Credible alliances and the balance of power[*]

Sam Jindani[†] and Shuo Yang[‡]

24 June 2026

We study a model of conflict and alliances in international relations. States are endowed with military strength and may form alliances to attack other states or defend themselves from attack. A key element of our theory is credibility: an alliance is credible if it is capable of sustaining peace among its own members after victory. The model yields novel predictions about the organisation of international systems: (*i*) systems of three states are stable if and only if the strongest power is checked by the other two combined; (*ii*) a more equal distribution of power can hamper stability; and (*iii*) stronger states may be more vulnerable to conquest. We relate these theoretical findings to historical evidence, especially from the Three Kingdoms periods in China and Korea, the contemporary hegemony of the United States, and the rise of China.

1 Introduction

The Three Kingdoms period in China (220–263 CE) followed the collapse of the Han dynasty into the rival states of Wei, Shu, and Wu. Despite periodic conflict, the

---

[*] We thank Tuan-Hwee Sng and Rui Xue Jia for comments.

[†] Department of Economics National University of Singapore; sam.jindani@nus.edu.sg.

[‡] Department of Economics, National University of Singapore; shuo.yang@u.nus.edu.

borders of the three states remained relatively unchanged; Wei, the strongest power, was kept in check by an uneasy alliance between Shu and Wu. Eventually, Shu and Wu were weakened by internal politics, and Wei launched a decisive invasion of Shu in 263 CE. Wu sought to divert Wei's resources by attacking it in response, but this was unsuccessful and Shu surrendered after a single pitched battle. Now unmatched, Wei conquered Wu in 280 CE, uniting China once more.[1]

A similar dynamic unfolded on the Korean peninsula, where the kingdoms of Koguryo, Paekche, and Silla coexisted for three centuries (391–676 CE). In the fifth century, Paekche and Silla were allied against the stronger Koguryo; although Paekche in particular suffered losses, Koguryo could not conquer its rivals. By the sixth century, Silla had developed its military power, and it betrayed Paekche following a joint campaign against Koguryo. The tripartite system ended when Silla allied with the Chinese Tang dynasty and defeated first Paekche and then Koguryo (Park, 2022).

Since the collapse of the Soviet Union in 1991 CE, the United States has dominated the world order. The period saw no overt war between major powers, despite regional and proxy conflicts. Today, this world order is increasingly challenged by the rise of China, and a central question is whether peace can be sustained in a new bipolar system (Tunsjø, 2018). More generally, the question of which configurations of power are stable and which are unstable is fundamental in international relations (Morgenthau, 1949).

We propose a new framework in which to approach this question. We set out a game-theoretic model of conflict and coalition formation. States are endowed with territories, which determine military strength. States may form alliances to attack other states or defend themselves from attack. Victorious alliances distribute conquered territories among their members. Crucially, we impose that any alliance must be *credible* in the sense that it must be capable of sustaining peace among its own members after victory. A configuration of states is *stable* if no credible alliance could benefit from attacking another. Stable configurations are thus defined recursively: whether a configuration of states is stable depends on which configurations of subsets of the states are stable. The notion of credibility, and in turn the notion of stability, have far-reaching implications about the organisation of international systems.

---

[1] For an overview of the period, see the first three chapters of Dien and Knapp 2019.

Three findings merit emphasising. First, three-player systems are stable if and only if the strongest power is sufficiently checked by the other two combined. No coalition of two states is credible, because of prospect of the stronger among the pair subsequently attacking the weaker. Second, equal distributions of power are neither necessary nor sufficient for stability. Equality is not sufficient because, for example, in a perfectly balanced four-state system, any coalition of three states could overpower the fourth and split its territory in a stable way. Neither is equality necessary: in fact, contrary to conventional wisdom, we show making a system more equal can render it unstable. Thus balance of power and equality of power are sometimes at odds. Intuitively, this is because although a more equal system makes it harder for a given coalition to dominate the remaining states, it may be easier to achieve credibility. Third, being the strongest power can be a disadvantage: in some configurations, smaller states can credibly attack the strongest state, but no alliance containing the strongest state is credible. This suggests that being a hegemon in decline is an especially vulnerable position. Taken together, our results show that the determinants of the balance of power are highly subtle.

