Causes of War*

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

February 2, 2024

Abstract

Even though war is not a Pareto optimal outcome, attempts to negotiate a peaceful settlement may be unsuccessful. Even if the ability to write and enforce contracts is unrestricted, the optimal bargaining strategy may involve a calculated risk of war, for example by creating or exploiting a *fait accompli*. In general, the scope for such opportunistic behavior depends on the bargaining procedure and on the information asymmetries. With asymmetric information about the balance of power, no bargaining procedure can guarantee peace. Enforcement problems increase the risk of war, since peace treaties become vulnerable to opportunistic behavior *ex post*. Lacking the assurance of permanent peace, a security-seeker may attack out of fear that the opponent will strike first (Hobbesian trap) or is becoming too powerful (Thucydides trap).

Keywords: War, Bargaining, Asymmetric information, Commitment, Incomplete contract, Hobbesian trap, Thucydides trap, Strategic complements, First-mover advantage, Self-enforcing allocation.

1 Introduction

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

*We are grateful to Sungmin Kang for helpful comments.
†Email: baliga@kellogg.northwestern.edu
‡Email: tsjostrom@economics.rutgers.edu

> "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).

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).

Thomas Schelling showed that game theory can serve as a formal framework for analyzing conflicts as bargaining problems. In his pioneering book *The Strategy of Conflict* (Schelling, 1960), he emphasized that bargaining involves both common and opposing interests: both sides can benefit if the expected (social) surplus is large, but each wants a bigger share at the other's expense. There is typically a risk-reward trade-off: aggressive bargaining tactics yield a big reward (a large share of the surplus) if the opponent backs down, but there is a risk of war (a reduced expected surplus) if the opponent does not. 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

[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 due to the unforseen actions of local decision makers, 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).

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 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). The players are unsure about whether commitment attempts will be successful, but they have no private information. A successful commitment may bring about a first-mover advantage, but if the opponent makes an incompatible commitment then it may lead to war. Elimination of dominated strategies leads to 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 considers the case where commitment costs can be large but are privately known. In equilib-

---

Similarly, when military forces of China and India began pushing into contested frontier regions, an unwanted war resulted from

> "the movement and stationing of Chinese and Indian security personnel. They acted in a competitive fashion, and incidents were bound to occur, particularly because jurisdictions and border markings had never been jointly defined. The Indians hardened their stance on the borders after a major incident occurred not in the Assam Himalaya but near the Aksai Chin. The conflict spiral possessed a momentum of its own and culminated in the Indian-Chinese border war of October-November 1962" (Hoffmann, 2006, p. 183).

rium, players with costs below a cut-off make a commitment, generating a strictly positive probability of war.

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).

It is useful to distinguish the fear of an immediate attack from the fear of a long-term shift in the balance of power. A war launched in order to preempt an immediate attack is a *preemptive war*; a war launched in order to prevent an adverse shift in the balance of power is a *preventive war*. A preemptive war requires the (perceived) existence of a first-strike (offensive) advantage, i.e., the technology of war must favor the offense.[3] First-strike advantages are considered in Section 4.2.

If there is a first-strike advantage then a security-seeker, who prefers the status quo to war, may attack preemptively. The decision to attack is caused by fear and distrust, like hunting rabbit in a game of stag hunt. World War I is often cited as an example:[4]

[3] A first-strike advantage may not be all-or-nothing. If the opponent strikes first, it may be crucial to respond as quickly as possible. Each side will therefore look for signs that the other side has decided to strike, or has already done so, and there is no time to resolve misunderstandings and misperceptions. A malfunctioning early warning system could cause a country to fire its missiles in the mistaken belief that a war has already started. This kind of accidental war is discussed in Acemoğlu and Wolitzky's chapter in this Volume.

[4] 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.

> "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).

In Section 4.2 we consider a Hawk-Dove game where Hawk represents a decision to attack. Actions are strategic complements due to a first-strike advantage. Each player's cost of choosing Hawk is his private information (his type), with a small chance of being a dominant strategy hawk (who surely attacks). This situation may lead to preemptive war among security-seekers, which we refer to as the *Hobbesian trap*.[5] Security-seekers with a sufficiently small cost-type choose Hawk for fear of dominant strategy hawks, and this triggers a cascade of security-seekers with higher and higher costs who all choose Hawk. Under some assumptions, *all* types will choose Hawk. This outcome is Pareto inefficient, since all types prefer (Dove,Dove) to (Hawk,Hawk). The outcome may be improved by costly signaling (Kydd, 1997), or by cheap talk (Baliga and Sjöström, 2004), but in general the Hobbesian trap cannot be avoided without a binding peace treaty. Even if binding agreements are feasible, however, opportunistic behavior may cause bargaining failure. With ultimatum bargaining, the first-mover makes a proposal which corresponds to a *fait accompli*. Since he does not know his opponent's type, he faces a risk-reward trade-off: an aggressive proposal generates a high expected reward at the cost of an increased risk of war. In equilibrium, war may occur with strictly positive probability, even if there are no obstacles to making a binding agreement. However, there are other

[5] If both players were commonly known to be dominant strategy hawks, then the game would be a prisoner's dilemma. But Hobbes did not consider the "war of every man against every man" in a state of nature to be the outcome of a prisoner's dilemma. In his view, most people would prefer peace if they had assurance that others would stay peaceful (Baliga and Sjöström, 2012a). The problem is that they cannot have such assurance, because they cannot know the intentions and motives of others, and wicked people (dominant strategy hawks) do exist:

> "For though the wicked were fewer than the righteous, yet because we cannot distinguish them, there is a necessity of suspecting, heeding, anticipating, subjugating, self-defending, ever incident to the most honest and fairest conditioned" (Hobbes, 1983, p. 100).

Thus, the Hobbesian trap is caused by the interaction of first-strike advantages and incomplete information.

bargaining games that guarantee peace. This raises the issue of the design of bargaining procedures, which will be discussed in Section 5.

In Section 5, we consider in more detail the case where a binding peace agreement can be negotiated before the players have a chance to launch an attack. There is asymmetric information, but for simplicity, no first-strike advantage. The study of bargaining with private (asymmetric) information was pioneered by Chatterjee and Samuelson (1983) and Myerson and Satterthwaite (1983). They showed how information asymmetries can prevent traders from finding a mutually acceptable price. By analogy, privately informed diplomats may struggle to find a mutually acceptable compromise. Blainey (1988) argued that this struggle may cause each side to conclude that it can gain more by fighting than by negotiating:

> "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 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).

It is important to distinguish between private information about preferences (e.g., the cost of fighting) and private information about military capabilities (e.g., the quality of weapons). By definition, military capabilities determine the likely outcome of a war (the balance of power), while preferences determine the evaluation of different outcomes.

In Section 5.1 we assume there is asymmetric information only about preferences. In the ultimatum game, the first-mover makes a proposal and the responder (second-mover) is forced to choose between war and appeasement.[6] With uncertainty about the responder's type, the proposer faces a risk-reward trade-off, and the optimal proposal may be such that war happens with strictly positive probability (Fearon, 1995).[7] War destroys social surplus, but it is required in order to screen the responder's types. Thus, even though war is not a Pareto optimal outcome in the standard sense, the ultimatum game is *incentive-efficient*.

If the balance of power is common knowledge, i.e., independent of the players' types, then there exists a compromise outcome that is individually rational for all types. That is, each type of each player prefers this compromise to his outside option (war). Specifically, resources should be allocated in proportion to military strength.[8] If an impartial mediator proposes this compromise, then both players will accept it. Thus, a simple bargaining procedure can guarantee peace, with no scope for opportunistic behavior (Fey and Ramsay, 2011). In fact, if there is no first-strike advantage then there is no need for a binding peace treaty. The compromise outcome is self-enforcing,

[6] What makes the threat of war credible? Perhaps the "proposal" corresponds to a strategic move where the first-mover occupies a contested territory, and backing down would be prohibitively costly. However, unlike in Section 3, in the ultimatum game the identity of the first-mover is determined exogenously.

[7] Blainey (1988) argued that an unintended war occurs when a risk-taker underestimates the opponent's willingness to fight:

> "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).

[8] 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 must be 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.

since each player prefers his share to war.[9]

In Section 5.2 we consider bargaining with asymmetric information about the balance of power (i.e., about military capabilities). In this case, there may be no compromise outcome that is individually rational for all types.[10] To be pacified, strong types (who expect to prevail on the battlefield) must get large shares of the available surplus. This means information about the true balance of power must be revealed during the negotiations. But to prevent weak types from exaggerating their strength, there must be a risk-reward trade-off: a strong type can get a larger share only by accepting a risk of war. It follows that no bargaining procedure can guarantee peace, even though it is feasible to sign a binding peace treaty and there is no first-strike advantage (Bester and Wärneryd, 2006, Fey and Ramsay, 2009, 2011). This supports Blainey's (1988) argument that wars are caused by disagreements about the balance of power.

In Section 6 we consider the interaction between commitment problems and anticipated power shifts. If long-term peace agreements cannot be enforced, then war may be inevitable, even if there are no first-strike advantages and no private information (Fearon, 1995, 1998). In International Relations theory, the most famous example is the *Thucydides trap*, a scenario where a status quo power faces a rising power. A similar trap may occur in a civil conflict, where a weak government must appease a rebel group in order to avoid a civil war. To do so, it must promise the rebels a large permanent share of the country's resources. However, if the government is expected to be stronger in the future (an exogenous power shift) and cannot make long-run commitments, then it has a credibility problem. If the rebels think the government will renege on the deal when it gets strong enough, then appeasement may be impossible.

Suppose instead that no exogenous power shift is expected, but resources make the rebels stronger because "wealth is power". If the rebels cannot commit to staying peaceful, then they have a credibility problem. Appeasement attempts may be self-defeating, because transferring resources to the rebels

[9] With a first-strike advantage it becomes a matter of timing: war may be unavoidable if it is possible to launch a surprise attack before a peace treaty has been signed. See Chassang and Padró i Miquel (2009) for a dynamic model with stochastic changes in the (opportunity) cost of fighting.

[10] Analogously, the Myerson and Satterthwaite (1983) impossibility theorem relies on the assumption that there is no price that is acceptable to each type of each trader.

will only make them a bigger threat. We refer to this as *Ethelred's trap*.[11] In this situation, the government may prefer war to appeasement (Jackson and Morelli, 2007, Beviá and Corchón, 2010). However, if power is a continuous function of wealth and the rebels can make short-term commitments, then the government may prefer to avoid war by a sequence of short-term peace deals (Fearon, 1996, Schwarz and Sonin, 2008).

In this survey, we make a number of simplifying assumptions: two unitary actors contest a perfectly divisible resource, with no misperceptions or bounded rationality, no guns versus butter trade-off, and with war as a game-ending costly lottery. These assumptions have all been relaxed in the literature. 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. More generally, Blattman (2022) argues that intangible incentives, such as pride or ideology, can eliminate any hope of compromise. If there is no mutually acceptable middle ground, then war is Pareto optimal, a case which is outside the scope of our survey.[12] In many cases, however, compromise alternatives exist. As Fearon (1995) argued, compromises can often be generated by creating linkages across issues, and settlements may involve transfers of both land and other resources.[13] Misperceptions are considered in Acemoğlu and Wolitzky's chapter, although those models do not integrate bargaining theory. In a world with more than two nations, aggression can be balanced by defensive alliances – but alliances can be offensive and destabilizing as well. Jordan (2006) used cooperative game theory to study alliance formation in a

[11] King Ethelred the Unready tried to appease the Danish invaders with hard cash:

> "And that is called paying the Dane-geld;
> But we've proved it again and again,
> That if once you have paid him the Dane-geld
> You never get rid of the Dane" (Kipling: Dane-Geld, in Fletcher and Kipling, 1911, p. 40).

