What Makes a World War?
A Structural Analysis of Integration

Joseph J. Ruggiero

September 3, 2024

Abstract

The size of a war, as well as the incentive to start one, depends on expectations about who will join. These expectations are in turn shaped by alignments and integration. I propose a model of conflict that accounts for these factors. Using data on international disputes from 1816-2014, I estimate the effect of integration on the incentives for conflict. The results suggest (i) network spillovers pacify potential joiners, reducing the scale of war and consequently triggering more attacks, and (ii) countries bandwagon, causing a correlation between power inequality and the expected size of war in equilibrium. Finally, the model facilitates counterfactual experiments, which is demonstrated by recomputing equilibrium behavior in potential disputes between China and Taiwan after the Cold War.

Ruggiero: Postdoctoral Fellow, Stanford University. 616 Jane Stanford Way, Encina Hall West 402, Stanford, California USA 94305. jr0@stanford.edu.

Seemingly small events, such as the assassination of an archduke, can lead to world wars that span continents. On the other hand, many disputes with global implications remain bilateral or never progress to war at all. Why do some conflicts escalate and expand, while others stay contained?

At its core, the size of a war, as well as the incentive to start one, depends on expectations about who will join. A country may be less willing to assist their ally in an attack if they expect a powerful rival to oppose them, which in turn makes their ally less willing to launch an attack in the first place. Therefore, the interests of all countries, as well as their willingness and ability to pursue and defend those interests by force, affect individual decisions to start and join wars. Adding further complication, a country's interests do not simply reflect their immediate rewards, but also account for how the resulting spoils and costs of war are expected to propagate across international connections. These direct and indirect incentives can interact in subtle ways, making it difficult to predict the trajectory of a dispute.

In this article, I propose a model of conflict where many countries are connected in an international system. One country may initiate war against a target of their choosing and, if they do, all remaining countries decide whether to join. The probability of victory, and therefore each country's expected payoff, is responsive to which countries fight. Moreover, an essential feature that distinguishes it from existing models is how preferences depend on the structure of the network. Country welfare comprises a direct payoff—potentially including gains from winning a war or costs of losing—and an indirect payoff based on the welfare of others, weighted by the strength of their relations. The direct payoff of one country therefore affects the payoffs of all their connections, and so on. In this way, integration plays a key role: small changes to the network reverberate throughout the system.

I then estimate the structural parameters of the model, including the benefits and costs associated with fighting in specific wars as well as the weights that govern international relations, using data on international disputes between 1816-2014. The marriage of theory and data facilitates a mapping from theoretical objects of interest to empirical quantities, enabling both a theoretically informed quantitative analysis and a data-driven theoretical investigation of mechanisms. The results show how the nature of international conflict changes significantly from over the past two centuries. I find that disputes in the 19th century typically involve public value goods, i.e., those where the benefits of victory are received by all countries on the winning side, not just those that fought. However, in the 1990s and onward, wars are fought over almost exclusively private value goods—only countries that fight recover the spoils of war. For this reason, countries tend to work in concert with their political and economic partners in the 19th century, whereas in the contemporary age, countries tend to act according to their own private interests, which often diverge from those of their partners.

Additionally, I recover the effect of integration on the propensity to start and join wars. This is achieved by comparing outcomes in the data to those in counterfactuals without network spillovers. I find that integration consistently pacifies potential joiners, reducing the expected scale of conflict given its outbreak. However, because countries typically join to defend an ally being attacked, less joining typically corresponds to less opposition for the initiator. As a result, integration typically makes countries more willing to launch initial attacks. The one exception is the post-Cold War period, where the effect of integration is consistently pacifying. Counterintuitively, this occurs because greater economic activity after the Cold War causes the interests of political and economic partners to diverge, *not* because it causes the interests of rivals to converge.

The model can also evaluate existing theories and weigh in on longstanding questions. In particular, the model produces natural endogenous measures of power and stability by quantifying equilibrium join and attack probabilities in addition to conditional probabilities of victory. Power—conceptualized as a country's ability to swing outcomes in their favor by the use of force—and stability—the ex-ante susceptibility of a system to large wars that involve many countries—each depend fundamentally on equilibrium behavior. I find that there is a positive and statistically significant relationship between power inequality and systemic instability in equilibrium. Moreover, I find that bandwagoning drives the correlation: as power inequality grows, countries on the stronger side become more willing to fight, causing larger wars on average. This occurs not only because potential joiners are more inclined to defend their allies when victory is probable, but also because there is little incentive for initiators to start such wars in the first place.

Finally, the model can be used to explore out-of-sample cases of interest. To demonstrate this, I conduct counterfactual experiments to recover the likelihood and expected scale of conflict between China and Taiwan for each year after the Cold War. I focus specifically on the case where Japan and the United States are aligned with Taiwan and Russia is aligned with China. Over time, the United States becomes more willing to fight a war with China over Taiwan, while Russia becomes less willing. In part a consequence of the changing international context and in part a response to the changing equilibrium behavior of others, China becomes gradually less willing to attack Taiwan. Furthermore, I recompute equilibrium strategies for all countries in the China-Taiwan example across a fine grid of trading relations, spanning from the true level of international trade given by the data to a setting without international trade flows. I find that, while reductions in international trade consistently increase the likelihood of war and its expected scale, the overall magnitude of the behavioral changes are marginal.

Related Literature

This paper contributes to work in international politics and political economy. Most relevant is research on integration and conflict. Centuries ago, Montesquieu (1748) declared that "peace is the natural effect of trade," yet even free trade champion Adam Smith (1776) worried that economic liberalization might create perverse incentives that lead to war (Paganelli and Schumacher 2019). Many empirical tests have explored the relationship between international commerce and war, but mixed results have made it difficult to establish concrete knowledge (Barbieri 2002; Hegre, Oneal, and Russett 2010; Martin, Mayer, and Thoenig

2008). Moreover, qualitative approaches that focus on specific cases are limited in their ability to provide general results (Copeland 1996, 2014). To break new ground, more recent work has begun to adopt analytical approaches that account for many interacting countries. Jackson and Nei (2015) provide a theoretical analysis of trade's potential to pacify and stabilize a network. Additionally, empirical work has begun to use networks to find conditional associations pertaining to economics and war (Kinne 2012; Lupu and Traag 2013; Dorussen, Gartzke, and Westerwinter 2016). My model contributes to this effort by providing a unified theoretical and empirical framework for the strategic decisions of many integrated countries to start and join wars.

To address limitations of standard bilateral conflict settings in understanding systemic interests (Braumoeller 2008; Cranmer and Desmarais 2016), recent scholarship has begun to adopt network techniques. König, Rohner, Thoenig, and Zilibotti (2017) is especially relevant as it uses networks to study how the structure of alliance groups affects conflict intensity. My paper differs from theirs by (i) taking individual countries as the main strategic agents as opposed to groups, (ii) accounting for war onset and escalation as sequential decisions to start and join as opposed to exerting effort in a simultaneous contest success function, (iii) incorporating economic ties that govern spillovers generally as opposed to alliance ties over the reward of the contest specifically, and (iv) connecting the model to disputes spanning 1816-2014 as opposed to focusing on the Great War of Africa. Also noteworthy is Xu, Zenou, and Zhou (2022), which presents a theoretical framework for studying conflict as a networked structure, with edges representing battles; Olivella, Pratt, and Imai (2022), which uses a dynamic stochastic block model to evaluate the conventional wisdom of democratic peace theory; and Kleinman, Liu, and Redding (2024), which establishes a causal relationship between economic dependence and political alignment.

Additionally, the model weighs in on longstanding questions about power and stability. One of the most central questions of international politics is whether stability is achieved by a balance of power (Thucydides 1954; Waltz 1964, 1979; Morgenthau and Thompson

1985) or by hegemonic imposition (Blainey 1973; Organski and Kugler 1980; Gilpin 1981; Kindleberger 1986). Statistical analyses of these theories, like those on integration and war, have found empirical relations that vary with model specification (Singer, Bremer, and Stuckey 1972; Siverson and Tennefoss 1984; Bueno de Mesquita and Lalman 1988; Kim 1989; Mansfield 1992). This paper provides a theoretically grounded empirical answer, as well as an investigation of behavioral sources from the model.

Lastly, this article belongs to a growing field that uses structural estimation to study the political economy of conflict (Carter 2010; Whang, McLean, and Kuberski 2013). Noteworthy recent examples includes Gibilisco and Montero (2022), which estimates the effect of major-power interventions on civil war onset; Crisman-Cox and Gibilisco (2018), which estimates audience costs in a dynamic model of crisis escalation; and Kenkel and Ramsay (2023), which uses a bargaining model of conflict to determine what makes a coalition powerful. The use of structural estimation in international politics is on the rise, but these methods have had an important place in conflict scholarship for at least two decades (Signorino 1999; Lewis and Schultz 2003).

1 Model

There are $n > 2$ countries, each indexed by $i \in N$ and endowed with characteristics, $m_i \in M$, as well as an alignment in the crisis situation, $s_i \in \{0, 1\}$. For simplicity in the theoretical exposition, it is useful to think about characteristics as military strength, though it may generally include other attributes in the empirical model. Additionally, aligned countries are said to have the same side and are referred to as allies, while countries on the other side are called rivals. Given the sequential structure of the game, the partitioning of countries into two sides is without loss of generality.

Countries care about the crisis outcome according to their alignment, but they are also integrated to others within and across their side. Each country's welfare depends on the welfare of the others, so that a direct gain to one country can be an indirect gain (or loss) to another. The nature of interdependence may but need not correspond to their geopolitical alignments in the crisis: it is possible that countries receive negative spillovers from allies and positive spillovers from rivals. This is modeled by representing countries by the vertices of a weighted digraph $\mathcal{G} = (N, \mathbf{\Phi})$, where $\mathbf{\Phi}$ is a real-valued $n \times n$ matrix reflecting the network connections.[1] In particular, $\phi_{ij} \in \mathbb{R}$, being the i, jth element of $\mathbf{\Phi}$, represents the extent to which country i internalizes country j's welfare.[2]

At the start of the game, a country is selected by Nature to be the first-mover. I will refer to the first-mover by h. Country h can choose to attack any rival of their choosing or remain peaceful, denoted by action $\tilde{a}_h \in \{\emptyset\} \cup R_h$ where R_h gives the set of h's rivals. If h remains peaceful, $\tilde{a}_h = \emptyset$, the game ends and peace payoffs are realized. However, if h chooses to start a war with a rival country $\ell \in R_h$, the remaining $n-2$ countries observe the outbreak of conflict between h and ℓ, and choose whether to join the fight or not, $a_i \in \{0, 1\}$ for all $i \in N \setminus \{h, \ell\}$, with $a_i = 1$ indicating that country i is joining the fight and $a_i = 0$ denoting a choice to stay out.

Letting $a_h = a_\ell = 1$ when originators h and ℓ break out into war, the n-vector $\boldsymbol{a}$ provides a full account of which countries fight and which do not. Once all decisions are made, a winning side $w \in \{0, 1\}$ and war payoffs are realized, with the initiator's victory determined by a probability that depends on actions and country characteristics, given by $p : M^n \times \{0, 1\}^n \to [0, 1]$.[3]

If a country i fights and wins, they will receive a direct gain equal to $\delta_i \geq 0$. Note that δ_i is not country i's expected gain of war ex ante, taking into account the probability

[1] The graph is defined by a set of edges E such that $(i, j) \in E$ for all $(i, j) \in N^2$ if and only if the corresponding weight $\phi_{ij} \neq 0$. By assumption, $\mathcal{G}$ does not have multiple edges or loops, i.e., ϕ_{ij} is well-defined for all $(i, j) \in N^2$ and $\phi_{ii} = 0$ for all $i \in N$.

[2] At slight abuse of terminology, I regularly refer to $\mathbf{\Phi}$ as the network rather than more accurately as the edge weights that define the network.

[3] I allow for the possibility that characteristics of countries that do *not* join the fight still affect the probability of victory. For example, alignments with the United States in the 21st century may improve the probability of victory even if the U.S. does not deploy troops, as they may nonetheless provide funding or get involved in other indirect ways.

of losing, but instead their realized direct gain conditional on winning (i.e., their "winner's reward"). If country i's allies win a conflict but i did not fight themselves, country i will only receive a share of the gain according to parameter $\beta \in (0,1)$. Here, β close to 1 reflects the circumstance where noncombatants share equally in the earned gain, whereas β close to 0 indicates that direct gains from war are only received by combatants. Hence, β measures the extent to which the gain from war is a public good for the alliance.

On the other hand, when side s_i loses, country i incurs a cost $\kappa \geq 0$ if they fought and zero otherwise. This could be thought of as a penalty the winning side imposes on the combatants of the losing side. Then, we can write each country i's direct war payoff as a function of their action and the outcome,

$$d_i(a_i \,|\, w = s_i) = \begin{cases} \delta_i & \text{if } a_i = 1 \\ \beta\delta_i & \text{if } a_i = 0 \end{cases} \qquad \text{and} \qquad d_i(a_i \,|\, w \neq s_i) = \begin{cases} -\kappa & \text{if } a_i = 1 \\ 0 & \text{if } a_i = 0\,. \end{cases}$$

Additionally, countries receive network spillovers from their international relations according to $\mathbf{\Phi}$. Call $u_i(\boldsymbol{a} \,|\, w)$ country i's total utility after action profile $\boldsymbol{a} \in \{0,1\}^n$ and realized war outcome $w \in \{0,1\}$. Then, weight $\phi_{ik} \in \mathbb{R}$ governs how a country k's welfare responds to country i's. The greater the magnitude of ϕ_{ik}, the more a country i internalizes country k's welfare and, consequently, the more other countries that are connected to country i internalize country k's welfare indirectly through their relation with i.

Lastly, countries have privately known costs of fighting, which take the form of i.i.d. mean-zero Type-I Extreme Value random utility shocks $\varepsilon_{i,a}$ for their actions a. Therefore, the realized war payoff for country i can be expressed as $u_i(\boldsymbol{a} \,|\, w) - \varepsilon_{i,a}$, where the n-vector of all realized utilities $\boldsymbol{u}(\cdot)$ is given implicitly by the system of equations

$$u_i(\boldsymbol{a} \,|\, w) = d_i(a_i \,|\, w) + \sum_{j \neq i} \phi_{ij} u_j(\boldsymbol{a} \,|\, w) \tag{1}$$

for all $i \in N$.-

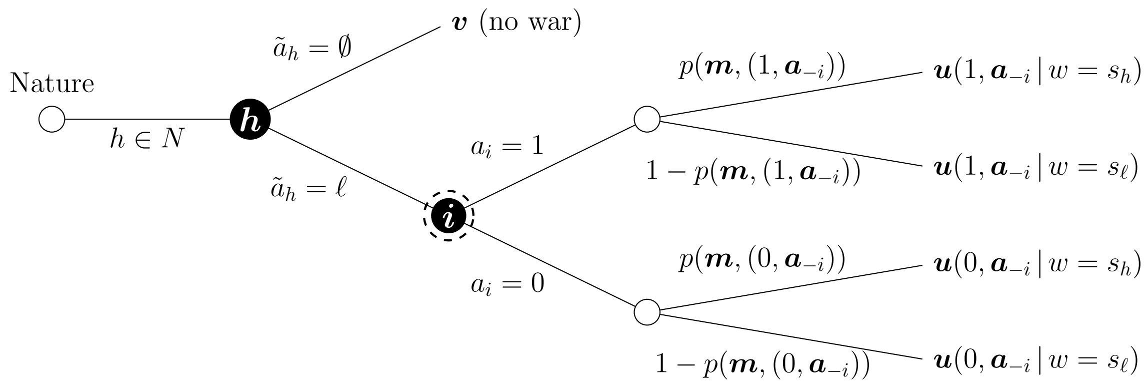


Figure 1: Graphical representation of the game

If, however, the first-mover chooses to keep the peace, all countries will receive peace payoffs given by the n-vector $\boldsymbol{v}$ that solves the system of equations

$$v_i = \alpha + \sum_{j \neq i} \phi_{ij} v_j \tag{2}$$

for all $i \in N$. The quantity $\alpha \geq 0$ is the direct public gain from total peace. For any action profile $\boldsymbol{a}$ and sufficiently small weights ϕ_{ij},[4] the systems of equations (1) and (2) have unique stable solutions. Both the direct war and peace payoffs propagate through the network, allowing for a country's willingness to fight to be influenced by the interests of those several degrees removed from them.

1.1 Timing

To recap, the game proceeds as follows. Figure 1 provides a graphical representation.

1. Nature selects a first-mover $h \in N$ and all countries realize their private costs of war.

2. The first-mover can either attack a target, $\tilde{a}_h = \ell \in R_h$, or keep the peace, $\tilde{a}_h = \emptyset$.

3a. If they choose peace, no war occurs and peace payoffs $\boldsymbol{v}$ are realized.

[4] Formally, the solutions are well-defined for any $\boldsymbol{a} \in \mathbb{R}^n$ if and only if $\varrho(\boldsymbol{\Phi}) < 1$, where $\varrho(\cdot)$ denotes the spectral radius. For more detail, refer to Theorem 1 and Lemma 1 in Appendix B.

3b. If they attack a rival $\ell \in R_h$, the remaining $n-2$ countries decide whether to join $a_i = 1$ or stay out $a_i = 0$ of the conflict. A war outcome is realized according to $p(\boldsymbol{m}, \boldsymbol{a})$ and war payoffs $\boldsymbol{u}(\boldsymbol{a} \,|\, w)$ are realized.

The solution concept is Perfect Bayesian Equilibrium. For proof of existence, refer to Proposition 1 in Appendix B.

1.2 Equilibrium

Upon observing that first-mover h has launched an attack on target ℓ, all countries must make a decision about whether to join the fight. For any action profile $\boldsymbol{a} \in \{0,1\}^n$, each country i can form an expected payoff conditional on the realized coalitions equal to $\tilde{u}_i(\boldsymbol{a}) - \varepsilon_{i,a_i}$ where $\tilde{u}_i(\cdot)$ is defined by

$$\tilde{u}_i(\boldsymbol{a}) := p(\boldsymbol{m}, \boldsymbol{a})\, u_i(\boldsymbol{a} \,|\, w = s_h) + (1 - p(\boldsymbol{m}, \boldsymbol{a}))\, u_i(\boldsymbol{a} \,|\, w = s_\ell). \tag{3}$$

Equation (3) therefore yields the expected utility from war given a realized action profile $\boldsymbol{a}$. Countries can form this expectation for any realized action profile, but they must form an additional expectation over all of these possibilities.