The paper contributes to the literature in international relations on the relationship between power distribution and peace (Morgenthau, 1949; Blainey, 1973; Waltz, 1979; Mearsheimer, 2001). Within this field, a growing literature uses formal game-theoretic approaches (Krainin and Wiseman, 2016; Dziubiński, Goyal, and Minarsch, 2021; Benson and Smith, 2023; Fernández-Villaverde, Koyama, Lin, and Sng, 2023). To the best of our knowledge, this is the first paper to study the role of credibility in alliance formation using a game-theoretic framework.

Our modelling approach builds on the game-theoretic literature on coalition formation (see Ray 2007 and Ray and Vohra 2015 for surveys). It is especially related to the branch of the literature that studies non-binding agreements (e.g., Ray 1989, Chwe 1994, and Genicot and Ray 2003). An important difference is that our model allows for the elimination of players (in our case, by conquest). We will see that this leads to a significantly altered analysis.

From a formal point of view, our paper is most closely related to Acemoglu, Egorov, and Sonin (2008), who study power struggles between political factions within a country, and who also focus on the credibility of coalitions. However, our work differs in several respects. First, whereas Acemoglu, Egorov, and Sonin (2008) offer both a cooperative and a noncooperative analysis, we focus entirely on the cooperative approach. This raises the important question of whether our results would hold in a suitable noncooperative framework. Second, motivated by our application, in our model the victory of a coalition allows it to capture and distribute the strength of the defeated coalition; in contrast, Acemoglu, Egorov, and Sonin (2008) assume that strengths following a victory are unchanged. This leads to qualitatively different results; in particular, in our model equality of power may lead to instability. Finally, we are able to fully characterise stable configurations up to eleven players, whereas Acemoglu, Egorov, and Sonin (2008) focus on the stability of approximately equal configurations.

It is important to note that our model abstracts away from a great number of considerations that matter in the real world. For instance, we assume that the outcome of war is deterministic, that there is no offensive or defensive advantage, that wars and conquests are not restricted by geography, that military strength can be rebuilt immediately after a conflict, that states purely seek to maximise their territory, that territory entirely determines military strength, that culture and history play no role in the formation of alliances, that states cannot commit to international policy, and so on. This is not to say that we think these considerations are not important. Our aim in this paper is to study the implications of credibility in alliance formation. To isolate these implications and to maintain tractability, we study a highly stylised model. Extending the model to incorporate more realistic assumptions is an important avenue for future research.

2 Model

Consider a set of players $N = \{1, 2, ..., n\}$ representing countries or polities. Each player i has a territory of size $x_i \geq 0$. The territory sizes corresponds to the strength of each player. A vector $x \in \mathbb{R}^n_+$ is a *state*. Whenever $x \in \mathbb{R}^n_+$ and $S \subseteq N$, we will let $x_S = \sum_{i \in S} x_i$.[2]

The set of *active* players in state x is $A(x) = \{i \in N : x_i > 0\}$ and the number of

---

[2] '$\subseteq$' and '$\subset$' denote weak and strict set inclusion, respectively.

active players is $a(x) = |A(x)|$. Given $x \in \mathbb{R}^n_+$, $x_{(k)}$ denotes the size of the kth largest territory in state x. When x is clear, (k) denotes the index of the player with the kth largest territory in state x, with ties being broken arbitrarily.

Stability is defined recursively. First, any state x with $a(x) = 1$ is *stable*. Second, a state x is *unstable* if there exist a subset of players $S \subset A(x)$ and a state $y \in \mathbb{R}^n_+$ such that:

i. (Dominance) $x_S \geq x_{N \setminus S}$,

ii. (Credibility) y is stable,

iii. (Rationality) $y_i \geq x_i$ for all $i \in S$, with strict inequality for at least one i, and

iv. (Non-disposal) $\sum_{i \in S} y_i = \sum_{i \in N} x_i$,

In this case, we say that y *blocks* x and that S is a *blocking coalition* at x. If the first condition holds, we say that S is *dominant* in x. Any state x that is not unstable is *stable*.

