The Strategy and Technology of Conflict

Sandeep Baliga, Northwestern University
Tomas Sjöström, Rutgers University

April 27, 2017

Abstract

Using a simple conflict bargaining game, we study how the strategic interaction is shaped by underlying preferences over a contested resource and by the technology of conflict. With risk-averse players, the game has strategic complements if the cost of conflict is small and there is a large first-mover advantage, and strategic substitutes otherwise. This characterization generates predictions about optimal bargaining tactics, the consequences of power imbalances, and the strategic advantages of *ex ante* tactics. It provides a theoretical foundation for important ideas of Thomas Schelling, as well as for more recent formal models inspired by his work.

1 Introduction

Russia's recent annexation of Crimea, and China's island-building in the South China Sea, caught the world off guard. Each situation presented the United States with a stark choice between concession or confrontation. Russia and China's "strategic moves" created first-mover advantages as envisioned by Schelling [21]. As Schelling explained, such 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

[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.

Hungary, a military confrontation would have been quite likely – the Soviets could not have retreated from these countries without a massive loss of reputation. 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 [21], 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 [21]). An East-West confrontation was avoided because the Soviets conceded Western Europe just as the West had conceded Eastern Europe.

Our simple two-player bargaining game aims to capture Schelling's [20] definition of a strategic move as a "voluntary but irreversible sacrifice of freedom of choice". In the model, a strategic move is a challenge to the status quo. A conflict occurs if both players challenge (since this implies mutually incompatible commitments), or if only one player challenges but the other refuses to concede. In the latter case, the challenger was the one who took the initiative, for example by placing his soldiers on the contested territory, which may give him a first-mover advantage. The optimal challenge is the largest demand the opponent would concede to. The bargaining game can then be represented by a two-by-two matrix game, with actions Hawk (the optimal challenge) and Dove (no challenge).

The model helps us understand when conflicts will be more likely. Consider the role of the cost of conflict. During the Cold War, the high cost of a nuclear war was thought to prevent the superpowers from challenging the status quo. On the other hand, some challenges did occur. For example, 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.[2] Similarly, 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. In our model, when the cost of conflict increases from an initially small level, each player is prepared to make larger con-

[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 [21], p. 46) – implying that it would be the West's decision to risk a war by entering East Germany which the Soviets already occupied.

cessions to avoid a conflict, which makes challenging the status quo more tempting. This makes challenges, and hence conflicts, more likely. However, when the cost of conflict is very large, further cost increases will makes challenges less likely. Therefore, the relationship between the cost of conflict and the equilibrium probability of conflict is non-monotonic. This may shed some light on the "stability-instability" paradox discussed by Hart [14].

A key distinction is between games of strategic complements and games of strategic substitutes. With strategic complements, conflicts are caused by an escalated lack of trust, as in Schelling's [20] analysis of the reciprocal fear of surprise attack. The generic game of this kind is the Stag Hunt. With strategic substitutes, conflicts are instead caused by a failure of deterrence, the generic example being Chicken. Scholars such as Jervis [15] have used such simple games as metaphors for conflict. Baliga and Sjöström ([2] and [4]) and Chassang and Padro-i-Miquel ([7] and [8]) studied the equilibria of such games when the players have incomplete information. The ability of strategic actors to manipulate the outcome turns out to depend critically on whether actions are strategic complements or substitutes. Baliga and Sjöström [2] found that communication between the two disputing parties may restore trust and prevent conflict in games of strategic complements. But Baliga and Sjöström [4] found that hawkish third parties ("terrorists") can trigger conflicts in such games by sending messages that create "fear spirals". With strategic substitutes, hawkish extremists cannot do this – instead dovish third parties ("pacifists") can prevent conflicts by persuading one party to back down. Clearly, the impact of policy choices will differ, depending on the nature of the underlying conflict. In this paper, we emphasize the importance of risk-aversion, the cost of conflict, and the first-mover advantage.[3]

If the cost of conflict is high, the game has strategic substitutes. This may explain why Khrushchev considered the Cold War to be a sequence of Chicken-type interactions (see footnote 2). If the cost of conflict is low, the analysis becomes more subtle. We show that the game has strategic complements if there is a significant first-mover advantage. This result may be helpful in understanding the important concepts of "offense dominance" and "defense dominance" in international relations theory (see Bueno de Mesquita [6]). The result is not obvious, because there are two opposing

[3]We can try to map the technology of conflict maps 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.

effects: when the first-mover advantage increases, the cost of choosing Dove when the opponent chooses Hawk increases, but so does the benefit from choosing Hawk when the opponent chooses Dove. The first effect tends to generate strategic complements, while the second effect does the opposite. The first effect dominates when the marginal utility of land is decreasing, so the cost of losing territory exceeds the benefit of acquiring the same amount. Therefore, in the expected-utility calculations, losses suffered by Doves who encounter Hawks outweigh gains enjoyed by Hawks who encounter Doves.