Taking countries h and ℓ as the conflict's originators, we can denote the corresponding set of possible combatant realizations by $\mathcal{C}^{h,\ell}$. For example, the elements of this set include the possibility that all remaining countries join, no remaining countries join, and everything in between. Likewise, denote by $\mathcal{C}^{h,\ell}_{-i}$ the set of potential combatant realizations that excludes country i. Then, the total expected payoff to country i taking action a when h and ℓ are at war is given by

$$U_i^{h,\ell}(a) = \sum_{C \in \mathcal{C}^{h,\ell}_{-i}} \Pr(\boldsymbol{a}_C)\, \tilde{u}_i(a, \boldsymbol{a}_C) - \varepsilon_{i,a} \tag{4}$$

where $\boldsymbol{a}_C$ is the action profile implied by the coalitions in C and $\Pr(\boldsymbol{a}_C)$ is the probability of that action profile being realized, which depends on equilibrium joining strategies of the remaining $n-3$ countries.

A country i then joins a fight between h and ℓ if and only if $U_i^{h,\ell}(1) - U_i^{h,\ell}(0) > 0$. Equivalently, using equation (4), country i joins if and only if

$$\sum_{C \in \mathcal{C}_{-i}^{h,\ell}} \Pr(\boldsymbol{a}_C) \left[\tilde{u}_i(1, \boldsymbol{a}_C) - \tilde{u}_i(0, \boldsymbol{a}_C)\right] > \varepsilon_i \tag{5}$$

where $\varepsilon_i := \varepsilon_{i,1} - \varepsilon_{i,0}$ represents country i's private costs of fighting in a war between h and ℓ. Equation (5) then has a clear interpretation: country i enters if and only if their private costs of war are less than the total expected surplus from joining.

Usually, increases to the value of winning will have the direct effect of increasing the probability of joining, and vice versa for the cost of losing. Further, the more the gain from war is public-valued, the less countries will be willing to fight: if a larger share of the gain from victory is retained without taking the risk of fighting, countries will not take the risk. However, network spillovers can disrupt this logic. For example, increases in the penalty for losing may encourage a country to join a war if they are closely connected to their originating ally and can help swing the outcome in their side's favor. Even in simplified special cases, the additional benefit of being a joiner is ambiguous in changes to the network. With three countries, we can obtain a closed-form expression for the joiner's threshold, which demonstrates that the effect of changes to the network are highly contingent on other parameters of the model. Appendix A provides formal comparative statics of these special cases, in addition to numerical simulations.

Finally, consider the decision of the first-mover h. Attacking a country $\ell \in R_h$ yields total expected utility

$$U_h(\ell) = \sum_{C \subseteq \mathcal{C}^{h,\ell}} \Pr(\boldsymbol{a}_C)\, \tilde{u}_h(\boldsymbol{a}_C) - \varepsilon_{h,\ell} \tag{6}$$

The intuition for equation (6) mirrors that of equation (4), except country h takes $|R_h|$ expectations (as opposed to two for the potential joiners), and that the expectations are formed over more possibilities, as the set of combatant realizations $\mathcal{C}^{h,\ell}$ is 2^{n-3} larger than each $\mathcal{C}^{h,\ell}_{-i}$. Further, the privately known shock, $\varepsilon_{h,\ell}$, can still be thought of as country h's private costs of fighting a particular target ℓ.

The option of keeping the peace, on the other hand, is given by $U_h(\emptyset) = v_h$, defined in equation (2). Then, it follows that first-mover h chooses the solution to

$$\tilde{a}_h \in \underset{\tilde{a} \in \{\emptyset\} \cup R_h}{\arg\max}\ U_h(\tilde{a}).$$

In the measure zero event that the first-mover is indifferent across $k > 1$ optimal actions, they choose each option with probability k^{-1}.

The effect of changes to parameters on attack probabilities depends crucially on how their effect on join probabilities. For example, because wars met with fierce opposition are less desirable than those without, changes in parameters that make rivals join with greater frequency relative to allies will have the net effect of reducing the incentive to attack. In that way, some parameters have clear expectations: a country's equilibrium attack probabilities are typically declining in the value their rivals receive from winning war and increasing in the value their allies receive. However, all forces are highly contingent on the network structure since changes to any parameter not only influences a change in behavior via their direct effect, but also through a complex web of indirect effects that reverberate through the network.

Among the forces at play, those that prevail depend on specifics of the game and resulting equilibrium behavior of others. For example, consider a close connection to an ally that is in turn closely connected to a rival. If rival is likely to join the war, this may reduce the incentive to fight. This is because the rival joining implies that victory in war would be met with an internalization of the rival's costs of losing. Then, this second-degree connection mitigates the desire to win the war. However, if the rival is not expected to join, the incentive to fight will be unchanged since the rival would no longer incur the penalty of losing and, therefore, it would not be internalized through their ally's connection.

In this same way, the interaction between the network and power plays an important role. Building on the previous example, the extent to which a closely-connected rival joining mitigates the incentive to fight depends on power. The effect will be strongest for very powerful countries, since those are the countries that can swing the outcome in their side's favor and doing so leads to indirect penalties from the losing rival. However, if the country is weak, they cannot considerably change the war's outcome and, therefore, their decision is not consequential on the likelihood of receiving the indirect penalty.

There are two forces that compel countries to fight: (i) improving the probability of their side's victory and (ii) the incentive to recover additional gains from a victory. The first channel is shut down for weak countries and, as a result, they act according to their resource incentive. It is possible, then, that weak countries in these situations effectively "hedge their bets," entering wars with greater frequency as their connections to rivals strengthen to reduce the pain of a military loss. Appendix A provides numerical simulations that demonstrate this behavior, in addition to analysis of special cases.

1.3 Discussion of Model Assumptions

There are several model features that deserve discussion. First, the model takes alliances as predetermined and, hence, exogenous to the strategic interaction. As a result, countries are not free to target any country of their choosing when selected as the first-mover, but are instead constrained to targeting nonallies. The assumption also restricts nonoriginators to two alternatives—to join the fight or not—as opposed to allowing them to pick sides. This modeling choice makes sense given our data, as international alignments generally *are* known prior to the outbreak of conflict. Further, accounting for a country's true ability to switch sides would not only require a model with much more complicated dynamics,but would obscure the strategic interaction of interest: entering conflicts.

Second, the model assumes that alignments divide all countries into one of only two groups. This assumption is fairly innocuous due to the sequential structure of the game: if there are more than two alliances, the outbreak of an initial dyadic war necessarily divides these groups into two sides. For example, if a war breaks out between two countries from two different alliances, a third country not belonging to either of their alliances would be aligned with the side they prefer to win (and, if indifferent, they would simply have no value for winning the war regardless of their assigned side). If countries know how these alignments divide, systems with many alliances easily fold into this two-sided version.

Third, whether there is war or peace is at the discretion of a first-mover that is chosen randomly by Nature. Albeit a theoretical construct, this effectively captures the long-held view in international relations that there are "windows of opportunity" for an attack (Lebow 1984). The incorporation of this feature enriches the model by facilitating an initiation stage without taking the outbreak of war for granted, as is standard in models of escalation. Alternative initiation processes that do not rely on selection are likely to yield unnecessary complications, creating arbitrary yet much more consequential modeling choices that can easily be avoided by imposing a randomly selected first-mover.

Fourth, countries cannot condition their choice to join a conflict on observing another country's decision, but must decide based on their expectations. Simultaneous moves allow for strategic best responding to the equilibrium actions of others without the need to impose an artificial ordering. Sequential move, on the other hand, would not only demand committing to a specific ordering, but would also require the more harsh assumption that early movers are committed to fighting regardless of the later decisions of others. It is also worth noting that the model, as is, allows for the implementation of what are effectively ad hoc sequential move "contingencies" via counterfactual analysis.

2 Data

Throughout my analysis, the main data I rely on is the Correlates of War (CoW) project's Militarized Interstate Disputes (Palmer, McManus, D'Orazio, Kenwick, Karstens, Bloch, Dietrich, Kahn, Ritter, and Soules 2020). These include 2,436 disputes from 1816 to 2014, ranging from petty fisheries-related disputes to large-scale global war. The data provide the level of hostility reached at the dispute level, which allows for differentiation between those that escalated to an attack and those that did not. Figure 2 presents the distribution by the number of involved countries, with shading to reflect the number of casualties associated with the dispute. The vast majority of disputes do not result in fatalities and involve only two countries; however, disputes that involve many countries are much more likely to be extremely fatal. The scarcity of nondyadic international wars makes it difficult to obtain enough sample data for standard methods of empirical analysis. A benefit of the structural approach of this paper is that disputes that do not escalate to war, in addition to those that do, can be used to make inferences about model parameters.

An interesting feature of the data pertains to the size of disputes over time. Figure 3 shows the average number of disputants per dispute by decade. There is high volatility prior to 1870 with no clear pattern. However, for the century from 1880-1980, there is a consistent decline in the average size of the dispute, with an increased rate of decline following World War I. The pattern stops abruptly in 1990, marking the end of a century long downward trend. This is in part due to smaller incidents included in the data for the years after 1993, as noted by the authors of the data set. To avoid issues with structural breaks, I divide the available data into four distinct corresponding time periods for structural estimation, divided by 1870, 1914, and 1989. These happen to correspond to substantively meaningful moments in history: respectively, the Franco-Prussian War and unification of Germany, World War I, and the end of the Cold War.

Additionally, the data provide a full account of participating countries that identifies originators and alignments. Unfortunately, this is a not a full account of the relevant coun-

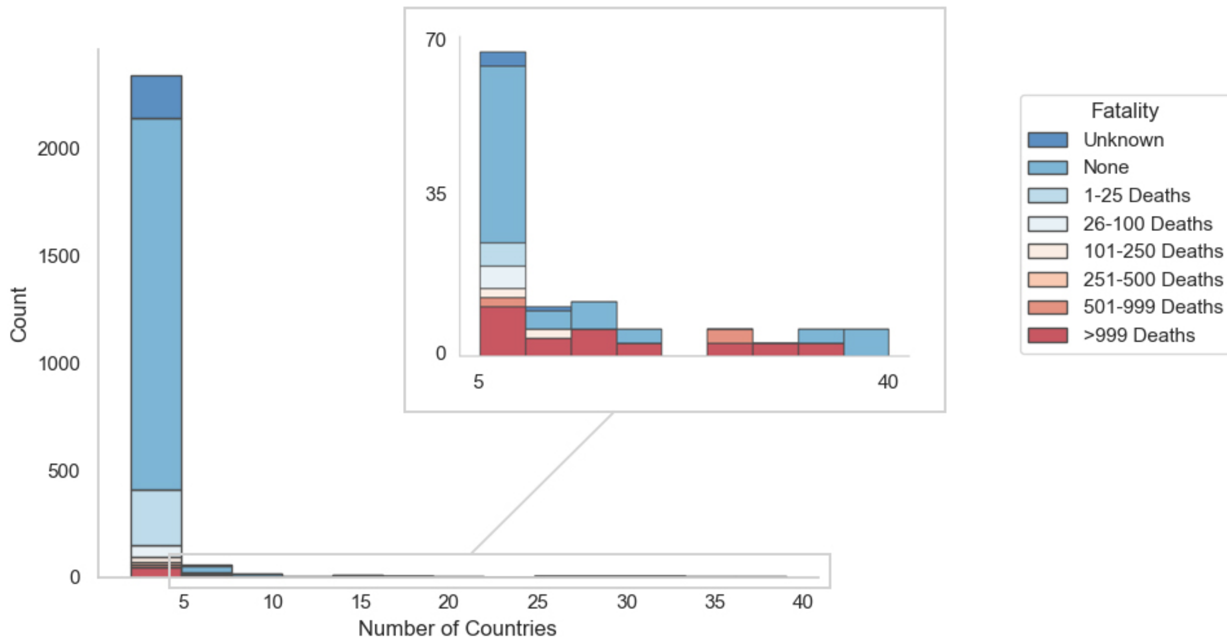


Figure 2: International disputes, 1816–2014

tries, and there is no clear way to distinguish irrelevant countries from relevant ones that simply chose not to fight. This is especially important since *all* relevant countries, including those who choose not to join the war, are critical to the strategic calculus and resulting equilibrium behavior. To overcome this issue, I follow previous scholarship in international relations (Maoz 1996; Lemke and Reed 2001) and employ the concept of a "politically relevant international environment," constituted by major powers and contiguous countries as identified by the CoW State System (2017) and Direct Contiguity (Stinnet, Jaroslav, Schafer, Diehl, and Gochman 2002) data sets.

Once the relevant set of countries are identified, we face an additional problem: to which side of the dispute do they belong? Alignments for active combatants are given by the dispute data but, for relevant countries that chose not to fight, it is not obvious which side they would have fought for if they had. I assign nonparticipants to the initiating side if and only if they are identified as a member of a formal alliance with the initiating country (i.e., they have an active defense pact or other arrangement of mutual military support) or as a strategic rival of the target country. For this, I rely on CoW's Formal Alliances

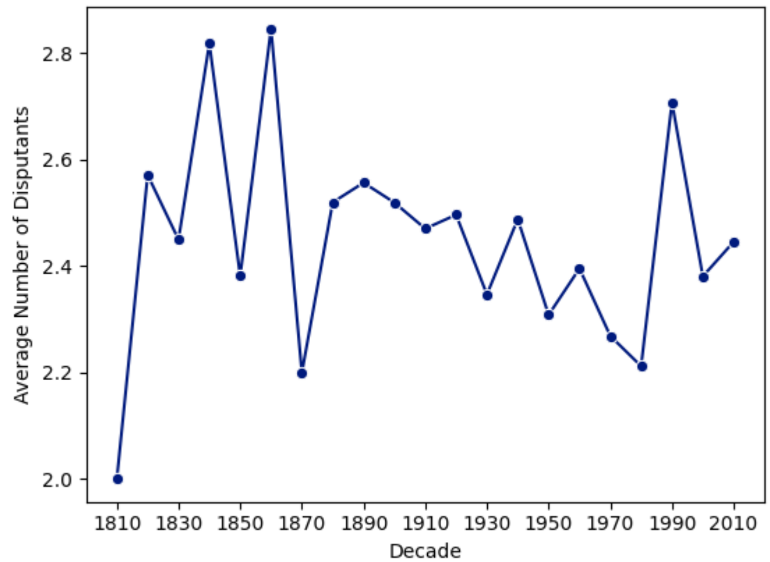


Figure 3: Average number of disputants per dispute over time

(Gibler 2009) data, in conjunction with data collected from Thompson and Dreyer (2012) on strategic rivalries (Miller 2022). Throughout the paper, "allies" is used to loosely reflect countries on the same side of a political issue, and does not imply the common and more strict notion of an alliance with formal agreement and obligation.

CoW Trade (Barbieri, Keshk, and Pollins 2009; Barbieri and Keshk 2016) data provide direct trade flows from 1870-2014 that inform integration. To avoid losing a large amount of already limited dispute data, I additionally impute missing trade data, including all trade prior to 1870, using a gravity model and a Poisson Pseudo-Maximum Likelihood (PPML) estimator (Silva and Tenreyro 2006, 2011). This approach is preferable to common alternatives of listwise deletion or the assumption of zero trade when missing. For more details on the gravity model and PPML estimator, refer to Appendix C.2.

The trade data reinforce the conventional wisdom that the world becomes significantly more integrated over the 20th century leading into 2014. One interesting consequence of this is how opposing disputants in the 2010s have a vastly different relation to each other than opposing disputants for all of the preceding history. Figure 4 shows average trade flows between dispute opponents by decade, demonstrating the unprecedented conditions

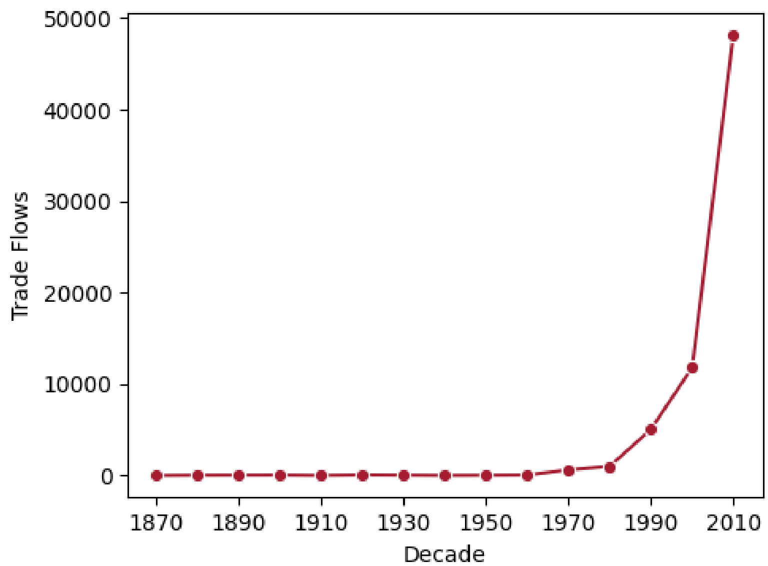


Figure 4: Average trade flows between dispute opponents in millions of current USD

of the post-Cold War period. In 1870, disputes had an average of $22 million in current U.S. millions flowing across sides. That number would not exceed $50 million until the 1960s. There is an exponential rise after the 1980s, with the 1990s seeing more trade across disputing sides than all previous decades combined. Average trade with dispute opponents in 2010-2014 exceeded $48 billion, two and a half times more than that spanning 1870-2010 combined.

For country-level characteristics, I use CoW's National Material Capabilities (Singer, Bremer, and Stuckey 1972; Singer 1987) data set, which provides military capacities and population. Penn World Table (Feenstra, Inklaar, and Timmer 2015) provides measures of country gross domestic product (GDP), for which I follow Kenkel and Ramsay (2023) in using the $RGDP^0$ measure to facilitate comparisons across country-years. Further, the Polity5 Project (Marshall and Gurr 2020) provides data on regime type, and Roser, Herre, and Hasell (2013) provide estimated stockpiles of nuclear weapons. I also use data from the Issue Correlates of War (ICOW) Historical State Names Data Set (Hensel 2016) to compute dyad-year geographic distance with Google's geocoding API according to the methodology detailed in Cooley (2018). These variables are also aggregated and used as system-level variables, such as total gross domestic product or average geographic distance. Multiple imputation by an Expectation-Maximization with Bootstrapping (EMB) algorithm (Honaker, King, and Blackwell 2011) is conducted using additional data sources, the details of which are available in Appendix C.1.

3 Estimation

In this section, I first demonstrate that the model parameters are uniquely pinned down with unlimited data. Then, I outline the estimation procedure that exploits heterogeneity in the available data.

3.1 Identification

Identification consists of two parts that correspond to each step in the subsequent two-step estimation process. First, I show that the join probabilities under each possible war, the attack probabilities under each possible recognition, and the probability of victory under each possible configuration of combatants are nonparametrically identified from the Militarized Interstate Disputes data directly. Second, I show that structural parameters $\boldsymbol{\theta} := (\boldsymbol{\delta}, \beta, \kappa, \alpha, \boldsymbol{\Phi})$ are uniquely consistent with these probabilities. Recovering equilibrium choice probabilities directly from data (Hotz and Miller 1993; Iaryczower and Shum 2012) improves tractability by turning the complex strategic problem into n simple optimization problems.

