model predicts a non-monotonic relation between democracy and peace, which Baliga, Lucca and Sjöström (2011) found empirical evidence for.

A country's political leader typically has information which is not available to the representative citizen (Downs and Rocke, 1994, Hess and Orphanides, 1995). Therefore, even in a fully democratic country, to induce the leader to act in the interests of the citizen is a complicated principal-agent problem. Given the information asymmetries, a leader may increase his chances of staying in power by making a decision which is bad for the citizen. For example, a leader who privately knows that he has good war leadership skills may start an unnecessary war is order to reveal these skills (Hess and Orphanides, 1995). Conversely, a leader who makes the best possible decision for the citizen, given the leader's information, may lose office. For example, if an ongoing war is unlikely to end in victory, the best decision may be to terminate it. But if this would cause the leader to lose office, he may decide to prolong the war and "gamble for resurrection" (Downs and Rocke, 1994).

[43] In the unitary actor model of Section 6.2.1, the status quo is both self-enforcing and individually rational if $\lambda(\omega_B) = \omega_B$ and $u_i(\omega_i) = \omega_i$ for each i.

7.3 How long will war last?

In this survey we considered war to be a *costly lottery* where the winner gets all of the contested resource. In reality, most wars end with a negotiated settlement. Wittman (1979) provided an early discussion of the conditions under which a settlement would be reached. More recent models of negotiation during war include Wagner (2000), Filson and Werner (2002), Fearon (2007), Smith and Stam (2004) and Heifetz and Segev (2005).

Fearon (1997) considered civil war in an infinite horizon model with power shifts and incomplete contracts. During a war period, the rebels may pause the fighting so that negotiations can take place during a period of peace. However, the government's bargaining power may increase during the peace period. If the cost of fighting is sufficiently small and the government is weak, then the rebels would rather continue fighting than risk a power shift. Intuitively, if negotiations require a cease-fire, then the side that has the momentum in the war may refuse to negotiate, because a cease-fire can ruin the momentum.[44] The war will continue until one side is defeated, which means a long war is possible.

Powell (2004b) considered the case of one-sided private information in a model where war is a sequence of indivisible battles of fixed length. Before each battle, uninformed player A makes repeated offers, and the war ends if an offer is accepted. If player A could commit to a sequence of offers, then he could present player B with an ultimatum: accept a low offer now, or fight a long time before getting a better offer. Just as in Section 5.1, player A would trade off rent-extraction against losses of social surplus, and long wars would be possible in equilibrium. However, Powell (2004b) assumed player A cannot commit to a sequence of offers. He found that if player B's private information is his cost of war, then a version of the *Coase conjecture* (Coase, 1972) holds: if negotiations are intense, in the sense that the interval between offers is very short, then the war will be very short. Since a rejection signals that player B's cost is low, player A has an incentive to quickly make a more generous offer. As the periods shrink, player A's commitment ability (which is the source of his bargaining power) vanishes, making it impossible to extract rent at the expense of social surplus.[45]

[44]In 1847, US troops halted their advance on Mexico City in order to negotiate a peace treaty, but they "paused too long while the enemy regrouped" (Schelling, 1966, p. 128).

[45]In economics, the Coase conjecture has been studied in the context of a buyer-seller model (Gul and Sonnenschein, 1988, Deneckere and Liang, 2006). Bargaining during

Powell (2004b) also considered the case where player B's private information is his military strength, formalized as the probability of collapse during battle. In this case, intense bargaining before the battle does not have the usual negative effect on player A's bargaining power, so long wars are possible. Fearon (2013) dropped the assumption of indivisible battles and instead assumed player B can collapse in any period of bargaining. Player B's private information is his probability of collapse. Fearon found that the Coase conjecture would not be valid if, in addition to the inability of player A to commit to a sequence of offers, he also cannot commit to any long-term peace agreement which specifies how the resource is to be shared. Baliga and Sjöström (2022) found that even if he can commit to a peace agreement, and negotiations are intense, the war may be long. If player B is militarily strong, then he is convinced that he can win on the battlefield, so he will require a very large share of the resource in order to settle. If the war is to end very quickly, then player A must offer this very large share very quickly. But then player B can get a very large share, without much fighting, even if he is militarily weak. In such an equilibrium, player A's equilibrium payoff must be very low. But if the prior probability that player B is weak is sufficiently large, then player A gets a high payoff from the feasible strategy of refusing to negotiate. Since this is a contradiction, in equilibrium wars cannot end quickly.

