The Strategy and Technology of Conflict*

Sandeep Baliga, Northwestern University
Tomas Sjöström, Rutgers University

December 16, 2016

Abstract

We consider a simple bargaining model where conflict occurs if the players cannot agree to share a resource peacefully. Each player can decide to challenge the status quo. A challenge is a *strategic move*, a commitment to start a conflict unless the opponent makes a concession. Uncertainty about the cost of making a challenge generates a unique equilibrium. Increasing the cost of conflict makes the players *more* hawkish (the "stability-instability" paradox) because challenges become more profitable. Actions are strategic substitutes if the cost of conflict is large or if there is a small first-mover advantage, and strategic complements if the cost of conflict is small and there is a large first-mover advantage. When inequality is large, reducing inequality decreases the probability of conflict but, when inequality is small, reducing inequality *increases* the probability of conflict (the "Thucydides trap"). We also study the incentives to make strategic investments *ex ante* to influence the cost of conflict or the payoff to resources.

1 Introduction

Russia's recent annexation of Crimea and China's island-building in the South China Sea caught the world off guard. The United States was left in a

*We thank Daron Acemoglu, Nemanja Antic, Ethan Bueno de Mesquita, Tiberiu Dragu, Drew Fudenberg, Eric Maskin, Jean Tirole, Pierre Yared and seminar audiences for comments. All remaining errors are our responsibility.

position where it faced a stark choice between concession or a confrontation that risked war. In effect, as Schelling [20] long argued, Russia and China's "strategic moves" created a first-mover advantage. These gambits have a rich history. For example, after World War II the Soviet Union gained the first-mover advantage in Eastern Europe, by occupying it in violation of the Yalta agreement.[1] If the West had not conceded, for example in Czechoslovakia or Hungary, a military confrontation would have been quite likely – the Soviets could not have retreated from these countries without a massive loss of reputation.[2] Conversely, US soldiers stationed in Western Europe represented "the pride, the honor, and the reputation of the United States government and its armed forces" (Schelling [20], p. 47). There would have been no graceful way for them to retreat, leaving "the Soviet Union in no doubt that the United States would be automatically involved in the event of any attack on Europe" (Schelling [20]). An East-West confrontation was avoided because the Soviets conceded Western Europe just as the West had conceded Eastern Europe.

In this paper, we provide a model of conflict based on strategic moves and their costs and benefits. The model allows us to address questions such as: As the costs of conflict increase, what happens to the probability of war? When might the strategic nature of conflict display escalation (deterrence) so the marginal incentive to become aggressive increases (decreases) in an opponent's aggressiveness?

In our conflict game, each player may challenge the status quo division of a contested territory. The challenge is a strategic move or commitment, "a voluntary but irreversible sacrifice of freedom of choice" (Schelling [19]), which means that a conflict is likely unless the opponent makes a concession. The optimal challenge is to make the largest claim the opponent would concede to. If both players commit to incompatible positions, there is a costly conflict. We show that the game can be represented as a two-by-two matrix, with strategies labelled Hawk (the optimal challenge) and Dove (no

[1] At the Yalta conference in February 1945, it was agreed that the Soviet Union would recover the territory it had lost after 1941. Elsewhere there were supposed to be free elections and democratic governments.

[2] During the Berlin crisis, Khrushchev told an American visitor that Berlin was not worth a war to the US. Khrushchev was then asked whether it was worth a war to the Soviet Union. "No", he replied, "but you are the ones that have to cross a frontier" (Schelling [20], p. 46). Berlin was strictly inside the territory already controlled by the Soviets, so it would be the West's decision to risk a war by entering East Germany.

challenge). With sufficient uncertainty about the opponent's cost of making a challenge, there is a unique Bayesian-Nash equilibrium.

The technology of conflict determines the costs of conflict and the magnitude of first-mover advantage. These two parameters map into properties of different weapons that have been used at different points in time. For example, nuclear weapons are very destructive and afford little first-mover advantage when there is second-strike capability. On the other hand, after the advent of cannons in the 1400s, forts provided little defense against attack and there was significant first-mover advantage.[3]

If the first-mover advantage increases (e.g., because new weapons encourage surprise attacks), so does the probability of conflict. The effect of an increase in the cost of conflict (e.g., because new weapons are more destructive) is more subtle. If the parameters are such that the second-mover would not surrender the whole resource, then an *increased* cost of conflict makes the players *more* likely to challenge the status quo (and hence makes a conflict more likely). The reason is that in this parameter region the second-mover will concede more, the more costly a conflict would be, so the first-mover advantage is increasing in the cost of conflict.

This seem to contradict the notion that the high cost of a nuclear war prevents countries from challenging the status quo. However, challenges did occur. Pakistan has employed terrorist groups to attack India under the safety of a nuclear umbrella, and North Korea has attacked South Korean assets after conducting nuclear tests. During the Cold War, Khrushchev assisted the Cuban revolution in 1960, in defiance of the "Truman doctrine". Apparently he was convinced that the U.S. would not risk a major war by invading Cuba after the Soviets had landed. That is, he may have felt it safe to challenge the status quo precisely *because* a conflict between the superpowers would have been so costly. Indeed, the "stability-instability" paradox is the observation that there is an increase in conventional conflict after a country acquires nuclear arms (Hart [10]). This paradox appears in our model when the cost of conflict increases but not so much that the second-mover would

[3]There are many other examples. Before the advent of accurate cannon, siege warfare created little first-mover advantage as forts could withstand attack for many years. While sieges reduced normal economic activity, they were not exorbitantly costly. In the nineteenth century, first-mover advantage was predominant because of the use of trained mass armies initiated by Napoleon and because of Prussian innovations in rapid, well-planned attack with breech-loading, long range guns. See McNeil [14] for a study of changes in the technology of war over time.

concede everything. After that, the natural effect appears and an increase in the cost of conflict reduces hawkish behavior. Hence, our model not only shows the stability-instability paradox can emerge but why and when it can emerge.

The magnitude of first-mover advantage and the destructiveness of conflict also determine whether actions are strategic complements or substitutes. Both scenarios are possible in practice and indeed World War I is often described as a war of aggression and World War II as failed deterrence.[4] We show that actions are strategic substitutes if the cost of conflict is high. If the cost of conflict is low, then actions are strategic complements if the first-mover advantage is large enough (e.g., if the first-mover is very likely to win a conflict). This result is not obvious because there are two opposing effects: with a large first-mover advantage, the cost of choosing Dove when the opponent chooses Hawk is high, but so is the benefit from choosing Hawk when the opponent chooses Dove. However, the first effect dominates if utility functions are concave, so that the cost of being caught out and losing territory exceeds the benefit of acquiring more territory. This formalization of the notion of first-mover advantage adds some clarity to concepts such as "offense dominance" and "defense dominance" which are frequently discussed in the literature on international relations (see Bueno de Mesquita [3]). Similarly, our results on strategic substitutes versus complements illuminate discussions of deterrence versus escalation.

Some contemporary commentary concerned with changes in relative wealth and what impact they have on conflict. For example, what will happen as China grows relative to the United States (Allison [1])? We allow endowments to differ and study the impact of decreasing inequality on the probability of conflict. (Inequality could be changing as the countries grow at different rates.) We find that the poorer player, "the rising power", is always more aggressive than the richer player, "the status quo power" because he risks a smaller endowment in conflict. Therefore, the rising power always becomes less hawkish as inequality decreases - he has more to lose from being hawkish as he becomes wealthier and this reduces the incentive to be aggressive.

[4]For example, Nye (p. 111, [18]) observes, "World Wars I and II are often cast as two quite different models of war.. World War I was an unwanted spiral of hostility... World War II was not an unwanted spiral of hostility-it was a failure to deter Hitler's planned aggression." Also, see Jervis [11] for many other interpretations of conflicts using Stag Hunt and Chicken as metaphors.

Whether this dovishness is reciprocated depends on inequality. If inequality is large to begin with, the status quo power also becomes less aggressive when inequality declines. He has little incentive to be hawkish when the other player is dovish as he is wealthy to begin with. Hence, actions are strategic complements for the status quo power and it responds to the rising power's declining aggressiveness in kind. When inequality is low, actions are strategic substitutes for the status quo power. Therefore, it turns aggressive as the rising power becomes more accommodating with declining inequality. In fact, the probability of conflict can go *up* with declining inequality. But this happens not because of the rising power but because of the status quo power.

Finally we consider the incentive to make *ex ante* strategic moves before the bargaining game is played out. For example, an improved defensive technology might reduce player A's cost from conflict. The analysis of such ex ante moves is similar Fudenberg and Tirole's [9] analysis of strategic investment.[5] The main difference arises from the fact that player A's investment can directly impact player B's payoffs in a natural way. For example, if player A's investment in defensive technology reduces player B's first-mover advantage and makes player B *soft*, i.e., more likely to choose Dove in the bargaining game, then the defensive technology confers a strategic advantage on player A. Player A will then choose a "Top Dog" strategy and *over-invest* in defensive technology. This is different from the Top Dog strategy in Industrial Organization where over-investment is optimal when actions are strategic substitutes and the investment makes the investor more aggressive. In other cases, the results depend on the parameters in an intuitive way. Recall that actions are strategic complements if the cost of conflict is low but the first-mover advantage is large. Then an investment in defensive technology which makes player A *tough* confers a strategic *dis*advantage on player A (it makes player B more likely to choose Hawk). Player A will therefore choose a "Puppy Dog" strategy, i.e., *under-invest* to become less threatening.

Our model of bargaining resembles the Nash demand game (Nash [16]) and Schelling's ([19]) informal analysis of two-sided commitment in bargaining.[6] Unlike these classic contributions, we explicitly model first-mover advantage as well as the cost of bargaining breakdown and these two properties

[5] Tirole [23] has recently exapanded the scope of this theory to include two-stage games with ex ante information acquisition.

[6] For example, Schelling(p. 26, [19]) describes haggling over the price of a house as

are fundamental to our results. Fearon [7] has studied a model of one-sided commitment via the ultimatum game. But this formulation cannot capture the fact that both players can make moves and that their moves are not perfectly coordinated. Ellingsen and Miettinen's [6] have also offered a model of bargaining with costly commitments, in turn building on Crawford [4]. Crawford's model had multiple equilibria, including efficient equilibria where the parties made compatible commitments. Ellingsen and Mietinen [6] showed that if making a commitment is costly, the number of equilibria is reduced. But while we share the focus on strategic moves, the underlying games and hence the results are different. More recently, Meirowitz, Morelli, Ramsay and Squintani [15] also employ the Nash demand to study conflict. In their model, there is a private arming decision, followed by communication and bargaining. In our model, arming and bargaining are one and the same so we end up with a quite different approach.