Intuitively, a state is unstable if there exists a subcoalition that can attack all other players, win the war, and subsequently peacefully co-exist in a division that benefits all members of the subcoalition. The Dominance condition reflects the assumption that territory size determines military strength, and that members of a coalition add up their territories to obtain the strength of the coalition. Two implicit assumptions should be noted. First, there are no neutral players: conflict is between an attacking coalition S on the one hand and all remaining players on the other hand.[3] Second, the attacking coalition only needs to match the defending coalition's strength, not surpass it. This assumption may reflect the advantage of surprise or of being able to choose conditions of engagement.[4]

The Credibility condition is the heart of the model: an attacking coalition must be able to coexist in peace following a victory. One can think of this as reflecting strong

---

[3] This is without loss of generality: if a player is willing to remain neutral, then it must either be the case that it will coexist peacefully with the winning coalition, or that it will subsequently be conquered but would not change the outcome by joining the defensive coalition. The former case is equivalent to the neutral player joining the offensive coalition, and the latter case to joining the defensive coalition.

[4] The theory could be developed under the alternative assumption that the attacking coalition must surpass the defending coalition's strength. This would lead to substantively different results.

risk aversion, in the sense that any player who faced the possibility of being conquered subsequently would not join an attacking coalition. The Rationality condition simply states that all members of an attacking coalition must weakly benefit from the new state. Finally, the Non-disposal condition implies that all losing players must be eliminated and their territory must be distributed between the victors. This reflects the idea that if territory were left unoccupied, a player would want to claim it.

As discussed in the introduction, our definition of stability and the assumptions about conflict and alliances that it implies are highly stylised. However, we will see that this simple model already yields subtle and far-reaching predictions.

3 Results

We now turn to our results. Given the recursive nature of stability, we work in increasing order of the number of active players in a state. The proofs are relegated to the appendix.

Proposition 1. *Any state with two active players is unstable.*

Our first proposition relates to the inherent instability of two-player systems. Recall that any state with only one active player is stable. When there are only two players, the stronger player will form a singleton blocking coalition and attain hegemony. Note the role of the assumption that the attacking coalition need only match the strength of the defending coalition in ensuring that a state with two equally strong players is unstable.

Proposition 2. *A state x with three active players is stable if and only if*

$$x_{(1)} < x_{(2)} + x_{(3)}. \tag{1}$$

Proposition 2 characterises stability in states with three active players: as long as the strongest power is jointly checked by the two weaker powers, the state is stable. It provides a formal explanation for the endurance of historical tripartite systems, including the Three Kingdoms periods in China and Korea. An interesting point is that small players can matter as much as large ones for the balance of power.

For instance, consider a state consisting of two equally sized large players and one small player. Proposition 2 implies that the state is stable no matter how small the smallest player is. Yet if that player were removed, Proposition 1 would imply that the resulting state would be unstable.

Proposition 3. *No state with four to six active players is stable.*

Proposition 3 is perhaps unexpected. The intuition is as follows: For a state to be stable, no singleton or credible triplet can be dominant. These two conditions are in tension, since no single player being large relative to the others makes it easier for a triplet to be credible. We show that when there are four to six active players, the two conditions are in fact incompatible, so that instability obtains. This provides our first clue that equality can be destabilising, as it tends to make coalitions more credible.

Proposition 4. *Any symmetric state with seven active players is stable.*

Proposition 4 raises an intriguing question: what is different between six and seven players that stability is impossible in one case but possible in the other? The difference is that any triplet is dominant in a symmetric state of six players, but not in a symmetric state of seven players. Here we see clearly the waterfall property of the recursive definition of stability: the stability of states with many players depends on the stability of states with few players.

Proposition 5. *No state with eight to ten active players is stable.*

For states of more than seven active players, we must consider blocking coalitions of sizes one, three, and seven. As in Proposition 3, there is a tension between these possible coalitions: just as weak singletons make it easier for triplets to be credible, non-credible triplets make it easier for coalitions of seven to be credible. For states of eight to ten active players, we show that these tensions cannot be resolved, yielding instability.

