Alliance Dynamics in the Shadow of Shifting Power*

Colin Krainin† Robert Schub‡

Forthcoming at *International Studies Quarterly*

Abstract

Alliances are costly to form and to terminate, and yet alliances change frequently. Scholars typically attributes these decisions to static factors, such as the power balance, and retrospective ones, such as past power shifts. We highlight another factor: prospective changes, particularly anticipated military strength shifts. We analyze a three-country bargaining model of alliances and war that incorporates forward-looking power dynamics. The model, unlike those restricting players to set roles, flexibly allows players to ally in any arrangement. We find that alliance arrangements which are optimal when power is static are often suboptimal when power fluctuates. Maintaining prior alliances despite expected power shifts may even lead to preventive war. States thus strategically look to the future to identify optimal alliances in the present. Quantitative analyses corroborate the expectation. As the anticipated size of power shifts increases, alliance changes become more common. Accordingly, states navigate expected changes in the international landscape by rearranging current alliance commitments which can help minimize the risk of conflict. When power balances are in flux, malleable institutional arrangements may prove preferable to rigid ones.

*We thank ISQ reviewers and editors, Thomas Wiseman, Kristopher Ramsay, Robert Powell, James Fearon, Andrew Coe, and participants of the 2016 Formal Models in International Relations Conference at USC for helpful comments and suggestions.

†Princeton University. Email: colinkrainin@gmail.com

‡Corresponding author: University of Nebraska-Lincoln. Email: robert.schub@unl.edu

International alliances frequently change as states regularly form new alliances and terminate existing ones. Diplomatic history is littered with rapid revisions to alliance structures—for example, the Diplomatic Revolution of the 18th century or the swift negotiation and even swifter violation of the Molotov-Ribbentrop Pact. On average, nearly 9% of alliances that exist at the start of the year no longer exist by the end. The new alliances formed in a given year are equivalent to roughly 12% of all alliances present at the start of the year.[1]

The empirical frequency of alliance changes is puzzling. Distilling a common contention, Smith (1998, p. 316) writes that "nations pay costs for forming alliances." Alliance formation costs have several sources including forfeiting autonomy (Leeds, 2003), potential entrapment (Snyder, 1997), or adopting policies an ally prefers rather than policies most consistent with a state's own interests (Morrow, 1994). Smith (1998) continues, "nations [also] pay costs for not honoring their commitments," which is the most common cause of alliance termination with 34% of alliances ending due to violations or abrogation (Leeds and Savun, 2007). States that do not fulfill alliance responsibilities suffer a reputation downgrade (Crescenzi et al., 2012). Thus, forming new alliances entails sinking new costs and generating new commitments. Terminating old ones forfeits the signaling value of costs already sunk and imposes reputation costs. Assuming that states prefer to avoid myopic choices with costly outcomes, the costs to alliance creation and destruction would seem to promote a stable alliance network characterized by inertia rather than flux.

What explains the dynamism of alliance arrangements given the costs associated with it? Past inquiries typically study formation (Waltz, 1979; Walt, 1987; Johnson, 2017) or termination (Morrow, 1991; Leeds and Savun, 2007) separately. They also typically emphasize *static* conditions such as the balance of power or threat, reputation, or aligned interests. More dynamic accounts focus on *retrospective* changes, such as in power or interests, to explain alliance formation and termination. In contrast, this paper offers a unified framework for alliance dynamics while emphasizing an often neglected factor: *prospective* changes. States look to the future to identify

[1]Calculations based on alliances with an offense or defensive component as coded in the Alliance Treaty Obligations and Provision (ATOP) data (Leeds et al., 2002).

optimal alliances in the present. Anticipated changes, particularly shifts in military strength, induce alliance dynamism and help explain the empirical frequency of alliance alterations. Alliance formation and termination depend upon the distribution of power *and* the dynamics of power.