2 The Bargaining Game

There are two players, A and B. In the status quo, player i controls a share $\omega_i \in (0,1)$ of a disputed territory, his *endowment*, where $\omega_1 + \omega_2 = 1$. Player i's utility of controlling a share x_i is $u_i(x_i)$, where u_i is an increasing, strictly concave and differentiable function on $[0,1]$. If a conflict occurs, then each player $i \in \{A, B\}$ suffers a cost $\phi_i > 0$.

The game has two stages. In stage 1, each player i can either make a claim x_i, where $\omega_i < x_i \leq 1$, or make no claim. A claim is a *challenge* (to the status quo) which incurs a cost c_i for the challenger.[7] To make no claim

follows:

> "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."

[7]A challenge may be a (non-revokable) instruction to player i's military to cross the status quo demarcation, which requires physical resources and manpower. The military

incurs no cost. The game ends after stage 1 if either no player makes a claim, or both make claims. Stage 2 is reached if only one player makes a claim, in which case the other chooses to concede or not to concede. The final outcome is determined by three rules.

Rule 1. If nobody challenges in stage 1, then the status quo remains in place.

Rule 2. If only player i challenges, and claims $x_i > \omega_i$ in stage 1, then we move to stage 2. In stage 2, if player j concedes then player i gets x_i and player j gets $1 - x_i$. If player j does not concede, there is a conflict: with probability σ, player i (the challenger) wins and takes all of the resource; with probability $1 - \sigma$, player j wins and takes all of the resource.

Rule 3. If both players challenge the status quo in stage 1 then there is a conflict. Each player i gets all of the resource with probability $1/2$.

We interpret these rules as follows. If neither player challenges the status quo, then there is no reason why either player should retreat from his initial position, and the status quo remains in place. If only player i challenges in stage 1 then he becomes the first-mover and player j the second-mover. The challenge is a commitment to start a conflict unless player j concedes the claim. If player j concedes, the challenger gets what he claims, and thus increases his share of the resource. If player j does not concede, there is a conflict which player i wins with probability σ; there is no way to "gracefully back down" and avoid a conflict at this point. The first-mover advantage is greater, the bigger is σ. But the first-mover advantage also depends on other things, such as the cost of conflict (if conflicts are very costly, player j is willing to concede more at stage 2). If *both* players challenge the status quo, a conflict occurs because both players are committed to increase their share of the resource.[8,9]

operation may be condemned by the international community, leading to a loss of reputation and goodwill, possible sanctions or embargoes, etc. These costs would be included in c_i. For example, Stalin lost goodwill in the West when he violated the Yalta agreement. Apparently, he did not value goodwill very highly – but this was not known in the West. Similarly, Russia has invaded Ukraine and faced sanctions as a result. The Russian economy is suffering but President Putin's popular support is sky high.

[8]Rule 2 resembles Fearon's [7] ultimatum game analysis of conflict. Unlike Fearon's model, our game is symmetric like Nash's [16] demand game or Schelling's [19] model of bargaining. Rule 3 reflects their idea that conflict arises when players make incompatible strategic moves.

[9]A more general formulation would be that if both choose to challenge, there is some

To reduce the complexity of the exposition, we have assumed the two players are symmetric in terms of fighting strength. That is, σ is the same for each player and each player's probability of winning under rule 3 is 1/2. (Generalizing the model with asymmetric fighting strength would be straightforward.) Further, we make two assumptions that eliminate some less interesting cases.

First, we assume that each player's endowment is large enough that they never want to actively invite conflict:

$$u_i(\omega_i) > \sigma u_i(0) + (1-\sigma)u_i(1) - \phi_i. \tag{1}$$

The left-hand side is player i's payoff from the status quo, and the right hand side is his expected payoff when he is the second-mover and does not concede (so there is a conflict which he wins with probability $1-\sigma$). Therefore, each player prefers the status quo to conflict if and only if $\sigma > \underline{\sigma}_i$ where

$$\underline{\sigma}_i \equiv \frac{u_i(1) - u_i(\omega_i) - \phi_i}{u_i(1) - u_i(0)}. \tag{2}$$

Hence, we assume:

Assumption 1 $\sigma > \underline{\sigma}_i$ for $i \in \{A, B\}$.

Note that if the cost of conflict is high enough, specifically $\phi_i > u_i(1) - u_i(\omega_i)$, then Assumption 1 is automatically satisfied because $\underline{\sigma}_i < 0$. Note also that strict concavity implies $u_i\left(\frac{1}{2}\right) > \frac{1}{2}u_i(1) + \frac{1}{2}u_i(0)$. Therefore, in the symmetric case where $\omega_i = 1/2$ we have $\underline{\sigma}_i < 1/2$, so Assumption 1 is satisfied whenever $\sigma \geq 1/2$. (Of course, Assumption 1 may hold even for $\sigma < 1/2$.)

We now identify the minimum payoff that must be guaranteed to the second-mover to concede without a fight and hence the maximum payoff the first-mover can capture without a fight. We also identify some basic

---

probability $\alpha > 0$ that player $i \in \{A, B\}$ becomes committed before player j, in which case we move to stage 2 where player j decides whether or not to concede. Thus, each player would have a probability α of getting the first mover advantage. With probability $1-2\alpha$, they both become committed, and there is a conflict. Similarly, following Crawford [4] and Ellingsen and Mietinen [6], we could assume that a challenge only leads to a successful commitment with probability $q < 1$. But the more general model produces similar results to our current model, which is the special case $\alpha = 0$ and $q = 1$, so we present the simpler model for the sake of exposition. In particular, unlike in the Ellingsen and Mietinen [6] model, there would be no dramatic change in the set of equilibria at $q = 1$.

properties of these payoffs. Consider the second mover's behavior if stage 2 is reached. If player i is the second-mover and concedes to the claim x_j he gets $u_i(1-x_j)$. If he doesn't concede, he gets

$$\sigma u_i(0) + (1-\sigma)u_i(1) - \phi_i.$$

Thus, player i prefers to concede if

$$u_i(1-x_j) \geq \sigma u_i(0) + (1-\sigma)u_i(1) - \phi_i. \tag{3}$$

This is satisfied for $x_j = 1$ if

$$\phi_i \geq (1-\sigma)\left(u_i(1) - u_i(0)\right). \tag{4}$$

In this case, player i would rather concede the whole territory than have a conflict. If instead

$$\phi_i < (1-\sigma)\left(u_i(1) - u_i(0)\right) \tag{5}$$

then the maximum claim he will concede to satisfies (3) with equality. Thus, if player j makes the maximal claim that player i will concede to in stage 2, player i's share of territory will be:

$$\eta_i \equiv \begin{cases} u_i^{-1}\left[\sigma u_i(0) + (1-\sigma)u_i(1) - \phi_i\right] & \text{if} \quad \phi_i < (1-\sigma)\left(u_i(1) - u_i(0)\right), \\ 0 & \text{if} \quad \phi_i \geq (1-\sigma)\left(u_i(1) - u_i(0)\right). \end{cases} \tag{6}$$

This means $x_j = 1 - \eta_i$ is the maximum share that player j, as first-mover, can get without a fight.

Notice that when η_i is interior as (5) holds, we must have

$$u_i(\eta_i) = \sigma u_i(0) + (1-\sigma)u_i(1) - \phi_i. \tag{7}$$

Equation (7) says that player i is indifferent between η_i and a conflict when he is the second-mover. It reveals that η_i is decreasing in ϕ_i. That is, the more costly a conflict would be, the more player i is willing to concede in stage 2. This property will play a key role in some of our results below. Also, when η_i is interior,

$$u_i(0) < \sigma u_i(0) + (1-\sigma)u_i(1) - \phi_i < u_i(\omega_i), \tag{8}$$

where the first inequality is equivalent to (5) and the second is equivalent to Assumption 1. In view of (8), if (5) and Assumption 1 hold then $0 < \eta_i < \omega_i$. In general, $0 \leq \eta_i < \omega_i$ since either the first or second row of (6) may apply.[10]

We assume neither player knows the opponent's cost of violating the status quo and there is significant uncertainty about this cost. The idea that players in conflict may have private information about their payoffs is natural. As we show in the next section, our assumption that there is sufficiently large uncertainty also plays a technical role in ensuring a unique equilibrium and simplifies our analysis.

We assume that for each $i \in \{A, B\}$, the cost c_i is independently drawn from a distribution F with support $[\underline{c}, \overline{c}]$ and density $f(c) = F'(c)$. Player $i \in \{A, B\}$ knows c_i but not c_j. We refer to c_i as player i's type. If either the support of F is very small, or the density of F is highly concentrated around one point in the support, then the uncertainty is unimportant, because the players in effect are fairly certain about each others' types. To avoid this, we assume (i) that the support is not too small, and (ii) that the density is sufficiently "flat":

Assumption 2 (Sufficient uncertainty about types) (i)

$$\underline{c} < \min\{u_i(1-\eta_j) - u_i(\omega_i),\ \frac{1}{2}u_i(0) + \frac{1}{2}u_i(1) - \phi_i - u_i(\eta_i)\}$$

[10]Notice that Assumption 1 does not imply $\phi_i \geq (1-\sigma)(u_i(1) - u_i(0))$. For example, suppose $\omega_i = 1/2$. We have

$$\sigma u_i(0) + (1-\sigma)u_i(1) - u_i(1/2) < (1-\sigma)(u_i(1) - u_i(0)) \tag{9}$$

This inequality follows from

$$\begin{aligned} & (1-\sigma)(u_i(1) - u_i(0)) \\ & - [\sigma u_i(0) + (1-\sigma)u_i(1) - u_i(1/2)] \\ = \quad & u_i(1/2) - u_i(0) > 0. \end{aligned}$$

Therefore, there is a range of ϕ_i, namely

$$\begin{aligned} & \sigma u_i(0) + (1-\sigma)u_i(1) - u_i(1/2) \\ < \quad & \phi_i < (1-\sigma)(u_i(1) - u_i(0)) \end{aligned}$$

such that (8) holds.

and

$$\bar{c} > \max\{u_i(1-\eta_j) - u_i(\omega_i),\ \frac{1}{2}u_i(0) + \frac{1}{2}u_i(1) - \phi_i - u_i(\eta_i)\}$$

for $i \in \{A, B\}$. (ii)

$$f(c) < \frac{1}{\left|\frac{1}{2}u_i(0) + \frac{1}{2}u_i(1) - \phi_i - u_i(\eta_i) - u_i(1-\eta_j) + u_i(\omega_i)\right|}$$

for all $c \in [\underline{c}, \bar{c}]$ and $i \in \{A, B\}$.