A situation where a "status quo power" faces a "rising power" may create a "Thucydides trap" where the probability of conflict increases.[4] This resonates with contemporary commentary concerned with the rise of China and the decline of Russia. Are such situations especially dangerous? What happens when the status quo does not reflect relative military strengths? We extend our model to allow for unequal endowments, but maintain the assumption of equal military strength. The inequality necessarily makes the "rising power" (with the smaller endowment) more hawkish, because he stands to lose less and gain more from a conflict. Whether this hawkishness is reciprocated depends on the amount of inequality and on whether the game has strategic substitutes or complements. If inequality is large to begin with, then actions are necessarily strategic complements for the status quo power. Reducing the inequality (by reallocating the endowments) will then make both sides less aggressive, unambiguously reducing the risk of conflict. In contrast, at low amounts of inequality, reallocating the endowment has little effect on the risk of conflict, because an increased hawkishness on one side is compensated for by reduced hawkishness on the other side. In between these extremes, no general result is possible. In fact, the probability of conflict can go *up* with declining inequality as suggested by the idea of the Thucydides trap. But this happens not because of the rising power as is commonly thought but because of the status quo power becoming more aggressive.

Finally we discuss the incentives to make strategic moves *ex ante* (before the bargaining). If the game has strategic substitutes, then having a low cost of conflict or placing a high value on the territory he controls, for example

[4] See Allison [1]. Thucydides famously described how power imbalances and strategic complementarities can lead to conflicts: "It was the rise of Athens and the fear that this inspired in Sparta that made war inevitable" (Thucydides [22]).

by building settlements on it, is clearly strategically advantageous.[5] Thus, a player would benefit by, for example, delegating decision-making to an agent who is less conflict-averse or puts more value on their territory. This is Schelling's commitment tactic as traditionally interpreted: become tough in order to make the opponent less so. If the game has strategic complements, one may think that the reverse strategy of looking weak should be optimal. However, a new effect is present in our model. A player who reduces his cost of conflict is willing to concede less. This has an external effect on his opponent who becomes less aggressive as he has less to gain. This is good whether actions are strategic complements or substitutes, so the model generates a strong incentive to reduce the cost of conflict. An ex ante move of more dubious strategic value is for a player to raise his valuation of the territory he controls. Since a conflict may lead to the loss of *all* the contested territory, building the settlements could make the player *more* willing to give up *some part* of it in return for peace. This weakness encourages the opponent to challenge the status quo as he has more to gain. This turns the settlements into a strategic disadvantage, regardless of whether the game has strategic complements or strategic substitutes.

The literature on commitment in bargaining traces its origins to Nash's [18] demand game and Schelling's [20] seminal discussion. Most closely related to our paper is Ellingsen and Miettinen [12], who in turn build on Crawford [9]. Crawford's model had multiple equilibria. Ellingsen and Mietinen showed that if commitments are costly then the number of equilibria is reduced. Specifically, they showed that if the cost of making a commitment is small enough, and a commitment attempt only succeeds with probability $q < 1$, then there is a unique equilibrium where both parties attempt to make the maximum commitment (demand all the territory) with probability 1. When $q = 1$, their model also has asymmetric equilibria where one player gets all the territory with probability one. In our model there is no such discontinuity – uniqueness is a consequence of private information, it is not dependent on an exogenous probability that commitments fail. Also, unlike them we do not assume an uncommitted player would always concede

[5] We are evaluating here only the *strategic effect* on the opponent's behavior. Reducing one's cost of conflict is strategically advantageous because the opponent becomes more likely to choose Dove. Fudenberg and Tirole [13] pioneered this kind of strategic analysis. Tirole [23] has recently exapanded the scope of this theory to include two-stage games with ex ante information acquisition.

to any commitment. Some parameter values do generate "corner solutions", where demanding the whole territory is an optimal challenge, and the opponent concedes to this.[6] In reality it sometimes happens that countries do take such extreme positions – China claims essentially *all* of the disputed South China Sea. However, a "reasonable compromise" may be proposed if a party fears that a more extreme proposal will be rejected. Meirowitz, Morelli, Ramsay and Squintani [17] also employ the Nash demand game 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. Finally, Fearon [11] has studied a model of *one-sided* commitment via the ultimatum game.

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 bargaining 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. To make no claim incurs no cost. We interpret player i's challenge as a non-revokable instruction to player i's military to cross the status quo demarcation.

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 player (having observed the claim) 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 to player i's claim then player i gets x_i and player j gets $1 - x_i$. If player j does not concede, there

[6] This can be contrasted with standard bargaining models, such as Binmore, Rubinstein and Wolinsky [5], where the parties always make "reasonable" compromise proposals.

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 wins, and takes 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 to player i's claim. If there is a concession then player i gets what he claimed, 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. If $\sigma > 1/2$ then there is a *first-mover advantage*, in the sense that the first-mover is more likely to win a conflict. Finally, if *both* players challenge the status quo, a conflict occurs because they have made mutually incompatible commitments.[7]

To simplify the exposition, we assume the two players are *ex ante* symmetric in terms of fighting strength: each player would therefore have the same first-mover advantage as a challenger under Rule 2, and 50% chance of winning under Rule 3.