8 References

Acharya, Avidit, and Kristopher Ramsay. 2013. "The Calculus of the Security Dilemma." *Quarterly Journal of Political Science* 8 (2): 183–203.

Acheson, D. (1958). *Power and Diplomacy.* Cambridge, MA: Harvard University Press

Akerlof, G. (1970) "The Market for Lemons: Quality Uncertainty and the Market Mechanism," *Quarterly Journal of Economics*, 84: 488–5

Azam, J.-P. (1995), "How to Pay for the Peace? A Theoretical Framework with References to African Countries", *Public Choice*, 83: 173-184

Acemoğlu, D. and J. Robinson (2000) "Why Did the West Extend the Franchise? Growth, Inequality and Democracy in Historical Perspective." Quarterly Journal of Economics CXV: 1167-1199.

conflict is different due to the "outside option" of fighting instead of bargaining, hoping or believing in victory on the battlefield.

Acemoğlu, D. and J. A. Robinson, (2001), "A Theory of Political Transitions," American Economic Review, vol.91, 938-963.

Baliga, S. and T. Sjöström, (2004): "Arms Races and Negotiations," Review of Economic Studies, 71, 351-369.

Baliga, S. and T. Sjöström, (2008), "Strategic Ambiguity and Arms Proliferation," Journal of Political Economy, vol.116, 1023-1058.

Baliga, S. and T. Sjöström, (2011), "Domestic Political Survival and International Conflict: Is Democracy Good for Peace?," Review of Economic Studies, April, 78,2:458-486.

Baliga, S. and T. Sjöström, (2012a), "The Hobbesian Trap," in The Oxford Handbook of the Economics of Peace and Conflict, edited by Michelle Garfinkel and Stergios Skaperdas. New York: Oxford University Press.

Baliga, S. and T. Sjöström. (2012b). "The Strategy of Manipulating Conflict" *American Economic Review* 102 (6): 2897–2922.

Baliga, S,. and T. Sjöström (2013) "Bargaining and War: A Review of Some Formal Models." *Korean Economic Review* 29 (Winter): 235-266.

Baliga, S. and T. Sjöström (2020): "The strategy and technology of conflict." Journal of Political Economy 128(8): 3186-3219.

Baliga, S. and T. Sjöström, (2022): "Long Wars," mimeo, Northwestern University and Rutgers University.

Bester, H. and K. Wärneryd, (2006), "Conflict and the Social Contract," Scandinavian Journal of Economics, vol.108, 231-249.

Beviá, C. and L. Corchón, (2010), "Peace Agreements without Commitment," Games and Economic Behavior, vol.68, 469-487.

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

Brito, D. and M. Intriligator, (1985), "Conflict, War and Redistribution," American Political Science Review, vol.79, 943-957.

De Mesquita, B. and R. Siverson (1995), "War and the Survival of Political Leaders: A Comparative Study of Regime Types and Political Accountability", *American Political Science Review,* 89, 841-855.

De Mesquita, B., J. Morrow, R. Siverson and A. Smith (1999). "An Institutional Explanation of the Democratic Peace," *American Political Science Review*, 93(4): 791-807.

Chassang, S. and G. Padro-i-Miquel (2010): "Conflict and Deterrence under Strategic Risk." *Quarterly Journal of Economics* 125 (4): 1821–58..

Chatterjee, K and W. Samuelson. "Bargaining under incomplete information." Operations research 31.5 (1983): 835-851.

Coase, R. H. (1972). Durability and Monopoly. The Journal of Law and Economics, 15(1), 143-149.

Corchón, L. (2007), "The Theory of Contests: a Survey," Review of Economic Design, vol.11, 69-100.

Crawford, Vincent. 1982. "A Theory of Disagreement in Bargaining." *Econometrica* 50 (3): 607-637.

Deneckere, R. and M.-Y. Liang (2006), "Bargaining with Interdependent Values," Econometrica, Vol. 74, No. 5, pp. 1309-1364..

Downs, G. and D. Rocke (1994), "Conflict, Agency and Gambling for Resurrection: The Principal-Agent Model Goes to War", *American Journal of Political Science*, 38, 362-80.