If F is uniform, then (ii) is redundant because (i) implies (ii). Indeed, the uniform distribution is maximally "flat". However, we do not restrict attention to the uniform distribution. In the non-uniform case, (ii) guarantees that the density is not highly concentrated at one point.

3 Equilibrium and Basic Comparative Statics Results

Our model has a rich strategy set for each player. But we show only two strategies are used in equilibrium and that there is a unique equilibrium. We then perform comparative statics exercises with the equilibrium and determine when the stability-instability paradox can arise.

3.1 Equilibrium Analysis

We show that there is only one challenge a player would ever make. His other option is not to challenge at all so we will show our bargaining game reduces to a simple 2×2 game.

We begin by showing that player i's best challenge is to claim $1 - \eta_j$. If both players challenge simultaneously, the outcome falls under Rule 3, and any challenge will give the same lottery over outcomes. So to distinguish between player i's challenges, we need to consider what happens if stage 2 is reached following a challenge by player i alone. In this case, sequential rationality implies that player j concedes if and only if player i's claim satisfies $x_i \leq 1 - \eta_j$.[11] So player i should certainly not claim strictly less than $1 - \eta_j$.

[11] If player i claims $1 - \eta_j$ then player j is indifferent between conceding and not conceding, but we may assume he concedes in this case.

If he claims exactly $1 - \eta_j$, there is no conflict, and player i's payoff is

$$u_i(1 - \eta_j). \tag{10}$$

If $\eta_j = 0$, player i's best challenge is certainly to claim $x_i = 1$. If $\eta_j > 0$, we must consider what happens if player i claims strictly more than $1 - \eta_j$. Then there will be a conflict which, by Rule 2(b), gives player i expected payoff

$$\sigma u_i(1) + (1 - \sigma)\, u_i(0) - \phi_i. \tag{11}$$

But (11) is strictly smaller than (10). To see this, note that, by definition of η_j, if player i claims $1 - \eta_j$ then player j's payoff is $u_j(\eta_j)$ whether there is a conflict or not (see (7)). But conflicts are inefficient since $\phi_i > 0$, so player i strictly prefers to not have a conflict and get $1 - \eta_j$ for sure. Thus, (11) is strictly smaller than (10), so claiming $1 - \eta_j$ is strictly better than claiming $x_i > 1 - \eta_j$. Thus, we define player i's *optimal challenge* to be to claim $x_i = 1 - \eta_j$. Notice that player i can compute his optimal challenge without knowing c_j, since η_j is independent of c_j.

For convenience, we will label the optimal challenge *Hawk* (or H). To not make any challenge is labelled *Dove* (or D). Thus, we obtain the following 2×2 payoff matrix. Player i chooses a row, player j a column, and only player i's payoff is indicated:

	Hawk (claim $x_j = 1 - \eta_i$)	Dove (no challenge)
Hawk (claim $x_i = 1 - \eta_j$)	$\frac{1}{2}u_i(0) + \frac{1}{2}u_i(1) - \phi_i - c_i$	$u_i(1 - \eta_j) - c_i$
Dove (no challenge)	$u_i(\eta_i)$	$u_i(\omega_i)$

(12)

Remark 1 *If only one player challenges the status quo, part of the resource is transferred to the challenger, but there is no conflict and no ex post inefficiency. An inefficient conflict occurs only when* both *players challenge the status quo.*

Remark 2 *We know that* $u_i(1 - \eta_j) > \sigma u_i(1) + (1 - \sigma)\, u_i(0) - \phi_i$, *and also* $\omega_i > \eta_i$ *so* $u_i(\omega_i) > u_i(\eta_i)$. *Therefore, if* $\sigma \geq 1/2$ *then*

$$u_i(1 - \eta_j) - c_i > \frac{1}{2}u_i(0) + \frac{1}{2}u_i(1) - c_i - \phi_i. \tag{13}$$

The inequality (13) also holds when $\omega_j \leq 1/2$, *since in this case* $1 - \eta_j > 1 - \omega_j \geq 1/2$. *Thus, if either* $\sigma \geq 1/2$ *or* $\omega_j \leq 1/2$ *(or both) then player*

i always prefers player j to choose Dove, whatever action player i himself chooses. If, however, $\sigma < 1/2$ and $\omega_j > 1/2$ then player i may actually want player j to choose Hawk if player i himself chooses Hawk.

Next, we show that under Assumption 2, the bargaining game has a unique equilibrium.

Player i is a *dominant strategy hawk* if Hawk (H) is his dominant strategy.[12] Player i is a *dominant strategy dove* if Dove (D) is his dominant strategy.[13] Assumption 2(i) implies that the support of F is big enough to include dominant strategy types of both kinds.

Suppose player i thinks player j will choose H with probability p_j. Player i's expected payoff from playing H is

$$-c_i + p_j \left(\frac{1}{2} u_i(0) + \frac{1}{2} u_i(1) - \phi_i \right) + (1 - p_j)\, u_i(1 - \eta_j),$$

while his expected payoff from D is

$$p_j u_i(\eta_i) + (1 - p_j) u_i(\omega_i).$$

Thus, if he chooses H instead of D, his *net* gain is

$$-c_i + p_j \left(\frac{1}{2} u_i(0) + \frac{1}{2} u_i(1) - \phi_i - u_i(\eta_i) \right) + (1 - p_j) \left(u_i(1 - \eta_j) - u_i(\omega_i) \right). \tag{14}$$

A *strategy* for player i is a function $g_i : [\underline{c}, \overline{c}] \to \{H, D\}$ which specifies an action $g_i(c_i) \in \{H, D\}$ for each type $c_i \in [\underline{c}, \overline{c}]$. In Bayesian Nash equilibrium (BNE), all types maximize their expected payoff. Therefore, $g_i(c_i) = H$ if (14) is positive, and $g_i(c_i) = D$ if (14) is negative. If (14) is zero then type c_i is indifferent, and for convenience we assume he chooses H in this case.

Player i uses a *cutoff strategy* if there is a *cutoff point* $x \in [\underline{c}, \overline{c}]$ such that $g_i(c_i) = H$ if and only if $c_i \leq x$. Because the expression in (14) is monotone in c_i, all BNE must be in cutoff strategies. Therefore, we can without loss of generality restrict attention to cutoff strategies. Any such strategy is identified with its cutoff point $x \in [\underline{c}, \overline{c}]$. If player j uses cutoff point x_j, the

[12] Formally, $u_i(1 - \eta_j) - u_i(1/2) \geq c_i$ and $\frac{1}{2} u_i(0) + \frac{1}{2} u_i(1) - \phi_i - u_i(\eta_i) \geq c_i$ with at least one strict inequality.

[13] Formally, $u_i(1 - \eta_j) - u_i(1/2) \leq c_i$ and $\frac{1}{2} u_i(0) + \frac{1}{2} u_i(1) - \phi_i - u_i(\eta_i) \leq c_i$ with at least one strict inequality.

probability he plays H is $p_j = F(x_j)$. Therefore, using (14), player i's best response to player j's cutoff x_j is the cutoff $x_i = \Gamma_i(x_j)$, where

$$\Gamma_i(x) \equiv F(x)\left(\frac{1}{2}u_i(0) + \frac{1}{2}u_i(1) - \phi_i - u_i(\eta_i)\right) + (1 - F(x))\left(u_i(1-\eta_j) - u_i(\omega_i)\right). \tag{15}$$

The function Γ_i is the best-response function for cutoff strategies.

The role of Assumption 2(i) is to rule out corner solutions, where all types do the same thing. Indeed.

$$\Gamma_i(\underline{c}) = u_i(1-\eta_j) - u_i(\omega_i) > \underline{c}$$

and

$$\Gamma_i(\bar{c}) = \left(\frac{1}{2}u_i(0) + \frac{1}{2}u_i(1) - \phi_i\right) - u_i(\eta_i) < \bar{c}$$

by Assumption 2(i), so the equilibrium cutoff point will be strictly between $\underline{c}$ and $\bar{c}$.

Since the function $(\Gamma_A(x_B), (\Gamma_B(x_A)) : [\underline{c}, \bar{c}]^2 \to (\underline{c}, \bar{c})^2$ is continuous, a fixed-point $(\hat{x}_A, \hat{x}_B) \in (\underline{c}, \bar{c})^2$ exists. This is a BNE (where player i uses cutoff $\hat{x}_i$). Thus, a BNE exists. The slope of the best response function is $\Gamma_i'(x) = \Omega_i f(x)$, where

$$\Omega_i \equiv \frac{1}{2}u_i(0) + \frac{1}{2}u_i(1) - u_i(\eta_i) - u_i(1-\eta_j) + u_i(\omega_i) - \phi_i. \tag{16}$$

A standard sufficient condition for the existence of a *unique* equilibrium is that the absolute value of the slope of each player's best response function is less than 1. Assumption 2(ii) guarantees this. Thus, while Assumption 2(i) guarantees that any BNE is interior, Assumption 2(ii) guarantees that there is a unique BNE.

3.2 First-Mover Advantage, The Cost of Conflict and the Stability-Instability Paradox

If the two players are symmetric ex ante (before they draw their types), in preferences over territory, costs of conflict and endowments, then we can drop the subscripts on u_i, ϕ_i, ω_i and η_i. The payoff matrix becomes:

	Hawk (claim $x_j = 1-\eta$)	Dove (no challenge)
Hawk (claim $x_i = 1-\eta$)	$\frac{1}{2}u(0) + \frac{1}{2}u(1) - \phi - c_i$	$u(1-\eta) - c_i$
Dove (no challenge)	$u(\eta)$	$u(\frac{1}{2})$

(17)

Also, the unique BNE must be symmetric, and the equilibrium cutoff $\hat{x}$ is the same for both and implicitly defined by the equation

$$\hat{x} - \Omega F(\hat{x}) = u(1-\eta) - u(1/2). \tag{18}$$

where

$$\Omega \equiv \frac{1}{2}u(0) + \frac{1}{2}u(1) - u(\eta) - u(1-\eta) + u\left(\frac{1}{2}\right) - \phi \tag{19}$$

The symmetry assumption allows us to easily derive comparative statics results. Consider first how η depends on ϕ and σ. If $\phi > (1-\sigma)(u(1) - u(0))$ then $\eta = 0$ and $d\eta/d\sigma = d\eta/d\phi = 0$. But if $\phi < (1-\sigma)(u(1) - u(0))$ then (7) holds the second-mover concedes more if σ and ϕ increase as he is a weaker position if he triggers a conflict by playing Hawk:

$$\frac{d\eta}{d\sigma} = -\frac{u(1) - u(0)}{u'(\eta)} < 0 \tag{20}$$

and

$$\frac{d\eta}{d\phi} = -\frac{1}{u'(\eta)} < 0. \tag{21}$$

While both greater first-mover advantage and a higher cost of conflict decrease a second-mover's payoff η, they have different equilibrium effects on the probability of conflict.