We develop the theory from a three-country, complete information, bargaining model of alliance and war that incorporates a collection of features that have not been jointly studied. In the model, countries observe the initial distribution of power and recognize that military strength may shift as in typical commitment problem models (Fearon, 1995; Powell, 2006). States anticipate strength shifts and optimize alliance choices accordingly. Unlike most existing approaches, states form alliances without pre-assigned roles. Once allied, countries commit to fight together and divide resources as they see fit. Individual countries then opt to initiate war or find a settlement through bargaining.

The model builds upon prior work in several ways. Models that allow for dynamics such as power shifts (Fearon, 2004; Powell, 2004, 2006) typically have two countries. Those with more than two countries often lack ongoing dynamics (Morrow, 1994; Powell, 1999; Werner, 2000; Jackson and Nei, 2015). Models with both lack bargaining (Niou, Ordeshook, and Rose, 1989), examine alliances in the limiting case with high patience (Krainin and Wiseman, 2016), or assign fixed roles to the countries (Powell, 2017).[2] The point about fixed roles is critical since many issues associated with alliances are precluded in models with assigned adversaries and potential partners.

The flexibility of the model produces new insights linking the distribution and dynamics of power to alliance decisions. Alliance choices when power is static are crucial to understanding the effect of power shifting. For a static system we can specify the optimal alliance configuration. When states anticipate a shift in state capabilities, optimal alliance arrangements may also shift, whether through alliance formation, termination, or both (alliance switching). War looms in the background, serving as an outside option for states, and creates the relationship between strength shifts and alliance shifts. Alliances enable members to aggregate capabilities which im-

[2]Powell (2017) is a relaxation of many previous models where potential alliance partners are pre-assigned. However, only one country is allowed the freedom of choice.

proves their expected war payoffs. Higher war payoffs affect bargaining payoffs which drive the alliance choices in a static setting. In some circumstances, remaining in existing alliances despite anticipated strength shifts could induce preventive war. Given the linkages between military capabilities and alliance decisions, theories of the latter ought to incorporate dynamics in the former. This paper suggests that states do not idly watch battlefield-relevant attributes change. Rather, they anticipate changes when possible and act accordingly. Altering alliances in anticipation of a strength shift enables states to preserve peace whereas if they remained locked into existing alliance arrangements, sub-optimal outcomes including war occur.

Results from quantitative analyses strongly support the model's implications. We analyze power and alliance dynamics within geographic regions. The outcome variable measures alliance changes—new ones forming or old ones ending. The explanatory variable measures anticipated shifts in the military capabilities of the region's states, building on the measurement approach in Bell and Johnson (2015). Consistent with the theory, moving from an instance where states expect the regional power balance to remain static to one where they anticipate large power shifts is associated with a 55% increase in alliance changes. Alliance volatility is a fixture of international politics; expected strength shifts help explain the volatility.

The paper makes several contributions beyond demonstrating that states alter alliance behavior prospectively. First, the formal model provides a flexible approach for studying numerous issues related to alliances. It helps generalize findings from qualitative studies that incorporate forward-looking alliance changes, but that privilege the strategic choices of a single actor, such as a declining hegemon (Gilpin, 1981; MacDonald and Parent, 2011). We, instead, simultaneously analyze the incentives of all three actors while not foreclosing possible alliance arrangements. Second, the paper contributes to recent efforts to move beyond the dyad given that international behavior is often multilateral (Fordham and Poast, 2016). This study shows that an anticipated power shift in any of several relevant parties can alter the alliance structure among them all. We thus build on work that incorporates attributes of three actors—such as poles in a tripolar system (Schweller, 1993). Third, it shows that canonical scholarship on institutions can occasionally mislead when

predicting institutions to be sticky (Keohane, 1984; Ikenberry, 1998). Our results specify when institutional dynamism trumps rigidity.