Proposition 6. *Any symmetric state with eleven active players is unstable. Moreover, there exists a stable asymmetric state with eleven active players.*

Conventional wisdom suggests that greater power parity promotes peace. However Proposition 6 shows that this is not necessarily the case. In a sufficiently symmetric eleven-player state, a coalition of seven players is both dominant and credible. Instead, stability can hold when power is asymmetric. For instance, we show that there exists a stable unipolar state with one large player and ten small players. Interestingly, the stability of this state rests on the possibility of blocking coalitions of sizes both seven and three. In particular, a coalition of seven players including the largest player is dominant but not credible, because any redistribution of the remaining territory would create a blocking triplet. In contrast, in the symmetric eleven-player state, redistribution among the members of a seven-player coalition can be achieved without creating such a blocking triplet.

Proposition 7. *There exists an unstable eleven-player state such that the largest player does not belong to any blocking coalition.*

Proposition 7 establishes that being stronger is not necessarily an advantage. We construct an eleven-player state with one large and ten small players such that any blocking coalition excludes the large player. As a result, the large player will necessarily be eliminated. The reason that the large player is excluded from any blocking coalition is that it would be too powerful and make any new state unstable, so that any coalition that includes it is not credible.

4 Conclusion

This paper has studied the stability of international systems through the lens of credible alliance formation. States differ in territorial size, which determines military strength. An alliance can displace the status quo only if it is both stronger than its opponents and credible, in the sense that it can sustain peace among its own members after victory.

We highlight three findings. First, systems of three states are stable if and only if the strongest power is checked by the other two combined. Second, a more equal distribution of power can hamper stability. Third, stronger states may be more vulnerable to conquest. These results speak to a fundamental question in international relations about the causes of war and peace. Taken together, they show that issues about the credibility of alliances can influence the balance of power in complex and unexpected ways.

A Proofs

A.1 Proof of Proposition 1

Fix a state x with $a(x) = 2$. By definition, $x_N > x_{(1)} \geq x_{(2)}$. Let $S = \{(1)\}$. Then S constitutes a blocking coalition at x.

A.2 Proof of Proposition 2

Suppose that $x_{(1)} \geq x_{(2)} + x_{(3)}$. Let $S = \{(1)\}$. Then S is a blocking coalition at x, so the state is unstable.

Conversely, suppose $x_{(1)} < x_{(2)} + x_{(3)}$. Since all two-player states are unstable, any blocking coalition must be a singleton. Yet for any $i \in \{1, 2, 3\}$, $x_i \leq x_{(1)} < x_{N\setminus\{(1)\}} \leq x_{N\setminus\{i\}}$. Thus no singleton coalition is dominant, and hence the state is stable.

A.3 Proof of Proposition 3

Fix a state x with $4 \leq a(x) \leq 6$. First, suppose $x_{(1)} \geq x_N/2$. Then (1) is dominant and the state is unstable.

Second, suppose $x_{(1)} < x_N/2$. Let $S = \{(1), (2), (3)\}$. Then, since $a(x) \leq 6$, $x_S \geq x_{N\setminus S}$, so S is dominant. Since $x_{(1)} < x_N/2$, there exists $\delta \in \mathbb{R}^6$ such that $x_i < x_i + \delta_i < x_N/2$ for each $i \in S$, $\delta_i = x_i$ for each $i \notin S$, and $(x + \delta)_N = x_N$. Let $y = x + \delta$. By Proposition 2, y is stable. Hence y blocks x, and so x is unstable.

A.4 Proof of Proposition 4

Let $x = \left(\frac{x_N}{7}, \frac{x_N}{7}, \ldots, \frac{x_N}{7}\right)$. By Propositions 1 and 3, any blocking coalition must either be a singleton or a triplet. Yet $\frac{3x_N}{7} < \frac{x_N}{2}$, so no singleton or triplet is dominant. Hence x is stable.

A.5 Ancillary definitions and results

Before proceeding to the proof of Proposition 5, we first introduce some definitions and lemmas.