Consider an increase in first-mover advantage, e.g. the advent of accurate siege cannon. By definition, the magnitude of σ affects payoffs when player i challenges and is lucky to catch player j by surprise (i.e. when the action profile is HD) or in the reverse situation when player j catches player i by surprise (i.e. when the action profile is DH). In the former case, higher σ allows player i to extract more resources from player j - $u(1-\eta)$ increases. This increases his incentive to play Hawk. In the latter case, player i concedes more when he is himself caught off guard - $u(\eta)$ decreases. This also increases his incentive to play Hawk. Once σ becomes so high that $\eta = 0$, there is no further impact on the payoff functions and the probability of conflict does not change. There is then no non-monotonicity or stability-instability paradox in changes in first-mover advantage: increasing first-mover advantage also increases the equilibrium probability of conflict.

Proposition 3 *Suppose the players are symmetric ex ante. An increase in first-mover advantage σ increases the probability of conflict if $\phi < (1-\sigma)(u(1)-u(0))$. It has no effect on the probability of conflict when $\phi > (1-\sigma)(u(1)-u(0))$.*

Proof. Totally differentiating (18) we obtain

$$(1-\Omega f(\hat{x}))\frac{d\hat{x}}{d\sigma} = -\left[u'(\eta)F(\hat{x}) + u'(1-\eta)(1-F(\hat{x}))\right]\frac{d\eta}{d\sigma} \qquad (22)$$

where $1-\Omega f(\hat{x}) > 0$ from Assumption 2. From (6), the expression in (22) vanishes if $\phi > (1-\sigma)(u(1)-u(0))$. In this case, the second-mover concedes everything, so an increased σ has no effect on behavior. But if $\phi < (1-\sigma)(u(1)-u(0))$ then (20) holds. From (22), the equilibrium cutoff increases, so each player becomes more likely to choose H when σ increases. ■

Consider an increase in the cost of conflict, e.g. the advent of nuclear weapons. What impact will this have on the probability of conflict? The obvious intuition is that players will shrink from aggression because the costs of conflict when they are both hawkish have increased. But when ϕ is low, an increase in the cost of conflict confers a first-mover advantage for the same reasons as an increase in σ : it increases the incentive to play Hawk when the opponent plays Dove as the opponent will concede more, and it increases the incentive to play Hawk when the opponent plays Hawk as a dovish player has to concede more. When ϕ is low, these two effects overcome the incentive to shrink from conflict when ϕ increases. Thus, the equilibrium probability of conflict actually *increases* with ϕ when ϕ is low as both players become more aggressive trying to exploit increased first-mover advantage. For a sufficiently high cost of conflict, players will concede everything when faced with a surprise hawkish move. Further increases in the cost of conflict do not increase first-mover advantage and the probability of conflict falls with higher ϕ. Therefore we identify the impact of increased costs of conflict on first-mover advantage as the source of the stability-instability paradox. Increasing ϕ causes "instability" at low costs of conflict and "stability" at high costs of conflict:

Proposition 4 ***Stability-Instability Paradox*** *Suppose the players are symmetric ex ante. An increase in the cost of conflict ϕ increases the probability of conflict if $\phi < (1-\sigma)(u(1)-u(0))$, but reduces the probability of conflict when $\phi > (1-\sigma)(u(1)-u(0))$.*

Proof. When $\phi < (1-\sigma)(u(1) - u(0))$, totally differentiate (18) with respect to ϕ and use (21) to obtain

$$\frac{d\hat{x}}{d\phi} = \frac{1}{1 - \Omega f(\hat{x})} \frac{u'(1-\eta)}{u'(\eta)} (1 - F(\hat{x})) > 0.$$

Since the equilibrium cut-off increases, the players become more likely to choose Hawk, the higher is the cost of conflict.

When $\phi > (1-\sigma)(u(1) - u(0))$, an increase in ϕ will have no effect on η, and therefore it will reduce the probability of conflict. Indeed, when η is fixed at 0 we get

$$\frac{d\hat{x}}{d\phi} = -\frac{1}{1 - \Omega f(\hat{x})} F(\hat{x}) < 0.$$

■

4 Strategic Complements and Substitutes

Two-by-two matrix games are often used as stylized models of conflicts. A key distinction is between games of *strategic complements*, such as stag hunt, and games of *strategic substitutes*, such as chicken. Intuitively, the game has strategic complements (substitutes) if the gain from becoming more hawkish is greater (smaller), the more hawkish is the other player. The stag hunt game is a model of Hobbes's "state of nature", where conflict is caused by lack of trust, while chicken is a model of preemption and deterrence. In previous work (e.g., Baliga and Sjöström [2]), the results depended crucially on whether actions were *assumed* to be strategic complements or substitutes. Now we will use the payoff matrix (12) to characterize situations where actions are likely to be substitutes or complements.[14]

If player j chooses Hawk, then if player i switches from Dove to Hawk player i's net gain is

$$\frac{1}{2}u_i(0) + \frac{1}{2}u_i(1) - \phi_i - c_i - u_i(\eta_i) \tag{23}$$

If instead player j chooses Dove, then if player i switches from Dove to Hawk player i's net gain is

$$u_i(1-\eta_j) - c_i - u(\omega_i). \tag{24}$$

[14] Historical studies (see Jervis [11] or Nye [18]) also emphasize the distinction between situations where toughness feeds on itself in a cycle of fear (as in stag hunt), and situations where a sufficient show of toughness might force an opponent to back down (as in chicken).

Actions are *strategic complements for player i* if (23) is greater than (24), which is equivalent to $\Omega_i > 0$, where Ω_i is defined by (16). They are *strategic substitutes for player i* if $\Omega_i < 0$. The game is said to have strategic substitutes (resp. complements) if the actions are strategic substitutes (resp. complements) for both players.

We begin by showing that when there is no first-mover advantage, the game must have strategic substitutes. In this case, there is little to nothing to gain from an unmet challenge and nothing to lose by facing an hawkish opponent. This implies actions are strategic substitutes. For example, besieged cities could survive for years before the advent of cannons and perhaps there was no first-mover advantage.

Proposition 5 *The game has strategic substitutes if $\sigma \leq 1/2$.*

Proof. If $\eta_i > 0$ then (7) holds so

$$\Omega_i \equiv \left(\sigma - \frac{1}{2}\right)(u_i(1) - u_i(0)) - u_i(1-\eta_j) + u_i(\omega_i) < \left(\sigma - \frac{1}{2}\right)(u_i(1) - u_i(0))$$

since $1 - \eta_j > \omega_i$. If $\eta_i = 0$ then

$$\begin{aligned} \Omega_i &\equiv \frac{1}{2}(u_i(1) - u_i(0)) - u_i(1-\eta_j) + u_i(\omega_i) - \phi_i < \frac{1}{2}(u_i(1) - u_i(0)) - \phi_i \\ &\leq (\sigma - \frac{1}{2})(u_i(1) - u_i(0)) \end{aligned}$$

where the first inequality is due to $1 - \eta_j > \omega_i$.and the second to $\phi_i \geq (1-\sigma)(u_i(1) - u_i(0))$. Thus, it is always true that

$$\Omega_i < (\sigma - \frac{1}{2})(u_i(1) - u_i(0)) \leq 0.$$

■

To simplify the exposition, for the remainder of this section we will return to the case where the players are ex ante symmetric. This implies $\eta_A = \eta_B = \eta$, $\Omega_A = \Omega_B = \Omega$ (as defined by (19)) and

$$\underline{\sigma}_A = \underline{\sigma}_B = \underline{\sigma} \equiv \frac{u(1) - u(1/2) - \phi}{u(1) - u(0)} \tag{25}$$

from (2). As was shown above, $\underline{\sigma} < 1/2$. The game has strategic substitutes if $\Omega < 0$ and strategic complements if $\Omega > 0$.

The payoff matrix (12) becomes

$$\begin{array}{lcc} & \text{Hawk} & \text{Dove} \\ \text{Hawk} & \frac{1}{2}u(0) + \frac{1}{2}u(1) - \phi - c_i & u(1-\eta) - c_i \\ \text{Dove} & u(\eta) & u(1/2) \end{array} \tag{26}$$

Totally differentiating Ω yields

$$\frac{d\Omega}{d\sigma} = -\left(u'(\eta) - u'(1-\eta)\right)\frac{d\eta}{d\sigma} \geq 0 \tag{27}$$

with strict inequality when $\eta > 0$, in view of (20) and concavity. Also,

$$\begin{aligned} \frac{d\Omega}{d\phi} &= -\left(u'(\eta) - u'(1-\eta)\right)\frac{d\eta}{d\phi} - 1 \\ &= \left(u'(\eta) - u'(1-\eta)\right)\frac{1}{u'(\eta)} - 1 = -\frac{u'(1-\eta)}{u'(\eta)} < 0. \end{aligned}$$

by concavity. Thus, actions are more likely to be strategic complements the bigger is σ and the smaller is ϕ. We confirm this intuition in various results below.

It is clear that if ϕ is large, then the most important consideration is to avoid a conflict, as in the classic chicken game. For example, if playing Hawk against an opponent who is playing Hawk brings a risk of nuclear war, there is a strong incentive to back off. But if an opponent is playing Dove, there is a strong incentive to play Hawk as the opponent will likely concede everything to avoid nuclear war. Thus, we have the following result.