Ellingsen, T. and T. Miettinen. 2008. "Commitment and Conflict in Bilateral Bargaining." *American Economic Review* 98 (4): 1629-35.

Ellingsen, Tore and Topi Miettinen. 2014. "Tough Negotiations: Bilateral Bargaining with Durable Commitments." *Games and Economic Behavior* 87: 353-366.

Esteban, J. and D. Ray (2001), "Social decision rules are not immune to conflict," Economics of Governance, Springer, vol. 2(1), pages 59-67.

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

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

Fearon, J. (1996), "Bargaining over Objects that Influence Future Bargaining Power," Manuscript, Stanford University.

James D. Fearon, 1997, "Signaling Foreign Policy Interests: Tying Hands versus Sinking Costs," The Journal of Conflict Resolution, Vol. 41, No. 1

James D. Fearon, 1998, Commitment Problems and the Spread of Ethnic Conflict, in David A. Lake and Donald Rothschild, eds, The International Spread of Ethnic Conflict.

Fearon, J. (2004), "Why Do Some Civil Wars Last so Much Longer than Others?" *Journal of Peace Research*, 41: 275–301

Fearon, J. (2013), "Fighting Rather than Bargaining," mimeo, Stanford University.

James D. Fearon and David D. Laitin, 1996, "Explaining interethnic cooperation," American Political Science Review (1996): 715-735.

James D. Fearon and David D. Laitin, 2003, "Ethnicity, insurgency, and civil war," American Political Science Review (2003): 75-90.

Fey, M. and K. Ramsay, (2009), "Mechanism Design Goes to War: Peaceful Outcomes with Interdependent and Correlated Types," Review of Economic Design,.vol.13, 233-250.

Fey, M. and K. Ramsay, (2011), "Uncertainty and Incentives in Crisis Bargaining: Game-Free Analysis of International Conflict," American Journal of Political Science, vol.55, 149-169.

Filson Darren and Suzanne Werner (2002): “A bargaining model of war and peace: anticipating the onset, duration, and outcome of war,” Am. J. Polit. Sci. 46(4):819–38.

Fischer, F. (1967), *Germany's Aims in the First World War*. New York: WW Norton,

Garfinkel, M. (1990), "Arming as a Strategic Investment in a Cooperative Equilibrium," American Economic Review, vol.80, 50-68.

Garfinkel, M. and S. Skaperdas (1990) “Conflict without misperceptions or incomplete information: how the future matters”, *Journal of conflict resolution* 44: 793-807.

Gehlbach, S., Sonin, K. and M. Svolik (2016) "Formal models of nondemocratic politics." Annual Review of Political Science 19 (2016): 565-584.

Gul, F. and H. Sonnenschein (1988) "On Delay in Bargaining with One-sided Uncertainty." Econometrica Vol. 56, No. 3: 601-611.

Gurantz, R. and A. V. Hirsch (2017) “Fear, Appeasement and Effectiveness of Deterrence.” *Journal of Politics* 79 (3): 1041-1056.

Hassner, Ron (2003) "To halve and to hold: Conflicts over sacred space and the problem of indivisibility." Security Studies 12.4 (2003): 1-33.

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

Herrera, Helios, Massimo Morelli and Salvatore Nunnari. 2019. “A Theory of Power Wars.” CEPR Discussion Paper No. DP13917.

HESS, G. and ORPHANIDES A. (1995), “War Politics: An Economic, Rational-Voter Framework”, *American Economic Review*, 85, 828-846.

Higonnet, Patrice, 1968, The Origins of the Seven Years' War, The Journal of Modern History , Vol. 40, No. 1, pp. 57- 90

Hirshleifer, J. (1988), "The Analytics of Continuing Conflict," Synthese, vol.76, 201-233.

Hirshleifer, J. (1991), "The Paradox of Power," Economics and Politics, vol.3, 177-200.

Hirshleifer, J. (1994), "The Dark Side of the Force," Economic Inquiry, vol.32, 1-10.

Hobbes, T. (1886), *Leviathan*, Second Edition. London: Ballantyne Press.

Johannes Hörner, Massimo Morelli and Francesco Squintani (2015): "Mediation and Peace," Review of Economic Studies, 82, 1483-1501.

Iklé, F.C. (2005), Every War Must End, Revised Edition, New York: Columbia University Press

Jackson, M. and M. Morelli, (2007), "Political Bias and War", American Economic Review, vol.97, 1353-1373.