[12] 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. Pride and ideology can also rule out using a coin to settle a dispute.

[13] For example, in 238 BC, Carthage avoided a war with Rome by handing 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).

model where wealth is power and war has no cost. Powell (1999) provided a non-cooperative model of alliance formation, where the cost of war plays an important role. In Section 7 we discuss models that either relax the unitary actor or costly lottery assumptions, or include guns versus butter trade-offs.

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 as a Hawk-Dove game with strategic substitutes. Hawk represents a commitment attempt, and war is a calculated risk. Section 4 introduces asymmetric information and finds conditions for uniqueness of equilibrium in a class of 2×2 Hawk-Dove 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. Fearful security-seekers try to preempt each other by attacking first, triggering a war that nobody wants (the Hobbesian trap). Section 5 considers the mechanism design approach. Section 5.1 shows that war is consistent with incentive-efficient bargaining when there is asymmetric information about preferences. However, there exist bargaining games that guarantee a peaceful outcome. Section 5.2 shows that no such bargaining game exists with asymmetric information about military strength. Section 6 focusses on dynamic models where long-term peace agreements cannot be enforced. War may be unavoidable if the balance of power is expected to change for exogenous reasons (the Thucydides trap) or endogenously due to appeasement attempts (Ethelred's trap). Finally, Section 7 discusses military build-ups, the role of institutions, and the duration of war.

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 $x_A + x_B = 1$. 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 $X^* \equiv X \cup \{war\}$.[14] 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 $i \in \{A, B\}$. In particular, 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*, one player (the first-mover) makes a take-it-or-leave-it proposal. The proposal represents a commitment, a *fait accompli*, so the second-mover faces a stark choice between war and appeasement. With complete information, there is a unique subgame perfect Nash equilibrium where the second-mover 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.[15] But

[14] 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.

[15] If $x_A + x_B < 1$ then they may divide the unclaimed amount $1 - x_A - x_B$ according to some fixed rule. With equal split, player $i \in \{A, B\}$ would get

$$x_i + \frac{1}{2}(1 - x_A - x_B).$$

if $x_A + x_B > 1$ then they have made incompatible commitments, so there is war. The Nash demand game has many Nash equilibria.[16]

In Crawford's (1982) modified Nash demand game, each player decides whether to attempt a commitment, and if so, how much to demand. In the terminology of Schelling (1960), a commitment attempt is a *strategic move*. Such a move could consist of the occupation of a contested territory, the strategic placement of soldiers and missiles, or simply announcements (that could turn into policies) and ceremonies.[17] The commitment attempt is successful if retreating would involve an intolerable loss of face, and the opponent realizes this. However, the cost of backing down, in terms of reputation and prestige, may be hard to predict in advance. Schelling (1966) argued that the cost is determined by

> "the psychological process by which particular things become identified with courage or appeasement or how particular things get included in or left out of a diplomatic package. Whether the removal of their missiles from Cuba while leaving behind 15,000 troops is a "defeat" for the Soviets or a "defeat" for the United States depends more on how it is construed than on the military significance of the troops, and the construction placed on the outcome is not easily foreseeable" (Schelling, 1966, p. 93-94).

In the simplest version of Crawford's (1982) model, a commitment attempt is successful with exogenously given probability 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. If

[16] If each player prefers war to getting nothing, then there exists a "bad" Nash equilibrium where each player demands all of the resource, and the equilibrium outcome is war. But perfectly peaceful equilibria, where the probability of war is zero, always exist in the complete-information Nash demand game. In this survey, we focus on games which have no perfectly peaceful equilibrium. Nevertheless, models that have both peaceful and non-peaceful equilibria can be very insightful (e.g., Slantchev, 2003).

[17] A demand need not be restricted to territory that has already been occupied. On September 30, 2022, President Putin used an elaborate ceremony, including a celebration on Red Square, to claim four regions of Ukraine, even though these regions were partially controlled by Ukrainian forces at the time (New York Times, 2022). If, after this ceremony, relinquishing his claim on these regions would be prohibitively costly for President Putin, then a successful commitment was made.

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 mutually inconsistent strategic moves.

3.1 Bargaining and war

The following game is based on Crawford (1982), Ellingsen and Miettinen (2008) and Baliga and Sjöström (2020). In stage 1, the players simultaneously 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. For example, a country could attempt to cross a status quo demarcation and take control of additional territory. Making the challenge costs $c_i > 0$, and the commitment attempt succeeds with probability q, where $0 < q < 1$. 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.[18] Referring to those two actions as Dove and Hawk,

[18] 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).

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:

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

where[19]

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

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

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. We can see this by modifying the two-stage bargaining game as follows. 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 $\theta \geq 0$ is the military advantage from the

[19] 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.

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:

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

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

and

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

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 for each i. Baliga and Sjöström (2020) verified that a large θ tends to make the actions strategic complements. For a range of commitment costs, it is then a stag hunt game with two pure Nash equilibria, (Dove,Dove) and (Hawk,Hawk).[20]

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 Padró i Miquel (2010) identified a

[20] 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.

risk-dominance condition under which peace cannot be sustained in an infinite horizon interaction. Kydd (1997) assumed preferences are uncorrelated, but each player observes a signal which is correlated with the opponent's preferences. Here we will assume there is no such signal. As in Baliga and Sjöström (2004), types simply 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 types c_A and c_B are independently drawn from an atomless distribution with a continuously differentiable cdf F. The support is $[\underline{c}, \overline{c}]$, where $\underline{c} \geq 0$.[21] For simplicity, the players are symmetric *ex ante* (i.e., identical in all respects except their cost-types), so we drop all subscripts except on the types. Endowments are $\omega_A = \omega_B = 0.5$. After each player has learned his own type, they simultaneously choose Hawk or Dove. In the payoff matrix, player i's choice is represented by a row, player j's choice by 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 $c_j > z_j$.[22] 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

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

[21] Following convention in economics, the players have common prior beliefs represented by F. Smith and Stam (2004) consider disagreement caused by non-common prior beliefs.

[22] 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.

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.[23] 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.

4.1 Bargaining and uniqueness

Consider again the bargaining game of Section 3.1. Hawk represents a strategic move, i.e., a commitment attempt, 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). If the costs c_A and c_B are private information, independently drawn from a distribution which satisfies Condition (5), then there is a unique equilibrium.[24]

For example, suppose F is uniform on $[\underline{c}, \overline{c}]$. 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:

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

With a unique equilibrium, it is straightforward to do comparative statics.

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

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

That is, the support must be sufficiently large.

[24] 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.

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 by Equation (6). 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 by attacking the opponent. Accordingly, there is war if at least one player chooses Hawk. For now, the only decision each player makes is to choose Hawk or Dove. There is no bargaining. In Sections 4.2.1, 4.2.2 and 4.2.3, we will introduce signaling, peace talks and bargaining.

For simplicity, the players are again *ex ante* symmetric. If both players choose Dove then they 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. All types strictly prefer (Dove,Dove) to (Hawk,Hawk).

Player i is a dominant strategy hawk if

$$c_i < v^{HD} - u(0.5) = 0.5 + \theta - \phi - u(0.5)$$

and a security-seeker if

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

Security-seekers are *coordination types* who 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

$$0 \leq \underline{c} < 0.5 + \theta - \phi - u(0.5) < \bar{c} < \theta \tag{9}$$

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 cascade of security-seekers with higher and higher cost-types who eliminate Dove out of fear that the opponent will choose Hawk. There are no dominant strategy doves to stop the cascade from reaching all types in the support of F. Thus, if Conditions (8) and (9) hold then all types must choose Hawk. There is war with probability one, even though the probability of encountering a dominant strategy hawk, which is $F\left(v^{HD} - u(0.5)\right) > 0$, could be very close to zero. This is the Hobbesian trap.

4.2.1 Signaling and the spiral model

Without a binding peace treaty, a security-seeker may be inclined to make military preparations to protect against an attack. Such preparations could, however, be mistaken for preparations to launch an attack. A security-seeking opponent may respond by making his own military preparations or even by launching a preemptive attack. This is the *security dilemma* or *spiral model* (Jervis, 1976). To break the spiral, a security-seeker may signal peaceful intentions by not mobilizing. This promotes peace if the opponent is also a security-seeker, but if the opponent is a dominant strategy hawk then the outcome may be military disaster. This dilemma was formalized by Kydd (1997).

Assume Conditions (8) and (9) hold, so as shown above there is war with probability one in the Hawk-Dove game. Now we change the game as follows. After each player has learned his own type, they simultaneously choose whether or not to mobilize. For simplicity, mobilizing has no direct cost.

After each player has observed the opponent's mobilization decision, they simultaneously choose Hawk or Dove. However, only players who have mobilized can choose Hawk. Not mobilizing is a unilateral commitment to choose Dove, a perfectly credible signal of peaceful intentions. Assume Conditions (8) and (9) hold, so that if all types mobilize in equilibrium then (as shown above) all types will choose Hawk.

If dominant strategy hawks are sufficiently rare, specifically if

$$F\left(v^{HD} - u(0.5)\right) < \frac{\theta}{u(0.5) - v^{DH}} \tag{10}$$

then there exists an equilibrium where security-seekers with sufficiently high cost-types refrain from mobilizing (so the probability of peace is strictly positive).[25] An equilibrium where all types mobilize (so the probability of peace is zero) exists if and only if Condition (10) is violated.

In a symmetric perfect Bayesian equilibrium, there must be a cutoff type z^* such that player $i \in \{A, B\}$ mobilizes if $c_i < z^*$ but not if $c_i > z^*$. Dominant strategy hawks will always mobilize and choose Hawk, so $z^* \geq v^{HD} - u(0.5)$. Security-seekers who have mobilized will choose Hawk if and only if the opponent has mobilized.[26] The benefit of not mobilizing is that it forestalls war if the opponent is a security-seeker who mobilizes. Any player who does not mobilize will live in peace and get $u(0.5)$ if the opponent is a security-seeker, but he will be attacked and get v^{DH} if the opponent is a dominant strategy hawk. Condition (10) implies that the highest cost-types are willing to gamble for peace (by refraining from mobilizing). The cutoff will therefore satisfy $z^* < \bar{c}$.[27]

[25] If $F\left(v^{HD} - u(0.5)\right)$ is close to zero then Condition (10) holds, because the right hand side will be close to $\theta / (2\theta - \underline{c}) > 0$. Note that Condition (10) is perfectly consistent with Conditions (8) and (9).

[26] If only types below z^* mobilize, then after both have mobilized they in effect play the Hawk-Dove game with a type distribution truncated at z^*. Conditions (8) and (9) imply that the unique equilibrium for the truncated distribution is for all types to choose Hawk. Intuitively, if not mobilizing signals peaceful intentions, then mobilization indicates the opposite, making war unavoidable.

[27] The expected payoff from not mobilizing is

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

A security–seeker with cost-type c_i who mobilizes (and then chooses Hawk if and only if the opponent has mobilized) gets

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

If Condition (10) holds, then there is peace if both players are security-seekers and $c_i > z^*$ for at least one player $i \in \{A, B\}$. However, if both players are security-seekers with a relatively low cost of choosing Hawk, so that $c_i < z^*$ for each $i \in \{A, B\}$, then both mobilize, and after observing that the opponent has mobilized, both choose Hawk. This sequence of events, where military preparations trigger war between security-seekers, captures the essence of the spiral model.