Proposition 6 *Suppose the players are symmetric ex ante. If $\phi > u(1/2) - \frac{1}{2}u(0) - \frac{1}{2}u(1)$ then actions are strategic substitutes.*

Proof. By concavity,

$$u(\eta) + u(1-\eta) \geq u(0) + u(1).$$

Therefore,

$$\Omega = \left(u(0) + u(1) - u(\eta) - u(1-\eta)\right) + \left(u\left(\frac{1}{2}\right) - \frac{1}{2}u(0) - \frac{1}{2}u(1)\right) - \phi < 0.$$

■

If ϕ is small, however, then avoiding a conflict is less important, and actions become strategic complements if the first-mover advantage is large enough. A large σ has two effects: it becomes more costly to be caught out and play Dove against Hawk, but it becomes more profitable to play Hawk against Dove. The first effect tends to make actions strategic complements, while the second effect does the opposite. The first effect dominates because of strict concavity: it is more important to preserve your own territory than to acquire the opponent's territory. Overall, our result is:

Proposition 7 *Suppose the players are symmetric ex ante. There exists $\bar{\sigma} \in (1/2, 1)$ such that, for any $\phi < u(1/2) - \frac{1}{2}u(0) - \frac{1}{2}u(1)$, there is $\sigma^*(\phi) \in (1/2, \bar{\sigma})$ such that actions are strategic substitutes if $\sigma < \sigma^*(\phi)$ and strategic complements if $\sigma > \sigma^*(\phi)$. Moreover, such that $\sigma^*(\phi) < \bar{\sigma}$ for all $\phi < u(1/2) - \frac{1}{2}u(0) - \frac{1}{2}u(1)$.*

Proof. Fix ϕ such that $\phi < u(1/2) - \frac{1}{2}(u(0) + u(1))$. From Proposition 5, we have $\Omega < 0$ if $\sigma \leq 1/2$. Now let $\bar{\sigma} \in (1/2, 1)$ be implicitly defined by

$$\phi = (1 - \bar{\sigma})(u(1) - u(0)).$$

Then $\eta = 0$ if and only if $\sigma \geq \bar{\sigma}$. When $\eta = 0$ we have

$$\Omega = u\left(\frac{1}{2}\right) - \frac{1}{2}u(0) - \frac{1}{2}u(1) - \phi > 0$$

so that $\Omega > 0$ when $\sigma \geq \bar{\sigma}$. Thus, there exists $\sigma^*(\phi) \in (1/2, \bar{\sigma})$ such that $\Omega = 0$. At $\sigma = \sigma^*(\phi)$ we have $\eta > 0$. It follows from (27) that $\Omega < 0$ if $\sigma < \sigma^*(\phi)$ and $\Omega > 0$ if $\sigma > \sigma^*(\phi)$. ■

From the mid-seventeenth century to the eighteenth, wars between great powers were long and bloody (see Levy [13] and Fearon [8]). Hence, they were costly both in terms of resources and men. Also, the *trace italienne* (fort walls reinforced with soft, absorbent earth and protected by a wide ditch) provided a defense against artillery (see Duffy [5]). Hence, we identify this period with low σ (i.e. defense dominance), high ϕ and a conflict game with strategic substitutes. By contrast in the nineteenth century and until

World War 1, wars are shorter and involve fewer battle deaths in absolute and relative terms. Moreover, the development of large, patriotic armies, railways and new weapons conferred first-mover advantage. Hence, we identify this period with high σ (i.e. offensive advantage), low ϕ and a conflict game with strategic complements.

Finally, it is easy to check that the function $\sigma^*(\phi)$ is decreasing in ϕ.[15] This observation and Propositions 7 and 6 are summarized in Figure 1.

[15] Since η depends on ϕ and σ, we can write $\eta = \eta(\phi, \sigma)$. The function $\sigma^*(\phi)$ identified in Proposition 7 is such that $\Omega = 0$ when $\eta = \eta(\phi, \sigma^*)$. Substitute $\eta = \eta(\phi, \sigma^*(\phi))$ in (16) to get

$$\frac{1}{2}u(0) + \frac{1}{2}u(1) - u(\eta(\phi, \sigma^*(\phi)))) - u(1 - \eta(\phi, \sigma^*(\phi))) + u\left(\frac{1}{2}\right) - \phi \equiv 0 \qquad (28)$$

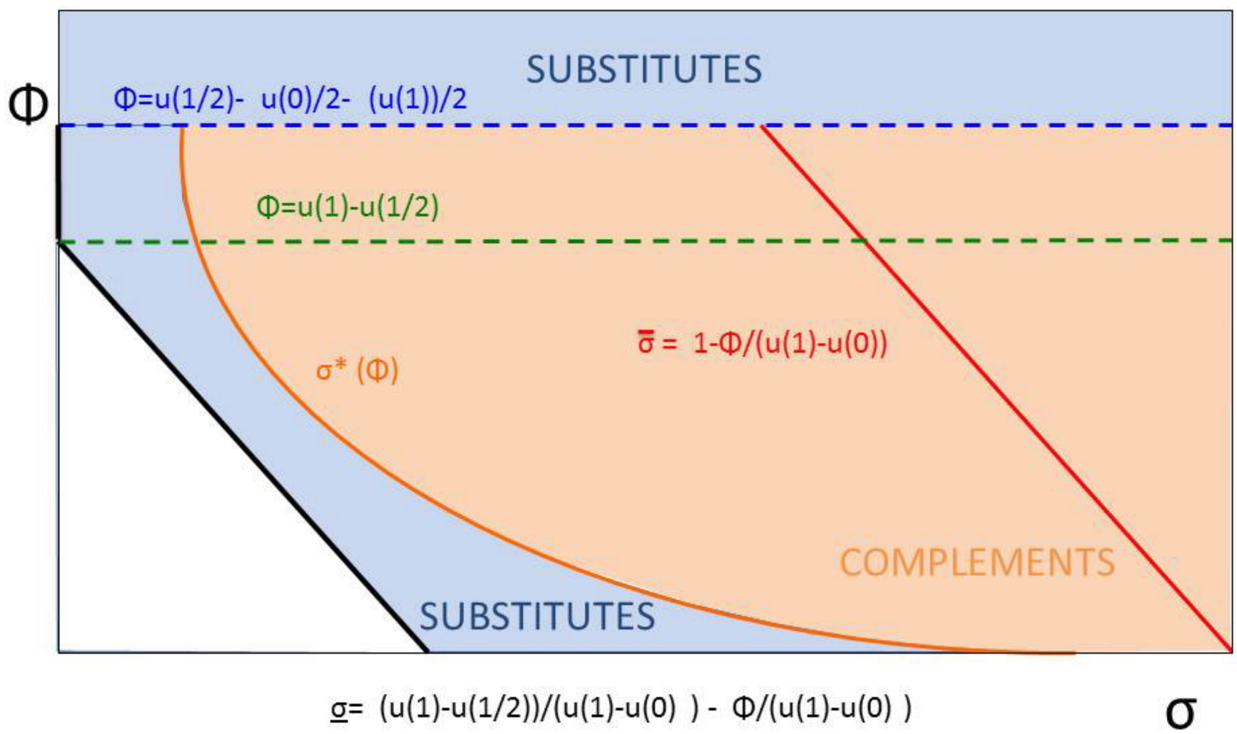
for all $\phi < u(1/2) - \frac{1}{2}u(0) - \frac{1}{2}u(1)$. The proof of Proposition 7 implies that

$$\sigma^* < \bar{\sigma} \equiv 1 - \frac{\phi}{u(1) - u(0)},$$

so $\eta(\phi, \sigma^*(\phi)) > 0$ satisfies (7). Using this fact, totally differentiating (28) yields

$$\frac{d\sigma^*(\phi)}{d\phi} = \frac{-u'(1-\eta)}{(u(1) - u(0))\,(u'(\eta) - u'(1-\eta))} < 0.$$

Figure 1
Strategic Substitutes and Complements



5 Power Imbalance and Aggression

Discussions of conflict often focus on the case where an incumbent "great power" faces a "rising power". For example, in his classic discussion of the war between Sparta and Athens, the Greek historian Thucydides [21] argued that "It was the rise of Athens and the fear that this inspired in Sparta that made war inevitable." Similarly, as China grows more rapidly that the Unites States, some ask whether war is inevitable. In fact, a war between a rising power and a incumbent great power has been called the "Thucydides trap" (Allison [1]). The key question is whether the overall probability of conflict increases as inequality falls, perhaps because one country grows faster than another. We discuss by explicitly introducing asymmetric endowments and studying the probability of conflict as inequality declines.

Suppose the two players are symmetric except that the status quo allocation favors player A. That is, $\omega_A = \frac{1}{2} + \varepsilon$ and $\omega_B = \frac{1}{2} - \varepsilon$, where $0 \leq \varepsilon < 1/2$. Player B's power is "rising" in the sense that equality is decreasing, i.e. $\Delta\varepsilon < 0$. This might happen because goods and services produced by player A are now produced by player B. Or it could be the case that player A and player B's economies are growing at different rates. If player i's endowment grows at rate g_i, players are risk neutral and costs scale by $g_A\omega_A + g_B\omega_B$ (so, for example, the cost of fighting is $\phi\,(g_A\omega_A + g_B\omega_B)$ and a fraction ϕ of total wealth is destroyed by conflict), we can set $\omega_i' = \frac{g_i\omega_i}{g_A\omega_A + g_B\omega_B}$ and $\Delta\varepsilon = \omega_A - \omega_A'$. Hence, we will refer to player A as the "status quo power" and player B as the "rising power" who has attained military but not economic parity. There is a Thucydides trap at ε iff

$$F(x_A)f(x_B)\frac{dx_B}{d\varepsilon} + F(x_B)f(x_A)\frac{dx_A}{d\varepsilon} < 0.$$

Recall that player i's best response to player j's cutoff x_j is the cutoff $x_i = \Gamma_i(x_j)$, where

$$\Gamma_i(x) \equiv F(x)\left(\frac{1}{2}u(0) + \frac{1}{2}u(1) - \phi - u(\eta)\right) + (1 - F(x))\,(u(1-\eta) - u(\omega_i))\,.$$

As player B has a smaller endowment than player A, he is always more aggressive in equilibrium: $x_B > x_A$. Also, in the terminology of Fudenberg and Tirole [9], an increase in the initial endowment ω_i makes player i *soft* (shifts his best-response curve down), because he now has more to lose from a conflict. A decrease in ω_i would instead make him *tough* (shift his best-response curve up), because he now has less to lose. So, the *direct effect* of decreasing inequality is that player B becomes *less* aggressive and player A becomes *more* aggressive. There are also *strategic effects* that depend on whether actions are strategic complements or substitutes and may reinforce or counterbalance the direct effects. For example, if actions are strategic substitutes for player A, the direct effect of reducing inequality on player A is clear: the softening of player B makes player A more aggressive. When actions are strategic complements, the strategic effect of reducing inequality will be a softening on the part of player A to meet the softening on the part of player B so the net effect on player A's aggressiveness is ambiguous. This implies that in general the impact of reducing inequality on player A and hence on the probability of conflict is ambiguous.

But we can say that player B becomes unambiguously *less* aggressive as inequality declines. Player B's incentive to turn dovish comes from the direct effect of retaining a larger endowment when player A is dovish, an event which occurs with probability $1 - F(x_A)$. Player B's incentive to turn hawkish comes from the strategic effect triggered by the incentive of player A to turn hawkish as he has a smaller endowment to lose from a surprise attack on a dovish player B, an event which occurs with probability $1 - F(x_B)$. But as $x_B > x_A$, $1 - F(x_A) > 1 - F(x_B)$ and for player B the direct effect of reduced inequality is greater than the strategic effect so he becomes more dovish. We now formalize these ideas and study if and when the Thucydides trap might arise.