I refer to the first-stage quantities as the "reduced-form" probabilities. Formally, the first-stage reduced-form probabilities include, for all $h \in N$ and $\ell \in R_h$,

(i) $\gamma_i^{h,\ell} := \Pr(a_i = 1 \,|\, h, \ell)$, the probability a country i will join a war between h and ℓ;

(ii) $\pi_h^\ell := \Pr(\tilde{a}_h = \ell \,|\, h)$, the probability that country h attacks country ℓ when recognized as a first-mover;[5] and

[5] While the structure of the data allows us to treat recognition as observed, an alternative approach may use the unconditional attack probability as a reduced-form probability and treat recognition probabilities as $n-1$ additional structural parameters to estimate.

(iii) $\rho_C := \Pr(w = s_h \,|\, C)$, the probability the initiating side wins a war given the realized combatants are $C \in \mathcal{C}$.

Assume that we have T observations of a country i's decision to join a war between h and ℓ. Denote by $a_{i,t}^{h,\ell} = 1$ the event that country i joins a war between h and ℓ in observation t, and 0 otherwise. By definition, the mean of the observations converges to i's equilibrium join probability for a war between h and ℓ, i.e., $\lim_{T\to\infty} \frac{1}{T}\sum_{t=0}^{T} a_{i,t}^{h,\ell} = \gamma_i^{h,\ell}$. The same argument holds for the other reduced-form probabilities.

Next, to see that the structural parameters are uniquely consistent with these probabilities, denote by $\rho_{i,C}$ the probability that country i's alliance wins given the realized combatants are C, so that $\rho_{i,C} = \rho_C$ if $s_i = s_h$, and $\rho_{i,C} = 1 - \rho_C$ otherwise. For notational convenience, I write that $\rho_{i,C}(a) := \rho_{i,C'}$ for any two coalitions $C \in \mathcal{C}_{-i}^{h,\ell}$ and $C' \in \mathcal{C}^{h,\ell}$ if and only if $\boldsymbol{a}_{C'} = (a_i = a, \boldsymbol{a}_C)$. Equation (5) now becomes

$$\begin{aligned}\Omega_i^{h,\ell}(\boldsymbol{\theta}) := \sum_{C\subseteq \mathcal{C}_{-i}^{h,\ell}} \Bigg(\prod_{r\in C} \gamma_r^{h,\ell} \prod_{k\notin C} (1-\gamma_k^{h,\ell}) \Big[&\rho_{i,C}(1)\Delta u_i(1, \boldsymbol{a}_C) + u_i(1, \boldsymbol{a}_C \,|\, w \neq s_i) \\ &- \rho_{i,C}(0)\Delta u_i(0, \boldsymbol{a}_C) - u_i(0, \boldsymbol{a}_C \,|\, w \neq s_i) \Big] \Bigg) > \varepsilon_i\end{aligned} \tag{7}$$

where $\Delta u_i(\cdot)$ denotes the total gain in utility from winning a war, i.e.,

$$\Delta u_i(a_i, \boldsymbol{a}_C) := u_i(a_i, \boldsymbol{a}_C \,|\, w = s_i) - u_i(a_i, \boldsymbol{a}_C \,|\, w \neq s_i).$$

Then, given (h, ℓ) as the originating couple and recalling that $\varepsilon_i \overset{iid}{\sim} F$ for all $i \in N \backslash \{h, \ell\}$, the structural parameters are pinned down by two sets of moments corresponding to the join probabilities and attack probabilities. For each, we can map data to its theoretical counterpart as a function of the structural parameters. First, join probabilities yield a system of equations,

$$\gamma_i^{h,\ell} = F(\Omega_i^{h,\ell}(\boldsymbol{\theta})) \tag{8}$$

for all $i, h \in N$ and $\ell \in R_h$. Second, attack probabilities yield a system of equations,

$$\pi_h^\ell = \sigma(\bar{\boldsymbol{U}}_h(\boldsymbol{\theta}))_\ell \tag{9}$$

for all $h \in N$ and $\ell \in R_h$, where $\sigma(\cdot)_\ell$ denotes the ℓth element of the normalized exponential transformation that takes a vector of total expected utilities as input and returns the choice probabilities.

Intuitively, we can expect $\boldsymbol{\delta}$ and $\boldsymbol{\Phi}$ to be separately identified due to their different effects on country preferences. On one hand, increases in an individual δ_i will result in a greater utility for country i and those connected to i, in proportion to the network weights, when s_i wins the war. When s_i loses, utilities are unaffected. Thus, equilibrium behavior will change to reflect an added incentive to join for closely connected allies of i and reduced incentive to join for closely connected rivals. The direct effect is only on country i, with all other countries receiving indirect effects according to the structure of the network.

On the other hand, increases to the elements of $\boldsymbol{\Phi}$ result in greater internalization of each other's welfare, regardless of who wins and whether there is war at all. By increasing a particular ϕ_{ij}, country i becomes more invested in the welfare for country j and, thus, those invested in country i become more invested in country j, as well, and so on. We can then think about δ_i as pushing the incentive to fight up and down for different countries depending on how they are connected to i, while ϕ_{ij} pulls the interests of all countries towards those of j, depending on their proximity to i. Crucially, changes to δ_i only affect the behavior of others insofar as country i intends on entering or if β is sufficiently high, while ϕ_{ij} always affects the behavior of others.

The best estimate of the structural parameters are those that minimize the distance between these probabilities from the data and their theoretical counterparts, as defined by equations (8) and (9). The parameters were successfully recovered in simulated experiments and Appendix B shows that the system is overdetermined. Here, a key advantage of the structural approach is made clear: the model overcomes data scarcity by using observations of multiple types of actions to make inferences about preferences. This means leveraging observations of attack decisions to learn about what countries would do if they were faced with a decision to join, and likewise using observations of joining decisions to inform us about what they would choose to do if faced with a decision to attack.

Additionally, these estimands correspond to well-defined theoretical objects of interest. Power, economics, and war are all highly endogenous, and "effects" estimated in a reduced-form way will be difficult if possible to interpret in a strategic setting, not to mention they do not facilitate extrapolation to new settings with behavioral implications. In this paper, by estimating deep parameters that govern preferences directly, we can speak in clear terms about how theoretical objects of interest relate to strategic behavior. Any subsequent effect that we care about, such as the effects of integration, can still be recovered within this framework.

3.2 Estimation Procedure

The model is estimated in a two-step procedure that mirrors the two-step argument for identification. However, where identification assumed a consistent strategic environment across settings, the quantitative analysis below exploits heterogeneity in the data to recover reduced-form probabilities as a function of observable characteristics of the countries $\mathbf{X}_i$ and the relevant international system $\mathbf{Z}_t$.

This approach requires specifying the relevant set of countries for any particular conflict. To define these relevant sets, I rely on the notion of "politically relevant international environment" from previous work in international relations (Maoz 1996; Lemke and Reed 2001), which include major powers in the year of the dispute and contiguous countries to the disputants. These groups are ideally large enough to satisfy the expectations of the research problem while remaining tractable. Computation of the model is costly as n gets large due to the complexity that arises when $n-2$ countries take expectations over 2^{n-3}

possible combatant realizations, each of which can result in winning or losing. By employing this method, inferences across different strategic settings are possible when data on country characteristics are available.

Reduced-form probabilities for γ and ρ are then estimated directly from the data. In particular, let each probability be a function of country-level $\mathbf{X}_i$ and system-level $\mathbf{Z}_t$ characteristics according to logistic models[6]

$$\rho(\mathbf{X}, \mathbf{Z}_t; \boldsymbol{\mu}_{\rho,t}) := \frac{\exp(\mathbf{X}_i'\boldsymbol{\eta}_\rho + \mathbf{X}_{-i}'\boldsymbol{\zeta}_\rho + \mathbf{Z}_t'\boldsymbol{\xi}_\rho + \tau_{\rho,t})}{1 + \exp(\mathbf{X}_i'\boldsymbol{\eta}_\rho + \mathbf{X}_{-i}'\boldsymbol{\zeta}_\rho + \mathbf{Z}_t'\boldsymbol{\xi}_\rho + \tau_{\rho,t})} \in (0, 1)$$

$$\gamma(\mathbf{X}, \mathbf{Z}_t; \boldsymbol{\mu}_{\gamma,t}) := \frac{\exp(\mathbf{X}_i'\boldsymbol{\eta}_\gamma + [\mathbf{X}_h' \;\; \mathbf{X}_\ell']\boldsymbol{\zeta}_\gamma + \mathbf{Z}_t'\boldsymbol{\xi}_\gamma + \tau_{\gamma,t})}{1 + \exp(\mathbf{X}_i'\boldsymbol{\eta}_\gamma + [\mathbf{X}_h' \;\; \mathbf{X}_\ell']\boldsymbol{\zeta}_\gamma + \mathbf{Z}_t'\boldsymbol{\xi}_\gamma + \tau_{\gamma,t})} \in (0, 1)$$

where reduced-form parameters are given by $\boldsymbol{\mu}_t := (\boldsymbol{\eta}, \boldsymbol{\zeta}, \boldsymbol{\xi}, \boldsymbol{\tau}_t)$, with $\boldsymbol{\tau}_t$ denoting time period fixed effects for period t, and h and ℓ representing the initiator and target in network t. Thus, the first step of the procedure is estimating the coefficients of the reduced-form models, $\boldsymbol{\mu}_t$.

At the second stage, I use predicted values of first-stage parameters and observed actions to recover estimates for the structural parameters, $\boldsymbol{\theta} = (\boldsymbol{\delta}, \beta, \kappa, \alpha, \boldsymbol{\Phi})$. The model can accommodate observed heterogeneity at the country, dyad, and network levels to improve estimation. Specifically, the value of winning can be parameterized in terms of country characteristics $\mathbf{X}_i$ and network-level characteristics $\mathbf{Z}_t$, while the network connections can be parameterized dyad-level characteristics $\mathbf{X}_{ij}$. In particular,

$$\boldsymbol{\delta} := \begin{bmatrix} \mathbf{1} & \mathbf{X}_i' & \mathbf{Z}_t' \end{bmatrix} \boldsymbol{\delta}_x$$
$$\boldsymbol{\phi}_i := \begin{bmatrix} \mathbf{S}_{ij}' & \mathbf{X}_{ij}' \end{bmatrix} \boldsymbol{\phi}_x$$

where $\mathbf{1}$ is a column vector of ones and $\mathbf{S}_{ij}$ is a $2 \times n$ matrix with a first row of indicator variables identifying i's allies and a second row identifying rivals. Then, we can recover

[6] I additionally estimate recognition probabilities and attack probabilities as reduced-form logistic models, but these are not necessary to recover estimates of structural parameters.

estimates for any δ_i and any ϕ_{ij}.

To implement this approach, I include country polity score, population, and national wealth, as well as network-level features in the distance between a coalition's average polity score and the relative size of the coalition measured by aggregate population. Further, exports and imports as a share of gross domestic product, across allies and rivals, are included as dyad-level characteristics. Refer to the Data section for source details.

We can then take advantage of having observed all actions pertaining to join and attack decisions. The probability of observing an action $a_{i,t}$ from country i in network t when structural parameters are $\boldsymbol{\theta}$ is therefore given by

$$\psi(\boldsymbol{\theta}; \hat{\boldsymbol{\rho}}, \hat{\boldsymbol{\gamma}}_{-i,t}, a_{i,t}) := F(\Omega_i(\boldsymbol{\theta}; \hat{\boldsymbol{\rho}}, \hat{\boldsymbol{\gamma}}_{-i,t}))^{a_{i,t}}(1 - F(\Omega_i(\boldsymbol{\theta}; \hat{\boldsymbol{\rho}}, \hat{\boldsymbol{\gamma}}_{-i,t})))^{1-a_{i,t}}.$$

Further, let $\boldsymbol{y}_t$ denote the vector of binary attack outcomes for the first-mover h in network t such that $y_{j,t} = 1$ if $\tilde{a}_h = j$ and 0 otherwise. Then, the probability of observing $y_{j,t}$ when structural parameters are $\boldsymbol{\theta}$ is given by

$$\tilde{\psi}(\boldsymbol{\theta}; \hat{\boldsymbol{\rho}}, \hat{\boldsymbol{\gamma}}_{-h,t}, y_{j,t}) := \left[\sigma(\boldsymbol{U}_h(\boldsymbol{\theta}; \hat{\boldsymbol{\rho}}, \hat{\boldsymbol{\gamma}}_{-h,t}))_j\right]^{y_{j,t}} \left(1 - \sigma(\boldsymbol{U}_h(\boldsymbol{\theta}; \hat{\boldsymbol{\rho}}, \hat{\boldsymbol{\gamma}}_{-h,t}))_j\right)^{1-y_{j,t}}$$

For each observed network t, there is a corresponding first-mover h and, for all disputes that become wars, observed combatants. The combatants include the observed target ℓ and countries that join, corresponding to a known vector of attack decisions $\boldsymbol{y}_t$ and an n-vector of actions $\boldsymbol{a}_t$, respectively. The estimator can then be defined as the solution to the pseudo log-likelihood function,

$$\hat{\boldsymbol{\theta}} = \arg\max_{\boldsymbol{\theta}} \sum_t \sum_{j \in R_{h,t}} \Big(\ln(\tilde{\psi}(\boldsymbol{\theta}; \hat{\boldsymbol{\rho}}, \hat{\boldsymbol{\gamma}}^{h,j}_{-h,t}, y_{j,t})) + y_{j,t} \sum_{i \in N_t} \ln(\psi(\boldsymbol{\theta}; \hat{\boldsymbol{\rho}}, \hat{\boldsymbol{\gamma}}^{h,\ell}_{-i,t}, a_{i,t})) \Big). \tag{10}$$

Estimator (10) has a natural interpretation: the best estimates for the structural parameters are those that maximize the likelihood of having observed the data. In particular, they are the parameters that simultaneously explain the observed decisions to start wars and the observed decisions to join wars.

4 Results

Table 1 presents the structural estimation results for each time period. Consistently, the value of winning is small compared to the cost of losing, reflecting the fact that countries typically do not start wars and potential joiners stay out of the war more often than not.

There are several patterns in the value of war estimates worth discussing. First is the abrupt shift for polity value and population in the 1989-2014 time period. For the 19th and most of the 20th century, larger and more democratic countries placed greater value on winning wars. In the post-Cold War period, on the other hand, smaller and less democratic countries place greater value on winning wars. This is likely in part the result of proxy warfare and nonmilitaristic methods of intervention that popularized among large democratic countries during this period (Berman and Lake 2019). Estimates in the final period therefore reflect this substantively meaningful change to the way certain types of countries choose to engage in war: from the sidelines.

Additionally, the value of war is increasing in polity distance between sides during the 1816-1869 and 1989-2014 periods, and decreasing otherwise. This aligns with what we would expect given the militarized interstate dispute data during the post-Napoleonic period. In particular, not only are disputes of this time that are included in the data more likely to involve European countries, but this period is marked by the Concert of Europe, where the monarchical powers collaborated to maintain peace and stability following the Napoleonic Wars. The results therefore suggest that the Concert of Europe was not a diplomatic facade, but that European monarchies of the time were genuinely not interested in fighting with each other (Taylor 1954).

Another interesting trend is the changes in the extent to which the value of war is a

Term	Parameter	Estimates			
		1816-1869	1870-1913	1914-1988	1989-2014
Value of winning	$\boldsymbol{\delta}$				
Mean value		0.461	0.004	0.609	0.334
		(0.062)	(0.163)	(0.034)	(0.210)
Polity value		0.245	0.429	0.419	−0.331
		(0.431)	(0.275)	(0.447)	(0.896)
Polity distance		0.354	−0.928	−0.370	0.723
		(1.449)	(0.482)	(1.240)	(3.612)
Population		0.639	0.003	0.501	−0.203
		(0.125)	(0.175)	(0.022)	(0.267)
Coalition size		−0.003	0.624	0.116	0.441
		(0.129)	(0.191)	(0.030)	(0.236)
National wealth		−0.001	0.030	0.045	0.032
		(0.007)	(0.006)	(0.002)	(0.000)
Public good for alliance	β	0.791	0.181	0.205	0.049
		(0.088)	(0.064)	(0.002)	(0.022)
Cost of losing	κ	9.095	18.992	2.145	4.965
		(0.079)	(0.161)	(0.058)	(0.061)
Value of peace	α	1.865	6.930	19.703	0.114
		(0.251)	(0.255)	(1.310)	(0.066)
Network	$\boldsymbol{\Phi}$				
Mean weight on allies		0.265	0.309	0.212	−0.267
		(0.005)	(0.006)	(0.001)	(0.000)
Mean weight on rivals		−0.182	−0.058	−0.232	−0.034
		(0.004)	(0.133)	(0.001)	(0.008)
Ally exports by wealth		0.233	0.289	−0.021	−0.104
		(0.470)	(0.302)	(0.093)	(0.005)
Ally imports by wealth		0.306	−0.360	0.258	−0.374
		(1.674)	(0.761)	(0.077)	(0.033)
Rival exports by wealth		0.202	0.276	−0.188	−0.135
		(1.627)	(0.227)	(0.034)	(0.041)
Rival imports by wealth		0.010	−0.468	0.471	−0.205
		(0.890)	(0.273)	(0.048)	(0.050)

Table 1: Structural parameter estimates $\hat{\boldsymbol{\theta}}$. Asymptotic standard errors in parentheses are adjusted to incorporate first-stage uncertainty, refer to Appendix C.5.

public good for the alliance or a private good for the winning combatants. There is a sharp decline from the post-Napoleonic period, where the value of war is typically a public good to the alliance, to the post-Cold War period, where winning a war yields virtually no public value to nonparticipating allies.

This feature goes hand in hand with the estimates on the network. During the post-Napoleonic period, network spillovers are positive across allies and trading partners and negative with rivals. However, by the post-Cold War period, all spillovers are negative, including those between allies. In fact, the spillovers between allies are even more negative than those between rivals.

This demonstrates how the nature of conflict has changed since the 19th century. When countries face a decision to start or join a war, other countries also have preferences over that country's decision. The negative coefficients should be interpreted as the extent to which those preferences align. In other words, in the post-Napoleonic period, countries typically act in a way that corresponds to how other countries wish they would act, unless they are rivals. By the post-Cold War period, countries typically act in a way that no other country wishes they would, especially their allies. This is directly a consequence of wars being fought over goods with almost exclusively private value. After 1989, countries join wars only when it is in their private interests, without regard for whether it harms others, even those who share their side in the conflict.