Let $y \in \mathbb{R}^n$ for some integer n. We say that y satisfies the *weak triplet constraint* ($WC3$) if $y_{\{(1),(2),(3)\}} \leq \frac{x_N}{2}$. We say that y satisfies the *strict triplet constraint* ($SC3$) if the inequality holds strictly.

Fix a state x and a non-empty coalition $S \subset N$. We say that S is *dominant* at x if $x_S \geq \frac{x_N}{2}$. Let $x|_S \in \mathbb{R}^S$ be the restriction of vector x to the indices contained in S.[5] We say that S is *credible* in x if there exists a vector $y \in \mathbb{R}^S$ such that (i) $y \geq x|_S$, (ii) $y_S = x_N$, and (iii) y is stable. Clearly, S is a blocking coalition at x if and only if S is both dominant and credible.

For any $v \in \mathbb{R}^7$ satisfying ($SC3$), define the *constrained water-filling algorithm* as follows:

1. Increase $v_{(7)}$ until either $v_{(6)} = v_{(7)}$ or ($WC3$) binds. If the latter, exit the algorithm; continue otherwise.

2. Increase $v_{(i)}$ for all $i \geq 6$ until either $v_{(5)} = v_{(6)} = v_{(7)}$ or the triplet constraint binds. If the latter, exit the algorithm; continue otherwise.

3. Repeat the steps inductively for all $i > 4$ and $i > 3$, etc., while applying the same stopping rule. Iterate until either the algorithm terminates or the state is symmetric.

Let $f(v) \in \mathbb{R}^7_+$ be the output of the algorithm with input $v \in \mathbb{R}^7_+$ and let $F(v) = \sum_{i=1}^7 f_i(v)$. Note that $f(v) \geq v$, and hence $F(v) > \sum_{i=1}^7 v_i$.

Intuitively, the water-filling algorithm maximises the total territory among Pareto-improving allocation consistent that satisfy the weak triplet constraint. Since stability among seven players requires that no triplet of players dominate the rest, this allocation maximises the total territory that can be assigned while maintaining stability.[6]

---

[5] For example, suppose $n = 3$ and $S = \{1, 3\}$, then $x|_S = (x_1, x_3)$.

[6] Water-filling algorithms originate in information theory. See, for instance, Cover and Thomas (2001, section 9.4).

Lemma 1. *For any v satisfying $(SC3)$,*

$$F(v) = \begin{cases} \frac{5}{2}x_N - 4\big(v_{(1)} + v_{(2)}\big) & \text{if } v_{(1)} + 2v_{(2)} > \frac{x_N}{2} \qquad (i) \\ \frac{3}{2}x_N - 2v_{(1)} & \text{if } v_{(1)} > \frac{x_N}{6} \text{ and } v_{(1)} + 2v_{(2)} \leq \frac{x_N}{2} \qquad (ii) \\ \frac{7}{6}x_N & \text{if } v_{(1)} \leq \frac{x_N}{6}. \qquad (iii) \end{cases} \tag{2}$$

Proof. Fix $v \in \mathbb{R}^7$ satisfying $(SC3)$ with x_N. Note that $v_{(1)} + 2v_{(2)} > x_N/2$ implies $v_{(1)} > x_N/6$, so that the cases are mutually exclusive and F is well defined.

First, since $(WC3)$ depends only on $v_{(1)}$, $v_{(2)}$, and $v_{(3)}$, the algorithm proceeds until the point where $v_{(i)}$ is increased for all $i > 3$ until $v'_{(4)} = ... = v'_{(7)} = v_{(3)}$.

Second, suppose $v_{(1)} + 2v_{(2)} \geq \frac{x_N}{2}$. Then $(WC3)$ does not bind at v' and we can further increase $v_{(i)}$ for all $i > 2$ until $v''_{(3)} = v''_{(4)} = ... = v''_{(7)} = \frac{x_N}{2} - v_{(1)} - v_{(2)}$, at which point the constraint binds. The algorithm terminates with $f(v) = (v_{(1)}, v_{(2)}, v''_{(3)}, v''_{(4)}, ..., v''_{(7)})$, yielding $F(v) = 5x_N/2 - 4(v_{(1)} + v_{(2)})$.