4.2.2 Cheap talk

When mutual suspicions arose between Greeks and Persians, Xenophon arranged a meeting that he hoped would remove the suspicions.[28] President Reagan said that the main objective of meeting with Mikhail Gorbachev was

---

If $z^* < \bar{c}$ then the cutoff type must be indifferent between mobilizing and not mobilizing. Setting $c_i = z^*$ in Equation (12) and subtracting it from Equation (11) yields the expression

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

This expression is strictly negative for $z^* = v^{HD} - u(0.5)$, and strictly positive for $z^* = \bar{c}$ if Condition (10) holds. By continuity, the expression equals zero for some z^* such that $v^{HD} - u(0.5) < z^* < \bar{c}$. This is an equilibrium cut-off point. Note that Condition (10) can be expressed in words as follows: the least aggressive type $\bar{c}$ strictly prefers not to mobilize against an opponent who always mobilizes. Thus, an equilibrium where all types mobilize exists if and only if Condition (10) is violated.

[28] The Greek leader said:

> "I observe that you are watching our moves as though we were enemies, and we, noticing this, are watching yours, too. On looking into things, I am unable to find evidence that you are trying to do us any harm, and I am perfectly sure that, as far as we are concerned, we do not even contemplate such a thing; and so I decided to discuss matters with you, to see if we could put an end to this mutual mistrust. I know, too, of cases that have occurred in the past when people, sometimes as the result of slanderous information and sometimes merely on the strength of suspicion, have become frightened of each other and then, in their anxiety to strike first before anything is done to them, have done irreparable harm to those who neither intended nor even wanted to do them any harm at all. I have come then in the conviction that misunderstandings of this sort can best be ended by personal contact, and I want to make it clear to you that you have no reason to distrust us" (Xenophon, 1972, p. 123-124).

"the elimination of suspicion and distrust" (New York Times, 1985). To see whether communication can create trust, even if it is impossible to send a perfectly credible signal of peaceful intentions as in Section 4.2.1, consider a cheap-talk extension of the Hawk-Dove game. After each player has learned his own type, they simultaneously send either a friendly or a tough message. Each player observes the opponent's message, and then they simultaneously choose Hawk or Dove. Since talk is cheap, a player who sends a friendly message can still choose Hawk. Nevertheless, the exchange of friendly messages can create trust among security-seekers, at least if dominant strategy hawks are sufficiently rare.

Assume Conditions (8) and (9) hold, so that all types will choose Hawk if the cheap talk is uninformative "babbling". To construct informative cheap-talk equilibria, Baliga and Sjöström (2004) used the fact that different types have different 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. In equilibrium, security-seekers with the highest cost-types send the friendly message, and if they also receive a friendly message then they choose Dove. Dominant strategy hawks also send the friendly message, but they always choose Hawk. Intermediate types – security-seekers with a relatively low cost of choosing Hawk – send the tough message, because this allows them to match the opponent's action. Under assumptions that guarantee that dominant strategy hawks are rare, Baliga and Sjöström (2004) found that the highest cost-types are willing to gamble for peace by choosing Dove after exchanging friendly messages. Those types escape the Hobbesian trap, because the "cascade" is blocked by a firewall of intermediate types who send tough messages. Thus, there are scenarios where cheap talk can create peace among security-seekers, at least among those who are sufficiently averse to war.[29]

4.2.3 Binding agreements

If the two players can sign a binding peace agreement before playing the Hawk-Dove game, will they do so? The answer depends on how they negotiate. Suppose Conditions (8) and (9) hold, so without a peace agreement there will be war with probability one. Now consider the following bargain-

[29] Cheap-talk cannot bring peace if all security seekers have a low cost of choosing Hawk. This is consistent with Acharya and Ramsay's (2013) negative results on cheap talk in a global games model.

ing game. After the players have learned their own types, player A makes a take-it-or-leave-it proposal x_B. Thus, player A has an exogenously given first-mover advantage. If player B accepts the proposal, then there is a binding peace treaty and the proposal is implemented: each player must choose Dove and player B gets the share x_B of the contested resource. If player B rejects, then there is war; each player i chooses Hawk and gets $v^{HH} - c_i = 0.5 - \phi - c_i$.

Player B is more willing to appease player A, the higher is c_B. Specifically, he will accept the proposal if $c_B \geq 0.5 - \phi - x_B$.[30] Since player A does not know the true c_B, he faces a risk-reward trade-off. A high offer will be overly generous if c_B is high, but a low offer may lead to war if c_B is low. If the cost of war is not too high, then low offers will be made and rejected with strictly positive probability in equilibrium. To show this, we simplify by assuming costs are uniformly distributed on $[\underline{c}, \overline{c}]$, and each player $i \in \{A, B\}$ has a linear utility function, $u_i(x_i) = x_i$ for all $x_i \in [0, 1]$.

Theorem 1 *There is war with strictly positive probability if and only if* $\phi < (\bar{c} - 3\underline{c})/2$.

Proof. Player B rejects the offer x_B with probability $F(0.5 - \phi - x_B)$. By assumption, $F(c) = (c - \underline{c}) / (\bar{c} - \underline{c})$ for $c \in [\underline{c}, \overline{c}]$. The offer x_B is accepted with probability 1 if $x_B \geq 0.5 - \phi - \underline{c}$, and with probability 0 if $x_B \leq 0.5 - \phi - \bar{c}$. Therefore, the optimal offer satisfies

$$0.5 - \phi - \bar{c} \leq x_B \leq 0.5 - \phi - \underline{c}. \tag{14}$$

Player A's expected payoff is

$$\begin{aligned} & (1 - F(0.5 - \phi - x_B))(1 - x_B) + F(0.5 - \phi - x_B)(0.5 - \phi - c_A) \\ = \quad & 1 - x_B + \frac{0.5 - \phi - x_B - \underline{c}}{\bar{c} - \underline{c}} (x_B - 0.5 - \phi - c_A). \end{aligned}$$

Maximizing this expression subject to (14) yields $x_B = 0.5 - \phi - \underline{c}$ if

$$c_A \geq \bar{c} - 2(\phi + \underline{c})$$

and

$$x_B = \frac{1 - (\bar{c} - c_A)}{2} < 0.5 - \phi - \underline{c}$$

[30] Player B's payoff is independent of c_A, whether he accepts or rejects. Therefore, player B's decision does not depend on his beliefs about player A's type.

if

$$c_A < \bar{c} - 2(\phi + \underline{c}).$$

Note that $\bar{c} - 2(\phi + \underline{c}) > \underline{c}$ if and only if $\phi < (\bar{c} - 3\underline{c})/2$. If this inequality holds, then with probability $F(\bar{c} - 2(\phi + \underline{c})) > 0$ player A makes a proposal $x_B < 0.5 - \phi - \underline{c}$ which is rejected with probability $F(0.5 - \phi - x_B) > 0$. ■

Thus, if one player makes a take-it-or-leave-it proposal, then there is war with strictly positive probability if war is not too costly. However, there are many other ways to bargain. Suppose an impartial mediator drafts a peace treaty which specifies equal division of the contested resource. After each player has learned his own type, they simultaneously choose whether or not to sign the treaty. If both sign, then both must choose Dove and take equal shares of the resource. Thus, signing is a commitment to choose Dove conditional on the opponent also signing. If at least one player did not sign, then each player is free to choose either Hawk or Dove. 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.[31] Since all types prefer (Dove,Dove) to (Hawk,Hawk), all types prefer to sign, and there is never any war.

To summarize, if the players can negotiate a binding peace treaty which specifies how the contested resource should be allocated, then whether war can be avoided depends both on the cost of war and on the details of the bargaining procedure. To rigorously study the design of bargaining procedures requires the theory of mechanism design.

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 the theory of mechanism design. A mechanism, or game form, specifies the "rules of the game": what actions the players can take

[31]The threat is credible because, by assumption, there are no dominant strategy doves in the support of F. The fact that dominant strategy doves don't exist makes it easier to implement a peace treaty.

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 (unverifiable) characteristics. The game starts by drawing the types from a commonly known distribution. A *type-contingent outcome* is a function from the set of possible type-profiles 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 types, since these are private information. The outcome can depend on the true types only indirectly, via the players' strategies. Each player's strategy specifies, for each of his possible types, what actions he will take. The equilibrium strategies will generate a type-contingent outcome. This type-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 (which is his "outside option"). These are *necessary* conditions which must be satisfied even if binding agreements are feasible.[32] 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 types and cannot prevent the players from declaring war. A type-contingent outcome is *incentive-efficient* (Myerson, 1979) if no type-contingent outcome which satisfies the IC conditions can make each type of each player better off. If a bargaining game has an equilibrium such that the type-contingent outcome is incentive-efficient, then it is an incentive-efficient bargaining game.

In Section 5.1, the private information concerns preferences, specifically, the cost of going to war. War is Pareto inefficient (in the standard sense), and there are bargaining games that guarantee peace. The ultimatum bargaining game cannot guarantee peace, yet the ultimatum mechanism is incentive-efficient. In Section 5.2, military capabilities are private information. If the cost of war is not too high, then no mechanism can guarantee peace.

For simplicity, in this section we will assume each player $i \in \{A, B\}$ has a linear utility function, $u_i(x_i) = x_i$ for all $x_i \in [0, 1]$.

[32] 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.

5.1 Private information about preferences

The bargaining game in Section 3 showed how two-sided commitment can cause war in a model with symmetric information. The ultimatum game shows how one-sided commitment can cause war when information is asymmetric (Fearon, 1995).[33] One player, say player A, has a built-in first-mover advantage and makes a take-it-or-leave-it offer. He is committed to fight if player B rejects. Suppose player B's privately known type is his cost of war, $\phi_B \in \{\phi_B^L, \phi_B^H\}$, where $0 < \phi_B^L < \phi_B^H < p_B$. The private information is *soft* in the sense that player B cannot produce any verifiable (non-falsifiable) evidence of his type. Player A thinks $\phi_B = \phi_B^H$ with probability h, where $0 < h < 1$.

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 therefore 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. If player A offers the share $x_B = p_B - \phi_B^L$ then both types of player B accept, so player A's payoff is $1 - (p_B - \phi_B^L) = p_A + \phi_B^L$. This (constant) type-contingent outcome is incentive-compatible, individually rational and Pareto efficient (since there is no war). However, this may not be player A's optimal offer, since type ϕ_B^H gets an information rent. That is, type ϕ_B^H is strictly better off than if there is war: $x_B > p_B - \phi_B^H$. To optimally exploit his first-mover advantage, player A must trade off the gain from extracting the information rent from type ϕ_B^H against the risk of war against type ϕ_B^L. 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

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

[33] As in Section 3, a first-mover advantage may be created by a strategic move, such as the occupation of a contested territory. War occurs if the occupier has underestimated the opponent's willingness to fight:

> "Joseph of Austria quietly sent troops into Bavaria in January 1778 in the faith that he could occupy territory without firing a shot. He predicted that old Frederick of Prussia, now grounded with gout, would not fight the Austrians; but his prediction erred and the War of the Bavarian Succession ensued" (Blainey, 1988, p. 43).

Adding asymmetric information to the bargaining game of Section 3 would compound the risk of war: they will risk war both to get the first-mover advantage and to exploit the advantage (via the risk-reward trade-off) if they get it.

where $\phi_A > 0$ is player A's cost of war. The optimal offer is $p_B - \phi_B^H$ if

$$h\phi_B^H - (1-h)\phi_A > \phi_B^L. \tag{15}$$

If this inequality holds, then in equilibrium there is war with probability $1 - h > 0$.[34] Player A prefers to eliminate his opponent's information rent, even though war destroys social surplus.