4.1 Network Effects

This section investigates the effects of the network on behavior. In particular, I recompute best responses when all countries place a weight of zero on each other's welfare (i.e., $\phi_{ij} = 0$ for all $i, j \in N$) and compare the resulting join probabilities to those with the estimated network.[7] Then, using the new join probabilities under the disconnected graph, I recompute

[7] I focus on direct effects for joining because they are significantly more costly to compute and trials indicate that additional equilibrium effects are marginal at the joining stage when there is no network, making the direct effects approximately equivalent to the total effects.

attack probabilities and compare those to those with the original estimated network, as well. Recall that $\hat{\boldsymbol{\theta}} = (\hat{\boldsymbol{\delta}}, \hat{\beta}, \hat{\kappa}, \hat{\alpha}, \hat{\boldsymbol{\Phi}})$ and define $\boldsymbol{\vartheta} := (\hat{\boldsymbol{\delta}}, \hat{\beta}, \hat{\kappa}, \hat{\alpha}, \mathbf{0})$ so that $\boldsymbol{\vartheta}$ gives the collection of estimated parameters with the edge weights being replaced by a matrix of zeros, reflecting a completely disconnected graph. Then, the direct effects of the network on joining and attacking are formally given by

$$\gamma_i(\hat{\boldsymbol{\gamma}}_{-i}; \hat{\boldsymbol{\theta}}) - \gamma_i(\hat{\boldsymbol{\gamma}}_{-i}; \boldsymbol{\vartheta}) \qquad \text{and} \qquad \pi_h^\ell(\hat{\boldsymbol{\gamma}}(\hat{\boldsymbol{\gamma}}_{-i}; \hat{\boldsymbol{\theta}}); \hat{\boldsymbol{\theta}}) - \pi_h^\ell(\hat{\boldsymbol{\gamma}}(\hat{\boldsymbol{\gamma}}_{-i}; \boldsymbol{\vartheta}); \boldsymbol{\vartheta})$$

respectively.

Table 2 displays the results for each time period. Typically, the network pacifies joiners. In the post-Napoleonic and pre-WWI periods, the average direct effect on joining is positive due to outliers: a few countries experience extremely large positive effects as a result of the network, while most experience smaller negative effects. Approximately 95 percent of all countries experience a negative effect in these periods. In the post-Cold War period, the network reduces the join probability for 99.8 percent of countries.

In part a consequence of the network pacifying joiners, countries are now more willing to launch initial strikes. For all periods except the post-Cold War, the majority of countries attack with greater likelihood as a consequence of the network. Attack probabilities are typically very small when countries are completely disconnected, and therefore the effect of the network in terms of percent change is substantial. The post-Cold War period is once again distinct from the other periods, as the network decreases attack probabilities in addition to join probabilities. The conventional wisdom that integration pacifies holds in the post-Cold War period, but for the opposite reasoning: peace is not the result of countries liking their rivals more, but their allies less.

Table 2: Effects of integration

		Join Probability		
Period	Avg. Effect	Std. Dev.	Pct. Positive	Avg. Pct. Change
1816-1869	0.009	0.084	0.041	−0.742
1870-1913	0.006	0.037	0.055	−0.786
1914-1988	−0.109	0.054	0.020	−0.914
1989-2014	−0.005	0.007	0.002	−0.485
		Attack Probability		
Period	Avg. Effect	Std. Dev.	Pct. Positive	Avg. Pct. Change
1816-1869	0.209	0.380	0.927	6.59e9
1870-1913	0.158	0.346	0.761	1.18e8
1914-1988	0.219	0.367	0.766	5.16e8
1989-2014	−0.002	0.009	0.018	−0.386

4.2 Power and Stability in Equilibrium

A central enterprise of international relations has been to understand the relationship between the balance of power and systemic stability. The results suggest that, over the past two centuries, power imbalances have been associated with less stable systems.

To clarify, *power* and *stability* need to be explicitly defined. The model provides two natural, endogenous measures of these concepts. First, the model provides a measure of stability in equilibrium: the ex-ante size of war. This is simply the number of countries that are expected to fight in a given system, prior to any specific realization. Any country can be recognized as a first-mover and, conditional on being recognized, may start a war with any of their rivals. For any such first-mover and any subsequent target, all remaining countries may join or stay out of the war. Aggregating over all of these possibilities according to equilibrium strategies yields the answer to how large we expect conflict to be, unconditionally. Formally, the ex-ante expected size of conflict is given by $\mathbb{E}\big[\sum_i a_i\big]$ or equivalently,

$$\sum_{h\in N}\lambda_h\sum_{\ell\in R_h}\pi_h^\ell\sum_{C\in\mathcal{C}^{h,\ell}}\prod_{r\in C}\gamma_r^{h,\ell}\prod_{k\notin C}(1-\gamma_k^{h,\ell})\,|C|. \tag{11}$$

Second, the model provides a measure of power in equilibrium: the extent to which a country can swing the odds of victory in their side's favor. War is modeled as a lottery and, by opting into or out of the conflict, countries affect which lottery takes place. Though it might be tempting to assert that the "strongest" countries (e.g., in terms of military capabilities) have the greatest ability to swing the odds in their favor, it may be the case that, due to the equilibrium behavior of others, the circumstances in which they would be able to affect the outcome rarely occurs.

For example, consider two international systems, 1 and 2, that are identical other than their strongest countries, A and B, respectively. Suppose A is stronger than B in terms of raw capacity so that if they were to fight the same opponents, A is more likely to succeed than B. Though seemingly reasonable, it would be incorrect to conclude that system 2 is more balanced than system 1. If B's presence in system 2 leads to different equilibrium behavior than that in system 1 (e.g., countries are less likely to join wars), it remains possible that A is less able to swing the odds than B. This is because the scale of the "expected war" that A and B would enter varies according to equilibrium strategies of the countries in their respective systems. Then, whether system 1 or 2 is more balanced in terms of each member's ability to swing the outcome depends crucially on equilibrium behavior of all countries in the system.

Then, taking power as the ability of a country to move to needle in their side's favor, one way to pin down power disparity in a system is to look at the net balance of those abilities. Recalling that $\rho(\boldsymbol{a})$ yields the probability of initiating side victory given action profile $\boldsymbol{a}$, an endogenous measure of how power is distributed in a system can be given by the equilibrium power advantage of the initiating side:

$$\sum_{i \in N} \mathbb{E}\big[\rho(a_i = 1, \boldsymbol{a}_{-i}) - \rho(a_i = 0, \boldsymbol{a}_{-i})\big] \tag{12}$$

where the expectation is formed similarly to quantity (11), relying on equilibrium behavior. When this quantity is close to zero, each side has roughly the same total ability to move the needle. The measure is also increasing in the ability of the initiating side to affect war outcomes, and decreasing in that of the target side, with positive values denoting a net advantage to the initiating side and negative values a net advantage to the target side. Hence, squaring or taking the absolute value of quantity (12) provides a measure of general power inequality.

Recall that, in principle, the model allows for a variety of relations to manifest between these two measures, depending on the configuration of structural parameters. The question, then, is: what happens at the parameters implied by the data?

Figure 5 demonstrates two empirical approaches to understand the relationship between power and stability. The first naive approach ignores equilibrium behavior by simply regressing the annual variance of military capabilities, as measured by CINC scores, on the observed number of combatants in wars from 1816 to 2014. As the left plot displays, there is no discernible relationship between the two, with a linear regression returning a negative and insignificant coefficient. The right plot, however, demonstrates the empirical relationship between quantities (11) and the square of (12) for each network, recovered from the same data. By looking at endogenous measures that account for behavior, a positive and statistically significant relationship is realized between the balance of power and systemic stability in equilibrium.

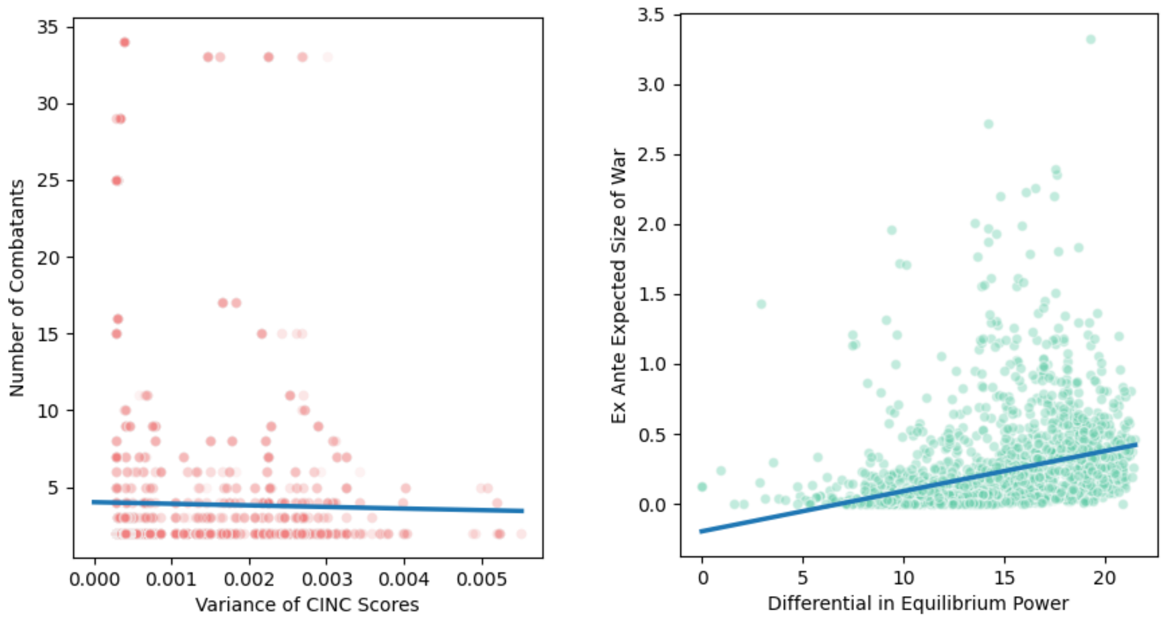


Figure 5: A naive approach (left) versus model equilibrium behavior (right)

4.2.1 Balancing or Bandwagoning?

There is a strong relationship between the unconditional expected size of war and power inequality, as measured by the differential in the aggregate ability of sides to swing the outcome in equilibrium. However, this alone does not explain the behavioral causes of this relationship. We may see this result as a product of *balancing*—countries on the weaker side join with greater frequency the more disadvantaged they are—or *bandwagoning*—countries on the stronger side join with greater frequency the more advantaged they are. A benefit of the structural approach in this paper is that we can investigate the theoretical model using the empirical estimates to discover which forces are at play in the data. Doing so provides strong evidence for bandwagoning.

Table 3 shows that the correlation between power inequality and the expected size of wars is driven by bandwagoning. Countries are more likely to join wars if the power balance favors their side, where power balance is given by quantity (12). As the findings demonstrate, the likelihood that allies of the initiator join the war is increasing in the power advantage of the

Table 3: Balancing or bandwagoning?

	Dependent Variable: Join Probability	
	Initiator Allies	Target Allies
Power Balance	5.886	−1.056
	(1.330)	(0.103)
Constant	0.187	0.059
	(0.009)	(0.002)
Observations	871	4,258
R^2	0.052	0.018
Adjusted R^2	0.051	0.018
Residual Std. Error	0.244	0.128
F Statistic	47.615	79.362

Note: Power Balance is given by quantity (12). Larger amounts reflect a power advantage for the initiating side, while smaller amounts reflect a power advantage for the target side.

initiating side, while the likelihood that allies of the target join the war is decreasing in the power advantage of the initiating side.

The result is intuitive. A country's ability to join a war is conditional on the outbreak of that war to begin with. In turn, initiators are more likely to launch an attack when their side is more powerful and they expect many of their allies will join to support them. Thus, the wars that would lead to balancing behavior instead of bandwagoning behavior are less likely to occur in equilibrium, not necessarily because countries would never join with the intention of balancing, but because, if any such wars did exist, initiators would not have incentive to start them in the first place. Similarly, while target allies also join with greater frequency when their side is more powerful, the effect is mitigated by discriminating initiators that consider the downstream effects when deciding whether to launch wars.

5 Counterfactuals

A key benefit of the structural approach is the ability to conduct counterfactual experiments across new settings with different behavioral implications. This would not be possible without a theory of behavior underlying the model specification, as estimating relationships between variables—causal or otherwise—does not provide an understanding of how behavior will respond to the new setting (Lucas 1976; Blundell 2017).

In this section, I explore what the model implies about the potential for war between China and Taiwan over the post-Cold War period. China has been ramping up its military presence in proximity of Taiwan, violating informal boundaries at the median line in the Taiwan Strait and intruding on the Taiwan Air Defense Identification Zone at much greater frequency than previous years.[8] If China goes to war with Taiwan, will other countries join the fight and, if so, who? Given the likelihood of others intervening, what does that imply about the likelihood that China attacks Taiwan? How do these probabilities change if we change the international setting? For example, how does the likelihood and expected scale of war between China and Taiwan change in counterfactual scenarios with less international trade? The model provides a quantitative answer to these questions.

To demonstrate, I take an international setting constituted by China, Japan, Russia, Taiwan, and the United States. Here, I impose that Russia is on the side of China and the U.S. and Japan are on the side of Taiwan. I use the estimates of structural parameters from the post-Cold War period to inform country preferences and the underlying network. Given this, we can consider the circumstance in which China has attacked Taiwan and ask: which countries will join?

Figure 6 displays computed equilibrium join probabilities $\gamma^{\text{China,Taiwan}}$ for each country-year during the post-Cold War period (1989-2014). The probabilities are small, never exceeding half of a percent, which reflects the empirical fact that most countries do not join

[8] Chris Buckley and Amy Chang Chien, "China's Military, 'Chasing the Dream,' Probes Taiwan's Defenses," *The New York Times*, 11 August 2023. https://www.nytimes.com/2023/08/11/world/asia/china-taiwan-military.html (Accessed 11 August 2023).

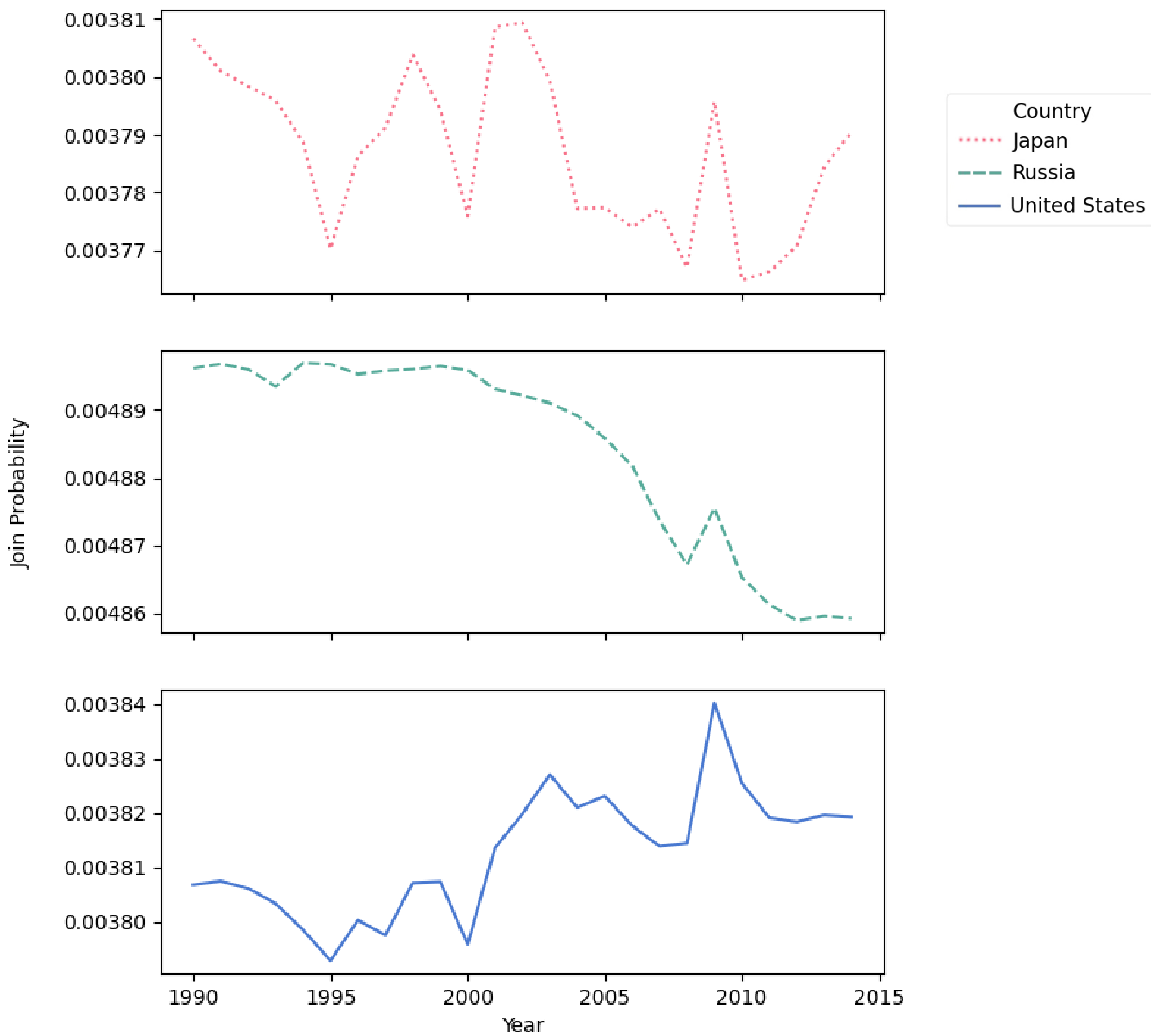


Figure 6: Equilibrium join probabilities $\gamma^{\text{China,Taiwan}}$, 1989–2014

wars most of the time. Moreover, it is even less common for major powers like the United States to intervene with military in a conflict happening across the world, where another major power is attacking a small contiguous country.

Over the period, Japan's willingness to join fluctuates, while the U.S. is generally becoming more willing and Russia is generally becoming less willing. The figure demonstrates that incentives of aligned countries do not always move in tandem. While the U.S. and Japan's willingness typically move together, with the U.S. trending up at greater rate, there are also moments where they move in opposite directions. For example, just after 2010, Japan's willingness is rising and the U.S.'s is falling, while the U.S.'s willingness is rising while Japan's is falling around 1997.

Additionally, movements in join probabilities across sides are not straightforward. There are periods of time where countries across sides are moving in opposite directions, such as the years following 2000, where the U.S. and Japan enter with more frequency and Russia is increasingly likely to stay out. However, there are also periods of time where all three countries move together. There is a noticeable peak around 2008 where all countries' join probabilities spike in response to a large contraction in international trade due to heightened protectionism during the global financial crisis.