Third, suppose that $v_{(1)} + 2v_{(2)} \leq \frac{x_N}{2}$ but $v_{(1)} > \frac{x_N}{6}$. Then $(WC3)$ does not bind at v'' and we can further increase $v_{(i)}$ for all $i > 1$ until $v'''_{(2)} = v'''_{(3)} = ... = v'''_{(7)} = (\frac{x_N}{2} - v_{(1)})/2$, at which point the constraint binds. The algorithm terminates with $f(v) = (v_{(1)}, v'''_{(2)}, v'''_{(3)}, ..., v'''_{(7)})$, yielding $F(v) = 3x_N/2 - 2v_{(1)}$.

Last, suppose that $v_{(1)} \leq \frac{x_N}{6}$. Then the algorithm proceeds to the final increment step where $v_{(i)}$ is increased for all i until $v''''_{(1)} = v''''_{(2)} = ... = v''''_{(7)} = \frac{x_N}{6}$. Then $f(v) = (\frac{x_N}{6}, \frac{x_N}{6}, ..., \frac{x_N}{6})$, yielding $F(v) = 7x_N/6$. □

Let $r(v) \in \{(i), (ii), (iii)\}$ denote the case that applies to v.

Lemma 2. *For any $v, w \in \mathbb{R}^7_+$ satisfying $(SC3)$, $v \leq w$ implies $F(v) \geq F(w)$.*

Proof. Observe that $r(v) \geq r(w)$. Consider three cases.

Case 1: Suppose $r(v) = r(w)$. Within each of the three cases, F is nonincreasing; hence $F(v) \geq F(w)$.

Case 2: Suppose $r(w) < r(v) = (iii)$. If $r(w) = (i)$, then

$$\begin{aligned} F(v) - F(w) &= \frac{7}{6}x_N - \frac{5}{2}x_N + 4(w_{(1)} + w_{(2)}) \\ &\geq -\frac{4}{3}x_N + 2w_{(1)} + \big(2w_{(1)} + 4w_{(2)}\big) \\ &> -\frac{4}{3}x_N + \frac{2}{6}x_N + x_N \\ &> 0. \end{aligned}$$

If $r(w) = (ii)$, then

$$\begin{aligned} F(v) - F(w) &= \frac{7}{6}x_N - \frac{3}{2}x_N + 2w_{(1)} \\ &= -\frac{x_N}{3} + 2w_{(1)} \\ &> 0. \end{aligned}$$

In each case, $F(v) \geq F(w)$.

Case 3: Suppose $(i) = r(w) < r(v) = (ii)$. Then

$$\begin{aligned} F(v) - F(w) &= \frac{3}{2}x_N - 2v_{(1)} - \frac{5}{2}x_N + 4(w_{(1)} + w_{(2)}) \\ &= -x_N + 2(w_{(1)} - v_{(1)}) + 2(w_{(1)} + 2w_{(2)}) \\ &> -x_N + x_N \\ &> 0, \end{aligned}$$

as desired. □

Lemma 3. *Let x be a state satisfying $(SC3)$. Let $S \subseteq A(x)$ be a coalition with $|S| = 7$. Then S is credible if and only if $F(x|_S) > x_N$.*

Proof. Suppose $F(x|_S) > x_N$. Let $v := x|_S \in \mathbb{R}^7_+$. Since $F(v) > x_N$, there exists $\delta \in \mathbb{R}^7_+$ such that

$$\sum_{i=1}^{7}(v_i + \delta_i) = F(v) > x_N,$$

where $v+\delta$ satisfies $(WC3)$. Since v satisfies $(SC3)$, there exists $\delta' \in \mathbb{R}^7_+$ such that

$$\sum_{i=1}^{7}(v_i+\delta'_i) > x_N,$$

where $v+\delta'$ satisfies $(SC3)$. Further, since $\sum_{i=1}^{7} v_i = x_S < x_N$, there exists $\delta'' \in \mathbb{R}^7_+$ such that

$$\sum_{i=1}^{7}(v_i+\delta''_i) = x_N,$$

where $v+\delta''$ satisfies $(SC3)$. Let $y = v+\delta''$. Then $y \geq x|_S$ and $y_S = x_N$. Furthermore, y satisfies $(SC3)$. Hence, by Propositions 1, 2, and 3, S is credible.