The results shed light on several contemporary policy matters including the lack of post-Cold War balancing targeting the US. We must look at the existing distribution of power *and* projected shifts in power—particularly anticipation of China's rise—to understand the dearth of balancing. Additionally, we illustrate the benefits of alliance flexibility. Theories without anticipatory alliance dynamics may lead to inaccurate predictions about war onset as preventive wars that erupt when states are locked into alliances can be avoided if states alter alliance arrangements. Wars can occur if alliances are immutable and only backward-looking. By contrast, forward-looking states prospectively alter alliance configurations to maximize their bargaining position which typically helps minimize the threat of war. Accordingly, alliance dynamism may prove crucial to peacefully navigating anticipated power shifts, whether they stem from nuclear proliferation or long-run changes in great power capabilities.

Literature: Static Alliances

This section reviews the power-centric literature on causes of alliance formation and termination. After discussing predictions when power is static, it turns to studies where power is shifting. It then discusses formal models of alliances, noting a common restriction on actors' roles which our model relaxes.

We build on the power-centric literature that views alliances primarily as a means of capability aggregation (Waltz, 1979). This represents only one of several strands of alliance theorizing. Motivations for alliances beyond those concerned with power relate to foreign policy interests (Morrow, 1994), regime type (Lai and Reiter, 2000), or reputation considerations (Crescenzi et al., 2012). States also employ alliances to achieve objectives unrelated to capability aggregation, such as gaining influence over allies' international and domestic behavior (Schroeder, 2004; Morrow, 1991). We acknowledge the relevance of other factors such as shared interests; our focus simply lies within the power-centric alliance literature. The size and influence of the power-centric literature suggests there is value to exploring prospective changes in power balances. In a power-centric framework,

states form alliances when capabilities incentivize these arrangements even when ideologically distant (Waltz 1979). Alliances between states with opposing ideologies occur infrequently but do exist, such as the UK-Russia alliance which persisted beyond World War 2.[3] Future work may consider the effects of anticipated changes in state attributes other than capabilities—including interests.

Scholarship within the capability aggregation framework necessarily emphasizes military strength. These accounts often share a static framework with fixed capabilities but differ on the predicted patterns of alliance behavior. In Waltz's (1979) theory of balancing, weaker states align against a stronger state. Walt (1987) offers a reformulation in which states that pose lower threats ally together to balance against more threatening, versus more powerful, states. Similarly operating in a static setting, others anticipate bandwagoning with weaker states allying with the predominant power for survival (Schroeder, 1994) or for profit (Schweller, 1994). Some theories specify conditions amenable to balancing and bandwagoning (Schweller, 1993; Powell, 1999). These too are static theories which fix the relevant variables. Many quantitative empirical analyses adopt this static approach (Lai and Reiter, 2000; Johnson, 2017).

Unfortunately, these studies rarely provide expectations for how *changes* in strength affect alliances. We can deduce predictions with some extrapolation. Because each theory takes the present characteristics as inputs, the intuitive implication is that *after* shifts occur, states reevaluate their alliance choices. Scholarship on alliances that does incorporate strength dynamics typically uses this backward-looking approach. Recent strength shifts and war conclusions correlate with alliance changes (Morrow, 1991; Bennett, 1997), as do battlefield results in civil conflicts (Christia, 2012). Additionally, large capability shifts since alliance formation are associated with states not fulfilling alliance obligations (Leeds, 2003) and terminating alliances (Leeds and Savun, 2007).

We similarly contend that changes in prevailing conditions increase the probability of alliance changes. In contrast, states are forward-looking in our account. Existing work on anticipatory

[3]The alliance targeting Germany endured until 1955 when the UK endorsed remilitarizing West Germany (Leeds et al., 2002).

alliance changes offers insights that we aim to generalize. Historically-informed studies analyze hegemonic transitions, emphasizing the waning hegemon's choices. Gilpin (1981, pp. 192-197) suggests that a declining hegemon pursues retrenchment via new alliances with states that do not pose imminent threats. He highlights the UK's efforts to accommodate US interests in the 1890s and develop cordial relations with Russia in the 1900s as alignment shifts prompted by the specter of Germany's rise. MacDonald and Parent (2011) find that retrenching hegemons form new alliances when their decline is moderately paced, with France forming several alliances in the 1920s exemplifying such behavior. Others note, with skepticism, that a declining hegemon may seek allies to check a rising power rather than launch a preventive attack (Levy, 1987, fn. 13). Alternatively, a declining hegemon seeks new allies when it is democratic and the rising state is autocratic, as fits the British case above (Schweller, 1992).