Given equilibrium join probabilities for all country-years with respect to a war between China and Taiwan, we can take the next step and compute equilibrium attack probabilities for China. Figure 7 demonstrates the results for each year, which indicates China is less likely to attack Taiwan over time. This matches our intuition: since they are increasingly likely to face opposition from the United States and cannot count on support from Russia, the prospects of initiating a war over Taiwan look worse each year. There are, however, some exceptions. Most notably, there is again a spike in the willingness to fight in 2008, where not only does China have a heightened direct incentive to start a war in response to the financial crisis, but they additionally have a heightened indirect incentive to start a war in response to the increase to Russia's join probability, which outweighs the increases in join probability

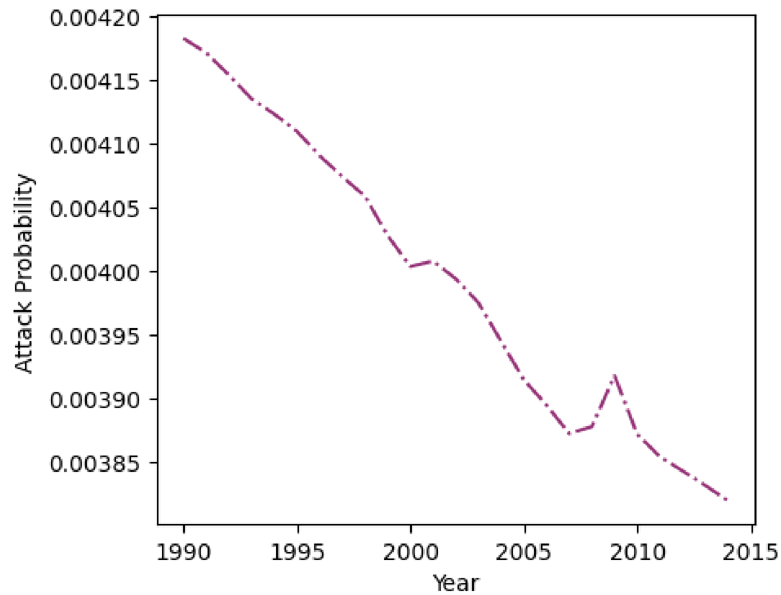


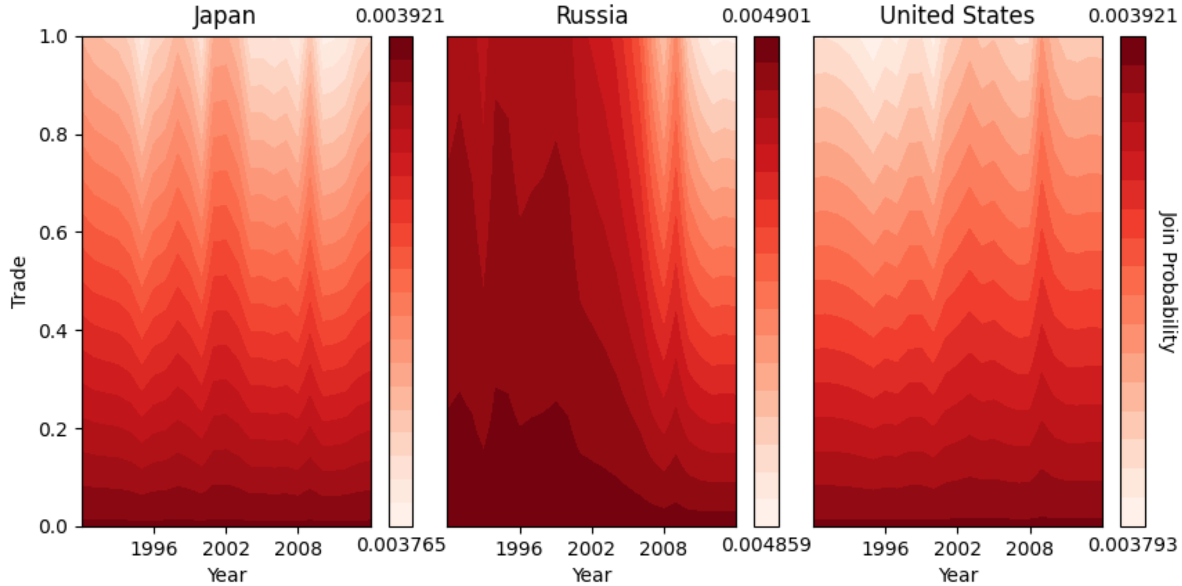
Figure 7: Equilibrium attack probability $\pi_{\text{China}}^{\text{Taiwan}}$, 1989–2014

for the U.S. and Japan.

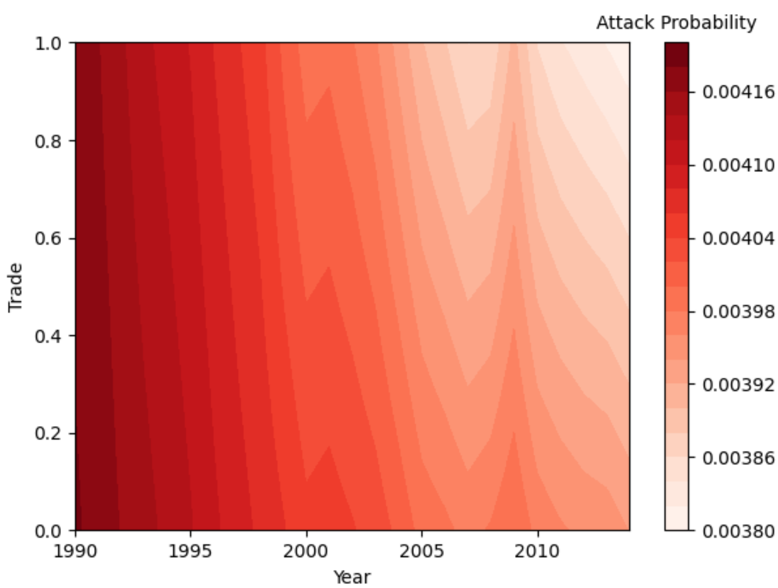
5.1 The Effect of Trade

We can take the counterfactual analysis further and explore how changes in the international setting affect behavior and outcomes. I now reconsider the extent to which trading relations influenced the findings of the previous section. In particular, for each year in the post-Cold War period, I recompute equilibrium strategies over a fine grid of parameters, each step corresponding to fewer trade flows until there is zero trade. The goal is to study how trade affects the propensity for conflict in this network. As we move towards a world where there is no trade, which countries are more willing to fight and which are less? How does that affect China's incentive to launch an attack?

Figure 8a demonstrates the results for the potential joiners. Reducing international trade flows increases the willingness of all countries to engage in war. However, there are several important observations. First, the rate of increase varies across countries and years. In



(a) Equilibrium join probabilities $\gamma^{\text{China,Taiwan}}$



(b) Equilibrium attack probability $\pi^{\text{Taiwan}}_{\text{China}}$

Figure 8: Behavior in trade. Trade equal to 1 reflects trade flows as given by the data while Trade equal to 0 reflects complete autarky.

particular, Russia's join probability only marginally increases as trade reduces to zero in in the first half of the period, until around the mid-2000s. On the other hand, the rate of increase for U.S. and Japan follows a largely comparable trend, peaking at comparable join probabilities under zero trade.

While it appears that trade consistently pacifies the network, the total effect of trade is fairly marginal. Specifically, trade appears to be reducing country join probabilities by approximately one hundredth of a percent. Even when dealing with these fairly low probabilities, that only amounts to several percentage points, at most. The findings, while consistent with previous work that argues trade reduces conflict, simultaneously suggest that the extent to which trade reduces conflict may not be considerable in scale.

Figure 8b shows the recomputed probabilities that China attacks Taiwan under various degrees of trade, further demonstrating the negligible effects of trade on war onset. While China's propensity for conflict is technically increasing as trade flows are reduced to zero, the absolute movement is minimal. In fact, for several years during the first half of the period in which Russia's join probability is relatively unyielding, China's attack probability fails to increase more than one thousandth of a percent.

6 Conclusion

In this paper, I provide a framework that allows us to explore what makes a world war, facilitating theoretical and empirical analysis through a singular lens. The model encompasses key elements of the international system—political alignments and integration—and connects them to the strategic decisions to start and join wars. Crucially, these decisions are shaped by expectations about which coalitions will form and how the spoils and costs of war will propagate through the system. Theoretical objects of interest are quantified by finding the parameter values that best align the model's predictions with observations from international disputes between 1816-2014. I further estimate the effects of integration and power asymmetries on the likelihood and scale of conflict.

The results demonstrate that the threat of systemic instability can be a stabilizing force for the international system. Since countries are less inclined to start wars if they expect to face additional opposition, the systems that are most likely to produce large wars can also be those that are least likely to produce initial attacks. The least stable systems are therefore not only those where potential joiners want to engage in fights, but also those where initiators want to launch attacks despite this fact. Expectations about which coalitions will form are thus essential to the probability and ultimate size of war.

I find that integration typically reduces the incentive to join wars and consequently increases the incentive to initiate them. However, integration reduces both incentives after the Cold War. This is because, while large wars in earlier periods are brought about by the desire to defend allies and economic partners, large wars after 1989 are driven by immediate self-interests. Since wars in the post-Cold War period are fought over goods with little public value to the alliance, self-interest counterintuitively drives the initiating side to join with greater frequency, improving the odds of winning for initiators and encouraging attacks. This is consistent with what we would expect given the dispute data, from the monarchical coordination of the Concert of Europe in the 19th century to the political realignments that occurred with unipolarity after the fall of the Soviet Union. Further, power imbalances correspond to less stable systems in equilibrium not only because power advantaged countries have a greater incentive to join wars, but also because initiators have a greater incentive to start wars when their allies are more likely to join. In all cases, outcomes depend critically on downstream equilibrium effects.

The analysis can help us understand the trajectories of disputes, shedding light on the underlying incentives and the influence of key factors. The complex relationship between power, economics, and war makes the problems facing standard methods particularly acute. An advantage to my approach is that, by explicitly incorporating a strategic model in the empirical analysis to recompute equilibrium behavior in new settings, counterfactual experiments are made possible. I demonstrate the added value of this ability with the potential conflict between China and Taiwan, finding that (i) China is becoming less likely to invade Taiwan as a result of an increasing likelihood of U.S. opposition and a decreasing likelihood of Russian support, and (ii) trade relations pacify but only marginally. The same procedure could be repeated for any system of countries.

In addition to serving as a tool to study international war, this framework can be used to make theoretically driven out-of-sample predictions that can inform policy and institutional design. For example, the model can be used to better understand the consequences of trade policies, military and economic investments, or alliance agreements. Direct quantitative competition of theoretical mechanisms can enable scholars to develop nuanced understandings that rely on objective performance metrics. The iterative process of refining, adapting, and competing these models is a productive path for future work in international conflict and political economy.

References

Barbieri, Katherine. 2002. *The illiberal illusion: Does trade promote peace?* University of Michigan Press.

Barbieri, Katherine, and Omar M. G. Keshk. 2016. Correlates of War Project Trade Data Set Codebook, Version 4.0. Technical report.

Barbieri, Katherine, Omar M. G. Keshk, and Brian Pollins. 2009. "Trading Data: Evaluating Our Assumptions and Coding Rules." *Conflict Management and Peace Science* 26(5):471–491.

Berman, Eli, and David A Lake. 2019. *Proxy wars: Suppressing violence through local agents.* Cornell University Press.

Blainey, Geoffrey. 1973. *The Causes of War.* New York: Free Press.

Blundell, Richard. 2017. "What have we learned from structural models?" *American Economic Review* 107(5):287–292.

Braumoeller, Bear F. 2008. "Systemic Politics and the Origins of Great Power Conflict." *American Political Science Review* 102(1).

Bueno de Mesquita, Bruce, and David Lalman. 1988. "Empirical support for systemic and dyadic explanations of international conflict." *World Politics* 41(1):1–20.

Carter, David B. 2010. "The Strategy of Territorial Conflict." *American Journal of Political Science* 54(4):969–987.

Cooley, Brendan. 2018. Calculating Historical Intercapital Distances. Technical report.

Copeland, Dale C. 1996. "Economic interdependence and war: A theory of trade expectations." *International security* 20(4):5–41.

Copeland, Dale C. 2014. *Economic interdependence and war.* Princeton University Press.

Cranmer, Skyler J, and Bruce A Desmarais. 2016. "A Critique of Dyadic Design." *Interna-*

tional Studies Quarterly 60(2):355–362.

Crisman-Cox, Casey, and Michael Gibilisco. 2018. “Audience Costs and the Dynamics of War and Peace.” *American Journal of Political Science* 62(3):566–580.

Dorussen, Han, Erik A Gartzke, and Oliver Westerwinter. 2016. “Networked international politics: Complex interdependence and the diffusion of conflict and peace.” *Journal of Peace Research* 53(3):283–291.

Feenstra, Robert C., Robert Inklaar, and Marcel P. Timmer. 2015. “The Next Generation of the Penn World Table.” *The American Economic Review* 105(10):3150–82.

Gibilisco, Michael, and Sergio Montero. 2022. “Do Major-Power Interventions Encourage the Onset of Civil Conflict? A Structural Analysis.” *The Journal of Politics* 84(2):944–959.

Gibler, Douglas M. 2009. “International Military Alliances, 1648-2008.” *Congressional Quarterly Press* .

Gilpin, Robert. 1981. *War and Change in World Politics*. Cambridge University Press.

Hegre, Håvard, John R Oneal, and Bruce M Russett. 2010. “Trade Does Promote Peace: New Simultaneous Estimates of the Reciprocal Effects of Trade and Conflict.” *Journal of Peace Research* 47(6):763–774.

Hensel, Paul R. 2016. ICOW Historical State Names Data Set, version 1.3. Technical report.

Honaker, James, Gary King, and Matthew Blackwell. 2011. “Amelia II: A Program For Missing Data.” *Journal of Statistical Software* 45:1–47.

Hotz, V. Joseph, and Robert A Miller. 1993. “Conditional choice probabilities and the estimation of dynamic models.” *The Review of Economic Studies* 60(3):497–529.

Iaryczower, Matias, and Matthew Shum. 2012. “The Value of Information in the Court: Get It Right, Keep It Tight.” *American Economic Review* 102(1):202–37.

Jackson, Matthew O, and Stephen Nei. 2015. “Networks of military alliances, wars, and international trade.” *Proceedings of the National Academy of Sciences* 112(50):15277–15284.

Kenkel, Brenton, and Kristopher W Ramsay. 2023. "The Effective Power of Military Coalitions: A Unified Theoretical and Empirical Model." *The Journal of Politics* (Forthcoming).

Kim, Woosang. 1989. "Power, alliance, and major wars, 1816-1975." *Journal of Conflict Resolution* 33(2):255–273.

Kindleberger, Charles P. 1986. *The World in Depression, 1929-1939*. Vol. 4 University of California Press.

Kinne, Brandon J. 2012. "Multilateral Trade and Militarized Conflict: Centrality, Openness, and Asymmetry in the Global Trade Network." *The Journal of Politics* 74(1):308–322.

Kleinman, Benny, Ernest Liu, and Stephen J Redding. 2024. "International Friends and Enemies." *American Economic Journal: Macroeconomics* (Forthcoming).

König, Michael D, Dominic Rohner, Mathias Thoenig, and Fabrizio Zilibotti. 2017. "Networks in conflict: Theory and evidence from the Great War of Africa." *Econometrica* 85(4):1093–1132.

Lebow, Richard Ned. 1984. "Windows of opportunity: do states jump through them?" *International Security* 9(1):147–186.

Lemke, Douglas, and William Reed. 2001. "The relevance of politically relevant dyads." *Journal of Conflict Resolution* 45(1):126–144.

Lewis, Jeffrey B, and Kenneth A Schultz. 2003. "Revealing preferences: Empirical estimation of a crisis bargaining game with incomplete information." *Political Analysis* 11(4):345–367.

Lucas, Robert E. 1976. "Econometric Policy Evaluation: A Critique." *Carnegie-Rochester Conference Series on Public Policy* 1:19–46.

Lupu, Yonathan, and Vincent A Traag. 2013. "Trading communities, the networked structure of international relations, and the Kantian peace." *Journal of Conflict Resolution* 57(6):1011–1042.

Mansfield, Edward D. 1992. "The concentration of capabilities and the onset of war." *Journal* *of Conflict Resolution* 36(1):3–24.

Maoz, Zeev. 1996. *Domestic sources of global change*. University of Michigan Press.

Marshall, Monty G, and Ted Robert Gurr. 2020. Polity5 Project: Political Regime Characteristics and Transitions, 1800-2018. Technical report.

Martin, Philippe, Thierry Mayer, and Mathias Thoenig. 2008. "Make trade not war?" *The Review of Economic Studies* 75(3):865–900.

Miller, Steven V. 2022. "peacesciencer: An R Package for Quantitative Peace Science Research." *Conflict Management and Peace Science* 39(6):755–779.

Montesquieu. 1748. *The Spirit of Law*.

Morgenthau, Hans Joachim, and Kenneth W Thompson. 1985. *Politics Among Nations: The Struggle for Power and Peace*. 6th ed. New York: Alfred A. Knopf.

Olivella, Santiago, Tyler Pratt, and Kosuke Imai. 2022. "Dynamic Stochastic Blockmodel Regression for Network Data: Application to International Militarized Conflicts." *Journal of the American Statistical Association* 117(539):1068–1081.

Organski, Abramo F. K., and Jacek Kugler. 1980. *The War Ledger*. University of Chicago Press.

Paganelli, Maria Pia, and Reinhard Schumacher. 2019. "Do not take peace for granted: Adam Smith's warning on the relation between commerce and war." *Cambridge Journal of Economics* 43(3):785–797.

Palmer, Glenn, Roseanne W McManus, Vito D'Orazio, Michael R Kenwick, Mikaela Karstens, Chase Bloch, Nick Dietrich, Kayla Kahn, Kellan Ritter, and Michael J Soules. 2020. "The MID5 Dataset, 2011-2014: Procedures, Coding Rules, and Description." *Conflict Management and Peace Science* 39(4):470–482.

Roser, Max, Bastian Herre, and Joe Hasell. 2013. "Nuclear Weapons." *Our World in Data* .

Signorino, Curtis S. 1999. "Strategic interaction and the statistical analysis of international conflict." *American Political Science Review* 93(2):279–297.

Silva, J.M.C. Santos, and Silvana Tenreyro. 2006. "The Log of Gravity." *The Review of Economics and Statistics* 88(4):641–658.

Silva, J.M.C. Santos, and Silvana Tenreyro. 2011. "Further Simulation Evidence on the Performance of the Poisson Pseudo-Maximum Likelihood Estimator." *Economics Letters* 112(2):220–222.

Singer, J. David. 1987. "Reconstructing the Correlates of War Dataset on Material Capabilities of States, 1816-1985." *International Interactions* 14:115–32.

Singer, J. David, Stuart Bremer, and John Stuckey. 1972. "Capability Distribution, Uncertainty, and Major Power War, 1820-1965." *Bruce Russett (ed) Peace, War, and Numbers, Beverly Hills: Sage* p. 19–48.

Siverson, Randolph M, and Michael R Tennefoss. 1984. "Power, alliance, and the escalation of international conflict, 1815-1965." *American Political Science Review* 78(4):1057–1069.

Smith, Adam. 1776. *An Inquiry into the Nature and Causes of the Wealth of Nations.* London: W. Strahan and T. Cadell.