5.1.1 Optimal screening

Even if there is war in equilibrium, the ultimatum game is an incentive-efficient bargaining game. Specifically, the type-contingent outcome maximizes player A's expected payoff subject to player B's IC and IR conditions. To see this, consider any type-contingent outcome. Forcing type ϕ_B^H to fight would serve no purpose, so we may eliminate this possibility. The expected payoff of type ϕ_B^H will simply be his expected share of the resource, denoted x_B^H. Forcing type ϕ_B^L to fight with some probability π serves the purpose of separating the two types. The expected payoff of type ϕ_B^L will be $(1-\pi)x_B^L + \pi(p_B - \phi_B^L)$, where x_B^L is his expected share if there is no war. The type-contingent outcome is represented as (π, x_B^H, x_B^L).

If type ϕ_B^H mimics type ϕ_B^L then he must fight with probability π, and his expected payoff will be $(1-\pi)x_B^L + \pi(p_B - \phi_B^H)$. If type ϕ_B^L mimics type ϕ_B^H then he avoids war and gets x_B^H. Therefore, the two IC conditions are

$$x_B^H \geq (1-\pi)x_B^L + \pi(p_B - \phi_B^H)$$

and

$$(1-\pi)x_B^L + \pi(p_B - \phi_B^L) \geq x_B^H.$$

The two IR conditions are

$$x_B^H \geq p_B - \phi_B^H$$

and

$$(1-\pi)x_B^L + \pi(p_B - \phi_B^L) \geq p_B - \phi_B^L$$

since each type's outside option is war.

[34] Since neither ω_i nor p_i appear in Condition (15), these parameters do not influence the probability of war. This is a special property of the ultimatum game. 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.

Player A's expected payoff is

$$h\left(1-x_B^H\right)+(1-h)\left((1-\pi)(1-x_B^L)+\pi(p_A-\phi_A)\right). \qquad (16)$$

Choosing (π, x_B^L, x_B^H) to maximize player A's expected payoff subject to the IC and IR conditions reproduces the equilibrium outcome of the ultimatum game. That is, if Condition (15) holds, then it is optimal to separate the two types: $\pi = 1$ and $x_B^H = p_B - \phi_B^H$ (and x_B^L is irrelevant since $\pi = 1$). If Condition (15) is reversed then pooling is optimal: $\pi = 0$ and $x_B^L = x_B^H = p_B - \phi_B^L$. The ultimatum bargaining game implements *optimal screening* in the sense that player A's expected payoff is maximized subject to the constraints implied by player B's private information and outside option. No type-contingent outcome that satisfies the IC and IR conditions can give a higher expected payoff for player A.

Thus, ultimatum bargaining is incentive-efficient even if Condition (15) holds. War destroys social surplus with probability $1 - h > 0$, but a credible threat of war is the only way to eliminate player B's information rent. It is well known that incentive-efficient mechanisms may produce outcomes that are not Pareto efficient in the standard sense (Myerson, 1979). From this perspective, war is not a puzzle. Brito and Intriligator (1985) may have been the first to make this point.

5.1.2 Mechanisms for peace

If the players' private information only concerns their preferences, e.g., the cost of going to war or the value of the contested resource, then the IC and IR conditions can be satisfied without fighting.[35] This implies that 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). The constant outcome $(p_A, p_B) \in X$ will be strictly preferred to war by each type of each player, so if (p_A, p_B) is common knowledge then it is a Bayesian Nash equilibrium of the Nash demand game for each type of player $i \in \{A, B\}$ to demand $x_i = p_i$. This is not the unique equilibrium, although sharing in proportion to military strength may be a natural focal point. There are other bargaining games such that in *every* equilibrium, the

[35] This is true even if preferences have a common value component, and each player has private information correlated with the true value. For example, the contested resource could be an oil field, and each player observes a private signal of the amount of oil it contains.

outcome is $(p_A, p_B) \in X$ with probability one. Let a mediator (the "mechanism designer") propose the allocation $(x_A, x_B) = (p_A, p_B)$, and the players sequentially either accept it or declare war. In perfect Bayesian equilibrium, both players accept with probability one.[36] Thus, with asymmetric information only about preferences, there exist bargaining games that guarantee a Pareto optimal peaceful outcome.

5.2 Private information about military power

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. It is plausible that some of these variables are private information.[37] Following Bester and Wärneryd (2006) and Fey and Ramsay (2009, 2011), we now show that war may be unavoidable when military strength is private information, even with unrestricted ability to sign binding peace agreements.

5.2.1 One-sided private information

Suppose the military capabilities of player A are common knowledge, but player A does not know the capabilities of player B. Player B can be militarily weak or strong. His privately known type is his probability of winning a war against player A, $p_B \in \{p_B^L, p_B^H\}$, where $0 < p_B^L < p_B^H$. Again, his private information is soft. Player A thinks $p_B = p_B^H$ with probability h, where

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

[37] Even if all these variables are common knowledge, there may be disagreement about how the probability of winning depends on these variables. Before World War I, England was the leading financial power, and its leaders based their belief in victory on their economic might; German leaders based their belief in victory on their superior military technology and strategy (Blainey, 1988, p. 40). This situation can be captured by the framework of this section, implying that war may be unavoidable. Suppose there are two variables, say financial strength and military technology. It is commonly known that one side has superior financial strength while the other side has superior military technology. Player i's type t_i is a private signal indicating which variable dominates in determining the probability of victory, which can be written as $p_i(t_1, t_2)$ as in Section 5.2.2.

$0 < h < 1$. Type p_B gets expected payoff $p_B - \phi_B$ from war, and player A gets $1 - p_B - \phi_A$ from war against type p_B. This is a common values environment: player A's expected payoff from war depends on player B's type. For simplicity, the costs of war ϕ_A and ϕ_B do not depend on player B's type.

There is a simple intuition for why war may be unavoidable even if it is possible to sign a binding peace treaty. If player B is strong, $p_B = p_B^H$, then he thinks he can win on the battlefield, so he will require a large share of the resource in order to sign a peace treaty. If the strong type can get a large share without a fight, then the weak type can pretend to be strong and get a large share without fighting as well. Thus, if the probability of war is zero, then the peace treaty must give player A a small share, regardless of player B's type. But if player A thinks player B is probably weak (i.e., if h is small), then player A may prefer to fight rather than take the small share. This means the probability of war cannot be zero. The peace treaty must give player B a bigger share when he is strong than when he is weak, and this requires that player B faces a risk-reward trade-off: he can only get a larger share by accepting a risk of war.

To see this formally, note that for type p_B^L to fight would serve no purpose, so we may disregard this possibility. This type's expected payoff will just be his expected share of the resource, denoted x_B^L. Suppose type p_B^H fights with probability π. His expected payoff will be $(1 - \pi)x_B^H + \pi(p_B^H - \phi_B)$, where x_B^H is his expected share if there is no war. A type-contingent outcome can be represented as (π, x_B^H, x_B^L).

If type p_B^L mimics type p_B^H then he must fight with probability π, and his expected payoff will be $(1 - \pi)x_B^H + \pi(p_B^L - \phi_B)$. To prevent player B from exaggerating his military capabilities, we need the IC condition

$$x_B^L \geq (1 - \pi)x_B^H + \pi(p_B^L - \phi_B). \tag{17}$$

(The IC condition that prevents player B from deprecating his own strength can be disregarded.) Since player B cannot be prevented from launching an attack when he is strong, we have the (interim) IR condition

$$(1 - \pi)x_B^H + \pi(p_B^H - \phi_B) \geq p_B^H - \phi_B. \tag{18}$$

(The IR condition for the weak type can be disregarded.)

Since player B is strong with probability h, player A's expected payoff is

$$h\left((1 - \pi)(1 - x_B^H) + \pi(1 - p_B^H - \phi_A)\right) + (1 - h)\left(1 - x_B^L\right).$$

If player A refuses to negotiate, and instead launches an attack, his expected payoff is

$$h(1-p_B^H)+(1-h)(1-p_B^L)-\phi_A.$$

Thus, player A's (interim) IR condition is

$$h\left((1-\pi)(1-x_B^H)+\pi(1-p_B^H-\phi_A)\right)+(1-h)\left(1-x_B^L\right)\geq h(1-p_B^H)+(1-h)(1-p_B^L)-\phi_A. \tag{19}$$

Define

$$A\equiv\frac{p_B^H-p_B^L}{\phi_A+\phi_B}.$$

Theorem 2 *(i) If $(1-h)A\leq 1$ then the following (constant) type-contingent outcome satisfies the IC and IR conditions: $\pi=0$ and $x_B^H=x_B^L=p_B^H-\phi_B$. (ii) If $(1-h)A>1$ then for any type-contingent outcome which satisfies the IC and IR conditions, the (ex ante) probability of war is at least*

$$h\frac{(1-h)A-1}{(1-h)A-h}>0.$$

Proof. (i) Suppose $(1-h)A\leq 1$. If $\pi=0$ and $x_B^H=x_B^L=p_B^H-\phi_B$, then regardless of player B's type, there is peace and player B gets the strong type's outside option. This clearly satisfies player B's IC and IR conditions. Player A's payoff is $1-p_B^H+\phi_B$. If player A declares war, his expected payoff is

$$h(1-p_B^H)+(1-h)(1-p_B^L)-\phi_A\leq 1-p_B^H+\phi_B$$

where the inequality is due to $(1-h)A\leq 1$. Thus, the outcome satisfies player A's IR condition as well.

(ii) Suppose $(1-h)A>1$. Adding Conditions (17), (18) and (19), we get

$$\pi\geq\frac{(1-h)A-1}{(1-h)A-h}>0.$$

Since player B is strong with probability h, the ex ante probability of war is $h\pi$. ■

Theorem 2 shows that, regardless of how they negotiate, war must happen with strictly positive probability if

$$\phi_A+\phi_B<(1-h)\left(p_B^H-p_B^L\right).$$

5.2.2 Two-sided private information

Suppose each player has private information about his own military strength, and this can be represented by a number between 0 and 1. 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. Thus, there is two-sided incomplete information. 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 strictly positive 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 $0 \leq p_A(t_A, t_B) \leq 1$, assume $\beta \leq \mu \leq 1 - \alpha$. Since a player's expected payoff from war depends on his opponent's type, this is a common values environment.

If the types are (t_A, t_B), then the expected payoff for player $i \in \{A, B\}$ is

$$V_i(t_A, t_B) = (1 - \pi(t_A, t_B))x_i(t_A, t_B) + \pi(t_A, t_B)(p_i(t_A, t_B) - \phi_i),$$

where $\pi(t_A, t_B)$ denotes the probability of war, and $x_i(t_A, t_B)$ is the expected share player i gets if there is no war. Since $x_B(t_A, t_B) = 1 - x_A(t_A, t_B)$, a type-contingent outcome can be represented by the functions $\pi(t_A, t_B)$ and $x_A(t_A, t_B)$.

Let $V_i(t_i)$ denote the interim expected payoff and $\pi_i(t_i)$ the interim probability of war for type t_i. That is,

$$V_A(t_A) \equiv \int_0^1 V_A(t_A, t_B) dt_B$$

and

$$\pi_A(t_A) \equiv \int_0^1 \pi(t_A, t_B) dt_B.$$

$V_B(t_B)$ and $\pi_B(t_B)$ are defined analogously.