Suppose 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 expected payoff

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

Thus, player i prefers to concede if

$$u_i(1-x_j) \geq \sigma u_i(0) + (1-\sigma)u_i(1) - \phi_i. \tag{2}$$

[7] A more general formulation would be that if both decide to challenge, there is some probability $\alpha > 0$ that player $i \in \{A, B\}$ can commit first, in which case player j must decide 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 [9] and Ellingsen and Mietinen [12], we could assume that a challenge only leads to a successful commitment with probability $q < 1$. But adding the generality does not seem to add any additional insights; we therefore focus on the case $\alpha = 0$ and $q = 1$ for the sake of exposition. (Unlike in Ellingsen and Mietinen [12], nothing dramatic happens to the set of equilibria at the point $q = 1$.)

This is satisfied for $x_j = 1$ if

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

If (3) holds then player i would rather concede the whole territory than have a conflict. If (3) is violated, i.e., if

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

then the maximum claim x_j player i will concede to satisfies (2) with equality, or

$$1 - x_j = u_i^{-1}\left[\sigma u_i(0) + (1-\sigma)u_i(1) - \phi_i\right]. \tag{5}$$

Thus, in general, the maximum claim player i would concede to in stage 2 is the claim $x_j = 1 - \eta_i$, where

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

Notice that if (4) holds then η_i is defined implicitly by

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

and satisfies $\eta_i > 0$. Equation (7) says that player i is indifferent between the share η_i and a conflict when he is the second-mover. Notice that η_i is decreasing in ϕ_i. The more costly a conflict would be, the more territory player i is willing to concede.

To make the problem interesting, we will assume:

Assumption 1 $\eta_i < \omega_i$ for $i \in \{A, B\}$.

Assumption 1 implies that if the first-mover's claim is sufficiently "modest", i.e., close to the status quo, then the second-mover prefers to concede. Assumption 1 rules out the less interesting case where the second-mover would never concede (even to an arbitrarily small change in the status quo). Assumption 1 is equivalent to the inequality

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

The left-hand side of (8) is player i's payoff from the status quo, and the right hand side is his expected payoff from Rule 2 when he is the second-mover and does not concede. Assumption 1 can also be re-written as $\sigma > \underline{\sigma}_i$ where

$$\underline{\sigma}_i \equiv \frac{u_i(1) - u_i(\omega_i) - \phi_i}{u_i(1) - u_i(0)}. \tag{9}$$

We then see that if the cost of conflict is high enough, specifically if $\phi_i > u_i(1) - u_i(\omega_i)$, then $\underline{\sigma}_i < 0$ so Assumption 1 is automatically satisfied. 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 $\sigma \geq 1/2$, i.e., as long as there is no *second*-mover advantage.

We assume all parameters of the game are commonly known, with one exception: neither player knows the opponent's cost of violating the status quo. 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.[8]

If either the support of F is very small, or the density of F is highly concentrated around one point, then the players are fairly certain about each others' types and private information is unimportant. To rule this out, 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)

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

and

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

[8] The cost c_i may partly be due to physical resources and manpower needed to cross the status quo demarcation. But it could also include the disutility of being condemned by the international community, leading to a loss of reputation and goodwill, possible sanctions or embargoes, etc. It may be difficult for outsiders to assess these costs: how much do Russian leaders suffer from sanctions and international condemnation after Russia's invasion of Ukraine? The Russian economy is suffering but President Putin's popular support is sky high.

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 $c \in [\underline{c}, \overline{c}]$ 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

We will first explain why player i's optimal challenge is to make the maximal claim player j would concede to: $x_i = 1 - \eta_j$. The exact size of player i's claim only matters if he challenges and player j doesn't. Now consider what happens if stage 2 is reached following a challenge by player i. Sequential rationality implies that player j concedes if and only if player i's claim satisfies $x_i \leq 1 - \eta_j$.[9] So player i should certainly not claim strictly less than $1 - \eta_j$. 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 the players are risk-averse and $\phi_i > 0$), 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

[9] 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.

$1-\eta_j$ is strictly better than claiming $x_i > 1-\eta_j$. Thus, player i's optimal challenge is to claim $x_i = 1-\eta_j$.

Eliminating all non-optimal challenges from consideration, we conclude that a conflict occurs only when *both* players challenge the status quo (because if only one player challenges, the other will concede). For convenience, we will label the optimal challenge *Hawk* (or H). To not make any challenge is to choose *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)

Claim 1 *If either $\sigma \geq 1/2$ or $\omega_j \leq 1/2$ (or both) then player i is better off when player j chooses Dove, whatever action player i himself chooses.*[10]

Proof. First note that $\omega_i > \eta_i$ implies $u_i(\omega_i) > u_i(\eta_i)$. Second, we showed above that $u_i(1-\eta_j) > \sigma u_i(1) + (1-\sigma)\, u_i(0) - \phi_i$ so if $\sigma \geq 1/2$ then

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

If $\omega_j \leq 1/2$ then the inequality (13) follows from $1-\eta_j > 1-\omega_j \geq 1/2$ and concavity of u_i. ■

Player i is a *dominant strategy hawk* if Hawk (H) is his dominant strategy.[11] Player i is a *dominant strategy dove* if Dove (D) is his dominant strategy.[12] 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

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

[10] Note that $\sigma \geq 1/2$ rules out *second*-mover advantages.