State System Membership List, v2016. 2017.

Stinnet, Douglas M, Tir Jaroslav, Philip Schafer, Paul F Diehl, and Charles Gochman. 2002. "The Correlates of War Project Direct Contiguity Data, Version 3." *Conflict Management and Peace Science* 19(2):58–66.

Taylor, A.J.P. 1954. *The struggle for mastery of Europe, 1848-1918.* United Kingdom: Clarendon Press.

Thompson, William R, and David Dreyer. 2012. *Handbook of International Rivalries.* CQ Press.

Thucydides. 1954. *The Peloponnesian War.* R. warner ed. New York: Penguin.

Waltz, Kenneth N. 1964. "The Stability of a Bipolar World." *Daedalus* p. 881–909.

Waltz, Kenneth N. 1979. *Theory of International Politics*. Reading, Massachusetts: Addison-Wesley.

Whang, Taehee, Elena V McLean, and Douglas W Kuberski. 2013. “Coercion, information, and the success of sanction threats.” *American Journal of Political Science* 57(1):65–81.

Xu, Jin, Yves Zenou, and Junjie Zhou. 2022. “Equilibrium Characterization and Shock Propagation in Conflict Networks.” *Journal of Economic Theory* 206:105571.

Appendix for “What Makes a World War? A Structural Analysis of Integration”

Contents

A Additional Model Details

In this section I analyze special cases of the model and provide numerical simulations to demonstrate the model's mechanisms.

A.1 Analysis

Consider the same model as outlined in the main text with the following two assumptions:

(i) each country has a common value for victory of $\delta > 0$ and

(ii) all countries place common weights ϕ on their within-alliance connections and ϕ_* on their cross-alliance connections.

Additionally, since these assumptions render each player on a given side as identical, it is worthwhile to focus on symmetric cases where all countries play the same threshold strategies in equilibrium.

Let each side be denoted by $s \in \{0, 1\}$, so that $-s$ refers to the side other than s, and n_s the number of countries on side s. Then, it is straightforward to define the probability of realizing a war with (c_s, c_{-s}) joining combatants as

$$\Gamma_s(c_1, c_2) := \binom{n_s - 2}{c_s}\binom{n_{-s} - 1}{c_{-s}} \gamma_s^{c_s}(1-\gamma_s)^{n_s-2-c_s}\gamma_{-s}^{c_{-s}}(1-\gamma_{-s})^{n_{-s}-1-c_{-s}}.$$

where γ_s gives the probability a country on side s joins the war in equilibrium.

Next, let $p_s(c_1, c_2)$ be the probability of side s winning given c_1 joining combatants for side 1 and c_2 for 2. Given a set of combatants fighting the war, the expected surplus of being a joiner on side s can then be expressed

$$\Lambda_s = \frac{1}{1+\phi}\sum_{c_s}\sum_{c_{-s}}\Gamma_s(c_s, c_{-s}) \cdot \Big[\delta(1-\beta)p_s(c_s, c_{-s}) - \kappa(1 - p_s(c_s, c_{-s}))\Big]. \tag{A1}$$

Finally, denote the weighted total probability of victory for side $s \in \{1, 2\}$ by

$$\Sigma_s := \frac{1}{1+\phi}\sum_{c_s}\sum_{c_{-s}}\Gamma_s(c_s, c_{-s})p_s(c_s, c_{-s}).$$

Then, we can conclude that

(i) $\frac{\partial \Lambda_s}{\partial \delta} = (1-\beta)\Sigma_s > 0$, i.e., the expected surplus of being a joiner is increasing in the value of winning;

(ii) $\frac{\partial \Lambda_s}{\partial \beta} = -\delta\Sigma_s < 0$, i.e., the expected surplus of being a joiner is decreasing in the premium for active combatants; and

(iii) $\frac{\partial \Lambda_s}{\partial \kappa} = \frac{1}{1+\phi}(\Sigma_s - 1) < 0$, i.e., the expected surplus of being a joiner is decreasing in the costs of losing.[9]

[9] While $\frac{1}{1+\phi}(\Sigma_s - 1) > 0$ when $\phi < -1$, this possibility is ruled out by the need for $\varrho(\boldsymbol{\Phi}) < 1$. Beyond the technical constraint, this circumstance is also empirically implausible as it implies such large negative spillovers from allies that countries prefer losing to winning.

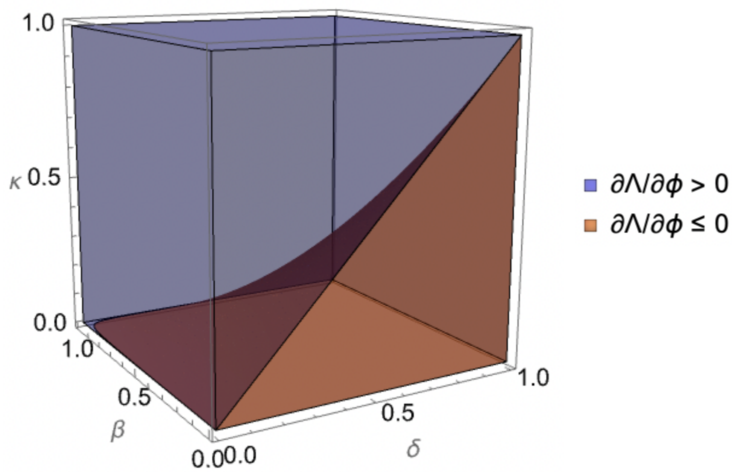


Figure 9: Expected surplus of joiner from within-alliance integration, $\partial\Lambda/\partial\phi$

As δ, β, and κ are all constants in this special case, the result can be attained by simply taking the first derivative of equation (A1) with respect to the parameter of interest.

Additionally, we can conclude that the joiner's expected surplus is decreasing in within-alliance integration when $\delta(1-\beta)$ is large and either κ is small or ϕ is large, and increasing otherwise. Formally,

$$\frac{\partial\Lambda_s}{\partial\phi} = \frac{1}{1+\phi}\left(\frac{\kappa}{1+\phi}(1-\Sigma_s) - \delta(1-\beta)\Sigma_s\right). \tag{A2}$$

The result can be attained by simply taking the first derivative of equation (A1) with respect to ϕ. When network connections are close to zero, this term approximates the expected loss from joining (i.e., the cost of losing κ times the total probability of losing, less the additional gain from participating and winning $\delta(1-\beta)$ times the total probability of winning).

As equation (A2) and Figure 9 demonstrate, it is not straightforward to understand how a joiner's expect surplus changes as they become more connected to their allies. When the gains of victory are contingent on fighting and the penalty of losing is small, joining countries' expected surplus will reduce as alliances becomes more integrated. This may seem counterintuitive, but has a straightforward reasoning: countries that are more disconnected from their allies will have greater incentive to join wars with large contingent payoffs and low costs, but this incentive is mitigated by positive spillover effects as countries integrate with their allies. Conversely, when the premium for fighting is small and costs of losing are large, disconnected countries will do better staying out of wars and, as allies integrate, countries face a greater incentive to join, reducing the probability of their alliance's military loss and the corresponding negative spillovers.

While the expected surplus of joining Λ_i appears to move in a sensible way, it is not the same as showing that the threshold on join probabilities, Ω_i, moves sensibly in parameters. Importantly, the expression for total expected surplus takes as given the number of combatants, therefore, it is correctly understood as the excess quantity that joiners can expect to receive over those who do not join. It is not the additional amount that an individual country would expect to get from joining, which governs their probability of joining in equilibrium.

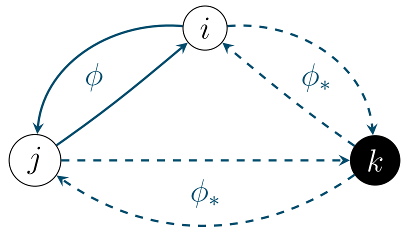


Figure 10: Special case with $n = 3$

Unfortunately, even in this simplified setting, characterizing conditions for a positive or negative sign for the derivative of Ω_i is complicated. Let us consider a further simplification of the game such that $n = 3$. Figure 10 provides an illustration of the network where, necessarily, one side has one country and the other side has two. Sides are represented by hollow and solid nodes, with countries i and j on side 1 and country k on side 2. Dashed arrows reflect cross-alliance relations by ϕ_* and solid arrows reflect within-alliance relations by ϕ. This setting has the benefit of allowing us to avoid complicated combinatorics of different country pairs starting different originating wars with different combinations of combatants, each of which require an understanding of expected outcomes in order to aggregate over them according to equilibrium behavior. Instead, with this setting, we can focus without loss of generality of the sole potential joiner on the majority side.

Since now there is only one potential joiner, the war can only be one of two lotteries: the one governed by p_L where the joiner stays out, and the one governed by $p_H > p_L$ where the joiner enters (where, without loss of generality, p_L and p_H give the probability of the joiner's side winning). Then, the threshold for joining can be given by

$$\Omega = \frac{(1-\phi_*)(p_H(\delta+\kappa)(1+\phi) - \kappa(1+\phi_*) - p_L(\delta\beta(1+\phi_*) + (\delta+\kappa)(\phi-\phi_*)))}{(1+\phi)(1-\phi-2\phi_*^2)} \tag{A3}$$

which in turn implies that

(i) $\frac{\partial\Omega}{\partial\delta} > 0$, i.e., the propensity to join wars increases in the war prize;

(ii) $\frac{\partial\Omega}{\partial\beta} < 0$, i.e., the propensity to join wars decreases in the amount retained by passive winners; and

(iii) $\frac{\partial\Omega}{\partial\kappa} > 0$ if and only if $\phi > \phi_*$, $p_H > \frac{1+\phi_*}{1+\phi}$, and $p_L < \frac{p_H(1+\phi)-1-\phi_*}{\phi-\phi_*}$.

The first two points are intuitive; however, the third is not. If allies are more integrated than rivals, the remaining country may be more likely to join as the penalty increases if their ability to swing the needle on winning is sufficiently large. This is because, as losing wars becomes very costly, a passive country will nonetheless pay this cost indirectly through a loss by their ally. Therefore, the remaining country is tempted to join specifically to prevent their ally's military loss.

Even in this simple setting with only three countries, the effect of the network on equilibrium behavior is highly contingent. Figure 12 demonstrates that increases in within- and cross-alliance integration can both encourage the third country to join more or less. The

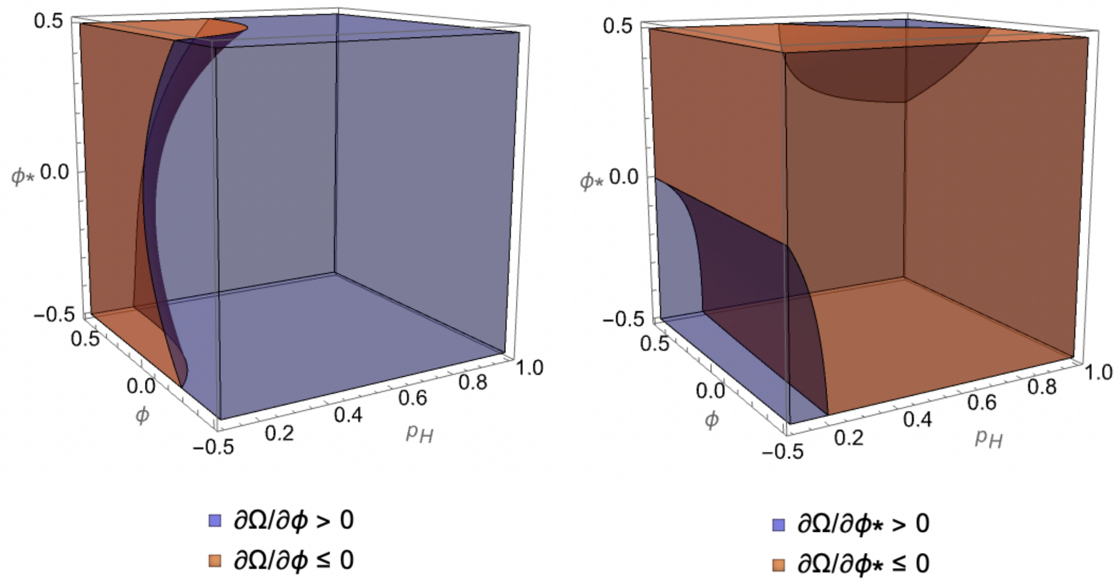


Figure 11: Probability of joining a war in integration

left panel shows how the probability of joining a war changes in within-alliance integration $\partial\Omega/\partial\phi$, while the right panel shows how it changes in cross-alliance integration $\partial\Omega/\partial\phi_*$. Both were computed with $\delta = \kappa = 1$, $\beta = \frac{1}{2}$, and with $p_L = \frac{1}{100}$ to extend the range of p_H.

The potential joiner is more likely to join as allies integrate as long as the probability of winning after entering is not very low, in which case spillovers between allies need to be negative (i.e., dislike for allies). On the other hand, integration with rivals tends to reduce join probabilities, unless either (i) the probability of victory when entering is low and cross-alliance spillovers are negative, or (ii) the winning when entering is nearly certain and all countries are strongly integrated.

The next step is to ask, given the equilibrium behavior of the potential joiner, how likely is it that we see attacks in the first place? Denote Ξ_s the equilibrium attack threshold of a country on side s, with Ξ_1 is for the majority side and Ξ_2 is for the minority side, without loss of generality. Figure 12 displays attack probabilities by the majority side (on the left) and the minority side (on the right) as in-group members integrate. In this case, within-alliance integration typically reduces attack probabilities for both sides. When a country is becoming more connected to an ally that could be a potential supporter in the war, they are less likely to initiate war in the first place. Likewise, when two rival countries are becoming more connected, a country is less inclined to attack that rival country. There remain, however, minor exceptions where the opposite is true. Although not displayed here, the effects of cross-alliance integration on attack probabilities look comparable to these figures, being almost entirely negative.

A.2 Simulations

This section presents the results of a variety of numerical simulations of the model that provide intuition on the underlying mechanisms. This approach is helpful since unilateral

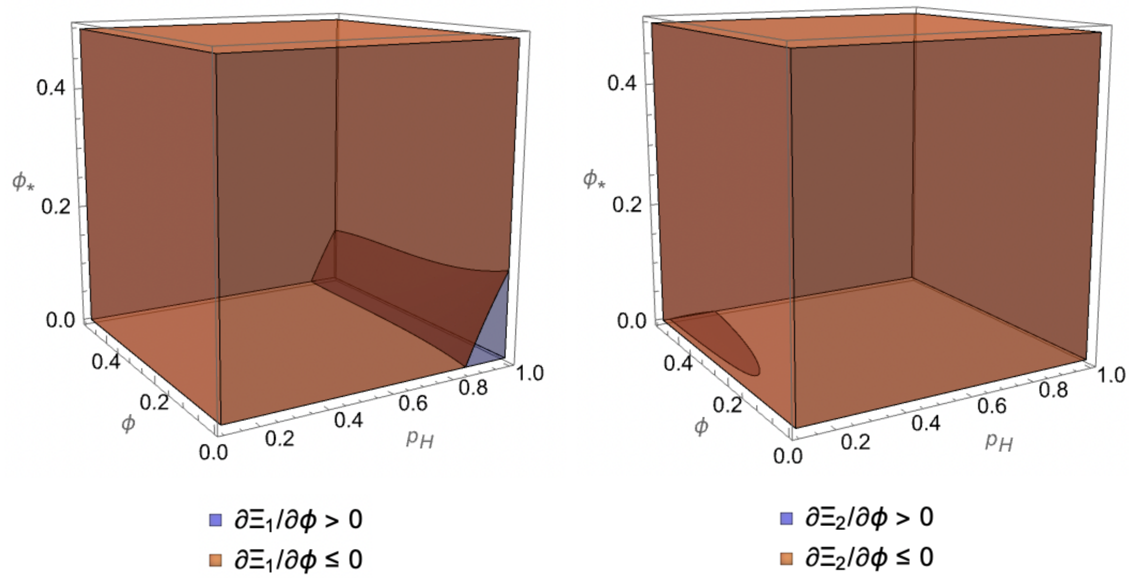


Figure 12: Probability of an initial attack, $\partial\Xi_s/\partial\phi$

changes in parameters will necessarily be difficult, if possible, to interpret analytically for the general model.

For example, consider an increase in δ_i for a particular $i \in N$. The direct effect on $u_i(\cdot)$ is clearly a weak increase, holding behavior constant. However, a true understanding of how increases in δ_i changes the interaction requires understanding not just the direct effect holding behavior constant, but also the (recursive) indirect effects on the entire n-vector $\boldsymbol{u}(\cdot)$ under every possible realization of actions and outcomes, implying potentially different behavior in equilibrium and a corresponding change to total expected utility. Simulations are therefore helpful as they allow for exploration of the model's features that would otherwise be incomprehensible analytically.

Simulations, in addition to creating opportunities for discovery on the model's implications, also provide a check that the model is appropriate for the purpose at hand. Behavior that deviates from expectations could mean one of two things: either the model is revealing something about the strategic interaction that is not straightforward, or the model is failing to capture the strategic interaction on the fundamentals. While most simulated results follow intuitively, those that do not reveal interesting behavioral implications.

First, I explore how changes in power results in changes to the equilibrium behavior of potential joiners. To understand the role of power, we need to first provide an explicit mapping from characteristics $\boldsymbol{m}$ (henceforth, military capabilities for simplicity) and actions $\boldsymbol{a}$ to the probability that the initiating side wins $p(\cdot)$. I assume, for the sake of these simulations only, that $M := \mathbb{R}_+$ and that the probability the initiating side s_h wins is given by the ratio of the total military capacities of country h's coalition to total military capacities among all combatants. Formally,

$$p(\boldsymbol{m}, \boldsymbol{a}) = \frac{\sum_{i \in N : s_i = s_h} m_i a_i}{\sum_{k \in N} m_k a_k}.$$

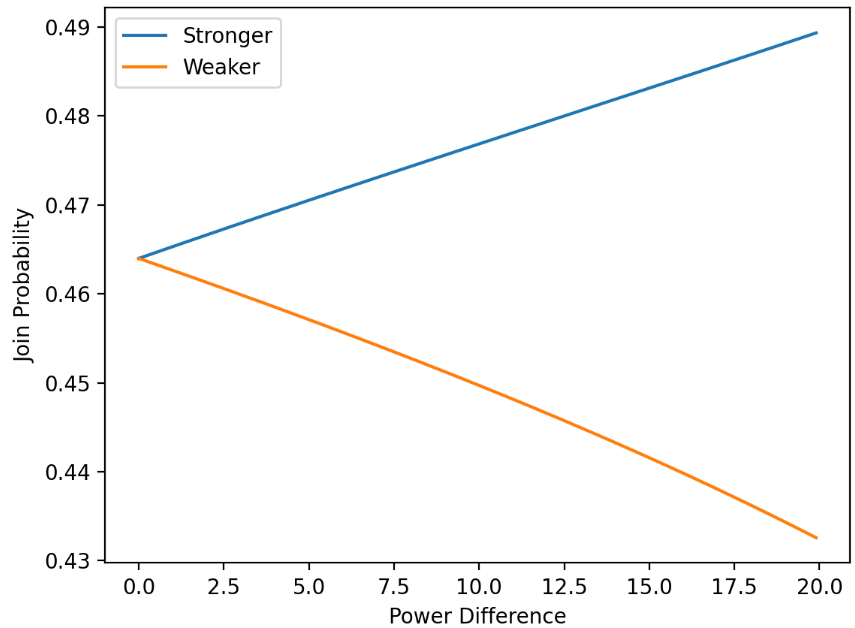


Figure 13: Probability of joining in shifting power

Then, we can impose changes in any particular country's military capacities m_i and assess its affect on their likelihood to enter the fight.