In equilibrium, x_A and x_B will satisfy

$$x_A = \Omega_A F(x_B) + u(1-\eta) - u(\omega_A) \tag{29}$$

and

$$x_B = \Omega_B F(x_A) + u(1-\eta) - u(\omega_B). \tag{30}$$

where

$$\Omega_i \equiv \frac{1}{2}u(0) + \frac{1}{2}u(1) - u(\eta) - u(1-\eta) + u(\omega_i) - \phi. \tag{31}$$

Notice $\Omega_A > \Omega_B$ as $\omega_A > \omega_B$. Totally differentiating (29) and (30) with respect to ε yields

$$dx_A = \Omega_A F'(x_B)dx_B - (1 - F(x_B))\, u'(1/2+\varepsilon)d\varepsilon$$

and

$$dx_B = \Omega_B F'(x_A)dx_A + (1 - F(x_A))\, u'(1/2-\varepsilon)d\varepsilon$$

We solve to obtain

$$\frac{dx_A}{d\varepsilon} = \frac{\Omega_A F'(x_B)\,(1-F(x_A))\,u'(1/2-\varepsilon) - (1-F(x_B))\,u'(1/2+\varepsilon)}{1-\Omega_A\Omega_B F'(x_A)F'(x_B)} \tag{32}$$

and

$$\frac{dx_B}{d\varepsilon} = \frac{(1-F(x_A))\,u'(1/2-\varepsilon) - \Omega_B F'(x_A)\,(1-F(x_B))\,u'(1/2+\varepsilon)}{1-\Omega_A\Omega_B F'(x_A)F'(x_B)}. \tag{33}$$

We have

$$\frac{dx_B}{d\varepsilon} - \frac{dx_A}{d\varepsilon} \tag{34}$$

$$= \frac{(1 - F(x_A))\,[1 - \Omega_A F'(x_B)]\, u'(1/2 - \varepsilon) + (1 - F(x_B))\,[1 - \Omega_B F'(x_A)]\, u'(1/2 + \varepsilon)}{1 - \Omega_A \Omega_B F'(x_A) F'(x_B)} > 0$$

as $|\Omega_A F'(x_A)| < 1$ and $|\Omega_B F'(x_B)| < 1$ by Assumption 2(ii). Because $x_A = x_B$ when $\varepsilon = 0$, (34) implies than $x_B > x_A$ for any $\varepsilon > 0$. That is, the rising power is always the more aggressive player, whether actions are strategic complements or substitutes. Moreover, as $F(x_A) < F(x_B)$, $|\Omega_B F'(x_B)| < 1$ and $u'(1/2 - \varepsilon) \geq u'(1/2 + \varepsilon)$ by concavity (33) is always strictly positive: The rising power becomes less aggressive as it becomes wealthier. *This implies that, if there is a Thucydides trap, it must arise from the increased aggressiveness of player A.*

As the denominator is positive, the sign of (32) is determined by the sign of the numerator, which is certainly negative if $\Omega_A < 0$. Even if $\Omega_A > 0$, (32) is negative if ε is small enough, since the numerator evaluated at $\varepsilon = 0$ is $[\Omega F'(x) - 1]\,(1 - F(x))\, u'(1/2) < 0$ where $x = x_A = x_B$ and $\Omega_A = \Omega_B = \Omega = \frac{1}{2}u(0) + \frac{1}{2}u(1) - u(\eta) - u(1 - \eta) + u\left(\frac{1}{2}\right) - \phi$. Thus, whether actions are strategic substitutes or complements, for small ε, the status quo power does becomes more aggressive as inequality declines. In aggregate though, for small ε, the increased hostility of the status quo power is met exactly by increased accommodation by the rising power so small asymmetries have no effect of the probability of conflict. Hence, if countries at the same level of development grow at slightly different rates, this has little effect on the probability of conflict.

Proposition 8 *A small amount of asymmetry in the status quo does not change the probability of a conflict so there is no Thucydides trap.*

Proof. Setting $\varepsilon = 0$ in (32) and (33) reveals that $dx_A/d\varepsilon = -dx_B/d\varepsilon$. The probability of a conflict is $F(x_A)F(x_B)$, and $dx_A/d\varepsilon = -dx_B/d\varepsilon$ implies that $F(x_A)F(x_B)$ is independent of ε, for a small change in ε evaluated at $\varepsilon = 0$.

■

Even if there is a large asymmetry and the status quo power considers actions to be strategic complements, there is no Thucydides trap because decreasing inequality causes both powers to become less aggressive which would be unambiguously good for peace. Indeed, the sign (32) is positive if and only if

$$\frac{u'(1/2+\varepsilon)}{u'(1/2-\varepsilon)} < \frac{1-F(x_A)}{1-F(x_B)}\Omega_A F'(x_B) \tag{35}$$

The left-hand side is decreasing in ε by concavity of u. If u satisfies the usual boundary condition $u'(x) \to \infty$ when $x \to 0$ then (35) is guaranteed to hold for ε close to 1/2 if actions are strategic complements for player A. Moreover, when ε is close to 1/2,

$$\Omega_A = \sigma(u(1) - u(0)) + u(1) - u(1-\eta) > 0$$

if there is a first-mover advantage ($\sigma > \frac{1}{2}$). Hence, actions are in fact strategic complements for the status quo power when inequality is extreme. We have the following result:

Proposition 9 *When inequality is large, there is a first-mover advantage and $u'(x) \to \infty$ when $x \to 0$, there is no Thucydides trap. In fact, the probability of conflict* declines *with reduced inequality.*

When inequality is large, the status quo power has nothing to gain by being aggressive when the rising power is dovish because its endowment is already large. Hence, the status quo power's incentives to be aggressive must be larger when the rising power is aggressive and actions are strategic complements for the status quo power. When its endowment is small and utility is concave, the rising power's incentives to be dovish increase dramatically with falling inequality as it greatly values any increase in endowment when it is very poor. As actions are strategic complements for the status quo power and the rising power is becoming much more dovish, the status quo power's incentive to be dovish outweighs any incentive to become hawkish because of a decreasing endowment. Therefore, the probability of conflict must fall. So, for example, if inequality between a status quo power and a poor rising power declines with trade, so will the chance of conflict.

So far, we have identified two opposite situations - when there is no inequality or there is large inequality - where conflict does not decrease with inequality. We will now show that in between these extremes there can be a Thucydides trap. Suppose types are uniformly distributed on $[0,1]$ and players are risk neutral. We begin with the case where $\eta = 0$ which requires that $\phi > 1 - \sigma$. Also, for Assumption 2(i) to be satisfied, we must have

$1-\sigma<\phi<\frac{1}{2}$ and $0\leq\varepsilon<\frac{1}{2}$. We have $\Omega_A=\varepsilon-\phi$ and $\Omega_B=-\varepsilon-\phi$ so actions are always strategic substitutes for player B and are substitutes for player A if and only if $\varepsilon<\phi$. When players are risk neutral and types are uniformly distributed, the change in the probability of conflict is given by

$$F(x_B)\frac{dx_A}{d\varepsilon}+F(x_A)\frac{dx_B}{d\varepsilon}. \tag{36}$$

There is more likelihood of a Thucydides trap if actions are strategic substitutes for player A and his hawkishness declines with inequality: $\frac{dx_A}{d\varepsilon}<0$. This is favored by low inequality and high costs of conflict so Ω_A is highly negative. Also when player B is poorer than player A, the probability that player B is hawkish, $F(x_B)$, must be much higher than the probability player A is hawkish, $F(x_A)$. This means that in (36) the fact that player A becomes more aggressive outweighs the fact that player B is becoming less aggressive and a Thucydides trap arises.

When inequality is high, the logic resembles the argument behind Proposition 9. We have the following result (the proof is in the Appendix):

Proposition 10 *Suppose players are risk neutral, types are uniformly distributed on* $[0,1]$ *and* $1-\sigma<\phi<\frac{1}{2}$*. Then, there is a Thucydides trap iff* $\varepsilon>0$ *and*

$$\varepsilon^2\leq\frac{6\phi-\phi^2(3+2\phi))-1}{(5-2\phi)}.$$

Now suppose $\phi<1-\sigma$ so $\eta>0$. For Assumption 2(i) to be satisfied, we must have $\sigma>\frac{1}{2}$ and $\sigma+\phi-\varepsilon>\frac{1}{2}$. We still have $\Omega_A=\varepsilon-\phi$ and $\Omega_B=-\varepsilon-\phi$. So, we have a similar result where low inequality and high costs of conflict favor a Thucydides trap (the proof is in the Appendix):

Proposition 11 *Suppose players are risk neutral, types are uniformly distributed on* $[0,1]$, $\phi<1-\sigma$, $\sigma>\frac{1}{2}$ *and* $\sigma+\phi-\varepsilon>\frac{1}{2}$*. Then, there is a Thucydides trap iff* $\varepsilon>0$ *and*

$$\varepsilon^2\leq\frac{(1-\phi)\,(5-6\sigma+\phi(-3+4\phi+2\sigma))}{(7-4\phi-2\sigma)}. \tag{37}$$

The result reflects the now familiar intuition that a Thucydides trap is more likely when actions are strategic substitutes for player A. This occurs when inequality ε is low, costs of conflict ϕ are high and first-mover advantage is low. For instance, (37) is impossible to satisfy for $\varepsilon > 0$ when σ is high.[16]

To return to our main question, declining inequality perhaps caused by differential growth does increase the chance of conflict if inequality is intermediate and costs of conflict are high. China and the United States have nuclear arms and inequality is now intermediate given China's fast growth rates. These countries might very well be subject to a Thucydides trap. Ironically, we have shown that a trap arises not because China becomes more aggressive with declining inequality but because the United States does so. To avoid this paradox, players might make an effort to alter the magnitude of first-mover advantage or the costs of conflict. This is the topic to which we now turn.

6 Strategic Investments

Player A might invest in "Star Wars" defensive technology to destroy incoming nuclear missiles and thereby reduce ϕ_A. He might publicly announce that his endowment is sacred; losing part of it then becomes more costly as it implies a loss of face.[17] Or he may invest on his portion of the endowment, for example by building settlements on his land. We can think of these moves as changing player A's payoff function directly by affecting the value of his endowment. What are the consequences of such an announcement? To study this kind of question, we suppose the parameters of the game depend on ex ante decisions. Our analysis closely follows Fudenberg and Tirole's [9] analysis of strategic ex ante investments in Industrial Organization theory. For convenience, we also refer to the ex ante decision as an investment.