[11] 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.

[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.

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 $g_i : [\underline{c}, \overline{c}] \rightarrow \{H, D\}$ 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 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, Assumption 2(i) implies that

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

and

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

so the equilibrium cutoff point will lie strictly between $\underline{c}$ and $\overline{c}$.

Since the function $(\Gamma_A(x_B), (\Gamma_B(x_A)) : [\underline{c}, \overline{c}]^2 \rightarrow (\underline{c}, \overline{c})^2$ is continuous, a fixed-point $(\hat{x}_A, \hat{x}_B) \in (\underline{c}, \overline{c})^2$ 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.

If the two players are symmetric ex ante (before they draw their types), in both preferences and endowments, then we can drop the subscripts on u_i, ϕ_i, ω_i and η_i. In this case, the unique BNE must be symmetric, and the equilibrium cutoff $\hat{x}$ is the same for both players and implicitly defined by the equation

$$\hat{x} - \Omega F(\hat{x}) = u(1-\eta) - u(1/2). \tag{17}$$

where

$$\Omega \equiv \frac{1}{2}u(0) + \frac{1}{2}u(1) - u(\eta) - u(1-\eta) + u\left(1/2\right) - \phi \tag{18}$$

Consider how η depends on ϕ and σ. If $\phi > (1-\sigma)\left(u(1) - u(0)\right)$ then $\eta = 0$ and $d\eta/d\sigma = d\eta/d\phi = 0$. But if $\phi < (1-\sigma)\left(u(1) - u(0)\right)$ then (7) holds, and the second-mover concedes more if σ or ϕ increases:

$$\frac{d\eta}{d\sigma} = -\frac{u(1) - u(0)}{u'(\eta)} < 0 \tag{19}$$

and

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

Now consider how an increase in σ, i.e., an increased first-mover advantage, affects equilibrium behavior. 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 2 *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 (17) we obtain

$$(1-\Omega f(\hat{x}))\frac{d\hat{x}}{d\sigma} = -\left[u'(\eta)F(\hat{x}) + u'(1-\eta)(1-F(\hat{x}))\right]\frac{d\eta}{d\sigma} \tag{21}$$

where $1-\Omega f(\hat{x}) > 0$ from Assumption 2. From (6), the expression in (21) 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 (19) holds. From (21), 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 3 ***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. Suppose $\phi < (1-\sigma)(u(1)-u(0))$. Totally differentiate (17) with respect to ϕ and use (20) to obtain

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

Thus, when ϕ increases $\hat{x}$ increases, making conflicts more likely.

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

Strategic complements (substitutes) means that the incentive to choose Hawk is greater (smaller), the more likely it is that the opponent chooses Hawk. If player j chooses Hawk, then if player i switches from Dove to Hawk player i's net gain, from the payoff matrix (12), is

$$\frac{1}{2}u_i(0) + \frac{1}{2}u_i(1) - \phi_i - c_i - u_i(\eta_i). \tag{22}$$

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

$$u_i(1-\eta_j) - c_i - u(\omega_i). \tag{23}$$

Actions are *strategic complements for player* i if (22) is greater than (23), 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 has strategic substitutes (resp. complements) if the actions are strategic substitutes (resp. complements) for both players.

We begin by showing that actions are strategic substitutes if there is no first-mover advantage.

Proposition 4 *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

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

■

To simplify the exposition, for the remainder of this section we will assume the players are ex ante symmetric in the sense that they have the same utility function, $u_A = u_B = u$, the same cost of a conflict, $\phi_A = \phi_B = \phi$, and the same initial endowment, $\omega_A = \omega_B = 1/2$. This implies $\eta_A = \eta_B = \eta$, $\Omega_A = \Omega_B = \Omega$ (as defined by (18)) and

$$\underline{\sigma}_A = \underline{\sigma}_B = \underline{\sigma} \equiv \frac{u(1) - u(1/2) - \phi}{u(1) - u(0)} \tag{24}$$

from (9). As was shown above, $\underline{\sigma} < 1/2$. The game has strategic substitutes if $\Omega < 0$ and strategic complements if $\Omega > 0$.

Under the symmetry assumption, the payoff matrix (12) becomes

$$\begin{array}{ccc} & \text{Hawk} & \text{Dove} \\ \text{Hawk} & \frac{1}{2}u(0) + \frac{1}{2}u(1) - \phi - c_i & u(1-\eta) - c_i \\ \text{Dove} & u(\eta) & u(1/2) \end{array} \tag{25}$$

Totally differentiating Ω yields

$$\frac{d\Omega}{d\sigma} = -\left(u'(\eta) - u'(1-\eta)\right)\frac{d\eta}{d\sigma} \geq 0 \tag{26}$$

with strict inequality when $\eta > 0$, in view of (19) and strict concavity. Also,

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

Thus, actions are more likely to be strategic complements the bigger is σ and the smaller is ϕ. It is intuitive that if ϕ is large, then the most important consideration is to avoid a conflict – just as in the classic Chicken game, where a collision would be disastrous. Thus, we have the following result.