Figure 13 demonstrates a case where the power differential between adversaries facing a decision to join is growing, holding constant the military capacities of the originating couple. As would be expected, the country that is gradually increasing in strength is growing more likely to enter the conflict, as their ability to make a greater difference in attaining a favorable war outcome makes their entry more worthwhile. Conversely, the country that is gradually weakening is becoming less likely to enter, as they are less capable of affecting the outcome and therefore have less incentive to involve themselves. Interestingly, the rate of increase in join probability when gaining strength is not equal to the rate of decline in join probability when losing strength; under the parameterization that generated Figure 13, the weaker country opts out of conflict at a greater rate than the stronger country opts in.

Next, I explore how the extent to which war is public-valued (or analogously the discount on the passive winners of war), β, changes equilibrium behavior. When β is close to 1, the war is public-valued and noncombatants still reap a majority of their winner's reward, while β close to zero implies the war is private-valued and noncombatants get close to none of it. Then, decreases in β should drive countries to fight more, since there is a greater incentive to actively participate with military force. Intuitively, high value countries enter more often than low value countries (given by δ_i) and increases in the noncombatant share drives them both to enter less often, with the rate of decline increasing in the value of winning.

Likewise, the effect of changing a particular country's direct gain from winning conflict (δ_i) produces the expected result of driving the country to join with greater likelihood. More interesting, however, is how countries may respond to changes in *another* country's value, holding their values constant. Figure 15 shows the case where the direct value of victory

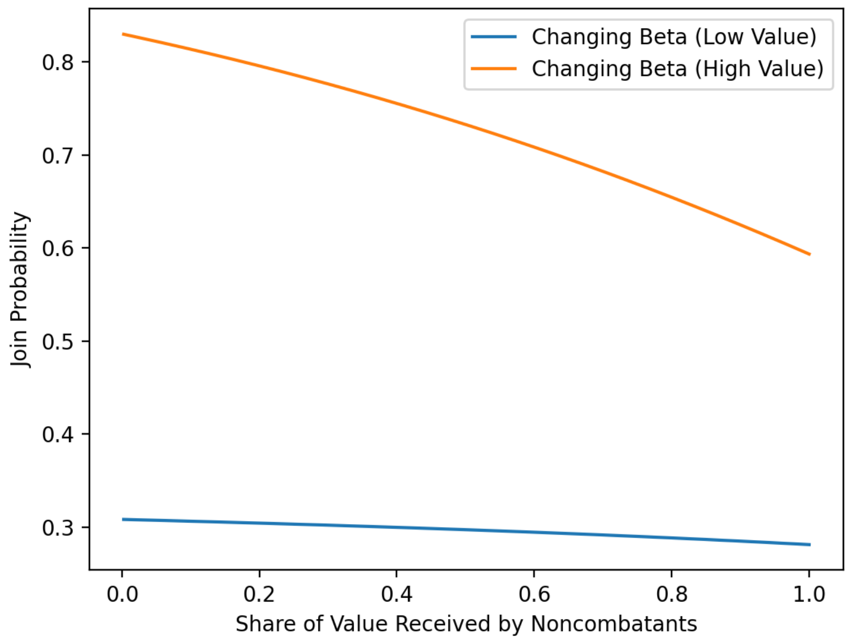


Figure 14: Probability of joining in the discount on passive winners, β

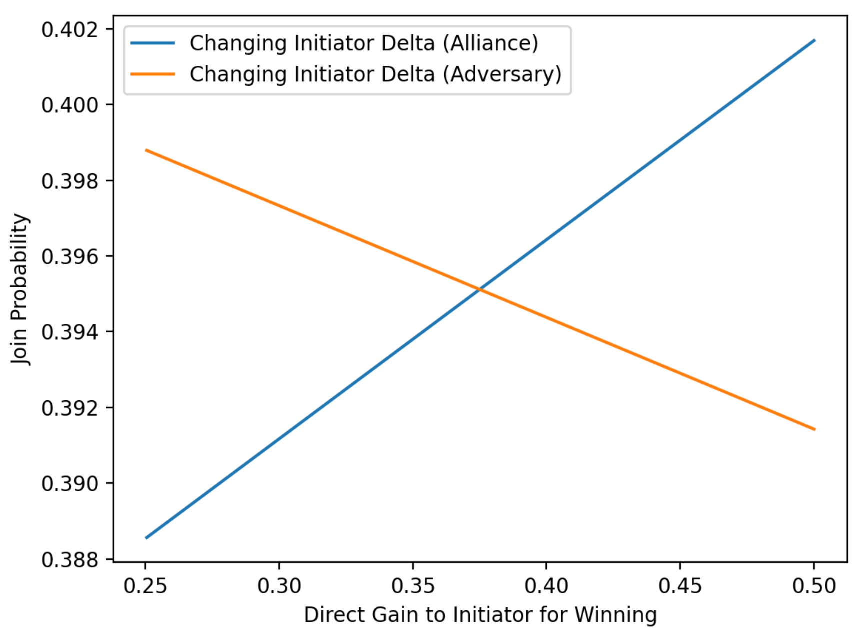


Figure 15: Probability of joining in initiator's direct value for winning, δ_h

is held constant for potential joiners as the initiator h's value of victory δ_h changes. As expected, increases in the initiator's value results in the allies entering more often and the adversaries entering less often.

However, simulations also reveal that the structure of the underlying network plays a significant role in the responsiveness of potential joiners to changes in the direct values of other countries. In particular, if both ally and adversary placed equal weight on the initiator, Figure 15 would look like a perfect "X" shape, with reductions in adversary join probability at the same rate as increases in the ally join probability. Instead, however, by assuming the network is correlated with alliance structures, so that the weight the ally places on the initiator is strictly larger than the weight the adversary places on the initiator, allies of the initiator are more responsive to changes in the initiator's direct value. Strong network ties intuitively make countries' equilibrium behavior more responsive to each other's interests.

Lastly, simulations demonstrate that effects of integration on the propensity for war are at times straightforward, yet occasionally counterintuitive. In particular, whether the network increases or decreases the propensity to fight typically depends on whether integration is driven by within-alliance or cross-alliance integration. If integration is driven by rival connections, then countries have less incentive to fight since they internalize a greater share of their adversaries' gains given a loss in war. This is the common intuition behind liberal theory. However, if integration is driven by ally connections, then countries have even more incentive to fight due to heightened stakes. The first plot of Figure 16 demonstrates this.

This tendency, however, does not always hold—most notably, it may depend on the distribution of power. While the first plot of Figure 16 demonstrates equilibrium behavior of a country that is sufficiently strong, following the logic of the previous paragraph, the second plot demonstrates equilibrium behavior of a weak country. Under this parameterization, not only does the weak country always respond to integration with increases in their join probability, but that the level and rate of increase is actually *greater* when integration is driven by their rival connections.

The divergence in behavior occurs because, unlike strong countries that can affect the outcome of war by contributing military strength to their alliance via $p(\cdot)$, weak countries cannot make much difference. When cross-alliance connections strengthen, all countries' welfare depends more heavily on the welfare of their adversaries and, therefore, countries are less likely to enter conflict to the extent they affect the probability of winning. Weak countries, however, do not affect the probability of winning much. For a weak country, then, joining the conflict serves a different function: hedging their losses. In the event their alliance wins, the weak country wants to ensure they reap the full winner's reward. Intuitively, this behavioral incentive is greatest for β close to zero and reduced as β goes to 1.

B Proofs

The first set of proofs concern conditions on the network that ensure the war and peace payoffs of all countries converge to a unique stable solution.

Theorem 1. *Let* $\mathbf{\Phi} \in \mathbb{R}^{n \times n}$ *with spectral radius* $\varrho(\mathbf{\Phi})$*. Then,* $\lim_{k \to \infty} \mathbf{\Phi}^k = 0$ *if and only if* $\varrho(\mathbf{\Phi}) < 1$*.*

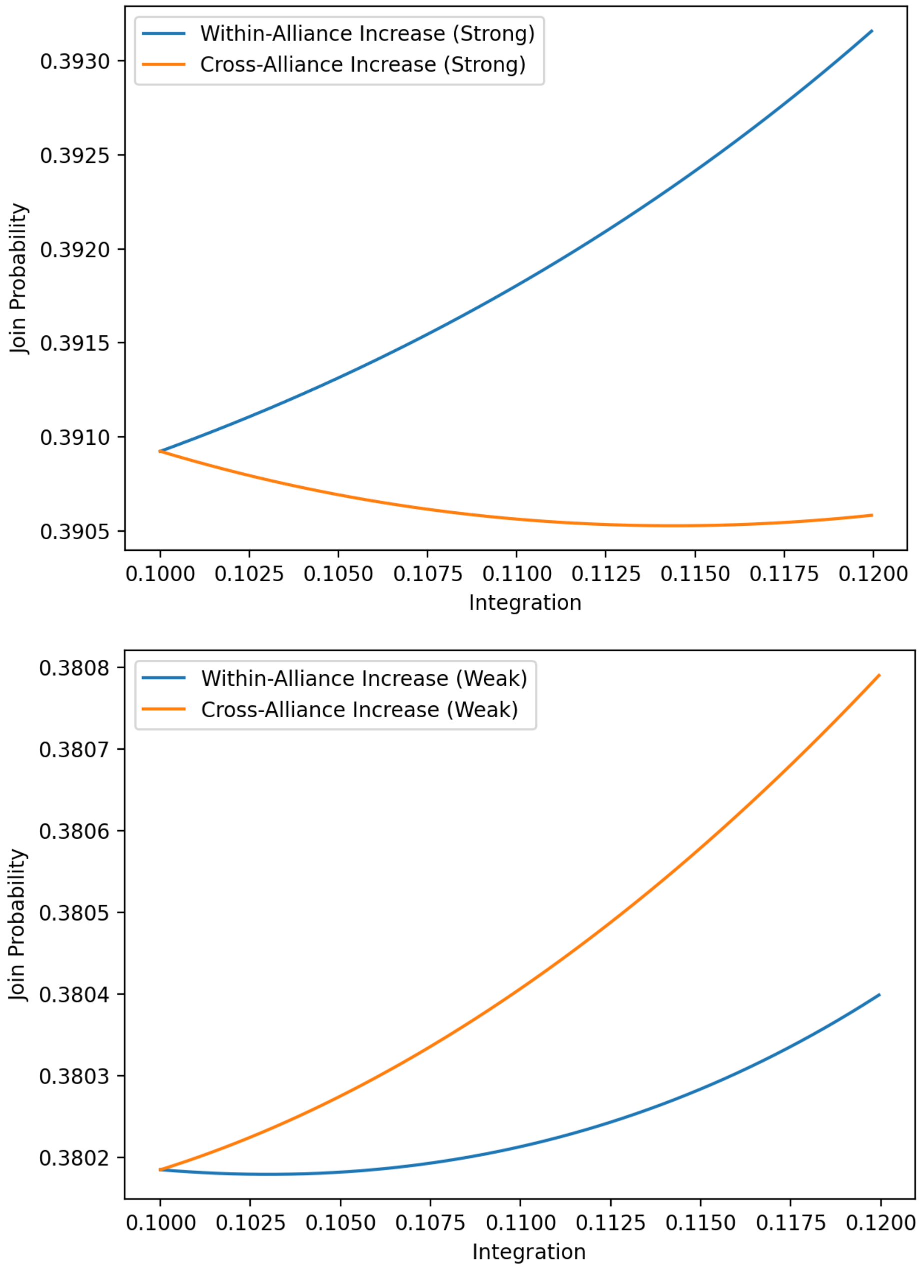


Figure 16: Probability of joining (strong vs. weak countries) in integration

Proof. Refer to Theorem 5.6.12 of Horn and Johnson (2012, p. 348-349). □

We can use Theorem 1 to establish the following lemma, which provides a necessary and sufficient condition on $\mathbf{\Phi}$ that guarantees payoffs are well-defined.

Lemma 1. *Let $\mathbf{\Phi} \in \mathbb{R}^{n \times n}$. Then, any mapping $T(\boldsymbol{u}) = \boldsymbol{b} + \mathbf{\Phi}\boldsymbol{u}$ yields a unique stable solution for any $\boldsymbol{b} \in \mathbb{R}^n$ and any initial guess $\boldsymbol{u}_0 \in \mathbb{R}^n$ if and only if $\varrho(\mathbf{\Phi}) < 1$, where $\varrho(\cdot)$ denotes the spectral radius.*

Proof. Define $\boldsymbol{\delta} := \boldsymbol{u} - \boldsymbol{u}_0$ for any initial guess $\boldsymbol{u}_0 \in \mathbb{R}^n$. Then, $T(\boldsymbol{u} + \boldsymbol{\delta}) = \boldsymbol{b} + \mathbf{\Phi}(\boldsymbol{u} + \boldsymbol{\delta})$ or equivalently $\boldsymbol{u} + \mathbf{\Phi}\boldsymbol{\delta}$. Iterating again yields $T(\boldsymbol{u} + \mathbf{\Phi}\boldsymbol{\delta}) = \boldsymbol{u} + \mathbf{\Phi}^2\boldsymbol{\delta}$. Therefore, it is straightforward to see that k iterations yields $T(\boldsymbol{u} + \mathbf{\Phi}^{k-1}\boldsymbol{\delta}) = \boldsymbol{u} + \mathbf{\Phi}^k\boldsymbol{\delta}$. By Theorem 1, $\lim_{k\to\infty} \boldsymbol{u} + \mathbf{\Phi}^k\boldsymbol{\delta} = \boldsymbol{u}$ if and only if $\varrho(\mathbf{\Phi}) < 1$. □

Therefore, it is necessary to make the following assumption.

Assumption 1. *$\varrho(\mathbf{\Phi}) < 1$, where $\varrho(\cdot)$ denotes the spectral radius.*

Although $\varrho(\mathbf{\Phi}) < 1$ guarantees theoretical convergence, computation with finite precision arithmetic may result in a failure to converge. Theorem 2 provides a sufficient condition to guarantee convergence of the floating point approximation.

Theorem 2 (Higham and Knight (1995))**.** *Let $\mathbf{\Phi} \in \mathbb{R}^{n \times n}$ with the Jordan canonical form, $\mathbf{\Phi} = \mathbf{X}\mathbf{J}\mathbf{X}^{-1}$ where $\mathbf{X}$ is nonsingular and*

$$\mathbf{J} = \operatorname{diag}(\mathbf{J}_1, \mathbf{J}_2, \ldots, \mathbf{J}_s), \qquad \mathbf{J}_i = \begin{bmatrix} \lambda_i & 1 & & \\ & \ddots & \ddots & \\ & & \lambda_i & 1 \\ & & & \lambda_i \end{bmatrix} \in \mathbb{C}^{n_i \times n_i}$$

where $\sum_{i=1}^{s} n_i = n$, have $\varrho(\mathbf{\Phi}) < 1$. A sufficient condition for $\lim_{k\to\infty} fl(\mathbf{\Phi}^k) = 0$ is

$$d_n u \kappa(\mathbf{X}) \|\mathbf{\Phi}\| < (1 - \varrho(\mathbf{\Phi}))^t \tag{A4}$$

for some p-norm, where $t = \max_i n_i$, u is the unit roundoff (i.e., machine epsilon), $\kappa(\mathbf{X}) = \|\mathbf{X}\|\|\mathbf{X}^{-1}\|$, and d_n is a modest constant depending only on n.

Proof. Refer to Higham and Knight (1995, p. 352). □

If $\mathbf{\Phi}$ is normal, then $\|\mathbf{\Phi}\|_2 = \varrho(\mathbf{\Phi}) < 1$, $t = 1$, and $\kappa_2(\mathbf{\Phi}) = 1$, so that the sufficient condition (A4) simply becomes

$$\varrho(\mathbf{\Phi}) < \frac{1}{1 + d_n u}.$$

The iterative algorithm hence allows for a gradual tightening on the assumption in the event of a failure to converge, which can but need not occur with finite precision.

Next, I establish the existence of an equilibrium to the game defined in the main text.

Proposition 1. *There exists a Perfect Bayesian Equilibrium.*

Proof. By Lemma 1, Assumption 1 guarantees $\boldsymbol{u}(\boldsymbol{a} \mid w)$ has a unique stable solution for any action profile $\boldsymbol{a} \in \{0,1\}^n$ and outcome $w \in \{0,1\}$. Further, since $p(\boldsymbol{m}, \boldsymbol{a}) \in (0,1)$, $\tilde{\boldsymbol{u}}(\boldsymbol{a})$ is also well-defined for any $\boldsymbol{a}$.

Suppose a war was initiated by h against ℓ and take joining strategies $\boldsymbol{\gamma}_{-i}$ as given for all players except some $i \in N \setminus \{h, \ell\}$. We know from the analysis above that country i's best response is to join if and only if

$$\Omega_i^{h,\ell}(\boldsymbol{\gamma}_{-i}) := \sum_{C \subseteq \mathcal{C}_{-i}^{h,\ell}} \left[\prod_{r \in C} \gamma_r^{h,\ell} \prod_{k \notin C} (1 - \gamma_k^{h,\ell}) \right] \Big(\tilde{u}_i(1, \boldsymbol{a}_C) - \tilde{u}_i(0, \boldsymbol{a}_C) \Big) > \varepsilon_i \tag{A5}$$

Given the joining behavior of others, country i determines whether they join on the basis of a unique threshold. Note that since the threshold $\Omega_i^{h,\ell}(\boldsymbol{\gamma}_{-i})$ is a polynomial function, it is also continuous. Likewise, we know from the definition of the logistic distribution that any country i's join probability is continuous in their threshold, $\gamma_i^{h,\ell} = F(\Omega_i^{h,\ell})$.

We can thus conclude the best response function is continuous. Since there are a finite number of payoffs, each of which is well-defined for all possible $\boldsymbol{a}$, $\boldsymbol{m}$, and w, the threshold Ω_i is bounded. From expression (A5), it is straightforward to see that a loose bound for any threshold is given by $2 \max_{\boldsymbol{a}} |\tilde{u}_i(\boldsymbol{a})|$. By Brouwer's fixed point theorem, there exists a vector of equilibrium join probabilities $\boldsymbol{\gamma}^{*h,\ell}$ such that $\boldsymbol{\gamma}^{*h,\ell} = F(\boldsymbol{\Omega}^{h,\ell}(\boldsymbol{\gamma}^{*h,\ell}))$.