Let x be a state such that $a(x) > 7$ and S be a credible seven-player coalition. Then there exists a vector $y \in \mathbb{R}^7_+$ such that $y \geq x|_S$, $y_S = x_N$, and y is stable. Proposition 2 implies that y satisfies $(SC3)$. By Lemma 2, $F(x|_S) \geq F(y) > y_S = x_N$, where the strict inequality follows from the observation that for any v satisfying $(SC3)$, $F(v) > \sum_{i=1}^{7} v_i$.[7] $\square$

A.6 Proof of Proposition 5

Fix a state x with $8 \leq a(x) \leq 10$. Consider three cases.

Case 1: Suppose $x_{(1)} \geq \frac{x_N}{2}$. Then (1) forms a singleton blocking coalition, so x is unstable.

Case 2: Suppose $x_{(1)} < \frac{x_N}{2}$ and $x_{\{(1),(2),(3)\}} \geq \frac{x_N}{2}$. Then the three largest players form a blocking coalition, so x is unstable.

Case 3: Suppose $x_{\{(1),(2),(3)\}} < \frac{x_N}{2}$. Let S be the coalition formed by the smallest seven players. Then

$$x_S = x_N - x_{\{(1),(2),\ldots,(a(x)-7)\}} \geq x_N - x_{\{(1),(2),(3)\}} \geq \frac{x_N}{2},$$

where the first inequality uses the fact that $a(x) \leq 10$. Thus S is dominant. Let $v = x|_S$. Consider two subcases.

Case (3a): If $v_{(1)} \leq \frac{x_N}{6}$, then $F(v) = \frac{7}{6}x_N > x_N$. By Lemma 3, S is credible.

---

[7] To see this, note that $f(v)$ satisfies $(WC3)$ by construction. Therefore, $f(v) > v$.

Case (3b): If $v_{(1)} > \frac{x_N}{6}$, then $a(x) = 8$. Indeed, if instead $9 \leq a(x) \leq 10$, then

$$v_{(1)} \leq x_{(3)} \leq \frac{x_{\{(1),(2),(3)\}}}{3} < \frac{x_N}{6},$$

where the first inequality uses the fact that S consists of the seven smallest players and the second inequality uses the assumption that $x_{\{(1),(2),(3)\}} < \frac{x_N}{2}$. Hence, $v_{(1)} + 2v_{(2)} = x_{(2)} + 2x_{(3)} \leq x_{\{(1),(2),(3)\}} < \frac{x_N}{2}$. By equation 2,

$$F(v) = \frac{3}{2}x_N - 2v_{(1)} > \frac{3}{2}x_N - \frac{x_N}{2} = x_N,$$

where the inequality follows from the fact that $x_{(2)} \leq \frac{x_N}{4}$ since $x_{\{(1),(2),(3)\}} < x_N/2$ Thus Lemma 3 implies that S is credible, and hence x is unstable.

A.7 Proof of Proposition 6

First, let x be a symmetric eleven-player state. Let $S = \{1, 2, \ldots, 7\}$. Then S is dominant since $x_S = \frac{7}{11}x_N \geq \frac{x_N}{2}$. Let $v = x|_S$. Then $v_{(1)} = \frac{x_N}{11} \leq \frac{x_N}{6}$. By equation 2, $F(v) = \frac{7}{6}x_N > x_N$. By Lemma 3, S is credible. Hence S is a blocking coalition at x, and x is unstable.