Theorem 3 *(i) If $\phi_A + \phi_B \geq (\alpha + \beta)/2$ then the following (constant) type-contingent outcome satisfies the IC and IR conditions: $\pi(t_A, t_B) = 0$ and $x_A(t_A, t_B) = \mu + \alpha - \beta/2 - \phi_A$ for all $(t_A, t_B) \in [0,1] \times [0,1]$. (ii) If $\phi_A + \phi_B < (\alpha + \beta)/2$ then for any type-contingent outcome that satisfies the IC and IR conditions, the (ex ante) probability of war is at least*

$$\frac{1}{2} - \frac{\phi_A + \phi_B}{\alpha + \beta} > 0.$$

Proof. (i) Suppose $\phi_A + \phi_B \geq (\alpha+\beta)/2$. If $\pi(t_A, t_B) = 0$ and $x_A(t_A, t_B) = \mu + \alpha - \beta/2 - \phi_A$ for all (t_A, t_B), then clearly the IC conditions are satisfied, because the outcome is type-independent. If type $t_A = 1$ refuses to negotiate, and declares war without knowing player B'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 IR condition holds for type $t_A = 1$, which means it holds for all $t_A \leq 1$. Similarly, type $t_B = 1$ expects to get

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

if he refuses to negotiate and declares war without knowing player A's type. But $\phi_A + \phi_B \geq (\alpha+\beta)/2$ implies

$$1 - \mu - \frac{\alpha}{2} + \beta - \phi_B \leq 1 - (\mu + \alpha - \beta/2 - \phi_A).$$

Therefore, the IR condition for type $t_B = 1$ holds, which means it holds for all $t_B \leq 1$.

(ii) Suppose $\phi_A + \phi_B < (\alpha+\beta)/2$. By a standard argument, V_i is non-decreasing, and

$$V_A(1) - V_A(0) = \alpha \int_0^1 \pi_A(t_A)dt_A = \alpha E\{\pi\}, \tag{20}$$

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

$$V_B(1) - V_B(0) = \beta E\{\pi\}. \tag{21}$$

Adding Equations (20) and (21) we get

$$(\alpha+\beta)E\{\pi\} = V_A(1) + V_B(1) - V_A(0) - V_B(0). \tag{22}$$

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

$$V_A(0) + V_B(0) \leq 1. \tag{23}$$

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

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

Therefore, the IR condition for type $t_A = 1$ is

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

Similarly, the IR condition for type $t_B = 1$ is

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

Substituting from Conditions (23), (24) and (25) into Equation (22) we obtain

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

■

Theorem 3 shows that, regardless of how they negotiate, 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 there is asymmetric information about military strength then war may be unavoidable. The problem is worse, the more uncertain is the balance of power, i.e., the bigger are α and β.

Theorem 3 formalizes Blainey's (1988) argument that wars are caused by disagreements about the balance of power, here represented by (p_A, p_B). It is sometimes further argued that the most dangerous situation is when the two sides have similar strength (Blainey, 1988, p. 276), but this is not supported by Theorem 3. If the two sides agree that their strengths are almost equal ($\mu = 0.5$, α and β very small), then they can negotiate an outcome close to equal split. If they agree that one player is very likely the strongest, but there is significant disagreement about *how much* stronger he is, then war may be unavoidable.

Theorem 3 applies to any bargaining game, but we can illustrate by a simultaneous-move Hawk-Dove game. There is war if at least one player chooses Hawk, otherwise the status quo remains in place. Suppose there exists a perfectly 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

$$\omega_A \geq \mu + \alpha - \frac{1}{2}\beta - \phi_A. \tag{26}$$

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

$$\omega_B \geq 1 - \mu + \beta - \frac{1}{2}\alpha - \phi_B. \tag{27}$$

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 (26) and (27) are satisfied. In this case, there exists a perfectly peaceful equilibrium.

If $\phi_A + \phi_B < (\alpha + \beta)/2$ then the Hawk-Dove game has no perfectly peaceful equilibrium. In fact, the game may unravel. Consider an equilibrium cutoff point $z_j^* < 1$ such that player j chooses Hawk if $t_j > z_j^*$ and Dove if $t_j < z_j^*$. Since peace must be voluntary on both sides, player i's choice only matters if player j chooses Dove, i.e., if player j is militarily weak, with t_j below z_j^*. This encourages player i to choose Hawk, which drives down the equilibrium cutoff point. A similar logic applies to markets with common values. Since trade must be voluntary on both sides, the buyer's decision only matters if the seller wants to sell, i.e., if the quality of the good is low (Akerlof, 1970). This discourages the buyer from buying. The no trade theorem (Milgrom and Stokey, 1982) says that purely speculative trade is impossible among risk-averse traders. Peace is possible in the Hawk-Dove game, because peace generates a social surplus (by not incurring the costs of war). With no cost of war, the situation would be analogous to Milgrom and Stokey (1982): the game would unravel completely, and peace would be impossible.

Blainey (1988, p. 51) pointed out that political leaders often experience worsening moods on the eve of war. This might suggest that the war was inadvertent rather than a deliberate choice (cf. Footnotes 2 and 7). However, if an imperfectly informed leader makes a deliberate choice to go to war, and then discovers that the opponent wants war, then he naturally becomes more pessimistic about the chances of victory. He may wish the coming war could be avoided, or he may still want war even after learning that his opponent wants war, but in any case, his mood will worsen.

For example, if $\phi_A = \phi_B = \phi$, $\mu = \omega_A = \omega_B = 1/2$ and $\alpha = \beta > 4\phi$, then the symmetric equilibrium cut-off point in the Hawk-Dove game is $z^* = 2\phi/\alpha < 1/2$. Consider type $t_i = 1$. His expected payoff from war conditional on $t_j > z^*$ is $1/2 + \alpha/2 - 2\phi > 1/2 = \omega_A$, so he still wants war even after learning that the opponent wants war. On the other hand, types slightly above z^* prefer peace if they discover that the opponent wants war. Communicating a desire for peace would, however, indicate weakness and therefore make the opponent even more intent on fighting.

5.2.3 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 in Section 4.2.2). As Fey and Ramsay (2009, 2011) pointed out, negotiations may reveal information which gives a 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) showed how this problem can be solved by a mediator. The mediator's recommendation hides information contained in the messages he has received, so that attack decisions cannot be made contingent on this information.

5.2.4 Hard information

Some variables that determine military strength 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). An unarmed country can be exploited by its enemies, but a well-armed country may be considered a threat to others. A country that refuses arms inspections may deter its enemies without provoking security-seekers (Baliga and Sjöström, 2008).

6 Commitment problems and power shifts

In this section we consider dynamic models without private information or first-strike advantages. The key assumption is incomplete contracting ability. As in the incomplete contracts literature (Grossman and Hart, 1986), a contract which specifies future outcomes would not be enforceable. Instead, future outcomes will be determined by future negotiations. Fearon (1995, 1998, 2004) and Powell (2004a, 2006) argued that this causes problems if the balance of power changes over time.

In Section 6.1, the balance of power shifts exogenously.[38] In order to avoid a preventive strike, a rising power should make a firm commitment not to challenge the status quo. A Thucydides trap occurs if such commitments are impossible. In Section 6.2, power depends on the control of resources. Ethelred's trap occurs if a player cannot commit to staying peaceful after getting more resources. By assumption, the players cannot influence the probability of a power shift by investing in military capabilities or new technologies. Models that relax this assumption are briefly discussed in Section 7.1.

6.1 Power and time: the Thucydides trap

There is an infinite number of periods, $t = 1, 2, 3...$ The contested resource is perfectly durable, e.g., a piece of land. In each period, as long as there was no war in the past, the two players either make a short-term deal on how to share the resource this period, or fight a war which ends the game. The share controlled by player B in period t is denoted $x_{Bt} \in [0, 1]$ and is referred to as the state of the world. Player A's share is $x_{At} = 1 - x_{Bt}$. The initial state is $x_{B0} = \omega_B$. Players A and B get utility x_{At} and x_{Bt}, respectively, in period t. The common discount factor is $\delta < 1$. The players have perfect foresight. By assumption, they cannot make binding long-term agreements.

Let p_{it} denote player i's exogenously given probability of winning a war in period t, assuming there has been no war in the past (which would have ended the game). There is no first-strike advantage, and $p_{At} + p_{Bt} = 1$ for each $t \geq 1$. 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.

In each period, each player can choose to fight rather than negotiate. Therefore, permanent peace requires that the infinite path $\{x_{B0}, x_{B1}, ..., x_{Bt}, ...\}$ is self-enforcing in the sense that no player has an incentive to start a war

[38] An exogenous power shift may be due to differences in economic growth rates (Krainin, 2017) or the introduction of a new military technology such as nuclear weapons (Bas and Coe, 2012).

in any period. In any period $\tau \geq 1$, player A's continuation payoff from following the path peacefully, expecting no war in any period, is

$$\sum_{t=\tau}^{\infty} \delta^{t-\tau} x_{At}.$$

If instead there is war in period τ, then player A's expected payoff is

$$\frac{p_{A\tau}}{1-\delta} - \phi_A$$

since he wins with probability $p_{A\tau}$, the gain from winning is $1/(1-\delta)$, and his cost of war is ϕ_A. Thus, player A prefers peace to war in period τ if and only if

$$\sum_{t=\tau}^{\infty} \delta^{t-\tau} x_{At} \geq \frac{p_{A\tau}}{1-\delta} - \phi_A. \tag{28}$$

Similarly, player B prefers peace to war if

$$\sum_{t=\tau}^{\infty} \delta^{t-\tau} x_{Bt} \geq \frac{p_{B\tau}}{1-\delta} - \phi_B. \tag{29}$$

The path $\{x_{B0}, x_{B1}, ..., x_{Bt}, ...\}$ is *self-enforcing* if Conditions (28) and (29) hold for all $\tau \geq 1$.[39]

There are many ways the players could negotiate the short-term deals: one player could make a take-it-or-leave-it offer, or they could use Nash demand bargaining or some other procedure. If we specify the procedure, we obtain an infinite-horizon game. The key assumption is that each player has the "outside option" of starting a war in any period. If this game has a subgame perfect equilibrium without war, then the equilibrium path $\{x_{B0}, x_{B1}, ..., x_{Bt}, ...\}$ must be self-enforcing. Military power is bargaining power, because it determines the value of the outside option. If a path is self-enforcing, then the stronger is a player, the bigger must be his share of the surplus. However, if the balance of power changes over time and the players are very patient, then no self-enforcing path exists.

Suppose $p_{B\tau+1} > p_{B\tau}$ for some period $\tau \geq 1$. Thus, player B is a rising power in the sense that a power shift in favor of him is expected to occur

[39] Conditions (28) and (29) correspond to the IR conditions in the static models discussed above.

between periods τ and $\tau + 1$. Assume δ is close enough to 1 so that

$$\frac{\delta p_{B\tau+1} - p_{B\tau}}{1-\delta} > \phi_A + \delta \phi_B. \tag{30}$$

Theorem 4 *If Condition (30) holds for some $\tau \geq 1$, then no path $\{x_{B0}, x_{B1}, ..., x_{Bt}, ...\}$ can be self-enforcing.*

Proof. Suppose, in order to derive a contradiction, that $\{x_{B0}, x_{B1}, ..., x_{Bt}, ...\}$ is a self-enforcing path. Then 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

$$\frac{1}{1-\delta} - \delta \left(\frac{p_{B\tau+1}}{1-\delta} - \phi_B \right). \tag{31}$$

Since $p_{A\tau} = 1 - p_{B\tau}$, player A's expected payoff from war in period τ is

$$\frac{1 - p_{B\tau}}{1-\delta} - \phi_A. \tag{32}$$

Condition (30) implies that Equation (32) exceeds Equation (31), a contradiction. ■

Theorem 4 implies that if δ is sufficiently close to 1, then no matter how the players negotiate the short-term deals, no subgame perfect equilibrium can have permanent peace. Thus, we have established circumstances where a Thucydides trap occurs. As emphasized by Schelling (1966), military power is bargaining power. Player B's outside option will become more valuable after period τ. He cannot make a credible commitment not to exploit this fact; his continuation payoff in period $\tau + 1$ cannot be less than the value of his outside option, which is $p_{B\tau+1}/(1-\delta) - \phi_B$. In view of this, Condition (30) implies that player A prefers a preventive war in period τ, before the adverse power shift has occurred.[40] Appeasement is impossible. The only way player B can transfer utility to player A in period τ is by giving him a bigger share of the resource, but the utility the resource generates in period τ is insignificant compared to the future stream of utility that is at stake.[41]

6.1.1 Commitment problems in intrastate conflicts

Commitment problems may be especially important in civil wars (Walter, 1997, Wagner, 2000, Fearon, 1998, 2004).[42] 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.[43]

If a settlement requires the rebels to disarm and join civil society, 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).