Proposition 5 *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 the players will be more concerned about territorial gains and losses than about avoiding a conflict. Actions then become strategic complements if the first-mover advantage is large enough. In fact, a large σ has two effects: it will be costly to be caught out and play Dove against Hawk, but it will be very beneficial 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. Thus, we have:

Proposition 6 *Suppose the players are symmetric ex ante and $\phi < u(1/2) - \frac{1}{2}u(0) - \frac{1}{2}u(1)$. There exists $\sigma^* \in (1/2, 1)$ such that actions are strategic substitutes if $\sigma < \sigma^*$ and strategic complements if $\sigma > \sigma^*$.*

Proof. Fix ϕ such that $\phi < u(1/2) - \frac{1}{2}(u(0) + u(1))$. From Proposition 4, we have $\Omega < 0$ if $\sigma \leq 1/2$. Now define

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

Then $\bar{\sigma} \in (1/2, 1)$, and $\eta = 0$ 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 $\sigma^* \in (1/2, \bar{\sigma})$ such that $\Omega = 0$. At $\sigma = \sigma^*$ we have $\eta > 0$. It follows from (26) that $\Omega < 0$ if $\sigma < \sigma^*$ and $\Omega > 0$ if $\sigma > \sigma^*$. ■

Note that σ^* from Proposition 6 depends on ϕ, so we can write $\sigma^* = \sigma^*(\phi)$. It is easy to check that the function $\sigma^*(\phi)$ is decreasing in ϕ.[13] This observation and Propositions 5 and 6 are summarized in Figure 1.

[13] Since η depends on ϕ and σ, we can write $\eta = \eta(\phi, \sigma)$. The function $\sigma^*(\phi)$ identified in Proposition 6 is such that $\Omega = 0$ when $\eta = \eta(\phi, \sigma^*)$. Substitute $\eta = \eta(\phi, \sigma^*(\phi))$ in (16) to get

$$\frac{1}{2}u(0) + \frac{1}{2}u(1) - u(\eta(\phi, \sigma^*(\phi)))) - u(1 - \eta(\phi, \sigma^*(\phi))) + u\left(\frac{1}{2}\right) - \phi \equiv 0 \tag{27}$$

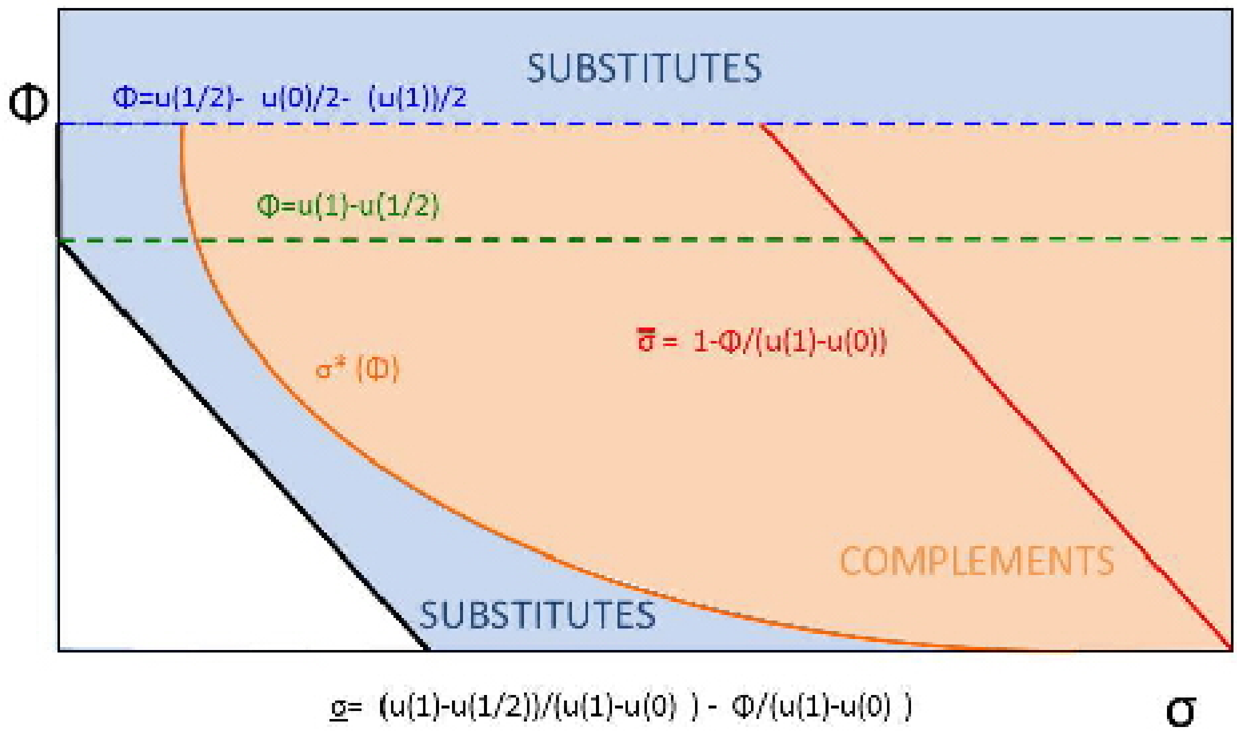
for all $\phi < u(1/2) - \frac{1}{2}u(0) - \frac{1}{2}u(1)$. The proof of Proposition 6 implies that