Jackson, M. and M. Morelli, (2009), "Strategic Militarization, Deterrence and Wars,".*Quarterly Journal of Political Science*, vol.4, 279–313.

Jervis, Robert. 1976. *Perception and Misperception in International Politics*, Princeton University Press: Princeton.

Jervis, Robert. 1978. "Cooperation Under the Security Dilemma." *World Politics* 30 (2): 167-214.

Kambe, Shinsuke. 1999. "Bargaining with Imperfect Commitment." *Games and Economic Behavior* 50: 321-342.

Kreps, D. and Wilson, R. (1982). "Reputation and Imperfect Information", Journal of Economic Theory 27, 253–279.

Kydd, A. (1997), "Game Theory and the Spiral Model", *World Politics* 10:371–400.

Lim, May, Richard Metzler, and Yaneer Bar-Yam (2007). "Global pattern formation and ethnic/cultural violence." Science 317.5844: 1540-1544.

Mas-Colell, A., Whinston, M. and J. Green (1995), *Microeconomic Theory.* New York: Oxford University Press

Meirowitz, Adam, Massimo Morelli, Kristopher W. Ramsay and Francesco Squintani. 2019. "Dispute Resolution Institutions and Strategic Militarization." *Journal of Political Economy* 127 (1): 378-418.

Milgrom, P., and Roberts, J. (1982). "Predation, Reputation and Entry Deterrence", Journal of Economic Theory 27, 280–312.

Myerson, R. and M. Satterthwaite, (1983). "Efficient Mechanisms for Bilateral Trading," Journal of Economic Theory, vol. 29, 265-281.

Nash, J. (1953), "Two-Person Cooperative Games", *Econometrica* 21: 128-140.

Neary, H. (1997), "Equilibrium Structure in an Economic Model of Conflict," Economic Inquiry, vol.35, 480-494.

Oye, K. (1986). *Cooperation Under Anarchy.* Princeton, N.J.: Princeton University Press.

Powell, R. (1993), "Guns, Butter and Anarchy," American Political Science Review, vol.87, 115-132.

Powell, R. (1996), "Uncertainty, Shifting Power, and Appeasement," American Political Science Review, vol.90, 749-764.

Powell, R. (1999.) *In the Shadow of Power.* Princeton, NJ: Princeton University Press.

Powell, R. (2002): "Bargaining theory and international conflict," Annu. Rev. Polit. Sci. 5:1–30

Powell, R. (2004a), "The Inefficient Use of Power: Costly Conflict with Complete Information," American Political Science Review, vol.2, 231-241.

Powell, R. (2004b), "Bargaining and Learning while Fighting," American Journal of Political Science, vol.48, 344-361.

Powell, R. (2006), "War as a Commitment Problem," International Organization, vol.60,169-203

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

Ramsay, Kristopher. 2017. "Information, Uncertainty, and War." *Annual Review of Political Science* 20: 505–27.

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

Schelling, T. (1966). *Arms and Influence.* New Haven: Yale University Press.

Schwarz, M. and K. Sonin (2008). "A theory of brinkmanship, conflicts, and commitments," *Journal of Law, Economics, & Organization*, 24(1): 163-183.

Skaperdas, S. (1992), "Cooperation, Conflict and Power in the Absence of Property Rights," American Economic Review, vol.82, 720-739.

Skaperdas, S. (1996), "Contest Success Functions," Economic Theory, vol.7, 283-290.

Skogh, G. and C. Stuart, (1982), "A Contractarian Theory of Property Rights and Crime," Scandinavian Journal of Economics, vol.84, 27-40.

Slantchev, B. (2003) "The Power to Hurt: Costly Conflict with Completely Informed States," *American Political Science Review*, .97: 123-133.

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

Wagner, R. Harrison (2000). "Bargaining and war." American Journal of Political Science (2000): 469-484

Walter, Barbara (1997). The critical barrier to civil war settlement. Int. Organ. 51(3):335-364

Walter, Barbara (2009). "Bargaining failures and civil war." Annual Review of Political Science 12.1 (2009): 243-261.

Walter, Barbara (2013) "Civil Wars, Conflict Resolution and Bargaining Theory." Handbook of International Relations, 99 656-672.

Williamson, O.E. (1985), The Economic Institutions of Capitalism. New York, NY: Free Press.

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