[40] In a standard repeated prisoner's dilemma, trigger strategies make cooperation possible when δ is close to 1. But here we do not have a repeated game: the side that loses a war is eliminated from the game. When δ is close to 1, the value of permanently eliminating the opponent dominates the one-time cost of war.

[41] The limited ability to make side-payments is important. If player B could borrow from a third party, he could use the loan to appease player A. However, since long-term contracts are ruled out, player B cannot credibly commit to repay the loan.

[42] Walter (2009) argued that informational asymmetries as well as commitment problems are especially pervasive in civil wars. 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.

[43] In a democratic country, the commitment/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).

A political partition would allow the rebels to stay armed when the war is over. However, it will be costly for each side to maintain its independent military force, and the expectation of future arms races may prevent a peaceful resolution (see Section 7.1). Moreover, 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-strike advantages (Posen, 1993).

More broadly, the interaction of power shifts and commitment problems plays an important role in the literature on political conflict and democratization. Acemoğlu and Robinson (2000, 2001) looked at class conflict as a commitment problem. At a moment when the poor are relatively strong, they have a good chance at staging a successful revolution. To prevent this, the rich may try to promise the poor that they will be treated well in the future. But such promises are not credible, because the rich will be expected to behave opportunistically in the future. Powell (2004a) considered the formal connections between theories of power shifts and commitment problems in international and domestic conflicts, and the conditions under which current resources are sufficient for a negotiated settlement.

6.2 Power and resources: Ethelred's trap

In this subsection, there are no exogenous power shifts. Instead, wealth (i.e., control of resources) is power.[44] 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.[45] Player B's initial power is $\lambda(\omega_B)$.

6.2.1 Self-enforcing allocations

We first consider the notion of self-enforcing allocation in a static (one-period) model. Following Jackson and Morelli (2007) and Beviá and Corchón (2010), suppose the players cannot make any binding commitments to stay peaceful.

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

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

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

$$u_A(x_A) \geq 1 - \lambda(x_B) - \phi_A. \tag{33}$$

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

$$u_B(x_B) \geq \lambda(x_B) - \phi_B. \tag{34}$$

The allocation (x_A, x_B) is *self-enforcing* if Conditions (33) and (34) both hold.[46]

The status quo is (ω_A, ω_B). To avoid war, the two players must negotiate a self-enforcing outcome. Since war is the outside option for each player, the negotiated outcome (x_A, x_B) must also satisfy individual rationality (IR) conditions: both sides must prefer (x_A, x_B) to war at the status quo. At the status quo, player A gets $1 - \lambda(\omega_B) - \phi_A$ from war, and player B gets $\lambda(\omega_B) - \phi_B$. Thus, (x_A, x_B) satisfies the IR conditions if

$$u_A(x_A) \geq 1 - \lambda(\omega_B) - \phi_A \tag{35}$$

and

$$u_B(x_B) \geq \lambda(\omega_B) - \phi_B. \tag{36}$$

However, there may not exist any self-enforcing allocation 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:

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

Player A is a *status quo power* who prefers peace to war at the status quo:

$$u_A(\omega_A) > 1 - \lambda(\omega_B) - \phi_A.$$

To appease player B, player A may agree to give him more of the contested resource. However, this will make player B more powerful, and by assumption

[46] With more than two players, an allocation is self-enforcing if no coalition can benefit from attacking any other coalition. Jordan (2006) considered a coalitional game where wealth is power and war has no cost. He found that if the players are myopic in sense of the core, then an allocation can be self-enforcing only if the world is unipolar (dominated by a hegemon) or bipolar (dominated by two superpowers). But if the players are farsighted in the sense of von Neumann and Morgenstern, then many kinds of multi-polar systems can be stable as well.

he cannot commit to staying peaceful. For the new allocation to be self-enforcing, player A has to concede a lot, perhaps all of the resource. This is Ethelred's trap. Realizing this, player A may prefer war to appeasement.

To illustrate, suppose for each $i \in \{A, B\}$, $\phi_i = 0.1$ and $u_i(x_i) = x_i$ for all $x_i \in [0, 1]$. 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 (34). This condition requires

$$x_B \geq \frac{0.4}{1 - \zeta}. \tag{37}$$

However, player A is only willing to appease player B if Condition (35) holds. This condition requires

$$x_B \leq 0.6 + \zeta \omega_B. \tag{38}$$

Thus, if (x_A, x_B) is self-enforcing and individually rational, then x_B must satisfy both Condition (37) and Condition (38). 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 wealth (ζ is big).

6.2.2 Short-term commitments: the continuous case

We now turn to the dynamic model. Fearon (1996) showed how war can be avoided if the players can make short-term commitments and power is a continuous function of wealth.[47] Consider again the infinite horizon model from Section 6.1. But now there are no exogenous power shifts; the balance of power changes endogenously because wealth is power. Suppose player A can make take-it-or-leave-it proposals. In each period $t \geq 1$, as long as there was no war in the past, player A either starts a war or demands a share x_{At}.

[47] 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.

If player B accepts, then $x_{Bt} = 1 - x_{At}$ becomes the new state of the world; players A and B get utility x_{At} and x_{Bt} in period t, and we move to period $t+1$. Rejection means war, which ends the game. Player B wins the war with probability $\lambda(x_{Bt-1})$, where x_{Bt-1} is the state at the beginning of period t. In this sub-section, the CSF λ is assumed to be increasing and continuous.

Since player i's cost of war is ϕ_i, and winning is worth $1/(1-\delta)$, the expected payoffs from war in period t are

$$\frac{1-\lambda(x_{Bt-1})}{1-\delta} - \phi_A \tag{39}$$

for player A, and

$$\frac{\lambda(x_{Bt-1})}{1-\delta} - \phi_B \tag{40}$$

for player B. For simplicity, assume

$$\frac{\lambda(0)}{1-\delta} > \phi_B \tag{41}$$

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 (40). 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 (39), 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

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

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

$$(1-\delta)\,x_{Bt} + \delta\lambda(x_{Bt}) = \lambda(x_{Bt-1}) - (1-\delta)^2\phi_B \tag{42}$$

with initial condition $x_{B0} = \omega_B$. Condition (41) implies that for each $x_{Bt-1} \in [0,1]$ there is $x_{Bt} \in (0,1)$ such that Equation (42) is satisfied.[48]

For example, suppose the CSF is $\lambda(x_B) = 0.5 + \zeta x_B$, with $0 < \zeta \leq 0.5$. Condition (41) requires $(1-\delta)\phi_B < \lambda(0) = 0.5$. The explicit solution to the difference equation is

$$x_{Bt} = x_B^* + (\omega_B - x_B^*)\, b^t \tag{43}$$

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 (43); 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. Thus, the equilibrium path is not self-enforcing in the sense discussed above. However, by accepting player A's proposal, player B has made a short-term commitment to stay peaceful until player A can offer another slice. In this way Ethelred's trap is avoided, and equilibrium is perfectly peaceful.[49]

[48] If Condition (41) did not hold, there would still exist an equilibrium without war, but player B might accept an offer of $x_{Bt} = 0$.

[49] 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 as in Section 6.2.2, but now the CSF is

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

where χ, λ_L and λ_H are constants such that

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

Assume δ is close enough to 1 so that

$$\frac{\lambda_L - \chi}{1-\delta} > \phi_B \tag{44}$$

and

$$\frac{\delta\lambda_H - \lambda_L}{1-\delta} > \phi_A + \delta\phi_B. \tag{45}$$

Theorem 5 *If Conditions (44) and (45) hold, then there is war in every subgame perfect equilibrium.*

Proof. 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 (44). 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

$$\frac{1}{1-\delta} - \left(\frac{\lambda_H}{1-\delta} - \phi_B\right) = \frac{1-\lambda_H}{1-\delta} + \phi_B \tag{46}$$

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

$$1 + \delta\left(\frac{1-\lambda_H}{1-\delta} + \phi_B\right). \tag{47}$$

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

$$\frac{1-\lambda_L}{1-\delta} - \phi_A. \tag{48}$$

Condition (45) implies that Equation (48) exceeds Equation (47), a contradiction. ■

Condition (44) implies that the status quo is unacceptable to player B. But appeasement would trigger a power shift which, by Condition (45), would be unacceptable to player A. In this discontinuous example, Ethelred's trap makes it impossible to find a mutually acceptable short-term deal.

6.2.4 Power, resources and deterrence

War may be unavoidable if a rising power cannot commit not to challenge the status quo. Conversely, it may be difficult for a status quo power to commit to resisting a challenge. At best, such a commitment would correspond to a highly incomplete contract:

> "We cannot have a clear policy for every contingency; there are too many contingencies and not enough hours in the day to work them all out in advance" (Schelling, 1966, p. 53).

This incompleteness means that the status quo power can find loopholes for getting out of commitments ex post.

Despite the difficulty of making firm commitments, Schelling (1960, 1966) argued that deterrence may be feasible. If the status quo is challenged, the next move will not be an all-or-nothing choice between war and appeasement. Instead, there will be an unforeseeable sequence of events. Even if neither side would deliberately choose war over peace, a competition in risk-taking (brinkmanship) may escalate and get out of hand:

> "It is the essence of a crisis that the participants are not fully in control of events; they take steps and make decisions that raise or lower the danger, but in a realm of risk and uncertainty" (Schelling, 1966, p. 97).

Schelling's idea of a "threat that leaves something to chance" was formalized by Nalebuff (1986) and Powell (1990).

More recently, Gurantz and Hirsch (2017) provided a theory of credible deterrence which combines incomplete information and power shifts. If a status quo power is expected to resist a challenge by all means, including war if necessary, then only a highly aggressive "wicked" type of opponent will make the challenge. A challenge therefore reveals that the status quo power faces a wicked opponent. The status quo power concludes that war must happen sooner or later, and he prefers war now in order to prevent a power shift. This makes it credible to resist the challenge by all means.[50]

For a formal model, let player A be the status quo power and player B the challenger. Wealth is power, and λ is a continuous and strictly increasing CSF. Assume

$$u_A(\omega_A) > 1 - \lambda(\omega_B) - \phi_A, \tag{49}$$

so player A is a status quo power in the sense of Section 6.2.1.

The following three-stage game is based on Gurantz and Hirsch (2017). In stage 1, player B makes a demand 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 stay peaceful, 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 $\lambda(x_B)$.

Since player B cannot pre-commit to staying peaceful in stage 3, Ethelred's trap applies. However, as in the ultimatum game, player B is committed to fight if player A rejects his demand.

Consider first a case of complete information. Suppose it is common knowledge that player B's cost of war is high, $\phi_B = \phi_B^H$, where

$$u_B(\omega_B) > \lambda(\omega_B) - \phi_B^H. \tag{50}$$

Inequalities (49) and (50) imply that the status quo (ω_A, ω_B) is strictly self-enforcing and strictly individually rational when $\phi_B = \phi_B^H$. That is, all

[50] According to Acheson (1958), the US can credibly threaten to use nuclear weapons if there is a conventional 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).

relevant inequalities, as defined in Section 6.2.1, are strict. But player B can use his first-mover advantage to improve his position. Let x_B^* be the largest x_B such that $(1-x_B, x_B)$ is both self-enforcing and individually rational when $\phi_B = \phi_B^H$. In the unique subgame perfect equilibrium, player B demands x_B^*, player A accepts, and player B stays peaceful in stage 3. Inequality (50) implies that player B prefers the status quo to war, yet he can use his first-mover advantage to increase his share to $x_B^* > \omega_B$. Player B is not deterred from challenging the status quo, even though his cost of war is high, because he knows that player A would rather acquiesce than fight. This is the problem of credible deterrence.