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

so $\eta(\phi, \sigma^*(\phi)) > 0$ satisfies (7). Using this fact, totally differentiating (27) 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 Imbalances

Discussions of conflict often focus on the case where an incumbent "great power" faces a "rising power". 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 $\phi\left(g_A\omega_A + g_B\omega_B\right)$ 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

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

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 [13], 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

$$x_A = \Omega_A F(x_B) + u(1-\eta) - u(\omega_A) \tag{28}$$

and

$$x_B = \Omega_B F(x_A) + u(1-\eta) - u(\omega_B). \tag{29}$$

where

$$\Omega_i \equiv \frac{1}{2}u(0) + \frac{1}{2}u(1) - u(\eta) - u(1-\eta) + u(\omega_i) - \phi. \tag{30}$$

Notice $\Omega_A > \Omega_B$ as $\omega_A > \omega_B$. Totally differentiating (28) and (29) 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

$$\frac{dx_A}{d\varepsilon} = \frac{\Omega_A F'(x_B)(1 - F(x_A)) u'(1/2 - \varepsilon) - (1 - F(x_B)) u'(1/2 + \varepsilon)}{1 - \Omega_A \Omega_B F'(x_A) F'(x_B)} \tag{31}$$

and

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

We have

$$\frac{dx_B}{d\varepsilon} - \frac{dx_A}{d\varepsilon} \tag{33}$$
$$= \frac{(1 - F(x_A))[1 - \Omega_A F'(x_B)] u'(1/2 - \varepsilon) + (1 - F(x_B))[1 - \Omega_B F'(x_A)] u'(1/2 + \varepsilon)}{1 - \Omega_A \Omega_B F'(x_A) F'(x_B)} > 0$$

as $|\Omega_A F'(x_A)| < 1$ and $|\Omega_B F'(x_B)| < 1$ by Assumption 2(ii). Because $x_A = x_B$ when $\varepsilon = 0$, (33) 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 (32) 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 (31) is determined by the sign of the numerator, which is certainly negative if $\Omega_A < 0$. Even if $\Omega_A > 0$, (31) 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 7 *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 (31) and (32) 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 (31) is positive if and only if

$$\frac{u'(1/2+\varepsilon)}{u'(1/2-\varepsilon)} < \frac{1-F(x_A)}{1-F(x_B)}\Omega_A F'(x_B) \tag{34}$$

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 (34) 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 8 *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 $0 \le \varepsilon < \frac{1}{2}$. 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

$$F(x_B)\frac{dx_A}{d\varepsilon} + F(x_A)\frac{dx_B}{d\varepsilon}. \tag{35}$$

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 (35) 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 8. We have the following result (the proof of this and the next result are in the Appendix):

Proposition 9 *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:

Proposition 10 *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*

$$\varepsilon^2 \leq \frac{(1 - \phi)\,(5 - 6\sigma + \phi(-3 + 4\phi + 2\sigma))}{(7 - 4\phi - 2\sigma)}. \tag{36}$$

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, (36) is impossible to satisfy for $\varepsilon > 0$ when σ is high.[14]

To return to our main question, declining inequality does increase the chance of conflict if inequality is intermediate and costs of conflict are high.

[14] If $\sigma > \sigma^*$, where $5 - 6\sigma^* + (1 - \sigma^*)(-3 + 4(1 - \sigma^*) + 2\sigma^*) = 0$ and $\sigma^* < 1$, the right hand side of (36) is negative and so there cannot be a Thucydides trap.

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 Ex Ante Strategic Moves

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. He may invest on his portion of the endowment, for example by building settlements on his land. Or he might delegate decision-making to an agent who considers it sacred. Therefore, we now consider the possibility that player A can make some *ex ante* move that changes the parameters of the bargaining game. This move is made before player A learns his type c_A. The move is observed by player B, and will have several effects. First, player A's move may directly influence his own best response function $\Gamma_A(x)$, which is a direct effect of his investment. Second, it may directly influence player B's best-response function $\Gamma_B(x)$, which is a new effect which does not arise in the typical models in industrial organization. Third, there will be indirect influence of the investment as player j responds to a change in the payoff function of player j, which is a strategic effect.

Recall player A's ex ante move makes player $i \in \{A, B\}$ tougher (softer) if the best response curve $\Gamma_i(x)$ defined by (15) shifts up (down), making player i more likely to Clearly, player A benefits if player B becomes softer (c.f. Claim 1). By the logic of Fudenberg and Tirole [13], player A benefits if he himself become tougher (softer) if the game has strategic substitutes (complements).