Second, let $x = \left(\frac{5}{16}, \frac{11}{160}, \frac{11}{160}, \ldots, \frac{11}{160}\right) x_N \in \mathbb{R}^{11}$. Suppose, seeking a contradiction, that x is unstable, so that there exists a blocking coalition S. By Propositions 3 and 5, $|S| \in \{1, 3, 7\}$. Since $x_{\{(1),(2),(3)\}} = \frac{72}{160}x_N < \frac{x_N}{2}$, it must be the case that $|S| = 7$. Moreover, $(1) \in S$, otherwise $x_S = \frac{77}{160}x_N < \frac{x_N}{2}$. Let $v = x|_S$. Note that $v_{(1)} = x_{(1)} > \frac{x_N}{6}$ and $v_{(1)} + 2v_{(2)} = x_{(1)} + 2x_{(2)} \leq \frac{x_N}{2}$. Then $F(v) = \frac{3}{2}x_N - \frac{5}{8}x_N = \frac{7}{8}x_N < x_N$. Hence, by Lemma 3, S is not credible. Thus x is stable.

A.8 Proof of Proposition 7

Let $x = (\frac{15}{56}, \frac{41}{560}, \frac{41}{560}, ..., \frac{41}{560}) \in \mathbb{R}^{11}$. Let $S = \{2, 3, ..., 8\}$. Then S is dominant because $7(\frac{41}{560}) > \frac{1}{2}$. By Proposition 4, S is also credible because $\frac{1}{7} > \frac{41}{560}$. Hence x is unstable. Next, let S' be any blocking coalition at x and suppose that $1 \in S'$. By Propositions 3 and 5, $|S'| \in \{1, 3, 7\}$. Since, $x_{\{1,2,3\}} = \frac{232}{560} < \frac{1}{2}$, it must be the case that $|S'| = 7$. Since $x_1 > \frac{1}{6}$, Lemma 1 implies that $F(x|_{S'}) = \frac{3}{2} - 2x_1 = \frac{27}{28} < 1$. By Lemma 3, S' is not credible. Hence, $1 \notin S'$.

References

Daron Acemoglu, Georgy Egorov, and Konstantin Sonin. Coalition formation in non-democracies. *Review of Economic Studies*, 75(4):987–1009, 2008.

Brett V Benson and Bradley C Smith. Commitment problems in alliance formation. *American Journal of Political Science*, 67(4):1012–1025, 2023.

Geoffrey Blainey. *The causes of war: From 1700 to today.* Macmillan, London, 1973.

Michael Suk-Young Chwe. Farsighted coalitional stability. *Journal of Economic Theory*, 63:299–325, 1994.

Thomas M. Cover and Joy A. Thomas. *Elements of Information Theory.* John Wiley & Sons, New York, NY, USA, 2001. doi: 10.1002/0471200611.

Albert E Dien and Keith N Knapp, editors. *The Cambridge History of China*, volume 2. Cambridge University Press, 2019.

Marcin Dziubiński, Sanjeev Goyal, and David EN Minarsch. The strategy of conquest. *Journal of Economic Theory*, 191:105161, 2021.

Jesús Fernández-Villaverde, Mark Koyama, Youhong Lin, and Tuan-Hwee Sng. The fractured-land hypothesis. *The Quarterly Journal of Economics*, 138(2):1173–1231, 2023.

Garance Genicot and Debraj Ray. Group formation in risk-sharing arrangements. *Review of Economic Studies*, 70(1):87–113, 2003.

Colin Krainin and Thomas Wiseman. War and stability in dynamic international systems. *The Journal of Politics*, 78(4):1139–1152, 2016.

John J. Mearsheimer. *The Tragedy of Great Power Politics.* W. W. Norton, 2001.

Hans J Morgenthau. *Politics among nations: The struggle for power and peace.* Alfred A Knopf, 1949.

Eugene Y Park. *Korea: A history.* Stanford University Press, 2022.

Debraj Ray. Credible coalitions and the core. *International Journal of Game Theory*, 18:185–187, 1989.

Debraj Ray. *A game-theoretic perspective on coalition formation.* Oxford University Press, Oxford, 2007.

Debraj Ray and Rajiv Vohra. Coalition formation. In H. Peyton Young and Shmuel Zamir, editors, *Handbook of game theory with economic applications*, volume 4, pages 239–326. North-Holland, Amsterdam, 2015.

Øystein Tunsjø. *The return of bipolarity in world politics: China, the United States, and geostructural realism.* Columbia University Press, 2018.

Kenneth N. Waltz. *Theory of international politics.* Random House, 1979.