Gurantz and Hirsch (2017) argued that uncertainty about the challenger's true type can facilitate credible deterrence.[51] Specifically, lets introduce a type of player B with a low cost of war, $\phi_B = \phi_B^L < \phi_B^H$, such that

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

Thus, type ϕ_B^L is a revisionist in the sense of Section 6.2.1. Suppose there is no allocation which is both self-enforcing and individually rational (as defined in Section 6.2.1) when $\phi_B = \phi_B^L$.

Suppose player B's true cost of war, $\phi_B \in \left\{\phi_B^L, \phi_B^H\right\}$, is his private information. We refer to type ϕ_B^H as the normal type, and type ϕ_B^L as the wicked type. The prior probability that player B is wicked is small but strictly positive. 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 player B is wicked and rejects the demand. The wicked type challenges the status quo by demanding $x_B > \omega_B$, so there is war. But the normal type 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 wicked type, so $(1 - x_B, x_B)$ cannot be both individually rational and self-enforcing. Either $u_B(x_B) \geq \lambda(x_B) - \phi_B^L$, 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 that if he accepts the demand then there will be war in stage 3 (Ethelred's trap), so player A prefers

[51] With repeated interactions, credible deterrence can also be facilitated by uncertainty about the status quo power's true type. That is, the status quo power may resist an initial challenge in order to create a reputation for toughness and discourage future challenges (Kreps and Wilson, 1982, Milgrom and Roberts, 1982).

to fight in stage 2 when player B's power is only $\lambda(\omega_B) < \lambda(x_B)$. Thus, if there is even a small chance that player B is wicked, then the normal type can be credibly deterred from challenging the status quo.

7 Other issues

7.1 Guns or butter?

This survey has not discussed the guns-or-butter problem, i.e., how to allocate resources between military and civilian uses. The formal analysis of this problem was pioneered by Haavelmo (1954). More recent contributions include Brito and Intriligator (1985), Garfinkel (1990), Grossman and Kim (1995), Hirshleifer (1988), Neary (1997), Skaperdas (1992) and Skogh and Stuart (1982). For a survey, see Garfinkel and Skaperdas (2007).

A basic guns-or-butter game has complete information and a single period with two stages. In the first stage, the two players simultaneously choose publicly observed arms levels.[52] Arms control agreements are ruled out by assumption. In the second stage, they bargain over the contested resource. The disagreement outcome is war, and a CSF determines the probability of winning as a function of the arms levels. There is no war in equilibrium. But they have an incentive to arm themselves in stage 1, because military power is bargaining power: a strong player gets a larger share of the resource. Thus, in equilibrium there is a costly arms race but no war.[53]

In a finitely repeated version of the basic guns-or-butter game, war may be unavoidable. In stage 2 of each period, war is the disagreement outcome

[52] In a more refined analysis, military strength requires capital and labor. In the Esteban and Ray (2008) model of social conflict, ethnic groups contain both rich and poor individuals. The rich can supply capital and the poor can supply labor. Ethnic groups can exploit these synergies, and conflicts are more likely to be ethnic in nature than class-based.

[53] If bargaining is impossible, and stage 2 simply consists of attack decisions, then there may not exist any peaceful equilibrium (even though there is no first-strike advantage). 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) studied an infinite horizon model where a similar logic holds.

as in the basic one-period game. By backward induction, there must be a costly arms race in every period, even if there is never any war. If a war causes the loser to be permanently disarmed, then there will be no more costly arms races, so the winner's military expenditures will be lower in the future.[54] If the future is sufficiently important, then in some period there may be no division of the resource that each side prefers to war (Powell, 1999, Garfinkel and Skaperdas, 2000).[55]

War may also be unavoidable if asymmetric information is added to the basic one-period game. Brito and Intriligator (1985) assumed player B privately knows his own cost of war (his type), which can be high or low. Guns-or-butter decisions are publicly observed, and player A can commit to a screening mechanism at the start of the game. Suppose player B prefers war to the status quo if and only if his cost of war is low. To avoid war, player A must makes a generous offer that the low cost type is willing to accept, but then the high cost type gets an information rent. To eliminate the rent, the screening mechanism will use the arming decision to separate the two types. Player B only arms himself if he is the low cost type. If player A observes that player B arms, then player A declares war with probability π; with probability $1 - \pi$ player A makes a generous offer. Incentive compatibility requires that π is large enough to deter the high cost type from arming. War thus emerges as a consequence of optimal screening, along the lines discussed in Section 5.1.1. But note that once player B has revealed his type by arming himself, war no longer serves any purpose; player A would strictly prefer to make a generous offer and avoid war ($\pi = 0$), thereby destroying incentive compatibility. Thus, if player A cannot pre-commit to a screening mechanism, then separating the two types by their arming decisions is impossible. Without commitment, the game will have a semi-separating equilibrium with a strictly positive probability that the high-cost type will bluff, i.e., arm himself in the hope that player A will make a generous offer. With strictly positive probability, player A calls the bluff, and if player B is not bluffing (i.e., his cost of war is low) then there is war (Baliga and Sjöström, 2013). The arms race is a game of brinkmanship, a poker game

[54] War is the only credible way to decide which player should be disarmed. A coin flip will not work if it is impossible to make a credible commitment to respect the outcome (Powell, 2006).

[55] If the game is infinitely repeated, however, an implicit arms control agreement may be supported by trigger strategies. This would remove the reason for going to war (Garfinkel, 1990).

with very high stakes. This model suggests that military preparations can lead to war for quite different reasons than in the spiral model discussed in Section 4.2.1.

Even if the players' types are common knowledge, unobserved guns-or-butter decisions generate asymmetric information and power shifts that can cause war. Meirowitz and Sartori (2008) pointed out that if there is no war in equilibrium, then neither player has any reason to acquire unobserved weapons. But if both are unarmed, then there may exist a profitable deviation: secretly arm and attack the unarmed opponent. In this case, there must be war in equilibrium. Bargaining cannot prevent war because an information asymmetry has been endogenously created by randomized arming. In mixed-strategy equilibrium, each player is uncertain of the opponent's true strength.

Meirowitz et al. (2019) assumed unobserved arming decisions in stage 1 and Nash demand game bargaining in stage 2. With randomized arming, each player is unsure of the opponent's military strength. Strong (i.e., well-armed) players are less afraid of war and therefore demand larger shares. If both players are strong, their aggressive demands are incompatible, which triggers war. They can reduce the risk of war by exchanging cheap-talk messages before making their demands. However, since the strong benefit from this, the incentive to arm increases. For this reason, cheap talk can actually make war more likely when arming decisions are taken into account. Meirowitz et al. (2019) showed that *mediated* communication can solve the problem. Cheap-talk messages are sent privately to a mediator whose proposals do not fully reveal these messages. This secrecy favors the weak, and therefore reduces the incentive to arm.

In Baliga and Sjöström's (2008) model, player B makes an unobserved arming decision (an investment in a military technology). Weapons inspections can generate hard (verifiable) information about his military strength. However, player B is not sure about player A's type. If player A is a security-seeker, then he may attack in self-defense if inspections reveal that player B is armed (as in the spiral model). 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 (and thus easily defeated). Player B's best option may be strategic ambiguity: he creates uncertainty by randomizing his arming decision,[56] and he refuses arms inspections both when he is armed and when he

[56] The randomization can be dispensed with if there is some other source of asymmetric

is unarmed.[57] Although this survey has discussed how war may be caused by incomplete information, providing more information is not always good for peace. Strategic ambiguity can prevent war because it is beneficial to hide the weakness of an unarmed player. A greedy player A can be "deterred by doubt", which reduces the incentive for player B to arm himself, thereby making a security-seeking player A feel more secure.

Debs and Monteiro (2014) considered an infinite horizon model where in each period, player B may invest in a new military technology. If he invests in period t then he will acquire the new technology in period $t+1$, which will cause a permanent power shift from then on. In period t, player A observes a noisy signal of the period t investment decision, and then either attacks or offers player B a division of the contested resource. If player B accepts, they make a short-term deal to share the resource. Long-term agreements are ruled out. If player A suspects that player B has invested, then an attack in period t can stave off the anticipated power shift. With an infinite horizon and patient players, one might expect a perfectly peaceful equilibrium with an implicit agreement: player B refrains from investing in return for generous offers. However, the more patient is player B, the more he gains from a permanent power shift, while the short-term cost of investment remains constant. Therefore, a very patient player B has a very large incentive to defect on any implicit agreement, so peace may not be attainable. Bas and Coe (2016) studied a related model where player B's weapons program must progress through a number of stages. Over time, player A imperfectly observes player B's investment decisions and the progress of the program. If player A believes the completion of the program is imminent, he may attack in order to prevent a power shift.[58]

information. Baliga and Sjöström (2012a) assumed player B makes an *observed* investment in a weapons program. The information asymmetry is generated by the stochastic success or failure of the program, not by deliberate randomization.

[57] For example, Saddam Hussein had weapons of mass destruction in the early 1990's, but not in the late 1990's. In neither situation did he reveal the truth. In the first situation, he may have wanted to avoid sanctions or preemptive strikes by a security-seeking U.S. president; in the second situation, he may have wanted to create "deterrence by doubt". The deterrence may have been directed towards a greedy U.S. president, or towards a third party such as Iran (the argument for ambiguity works the same way in either case).

[58] In these models, player A cannot counter player B's investment with his own military build-up. If he could, then he might prevent large shifts in the distribution of power without a preventive war (see Powell, 2012a).

7.2 Is democracy good for peace?

In this survey, each side of the conflict was considered a *unitary actor*. In reality, there may be differences of opinion within each side, and the top political leader who decides between war and peace must take domestic politics into account. Domestic political institutions are important mediating factors, and special interests, such as multinational corporations, may play important roles. An important question is whether democratization makes war less likely.

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. The smaller is the winning coalition, the larger is its *per capita* share of the country's resources. The disutility of a foreign policy failure is, however, the same for everyone. The trade-off between the winning coalition's private consumption and foreign policy outcomes therefore depends on the size of the winning coalition. In an autocratic country with a small winning coalition, the elite cares mainly about their own private consumption, and the leader is unlikely to lose power following a foreign policy failure. 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.[59] Democracies therefore avoid fighting each other, the so-called *democratic peace*, and they are likely to win wars against autocracies.[60]

The alternative to fighting is negotiating. In the Bueno de Mesquita et al. (1999) model, the size of the winning coalition does not influence the outcome of negotiations. However, if an increased share of the contested resource will boost the private consumption of the winning coalition, then the *per capita* boost is more significant when the winning coalition is small. An autocrat may be more willing to risk war in order to obtain an increased share, giving him an advantage when negotiating with a democratically elected leader (Bonneton, 2023). On the other hand, Fearon (1994) argued that

[59]Each member of a small winning coalition must pay a big share of the cost of war. On the other hand, capturing a contested resource (and dividing up he spoils) would yield a big per capita benefit for a small winning coalition. In this case, the "political bias" in favor of a war effort (Jackson and Morelli, 2007) would depend on the relative strength of these two effects.

[60]Since World War II, democratic states have avoided fighting each other, but they have been prone to conflict with autocracies (Oneal and Russet, 1997).

democratically elected leaders face more powerful domestic audiences, and therefore find it easier to make commitments by creating audience costs. This suggests that democratically elected leaders will be less likely to back down, giving them an advantage over autocrats.