To simplify the analysis of changes in the valuation of territory, in this section we assume the utility function is piecewise linear. Each unit of player i's own endowment is worth v_i to him, but each unit of player j's endowment is worth only $g_i < v_i$ to player i. Setting the status quo equal to $\omega_A = \omega_B =$

1/2 and normalizing the status quo utility to zero, we get the utility function

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

This utility function is concave but not strictly concave. However, strict concavity has only been 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 holds here as the two constants g_i and v_i satisfy $g_i < v_i$.[15] Player i's best-response functions is

$$\Gamma_i(x) = F(x)\left(\frac{v_i + g_i}{4} - \phi_i - v_i\eta_i\right) + (1 - F(x))\, g_i(\frac{1}{2} - \eta_j). \tag{37}$$

We will contrast an ex ante move that lowers ϕ_A (player A's cost of conflict) with one that increases v_A (player A's valuation of his endowment).

Consider first the effect of changes in ϕ_A and v_A on $\Gamma_B(x)$. From (37), we see that this depends on their effect on η_A. The more player B can demand from player A if he is hawkish while player A is dovish the greater his incentive to be aggressive. Hence, changes in ϕ_A and v_A will make player B tougher (softer) if they make player A more (less) willing to concede territory. Now we have

$$\eta_A = \begin{cases} (1-\sigma)\,\frac{v_A + g_A}{2v_A} - \frac{\phi_A}{v_A} & \text{if} \quad \phi_A < (1-\sigma)\frac{v_A + g_A}{2} \\ 0 & \text{if} \quad \phi_A \geq (1-\sigma)\frac{v_A + g_A}{2} \end{cases} \tag{38}$$

This implies that if $\phi_A \geq (1-\sigma)(v_A + g_A)/2$ then a small reduction in ϕ_A or increase in v_A would have no effect $\Gamma_B(x)$. But note from (37) that these changes would shift $\Gamma_A(x)$ up (make player A tougher) as conflict is less costly or keeping the endowment is more valuable.

If the game has strategic substitutes (i.e. $\phi_i > u_i(1/2) - \frac{1}{2}u_i(0) - \frac{1}{2}u_i(1) = \frac{v_i - g_i}{2}$ $i = \{1,2\}$), these effects work in player A's favor, so there is a *strategic advantage* for player A to have a low cost of conflict.[16] In the terminology of Fudenberg and Tirole [13], player A should use a "top dog" strategy of

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

[16] We are of course considering here only the *strategic effect* on player B's behavior in the Hawk-Dove game: player A benefits because player B becomes more likely to choose Dove (cf. Claim 1).

over-investing in cost-reduction. Also, player A would benefit by delegating decision-making to an agent who is less conflict-averse than he is. Alternatively, he would gain by investing in valuable settlements or businesses on his land. These intuitions are familiar from Schelling's seminal work on the value of commitment. If the game has strategic complements (i.e. $\phi_i < \frac{v_i - g_i}{2}$ $i = \{1, 2\}$), the strategic effect on player A becoming tougher is that player B also becomes tougher, which lowers player A's payoff. Then, player A should use the "puppy dog" strategy and underinvest in cost-reduction.

More novel effects arise if $\phi_A < (1-\sigma)(v_A + g_A)/2$. Here we consider reductions in ϕ_A and increases in v_A separately.

Since $\partial \eta_A / \partial \phi_A < 0$, a small reduction in ϕ_A increases η_A and shifts $\Gamma_B(x)$ down (make player B softer). Substituting from (38) into $\Gamma_A(x)$ we find that small changes in ϕ_A have no effect on $\Gamma_A(x)$ (because player B's optimal challenge makes player A indifferent between a conflict and a concession) - so actually there is no direct effect on player A's investment on his own payoff function. But there is an unambiguous externality that makes player B softer and player A always gains from this whether actions are strategic complements. Hence, we have an unambiguous result: if $\phi_A < (1-\sigma)(v_A + g_A)/2$, player A should overinvest in reducing ϕ_A. This result is driven by the external effect of player A's investment on player B's payoff function and hence does not have an analog in the work of Fudenberg and Tirole [13].

Turning to v_A,

$$\frac{\partial \eta_A}{\partial v_A} = -\frac{\frac{(1-\sigma)}{2} g_A - \phi_A}{v_A^2} \tag{39}$$

so an increase in v_A would make player B softer if $\phi_A > (1-\sigma)g_A/2$ but tougher if $\phi_A < (1-\sigma)g_A/2$. Substituting from (38) into $\Gamma_A(x)$ we find that an increase in v_A makes player A tougher when $\sigma > 1/2$ and softer when $\sigma < 1/2$.