Given equilibrium behavior in the second stage, suppose a random country $h \in N$ is chosen as a first-mover. Since it is already established that $\tilde{\boldsymbol{u}}(\cdot)$ is well-defined for any $\boldsymbol{a}$ and $\boldsymbol{v}$ has a unique stable solution given Assumption 1, the initiator simply chooses the action that yields the greatest total expected utility, given the beliefs of the initiator in the first stage and of the potential joiners in the second stage are consistent with equilibrium strategies of potential joiners. □

Lastly, I demonstrate that the sets of moments corresponding to join probabilities and attack probabilities yield an overdetermined system of equations.

Lemma 2. *The system of equations given by equalities (8) and (9) is overdetermined.*

Proof. First, equality (8) yields a system of $(n-2)(n-n')n'$ equations, one for each of $n-2$ potential joiners across $(n-n')n'$ potential originating pair $(h,\ell) \in (N, R_h)$ for all $h \in N$, where n' is the number of countries on an arbitrary side. Second, equality (9) yields a system of $2n'(n-n')$ equations, one for each country $h \in N$ that could be recognized and a corresponding target $\ell \in R_h$.

Equations (8) and (9) match the reduced-form join and attack probabilities to the structural model. There are $(n-1)^2 + n + 3$ structural parameters in $\boldsymbol{\theta}$: n direct values each country receives for winning a war, each given by δ_i; there is 1 discount on the value of winning for being a passive ally, β; there is 1 cost of losing a fight, κ; and there are $(n-1)^2$ parameters governing the network, ϕ_{ij} for all $i, j \in N^2$, given that $\phi_{ii} = 0$ for all i; there is 1 gain from peace, α. Thus, there are equations than parameters if and only if

$$(n-2)(n-n')n' + 2n'(n-n') > (n-1)^2 + n + 3$$

which holds for any $n > 2$, $n' \geq 1$. □

C Additional Estimation Details

C.1 Multiple Imputation

Table 4: Summary Statistics for Variables With Missingness, 1816-2014

Statistic	N	Mean	Min	Max	NAs
Population	14,986	27,848.69	16.00	1,390,110	3
Military (CINC)	14,986	0.01	0.00	0.38	3
Gross domestic product	8,211	308,049.50	250.48	18,244,220	6,778
Polity	13,763	−0.031	−1	1	1,226

Table 4 provides summary statistics for the four country-level characteristics subject to missingness that are used for the first step of the estimation process. These same covariates are also used to create aggregate, network-level characteristics (e.g., aggregate military capacity for each side of the dispute). To account for missingness, I conduct multiple imputation by an Expectation-Maximization with Bootstrapping (EMB) algorithm (Honaker, King, and Blackwell 2011).

First, I construct a data set of all country-years from 1816 to 2014. I include additional country-level covariates from the Maddison project (Bolt and Luiten van Zanden 2020) to improve the quality of the imputations. Strictly positive variables are log transformed and weakly positive variables are inverse hyperbolic sine transformed to best fit a joint normal distribution. Second, ten imputed data sets are created according to the methods in Honaker, King, and Blackwell (2011). Each imputation corresponds to a gravity model to impute missingness for trade, as outlined in the subsequent section, and first-stage estimation. Upon recovering first-stage estimates for each imputed data set, the first-stage models can be compiled to an implied distribution such that the point estimate is the average of all point estimates and the variance is given by accounting for both within- and between-imputation variance, as outlined in equation (A6) of Appendix C.5.

C.2 Gravity Model

Correlates of War project's data set on Trade (Barbieri, Keshk, and Pollins 2009; Barbieri and Keshk 2016) includes data from 1870 to 2014. Among available data, approximately 27 percent of trade data is missing. Further, data prior to the year 1870 is unavailable.

I employ a simple gravity model to estimate the missing values. In particular, I allow for trade flows from country i to j to be given by

$$T_{ij} = Gv_{ij} \cdot \frac{M_i^{\nu_1} M_j^{\nu_2}}{D_{ij}^{\nu_3}}$$

where M_i and M_j denote the "masses" of countries i and j as given by i and j's respective gross domestic products, D_{ij} is the geographic distance between i and j, G is a constant, and υ_{ij} is a mean 1 error term. The expression

$$T_{ij} = \upsilon_{ij} \exp(\nu_0 + \nu_1 \ln(M_i) + \nu_2 \ln(M_j) - \nu_3 \ln(D_{ij}))$$

can then be estimated using a Poisson Pseudo-Maximum Likelihood (PPML) estimator (Silva and Tenreyro 2006, 2011). For more information on the method generally, refer to Gourieroux, Monfort, and Trognon (1984).

C.3 Network Implementation

It is necessary to define the vertices of the network for estimation. Following the international relations literature, I choose to employ the notion of a politically relevant international environment (Maoz 1996; Lemke and Reed 2001).

The decision rule for inclusion in the network is as follows. First, I take each dispute that is provided in the CoW Militarized Interstate Disputes (MIDs) data set. The first step is to identify the initiator and target of the dispute to ensure they are included in the network. The data provides originators to each dispute, including those that do not escalate to war. If there are two originators to the dispute, I identify the initiator and the target as those on each respective side with the largest military capabilities, and consider the remaining originators to be joiners to the conflict.

The second step is to set a threshold $\bar{n} > 2$ such that if there exists a dispute in the MIDs data set with $n > \bar{n}$ participants, then I remove $n - \bar{n}$ of the participating countries with the lowest military capabilities from the network. On the other hand, if $n < \bar{n}$, then I can choose an additional $n - \bar{n}$ countries that did not join the dispute to include in the network. In this way $\bar{n}$ sets a consistent number of vertices. For the results in this paper, I select $\bar{n} = 5$ to improve computational efficiency and because trials with $\bar{n} = 7$, 9, and 11 have not led to significantly different results. This is important because, ex ante, we do not know which disputes will escalate and which will remain contained. Further, this is essential to consistent estimation, as the effect of network connections on equilibrium behavior is increasing in the total number of connections (e.g., as opposed to a country's total network influence remaining constant in the number of connections, with individual influence reflecting an average).

C.4 Reduced-Form Estimation

The estimation procedure calls for two reduced-form probabilities: the join probabilities for all countries and the initiator's probability of victory for each dispute. To recover the probability of the initiator's victory, I use total geographic distance across all countries, total geographic distance across initiating side countries, total geographic distance across rivals, total military capabilities, initiating side military capabilities, total gross domestic product, initiating side gross domestic product, total population for all countries, total population for initiating side countries, total nuclear stockpiles, initiating side nuclear stockpiles, total economic centrality, initiating side economic centrality, initiator gross domestic product, initiator total population, initiator military capabilities, initiator nuclear stockpiles, initiator polity score, initiator economic centrality, target gross domestic product, target total population, target military capabilities, target nuclear stockpiles, target polity score, target economic centrality, and period fixed effects.

To recover the join probability for country i, I use country i's military capabilities, gross domestic product, polity score, economic centrality, the total geographic distance across all countries, total geographic distance across countries on i's side, total geographic distance across rivals, total military capabilities, total gross domestic product, total population for all countries, total nuclear stockpiles, total military capacities for countries on i's side, total gross domestic product for countries on i's side, total population for countries on i's side, total nuclear stockpiles for countries on i's side, initiator gross domestic product, initiator total population, initiator military capabilities, initiator nuclear stockpiles, initiator polity score, initiator economic centrality, target gross domestic product, target total population, target military capabilities, target nuclear stockpiles, target polity score, target economic centrality, and period fixed effects.

There are also two other relevant reduced-form probabilities we can recover: the probability country h attacks country ℓ and the probability that country h is recognized. For the former, I use the same covariates as for the join probabilities in addition to country ℓ's military capabilities, gross domestic product, population, polity score, nuclear stockpiles, and economic centrality. For the latter, I use the same as the join probabilities excluding the initiator characteristics. All data sources are provided in the main text, with multiple imputation conducted according to Appendix C.1.

C.5 Standard Errors

For each data set d, I recover a vector of "reduced-form" estimates

$$\hat{\boldsymbol{\mu}}_d = (\hat{\boldsymbol{\eta}}_{\gamma,d}, \hat{\boldsymbol{\eta}}_{\rho,d}, \hat{\boldsymbol{\zeta}}_{\gamma,d}, \hat{\boldsymbol{\zeta}}_{\rho,d}, \hat{\boldsymbol{\xi}}_{\gamma,d}, \hat{\boldsymbol{\xi}}_{\rho,d}, \hat{\boldsymbol{\tau}}_{t,d}, \hat{\boldsymbol{\tau}}_{t,d}).$$

Little and Rubin (2020) show that averaging over $D \geq 2$ imputations yields the mean

$$\bar{\boldsymbol{\mu}}_D = \frac{1}{D}\sum_{d=1}^{D}\hat{\boldsymbol{\mu}}_d$$

and that the estimate's variability has two components: first, the average within-imputation variance

$$\bar{\mathbb{V}}_D^W(\hat{\boldsymbol{\mu}}) = \frac{1}{D}\sum_{d=1}^{D}\mathbb{V}(\hat{\boldsymbol{\mu}}_d),$$

and second, the between-imputation variance

$$\mathbb{V}_D^B(\hat{\boldsymbol{\mu}}) = \frac{1}{D-1}\sum_{d=1}^{D}(\hat{\boldsymbol{\mu}}_d - \bar{\boldsymbol{\mu}}_D)(\hat{\boldsymbol{\mu}}_d - \bar{\boldsymbol{\mu}}_D)'.$$

Hence, the vector of total variances given D imputations can be expressed

$$\mathbb{V}_D(\hat{\boldsymbol{\mu}}) = \bar{\mathbb{V}}_D^W(\hat{\boldsymbol{\mu}}) + \frac{D+1}{D}\mathbb{V}_D^B(\hat{\boldsymbol{\mu}}). \tag{A6}$$

Asymptotic standard errors are then computed according to standard two-step estimation (Murphy and Topel 1985; Greene 2018). Specifically,

$$\hat{\mathbb{V}}(\hat{\boldsymbol{\theta}}) = \hat{\mathcal{J}}_{\boldsymbol{\theta}}^{-1} + \hat{\mathcal{J}}_{\boldsymbol{\theta}}^{-1}\hat{\mathcal{J}}_{\boldsymbol{\mu}}\hat{\mathbb{V}}(\hat{\boldsymbol{\mu}})\hat{\mathcal{J}}_{\boldsymbol{\mu}}'\hat{\mathcal{J}}_{\boldsymbol{\theta}}^{-1}$$

where

$$\begin{aligned}
\hat{\mathcal{J}}_{\boldsymbol{\theta}} &:= \mathbf{J}_{\boldsymbol{\theta}}(\hat{\boldsymbol{\theta}} \,|\, \hat{\boldsymbol{\mu}}, \mathbf{X}, \boldsymbol{Z})'\mathbf{J}_{\boldsymbol{\theta}}(\hat{\boldsymbol{\theta}} \,|\, \hat{\boldsymbol{\mu}}, \mathbf{X}, \boldsymbol{Z}) \\
\hat{\mathcal{J}}_{\boldsymbol{\mu}} &:= \mathbf{J}_{\boldsymbol{\theta}}(\hat{\boldsymbol{\theta}} \,|\, \hat{\boldsymbol{\mu}}, \mathbf{X}, \boldsymbol{Z})'\mathbf{J}_{\boldsymbol{\mu}}(\hat{\boldsymbol{\mu}} \,|\, \hat{\boldsymbol{\theta}}, \mathbf{X}, \boldsymbol{Z}).
\end{aligned}$$

Hence, $\hat{\mathcal{J}}_{\boldsymbol{\theta}}$ and $\hat{\mathcal{J}}_{\boldsymbol{\mu}}$ denote the inverse of the outer product of gradients estimator, with $\mathbf{J}_{\boldsymbol{\theta}}(\cdot)$ and $\mathbf{J}_{\boldsymbol{\mu}}(\cdot)$ yielding the Jacobian with respect to $\boldsymbol{\theta}$ and $\boldsymbol{\mu}$, respectively.

Appendix References

Barbieri, Katherine, and Omar M. G. Keshk. 2016. Correlates of War Project Trade Data Set Codebook, Version 4.0. Technical report.

Barbieri, Katherine, Omar M. G. Keshk, and Brian Pollins. 2009. "Trading Data: Evaluating Our Assumptions and Coding Rules." *Conflict Management and Peace Science* 26(5):471–491.

Bolt, Jutta, and Jan Luiten van Zanden. 2020. "Maddison style estimates of the evolution of the world economy. A new 2020 update.".

Gourieroux, Christian, Alain Monfort, and Alain Trognon. 1984. "Pseudo Maximum Likelihood Methods: Applications to Poisson Models." *Econometrica* p. 701–720.

Greene, William H. 2018. *Econometric Analysis*. 8th edition ed. London: Pearson Education Limited.

Higham, Nicholas J, and Philip A Knight. 1995. "Matrix Powers in Finite Precision Arithmetic." *SIAM Journal on Matrix Analysis and Applications* 16(2):343–358.

Honaker, James, Gary King, and Matthew Blackwell. 2011. "Amelia II: A Program For Missing Data." *Journal of Statistical Software* 45:1–47.

Horn, Roger A, and Charles R Johnson. 2012. *Matrix Analysis*. Second edition ed. Cambridge University Press.

Lemke, Douglas, and William Reed. 2001. "The relevance of politically relevant dyads." *Journal of Conflict Resolution* 45(1):126–144.

Little, Roderick J.A., and Donald B. Rubin. 2020. *Statistical Analysis With Missing Data*. Third edition. ed. Hoboken, NJ: Wiley.

Maoz, Zeev. 1996. *Domestic sources of global change*. University of Michigan Press.

Murphy, Kevin M, and Robert H Topel. 1985. "Estimation and Inference in Two-Step Econometric Models." *Journal of Business & Economic Statistics* 3(4):370–379.

Silva, J.M.C. Santos, and Silvana Tenreyro. 2006. "The Log of Gravity." *The Review of Economics and Statistics* 88(4):641–658.

Silva, J.M.C. Santos, and Silvana Tenreyro. 2011. "Further Simulation Evidence on the Performance of the Poisson Pseudo-Maximum Likelihood Estimator." *Economics Letters* 112(2):220–222.

Software References

Arel-Bundock, Vincent, Nils Enevoldsen, and CJ Yetman. 2018. "countrycode: An R package to convert country names and country codes." *Journal of Open Source Software* 3(28):848.

Bates, Douglas, Martin Maechler, and Mikael Jagan. 2022. *Matrix: Sparse and Dense Matrix Classes and Methods*. R package version 1.5-3.

Csardi, Gabor, and Tamas Nepusz. 2006. "The igraph software package for complex network research." *InterJournal* Complex Systems:1695.

Harris, Charles R., K. Jarrod Millman, Stéfan J. van der Walt, Ralf Gommers, Pauli Virtanen, David Cournapeau, Eric Wieser, Julian Taylor, Sebastian Berg, Nathaniel J. Smith, Robert Kern, Matti Picus, Stephan Hoyer, Marten H. van Kerkwijk, Matthew Brett, Allan Haldane, Jaime Fernández del Río, Mark Wiebe, Pearu Peterson, Pierre Gérard-Marchant, Kevin Sheppard, Tyler Reddy, Warren Weckesser, Hameer Abbasi, Christoph Gohlke, and Travis E. Oliphant. 2020. "Array programming with NumPy." *Nature* 585(7825):357–362.

Hijmans, Robert J. 2022. *geosphere: Spherical Trigonometry*. R package version 1.5-18.

Honaker, James, Gary King, and Matthew Blackwell. 2011. "Amelia II: A Program for Missing Data." *Journal of Statistical Software* 45(7):1–47.

Hunter, J. D. 2007. "Matplotlib: A 2D graphics environment." *Computing in Science & Engineering* 9(3):90–95.

Inc., Wolfram Research,. N.d. "Mathematica, Version 13.3.". Champaign, IL, 2023.

Joblib Development Team. 2020. "Joblib: running Python functions as pipeline jobs.".

McKinney, Wes et al. 2010. Data structures for statistical computing in python. In *Proceedings of the 9th Python in Science Conference*. Vol. 445 Austin, TX p. 51–56.

Miller, Steven V. 2022. "peacesciencer: An R Package for Quantitative Peace Science Research." *Conflict Management and Peace Science* 39(6):755–779.

R Core Team. 2022. *R: A Language and Environment for Statistical Computing*. Vienna, Austria: R Foundation for Statistical Computing.

Seabold, Skipper, and Josef Perktold. 2010. statsmodels: Econometric and statistical modeling with python. In *9th Python in Science Conference*.

Van Rossum, Guido, and Fred L. Drake. 2009. *Python 3 Reference Manual*. Scotts Valley, CA: CreateSpace.

Virtanen, Pauli, Ralf Gommers, Travis E. Oliphant, Matt Haberland, Tyler Reddy, David Cournapeau, Evgeni Burovski, Pearu Peterson, Warren Weckesser, Jonathan Bright, Stéfan J. van der Walt, Matthew Brett, Joshua Wilson, K. Jarrod Millman, Nikolay Mayorov, Andrew R. J. Nelson, Eric Jones, Robert Kern, Eric Larson, C J Carey, İlhan Polat, Yu Feng, Eric W. Moore, Jake VanderPlas, Denis Laxalde, Josef Perktold, Robert Cimrman, Ian Henriksen, E. A. Quintero, Charles R. Harris, Anne M. Archibald, Antônio H. Ribeiro, Fabian Pedregosa, Paul van Mulbregt, and SciPy 1.0 Contributors. 2020. "SciPy 1.0: Fundamental Algorithms for Scientific Computing in Python." *Nature Methods* 17:261–272.

Waskom, Michael L. 2021. "seaborn: statistical data visualization." *Journal of Open Source Software* 6(60):3021.

Wickham, Hadley, Evan Miller, and Danny Smith. 2023. *haven: Import and Export 'SPSS', 'Stata' and 'SAS' Files*. R package version 2.5.2.

Wickham, Hadley, and Jennifer Bryan. 2023. *readxl: Read Excel Files*. R package version 1.4.2.

Wickham, Hadley, Romain François, Lionel Henry, Kirill Müller, and Davis Vaughan. 2023. *dplyr: A Grammar of Data Manipulation*. R package version 1.1.1.

Woelwer, Anna-Lena, Jan Pablo Burgard, Joshua Kunst, and Mauricio Vargas. 2023. *gravity: Estimation Methods for Gravity Models*. R package version 1.1.

Zeileis, Achim, and Gabor Grothendieck. 2005. "zoo: S3 Infrastructure for Regular and Irregular Time Series." *Journal of Statistical Software* 14(6):1–27.