To be specific, suppose player A makes an ex ante investment which influences his utility function and/or cost of conflict. This is the direct effect of his investment. To avoid having to deal with issues of signaling, we assume the investment is made before player A learns how costly it would be to challenge the status quo. Thus, player A's investment is independent of his type. The investment is observed by player B. Therefore, even if A's investment does not change B's utility function or cost of conflict, it will

[16] If $\sigma > \sigma^*$, where $5 - 6\sigma^* + (1-\sigma^*)(-3 + 4(1-\sigma^*) + 2\sigma^*) = 0$ and $\sigma^* < 1$, the right hand side of (37) is negative and so there cannot be a Thucydides trap.

[17] Alternatively, if player A is a political leader who makes the key decisions in international affairs, he might be replaced by someone who assigns higher value to the endowment.

influence what player B thinks A will choose in the bargaining game (H or D). This strategic effect may change B's behavior in the bargaining game. Fudenberg and Tirole [9] considered whether the strategic effect will lead to *over-investment* or *under-investment*, as compared to the situation where no strategic effect exists. For example, if a monopolist acquires additional production capacity for the sole purpose of deterring entry, this would be classified as over-investment caused by a strategic effect.

Of course, player A's investment might also have a direct effect on player B. For example, it might increase player B's cost of conflict. In our bargaining game there is also a more subtle effect: if player A's investment makes him less willing to concede, it will reduce the amount $1-\eta_A$ that player B can extract by challenging. This has no direct analogy in the Fudenberg and Tirole [9] model. Thus, the consequences of ex ante strategic moves are somewhat more complicated here than in traditional Industrial Organization models.

Recall player A's investment makes player $i \in \{A, B\}$ tough if the best response curve $\Gamma_i(x)$ defined by (15) shifts up, making player i more likely to choose Hawk in the bargaining game. The investment makes player $i \in \{A, B\}$ soft if the best response curve $\Gamma_i(x)$ defined by (15) shifts down, making player i more likely to choose Dove in the bargaining game. Fudenberg and Tirole [9] showed that over or under investment will occur depending on whether player A's investment makes A soft or tough, and whether actions are strategic substitutes or complements. Here we must also account for whether player A's investment makes player B soft or tough.

Consider first the scenario where player A invests in a defensive technology that reduces his cost of conflict ϕ_A. We must distinguish several cases.

Case 1: $\phi_i > u_i(1/2) - \frac{1}{2}u_i(0) - \frac{1}{2}u_i(1)$ for $i \in \{A, B\}$. In this case, actions are strategic substitutes. There are two sub-cases; both lead to the conclusion that A will over-invest (but for different reasons).

The first sub-case occurs when $\phi_A > (1-\sigma)(u_A(1) - u_A(0))$ so $\eta_A = 0$ and $\frac{d\eta_A}{d\phi_A} = 0$. From (15), player A's investment causes his best response curve Γ_A to shift up. Intuitively, the fall in player A's cost of conflict makes it less costly for him to choose H, so player A becomes tough. It has no effect on Γ_B. As actions are strategic substitutes, player A will over-invest to persuade player B to choose D, corresponding to Fudenberg and Tirole's [9] "top dog" strategy. This is the traditional interpretation of Schelling's commitment tactic – become tough in order to deter aggression.

The second sub-case occurs when $\phi_A < (1-\sigma)(u_A(1) - u_A(0))$, so $\eta_A > 0$ and, from (6),

$$\frac{d\eta_A}{d\phi_A} = -\frac{1}{u'_A(\eta_A)} < 0. \tag{38}$$

Now there are two effects on Γ_A. On the one hand, the fall in ϕ_A makes it less costly to choose H, as before. But on the other hand, (38) implies that player A will concede less when challenged, so choosing D becomes less costly as well. Equation (15) reveals that those two effects cancel out, so Γ_A is unaffected by the change in ϕ_A. Thus, player A neither becomes soft nor tough. However, player B's best response curve shifts down when ϕ_A falls, because $1 - \eta_A$ falls (as player A will concede less). Therefore, *player A's investment makes player B soft.* This is always beneficial for A. Hence, player A will again over-invest when $\phi_i > u_i(1/2) - \frac{1}{2}u_i(0) - \frac{1}{2}u_i(1)$. Player A achieves deterrence by overinvesting not because he becomes tougher but because player B becomes softer. This is the kind of external effect of investment that arises naturally in our model but is not studied in industrial organization.

Case 2: $\phi_i < u_i(1/2) - \frac{1}{2}u_i(0) - \frac{1}{2}u_i(1)$ for $i \in \{A, B\}$. There are two sub-cases to consider, and they lead to different conclusions.

The first sub-case occurs when $\phi_i > (1-\sigma)(u_i(1) - u_i(0))$ for $i \in \{A, B\}$, so actions are strategic complements. Also, as $\eta_A = 0$ in this case, the investment makes player A tough. By strategic complements, player B becomes more likely to choose H, which is disadvantageous for player A. Therefore, player A will under-invest, to make B feel less threatened and more likely to choose D. This corresponds to Fudenberg and Tirole's [9] "puppy dog" strategy.

The second sub-case occurs when $\phi_i < (1-\sigma)(u_i(1) - u_i(0))$ for $i \in \{A, B\}$, in which case actions may be complements or substitutes. As $\eta_A > 0$ in this case, as discussed above player A's investment has no effect of Γ_A but makes player B soft. Player A necessarily benefits from this, and hence will over-invest. This is independent of whether actions are strategic substitutes or complements, since in both scenarios player A wants the opponent to choose D. Hence, again, we have the same strategic effect operating via the impact of player A's investment on player B.

We summarize the discussion so far:

Proposition 12 *Suppose investment by player A reduces ϕ_A.*

(i) *(High cost of conflict.) If*

$$\phi_i > u_i(1/2) - \frac{1}{2}u_i(0) - \frac{1}{2}u_i(1)$$

then player A over-invests.

(ii) (Intermediate cost of conflict.) If

$$(1-\sigma)\left(u_i(1) - u_i(0)\right) < \phi_i < u_i(1/2) - \frac{1}{2}u_i(0) - \frac{1}{2}u_i(1)$$

then player A under-invests.

(iii) (Low cost of conflict.) If

$$\phi_i < \min\{u_i(1/2) - \frac{1}{2}u_i(0) - \frac{1}{2}u_i(1), (1-\sigma)\left(u_i(1) - u_i(0)\right)\}$$

then player A over-invests.

Consider now briefly the case where player A's investment increases player B's cost of conflict ϕ_B. This has no effect on Γ_B if $\eta_B > 0$. It does, however, shift Γ_A up, because B will concede more.[18] That is, player A's investment makes player A tough. Hence, the optimal strategy is either under or over-investment depending on whether actions are strategic complements or substitutes. If $\eta_B = 0$, however, then an increase in ϕ_B makes player B soft, so A will overinvest.

A more complex situation occurs when player A's investment changes player A's valuation of the resource. To make this tractable, suppose utility functions are piecewise linear. Suppose the status quo is $(\omega_A, \omega_B) = (1/2, 1/2)$ and the value of the status quo endowment is normalized to zero. Moreover, each unit of player i's own endowment that he loses reduces his payoff by v_i, while each unit of player j's endowment that player i acquires increases player i's payoff by $w_i < v_i$. Then,

$$u_i(x_i) = \begin{cases} v_i\left(x_i - 1/2\right) & \text{if} \quad x_i - 1/2 \leq 0 \\ w_i(x_i - 1/2) & \text{if} \quad x_i - 1/2 \geq 0 \end{cases}$$

[18]Making it easier for the opponent to concede is an old tactic. For example, Sun Tzu [22], Chapter 7.36: "When you surround an army, leave an outlet free. Do not press a desperate foe too hard."

and

$$\eta_i = \begin{cases} (1-\sigma)\frac{v_i+w_i}{2v_i} - \frac{\phi_i}{v_i} & \text{if} \quad \phi_i < (1-\sigma)\frac{v_i+w_i}{2} \\ 0 & \text{if} \quad \phi_i \geq (1-\sigma)\frac{v_i+w_i}{2} \end{cases} \tag{39}$$

This example does not satisfy strict concavity. However, strict concavity is only used to guarantee that a unit of a player's own endowment is more valuable to him than a unit of the opponent's endowment, and this is satisfied by this example as $w_i < v_i$.[19] Suppose player A's ex ante investment increases v_A (without affecting player B's utility function). The easiest case is where the cost of conflict is high, $\phi_A > (1-\sigma)\frac{v_A+w_A}{2}$, so the first-mover has the maximum advantage: $\eta_A = 0$. In this case, the investment makes player A tough (and has no effect on player B's best response function). Intuitively, player A becomes tough because he values his endowment a lot and wants to avoid giving the first-mover advantage to player B. Formally, player A's best response function is

$$\Gamma_A(x) \equiv F(x)\left(\frac{1}{2}u_A(0) + \frac{1}{2}u_A(1) - \phi_A - u_A(0)\right) + (1-F(x))\left(u_A(1) - u_A(1/2)\right).$$

Since $u_A(0) = -v_A/2$ this curve shifts up when v_A increases. Since the investment makes player A tough, he will under-invest if actions are strategic complements, but over-invest if actions are strategic substitutes. From Proposition 6, we know that actions tend to be strategic substitutes when conflicts are very costly, so player A will over-invest. Intuitively, when the cost of conflict is large, the first-mover advantage is very valuable, but both players also want to avoid a simultaneous challenge. In this case, player A gets a strategic advantage from valuing his own endowment highly, since it makes him tougher and player B more cautious.

If the cost of conflict is low, $\phi_A < (1-\sigma)\frac{v_A+w_A}{2}$, then the situation is more complex. We have

$$\eta_A = (1-\sigma)\frac{v_A+w_A}{2v_A} - \frac{\phi_A}{v_A} > 0$$

and

$$\frac{\partial \eta_A}{\partial v_A} = -\frac{\frac{(1-\sigma)}{2}w_A - \phi_A}{v_A^2},$$

[19] That is, the results of the paper go through if Assumption 1 is replaced by the weaker assumption: if $0 < x < \omega_i < y < 1$ then $u_i'(x) > u_i'(y)$.

which is negative if ϕ_A is small enough. Thus, player A will concede more, the more he values his endowment. The reason is that if he does not concede, there will be a conflict, and he risks losing everything. Player A's best response function is affected via the term

$$\begin{aligned}\frac{1}{2}u_A(0) - u_A(\eta_A) &= \frac{1}{2}u_A(0) - (\sigma u_A(0) + (1-\sigma)u_A(1) - \phi_A) \\ &= -\left(\frac{1}{2} - \sigma\right)\frac{v_A}{2} - (1-\sigma)\frac{w_A}{2} + \phi_A\end{aligned}$$

The derivative with respect to v_A is positive as long as $\sigma \geq 1/2$. Thus, an increase in v_A makes player A tough. Intuitively, if σ is large, so there is a large first-mover advantage, the situation is as in the previous paragraph: player A becomes tougher as he becomes more averse to losing his endowment. However, the reduction in η_A also shifts up Γ_B, so B becomes tougher. From the previous section, we know that σ large tends to make the actions strategic complements. In this case, as v_A increases player B becomes more likely to choose H. Intuitively, when σ is large but conflicts are not very costly, the competition over the territory becomes more intense as the territory becomes more valuable. Since this hurts player A, he will under-invest to become soft, the "puppy dog" strategy. If instead σ is small, actions are strategic substitutes. As both sides become tougher, the impact on player A's welfare is ambiguous.