Therefore, no clear-cut result can be given for an increase in v_A, because as we have seen, player B could become either tougher or softer depending on the parameters. To understand this, let dv_A denote a small increase in v_A. If player B chooses Hawk and claims $1 - \eta_A$, the change in v_A has two effects on player A's response. If player A concedes, he must give $\frac{1}{2} - \eta_A$ units of his endowment to player B. Thus, increasing v_A by dv_A increases player A's cost of conceding by $\left(\frac{1}{2} - \eta_A\right) dv_A$. On the other hand, if player A does *not* concede, he will lose *all* of his endowment (which is of size 1/2) with probability σ. Thus, increasing v_A by dv_A increases player A's expected cost of *not* conceding by $(\sigma/2)dv_A$. The first effect dominates, making player A less willing to concede, when $\eta_A < (1-\sigma)/2$. From (38) this inequality is equivalent to $\phi_A > (1-\sigma)g_A/2$. In this case player B becomes softer. But if $\phi_A < (1-\sigma)g_A/2$ the second effect dominates and player B becomes tougher. Intuitively, since a conflict may lead to the loss of *all* the contested territory, the settlements can make player A *more* willing to give up *some part* of it in return for peace. This weakness would encourage player B to challenge the status quo, turning player A's settlements into a strategic disadvantage. We summarize our findings:

Proposition 11 *(1) Suppose* $\phi_A \geq (1-\sigma)(v_A + g_A)/2$. *Then if the game has strategic substitutes, player A should over-invest in reducing the cost of conflict and in increasing the value of resources. If the game has strategic complements, he should underinvest. (2) Suppose* $\phi_A < (1-\sigma)(v_A + g_A)/2$. *Then player A should* always *overinvest in reducing the cost of conflict. Player should overinvest in increasing the value of his endowment if* $\phi_A > (1-\sigma)g_A/2$ *but underinvest if* $\phi_A < (1-\sigma)g_A/2$.

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

7 Concluding Comments

Some conflicts are due to a failure of deterrence, others can be blamed on mutual fear and distrust.[17] Chicken and Stag Hunt are often used metaphors for these types of conflicts (Jervis [15]). The more general theoretical distinction is between strategic substitutes and strategic complements. Our simple bargaining game allows us to study how fundamental factors – preferences and technology – determine which property is likely to be satisfied. When first-mover advantages are significant, the cost of being caught out and losing territory to an aggressive opponent is large, but so is the benefit from being aggressive against a peaceful opponent.[18] For risk-averse players the first

[17] "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" (Nye [19], p. 111).

[18] McNeil [16] discusses how it shifted over time. During the era when forts could withstand a siege for many years, the offensive advantage was not large. Mobile and powerful

effect dominates, producing a Stag Hunt type interaction of strategic complements.[19] In such interactions, conflicts are triggered by fear and distrust, but communication ("cheap-talk") can be surprisingly efficient in reducing fear and building trust (Baliga and Sjöström [2]).

If conflicts are very costly, then the main concern is to avoid a conflict, and we have a game of strategic substitutes. In such Chicken-type interactions, communication will be aimed at convincing the opponent that he *does* have something to fear. Thus, the nature and value of communication depends critically on whether the game has strategic substitutes or complements. With strategic substitutes, one sees conflict through the lens of deterrence, and looking weak is dangerous. With strategic complements, diplomacy can create cooperation – the danger is the escalation that comes from grandstanding. It is noteworthy that in our model, not only does a large cost of conflict lead to a game of strategic substitutes but the fact that challengers expect large concessions tends to boost the incentive to challenge the status quo. A new technology which increases the cost of conflict may be dangerous for many reasons.

In our model, the only private information is the cost of making a challenge. It would be fairly straightforward to introduce private information about other parameters, such as the cost of a conflict or the valuation of territory. In reality, information problems may be more fundamental. In a cyberattack, the antagonist may not even be verifiable. How can cooperation be realistically enforced when there is so much noise? In many cases, the players may not even know their own cost of conflict, or the probability of winning - some analysts thought the Iraq War would be short, and Germany entered World War I expecting it to be short. One way to capture this would be to add an additional layer of uncertainty about the technology of war, and perhaps allow the players to learn the parameters as the game unfolds.

---

siege cannon, developed in France in the late 1400s, gave the advantage to the attacker. This was later neutralized by the *trace italienne* (Duffy [10]). The offense again gained an advantage with Napoleon's use of trained mass armies and Prussia's rapid, well-planned attacks with breech-loading, long range guns.

[19]If for some religious or psychological reason the territory is considered essentially indivisible, the players become risk-seekers, and first-mover advantages would produce "the winner takes it all" Chicken races instead of Stag Hunts.

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 (2004): "Arms Races and Negotiations", *Review of Economic Studies*, 71: 351-369.

[3] Sandeep Baliga and Tomas Sjöström (2012) "The Hobbesian Trap", in *Oxford Handbook of the Economics of Peace and Conflict*, edited by M. Garfinkel and S. Skaperdas. Oxford University Press

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

[5] Ken Binmore, Ariel Rubinstein and Asher Wolinsky (1986): "The Nash Bargaining Solution in Economic Modelling", *Rand Journal of Economics* 17: 176-188

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

[7] Sylvain Chassang and Gerard Padro-i-Miquel (2008): "Conflict and Deterrence under Strategic Risk," mimeo, Princeton.

[8] Sylvain Chassang and Gerard Padro-i-Miquel (2009): "Economic Shocks and Civil War," mimeo, Princeton.

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

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

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

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

[13] 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.

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

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

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

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

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

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

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

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

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

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