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.[61] 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.

In Baliga, Lucca and Sjöström's (2011) model, the leaders of two countries play a Hawk-Dove game with strategic complements as in Section 4.2. Since the representative citizen of each country prefers to live in peace, peace is likely to obtain between two democracies. However, if an autocracy has a winning coalition consisting of hawkish individuals who will punish foreign policy failures, then the autocratic leader is highly fearful of being attacked, and this makes him hawkish. The representative citizen of a democracy may therefore support preemptive attacks on autocracies. If a dictator thinks he can stay in power whatever happens, then he is less fearful and therefore less threatening to other countries. 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 may have information which is not available to the representative citizen. Therefore, even in a fully democratic country, to induce the leader to act in the interests of the citizen is a principal-agent problem. In some situations, a democratically elected 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,

[61] 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.

if an ongoing war is unlikely to end in victory, the best decision would 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). In Levy and Razin's (2004) model, each country has a privately informed leader who can communicate both with the other leader and with the representative citizen of his own country. After the communication, the two countries play a 2×2 conflict game. With a high degree of strategic complementarity, a dyad consisting of two democracies has a higher probability of peace than other dyads.

7.3 How long will war last?

This survey treated war as a *costly lottery* which ends the game, and the winner gets all of the contested resource. In reality, most interstate wars end with a negotiated settlement, not with unconditional surrender.[62] If such a settlement is reached quickly, then the war will be short. 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), Smith and Stam (2004) and Heifetz and Segev (2005).

Fearon (2004) and Powell (2012b) considered infinite horizon models of civil war with power shifts and contracting/commitment problems. Fearon (2004) assumed the rebels may temporarily stop fighting so that negotiations can take place. However, the rebels might prefer to keep fighting if they expect that the government will consolidate its power during the cease-fire.[63] Fighting will then continue until one side has won a total victory. In Powell's (2012b) model, there are stages of more or less rapid power shifts in favor of the government. Fighting continues while there are rapid power shifts. Once the power shifts slow down, it becomes possible to reach a negotiated settlement.

Asymmetric information models provide a different account of how wars

[62] In contrast, civil wars often continue until one side is utterly defeated (Walters, 1997).

[63] The side that has the momentum in a war may refuse to negotiate if negotiations require a cease-fire, because a cease-fire can ruin the momentum. 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). The two sides cannot agree not to regroup during negotiations, because such an agreement would itself have to be negotiated, and in any case it may not be enforceable.

end. During war, events on the battlefield and at the bargaining table provide information about the true state of the world. When beliefs have become sufficiently aligned, a diplomatic solution becomes possible. Absent negotiations, this alignment may require a lot of intense fighting. For example, two years after Russia attacked Ukraine there are still sharply diverging opinions about how the war will end. However, if negotiations take place while the fighting goes on, then the *Coase conjecture* (Coase, 1972) suggests that agreement will be reached quickly, so the war will be short.

In a static buyer-seller model with one-sided private information and private values, the static IC and IR conditions are consistent with Pareto optimal trade. For example, if the informed buyer makes a take-it-or-leave-it offer, the outcome will be Pareto optimal. Gul and Sonnenschein (1988) showed that the Coase conjecture is valid in this case: if the uninformed seller makes a sequence of offers, and the interval between offers goes to zero, then trade will occur quickly and the outcome is (almost) Pareto optimal. With common values, however, the static IC and IR conditions can preclude Pareto optimality. In this case, regardless of how they negotiate, social surplus must be destroyed (by Pareto inefficient trade). Deneckere and Liang (2006) showed that the Coase conjecture is false in this case. Recall from Section 5.2.1 that the (static) IC and IR conditions can preclude Pareto optimality in a static conflict with one-sided private information about military strength. Regardless of how they negotiate, social surplus must be destroyed (by war). This suggests that the Coase conjecture must be false in this case as well.

Powell (2004b) considered the case of one-sided private information in a model where war is a sequence of indivisible battles, each of fixed length. Before each battle, uninformed player A makes a sequence of offers, and the war ends if an offer is accepted. Powell (2004b) found that the Coase conjecture is false if player B's private information is his military strength, formalized as the probability of collapse during battle. Fearon and Jin (2021) and Baliga and Sjöström (2022) found that war may last a long time even if the length of each battle is not held fixed, but instead equals the interval between offers which goes to zero. The intuition is analogous to the intuition given in Section 5.2.1. Thus, the same logic that explains why war must occur can explain why it cannot end quickly. If player B is militarily strong, then he is convinced that he can win on the battlefield, so he will require a large share of the resource in order to settle. But if player A offers this large share very quickly, then player B can get a large share without much fighting even if he is militarily weak. In this case, player A gets a small share regardless of player B's type. But if the prior probability that player B is weak is sufficiently large, then player A gets a higher expected payoff from the feasible strategy of continued fighting, hoping or believing in victory on the battlefield. This is a contradiction. Player B's true strength must be revealed on the battlefield before player A is willing to make a generous offer that player B's strong type is willing to accept. Since this may require a lot of fighting, war may last a long time. Eventually, a negotiated settlement becomes possible and the war ends. After a period of peace, beliefs may again start to diverge, triggering new diplomatic crises and war:

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

8 References

Acharya, A. and Ramsay, K.W., 2013. The calculus of the security dilemma. The Quarterly Journal of Political Science, 8(2), pp. 183-203.

Acheson, D., 1958. Power and Diplomacy. Harvard University Press, Cambridge, MA.

Akerlof, G.A., 1970. The market for "lemons": Quality uncertainty and the market mechanism. The Quarterly Journal of Economics, 84(3), pp. 488-500.

Azam, J.P., 1995. How to pay for the peace? A theoretical framework with references to African countries. Public Choice, 83(1-2), pp. 173-184.

Acemoglu, D. and Robinson, J.A., 2000. Why did the West extend the franchise? Democracy, inequality, and growth in historical perspective. The Quarterly Journal of Economics, 115(4), pp.1167-1199.

Acemoglu, D. and Robinson, J.A., 2001. A theory of political transitions. The American Economic Review, 91(4), pp.938-963.

Acemoğlu, D. and Wolitzky, A., 2022, Mistrust, misperception, and misunderstanding: The role of imperfect information in conflict dynamics. This Volume

Baliga, S. and Sjöström, T., 2004. Arms races and negotiations. The Review of Economic Studies, 71(2), pp.351-369.

Baliga, S. and Sjöström, T., 2008. Strategic ambiguity and arms proliferation. Journal of Political Economy, 116(6), pp.1023-1057.

Baliga, S., Lucca, D.O. and Sjöström, T., 2011. Domestic political survival and international conflict: is democracy good for peace?. The Review of Economic Studies, 78(2), pp.458-486.

Baliga, S. and Sjöström, T., 2012a. The Hobbesian Trap, in: Garfinkel, G. and Skaperdas, S. (Eds.), The Oxford Handbook of the Economics of Peace and Conflict. Oxford University Press, New York, pp. 91–109.

Baliga, S. and Sjöström, T., 2012b. The strategy of manipulating conflict. The American Economic Review, 102(6), pp.2897-2922.

Baliga, S. and Sjöström, T., 2013. Bargaining and war: A review of some formal models. The Korean Economic Review, 29(2), pp.235-266.

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

Baliga, S. and Sjöström, T. 2022. Long wars. Working Paper, Northwestern University.

Bas, M. and A. Coe (2012): "Arms diffusion and war," *Journal of Conflict Resolution*, 56(4), pp.651-674.

Bas, M. and A. Coe (2016): "A dynamic theory of nuclear proliferation and preventive war," International Organization, 70(4), pp.655-685.

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

Blattman, C. (2022). Why We Fight: The Roots of War and the Paths to Peace. Penguin.

Bonneton, N. (2023), "Nuclear Brinkmanship and Institution," Working Paper, University of Mannheim

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

Bueno 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.

Bueno 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. Padró i Miquel (2009): "Economic Shocks and Civil War", Quarterly Journal of Political Science: Vol. 4: No. 3, pp 211-228

Chassang, S. and G. Padró 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.

Chwe, M. (1994) “Farsighted coalition stability,” Journal of Economic Theory, 63: 299-325

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.

Debs, A. and N.P. Monteiro (2014), “Known Unknowns: Power Shifts, Uncertainty, and War,” International Organization, 68(1), 1–31.

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.

Esteban, J. and D. Ray (2008). "On the Salience of Ethnic Conflict," *American Economic Review* 98(5): 2185–2202.

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*. Princeton, N.J.: Princeton University Press, pp. 269-305.

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

Fearon, J. and X. Jin (2021), "Armed Conflict 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

Fletcher, C. R. L. and R. Kipling (1911), *A School History of England*. Oxford: Clarendon Press

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

Garfinkel, M. and S. Skaperdas (2000) "Conflict without misperceptions or incomplete information: how the future matters", *Journal of conflict resolution* 44: 793-807.

Garfinkel, M. and S. Skaperdas (2007). "Economics of conflict: An overview." in: Sandler, T. and K. Hartley, eds. *Handbook of Defense Economics*, Amsterdam: North-Holland.

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

Grossman, S. and O. Hart (1986). “The costs and benefits of ownership: A theory of vertical and lateral integration,” *Journal of political economy*, 94(4), 691-719.

Grossman, H. and M. Kim (1995) “Swords or plowshares? A theory of the security of claims to property.” Journal of Political Economy 103(6):1275–1288.

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.

Haass, R. (2010): *War of Necessity, War of Choice: A Memoir of Two Iraq Wars*, First Edition. New York: Simon and Schuster

Haavelmo, T. (1954) *A study in the theory of economic evolution.* Amsterdam: North-Holland

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 A. Orphanides (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.

Hobbes, T. (1983): *De Cive.* Edited by Howard Warrender. Oxford: Oxford University Press.

Hoffmann, S. 2006.“Rethinking the Linkage between Tibet and the China-India Border Conflict: A Realist Approach.” *Journal of Cold War Studies* 8 (3): 165-194.

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, N.J.: Princeton University Press

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

Jordan, J. (2006) "Pillage and property," *Journal of economic theory* 131: 26-44.

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

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

Krainin, C. (2017). "Preventive war as a result of long-term shifts in power," Political Science Research and Methods, 5:103-121.

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.

Levy, G. and R. Razin (2004). "It takes two: an explanation for the democratic peace." Journal of the European Economic Association 2.1: 1-29.

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

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.

Meirowitz, A., and Sartori, A. E. (2008). Strategic uncertainty as a cause of war. Quarterly Journal of Political Science, 3(4), 327-352.

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

Milgrom, P., and N. Stokey (1982). "Information, trade and common knowledge." Journal of Economic Theory 26: 17-27.

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

Myerson, Roger B (1979) "Incentive Compatibility and the Bargaining Problem." Econometrica, 47: 61-73.

Nalebuff, B. (1986). "Brinkmanship and nuclear deterrence: the neutrality of escalation," *Conflict Management and Peace Science*, 9: 19-30.

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.

New York Times (1985), "Reagan Expecting Follow-up Talks," November 13, page A10.

New York Times (2022), "In Defiant Speech, Putin Casts West as Russia's Enemy: Claims Annexations Will Be Forever," October 1, page A1

Oneal, J.R. and Russet, B.M., 1997. The classical liberals were right: Democracy, interdependence, and conflict, 1950–1985. International Studies Quarterly, 41(2), 267-293.

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

Powell, R. (1990), *Nuclear Deterrence Theory: The Search for Credibility.* Cambridge 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. (2012a), "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.

Powell, R. (2012b). Persistent fighting and shifting power. American Journal of Political Science, 56(3), 620-637.

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

Xenophon (1972) *The Persian Expedition.* Penguin Classics.