We have studied just a few policies and their impact on conflict. Many others might be analyzed using this framework.

7 Conclusion

In this paper, we identify the size of first-mover advantage and the cost of conflict as key parameters determining the strategic nature of conflict. The formalism has helped to identify the impact of the cost of conflict and rising power on the probability of war. It has also helped to determine when actions might be strategic complements or substitutes in a conflict game. Much remains to be done. For example, key players often enter a conflict with expectations that are falsified - some analysts thought the Iraq war was going to be short and Germany entered World War I expecting it to be over rapidly (see Jervis [12] on the role of misperception in international relations). One way to capture this kind of phenomenon is to add an additional layer of uncertainty about the technology of war and to study dynamics. Also, there is a historical record of the incidence of war (see Iyigun, Nunn and Qian [17] for instance) and of the technology of war (see McNeil [14] for instance). Our model and results provide a lens through which to examine this record.

References

[1] Graham Allison (2015): "The Thucydides Trap: Are the U.S. and China Headed for War?," *The Atlantic*, Sept. 24, 2015.

[2] Sandeep Baliga and Tomas Sjöström (2012): "The Strategy of Manipulating Conflict", *American Economic Review* 102: 2897–2922

[3] Bruce Bueno de Mesquita (2013): *Principles of International Politics*, CQ Press.

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

[5] Christopher Duffy (1997): *Siege Warfare: The Fortress in the Early Modern World 1494-1660*, Routledge

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

[7] James Fearon (1995): "Rationalist Explanations for War," *International Organization*4, Vol. 49, pp 379-414.

[8] James Fearon (1997): "The Offense-Defense Balance and War Since 1648," mimeo, Stanford University.

[9] Drew Fudenberg and Jean Tirole (1984): "The Fat-Cat Effect, the Puppy-Dog Ploy, and the Lean and Hungry Look," *American Economic Review*, 74 (2) :361-366.

[10] Basil Henry Liddell Hart (1960): *Deterrent or Defence*, Stevens and Sons, London.

[11] Robert Jervis (1978): "Cooperation Under the Security Dilemma," *World Politics*, Vol. 30, No. 2., pp. 167-214.

[12] Robert Jervis (1976): *Perception and Misperception in International Politics*. Princeton University Press: Princeton.

[13] Jack Levy (1983): *War in the Modern Great Power System, 1495-1975.* Lexington KY: University of Kentucky Press.

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

[15] Adam Meirowitz, Massimo Morelli, Kristopher W. Ramsay and Francesco Squintani (2016): "Dispute Resolution Institutions and Strategic Militarization," mimeo, Bocconi University.

[16] John Nash (1953): "Two-Person Cooperative Games," *Econometrica*, Vol. 21, pp 128-140.

[17] Murat Iyigun, Nathan Nunn and Nancy Qian (2015): "The Long-run Effects of Agricultural Productivity on Conflict: Evidence from Potatoes, 1400-1900", mimeo.

[18] Joseph Nye (2007): *Understanding International Conflict (6th Edition). Longman Classics in Political Science.* Longman: New York City.

[19] Thomas Schelling (1960): *The Strategy of Conflict.* Cambridge: Harvard University Press.

[20] Thomas Schelling (1966). *Arms and influence.* New Haven: Yale University Press.

[21] Thucydides (1972): *History of the Peloponnesian War.* London: Penguin Classics

[22] Sun Tzu : *The Art of War.* New York: Simon and

[23] Jean Tirole (2016): "Cognitive Games and Cognitive Traps," mimeo, Toulouse School of Economics.

8 Appendix

We record proofs that are not in the main text in the Appendix.

Proof of Proposition 10

We can show

$$x_A = \frac{1 + (2\varepsilon - 1)\,\varepsilon - \phi\,(1 + 2\varepsilon)}{2(1 + \varepsilon^2 - \phi^2)}$$

and

$$x_B = \frac{1 + (2\varepsilon + 1)\,\varepsilon - \phi\,(1 - 2\varepsilon)}{2(1 + \varepsilon^2 - \phi^2)}.$$

Then, the change in the probability of conflict with inequality ε is given by

$$\varepsilon\,(1 + 2\phi)\,\frac{(1 + \varepsilon^2(5 - 2\phi) - 6\phi + \phi^2(3 + 2\phi))}{2(1 + \varepsilon^2 - \phi^2)^3}.$$

Hence, there is a Thucydides trap if and only if there is non-negligible inequality (i.e. $\varepsilon > 0$) and

$$\varepsilon^2 \leq \frac{6\phi - \phi^2(3 + 2\phi)) - 1}{(5 - 2\phi)}.$$

The right hand side is increasing in ϕ. Thus, low inequality and high costs of conflict favor the Thucydides trap. For example, when $\phi = 0.4$, the probability of conflict changes with ε as depicted below:

Proof of Proposition 11

We can show:

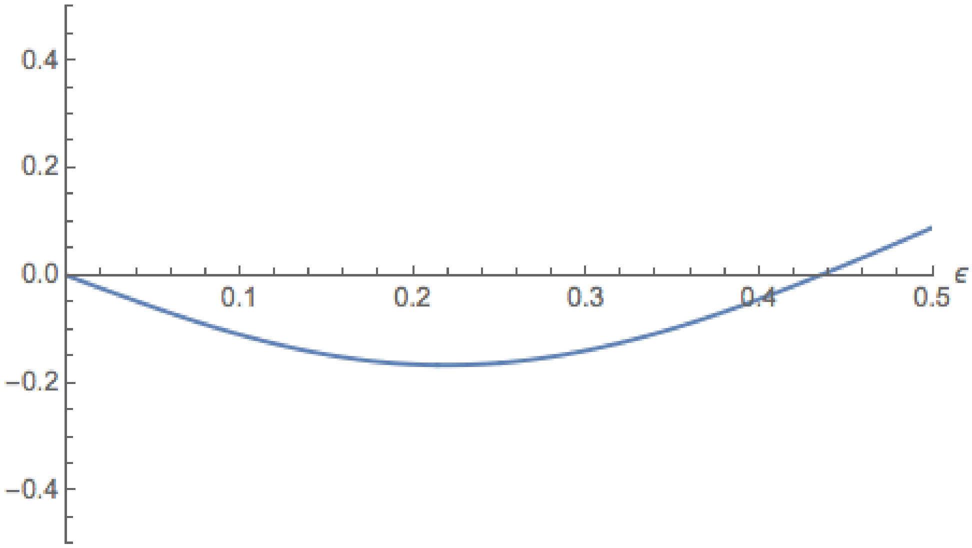
$$\begin{aligned} x_A &= \frac{2\varepsilon^2 + \varepsilon\,(2\sigma - 3) + (1 - \phi)\,(2\phi + 2\sigma - 1)}{2(1 + \varepsilon^2 - \phi^2)}, \text{ and} \\ x_B &= \frac{-1 + \varepsilon\,(3 + 2\varepsilon) + (3 - 2\phi)\,\phi + 2\sigma - 2(\varepsilon + \phi)\sigma}{2(1 + \varepsilon^2 - \phi^2)}. \end{aligned}$$

Then, the change in the probability of conflict with inequality ε is given by

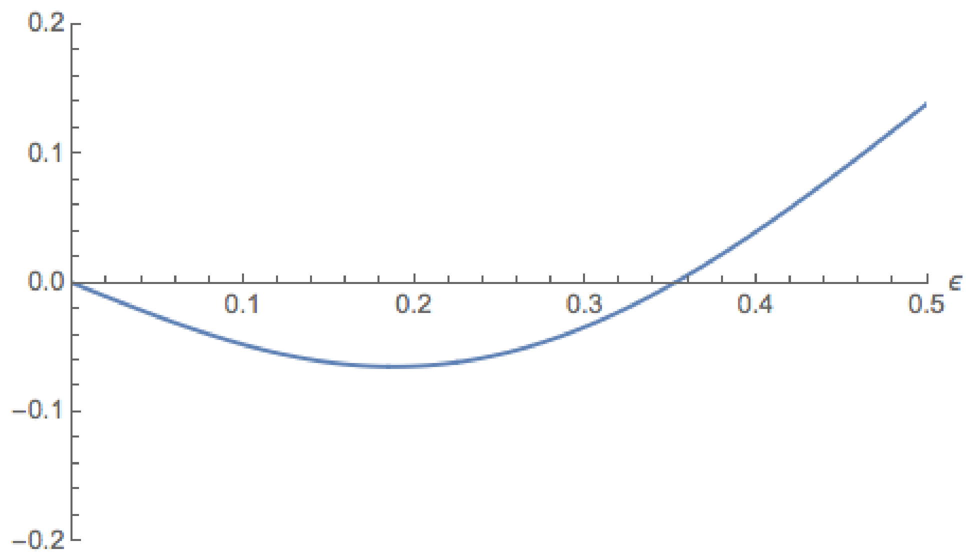
$$\varepsilon\frac{(2\sigma - 3)\,[5 + 2\sigma\,(\varepsilon^2 - (-3 + \phi)\,(-1 + \phi)) - 8\phi + (\varepsilon - \phi)\,(\varepsilon + \phi)\,(-7 + 4\phi)]}{2(1 + \varepsilon^2 - \phi^2)^3}.$$

Hence, as $\sigma < 1,\;\; \phi < \frac{1}{2}$ and $0 < \varepsilon < \frac{1}{2}$, there is a Thucydides trap if there is non-negligible inequality and

$$\varepsilon^2 \leq \frac{(1 - \phi)(5 - 6\sigma + \phi(-3 + 4\phi + 2\sigma)}{(7 - 4\phi - 2\sigma)}. \tag{40}$$



For example, when $\sigma = 0.7$ and $\phi = 0.2$, the probability of conflict changes with ε as depicted below:


0.2
0.1
0.0
−0.1
−0.2
0.1
0.2
0.3
0.4
0.5
ϵ