Others emphasize alliance choices of third parties rather than the hegemon. Weaker states allied with the declining hegemon look to the hegemon's capabilities and reassurances to determine whether making an alliance shift—toward neutrality or the rising state—is optimal (Castillo and Downes, 2020). Belgian defection from France in response to Germany's rise and reoccupation of the Rhineland typifies this behavior. For Kydd (2020), timing is key. Weaker states initially ally with the rising state to deter a preventive strike by the declining state. Weaker states then switch to allying with the declining state to deter the strengthened state from aggressively consolidating its position.

We expand upon this qualitative work to provide a general theoretical framework which enables a broader empirical test. Past work restricts states to prescribed roles—such as the declining hegemon—and preclude certain alliance arrangements. We instead allow any alliance structure to prove stable.

The formal model, which introduces power shifts into a three-player game with alliances, war, and bargaining, builds on two lines of scholarship: a contest approach (Tullock 1980) and a bargaining approach. Garfinkel (2004) exemplifies the contest tradition. She models countries as engaging in a three stage game. First, alliances form. Second, countries make a guns versus but-

ter decision over arming and consumption. This is followed by automatic winner-take-all conflict between alliances. Third, alliance partners compete over the spoils of war. A large number of possible equilibria result. The main limitation with the contest approach is the notion of automatic conflict. We depart from this approach by incorporating peaceful bargaining.

We draw upon two strands of the bargaining approach. First, shifting power in a two player model can lead to inefficient war since rising countries cannot commit to not exploit future increased bargaining power (Fearon, 1995; Powell 2006). Second, some work brings the bargaining approach to study alliances. Alliances may serve as signaling devices to demonstrate credibility (Morrow, 1994) with signaling mechanisms including statements emboldening allies (Trager, 2015) or risking entrapment by allies (Snyder, 1997). Others develop and test related ideas (Yuen, 2009; Wolford, 2014; Fang, Johnson, and Leeds, 2014; Smith, 2016). These models, however, assign actors to set roles within the strategic interaction. We relax this restriction to capture the full array of alliance dynamics.

Model

We develop a basic shifting power model (Fearon 1995) with the addition of a third player. Including a third country requires assumptions about what alliances do, how bargaining works inside and outside of alliances, and how war alters the power balance. The Supporting Information (SI) contains technical details, proofs, and several extensions. This section introduces the players and their payoffs, describes the nature of strength shifts, alliances, war, and bargaining, and finishes with the game sequence and equilibrium concept.

A set of 3 countries, denoted by N, interact at discrete times $t \in \{0, 1, 2, ...\}$. All countries discount the future at a common rate $\delta \in (0, 1)$. Countries have resources for consumption and for military purposes. The total amount of resources available for consumption each period is $X > 0$, and the amount controlled by country i (which may change after war) is denoted x_i. A nonnegative vector $\mathbf{x} = (x_1, x_2, x_3)$ specifies consumption resources for all countries, where $x_i > 0$ when a country is active, $x_i = 0$ if a country is inactive (defeated in war), and $\sum_{i=1}^{3} x_i = X$. For simplicity, we assume that $X = 1$ and that countries initially control equal shares of consumption

resources, $\mathbf{x} = (1/3,\ 1/3,\ 1/3)$. The results generalize beyond this simplification. In addition to consumption resources, each country i has current military resources or strength, s_i. There is a finite set $\hat{S} = S \cup \{0\}$ of possible states, where an element s of $S = \{1, 2, ..., K\}$, $K \geq 1$ represents an active country's strength in war. A country defeated in war has a strength of 0 and becomes inactive. Let $\mathbf{s} \in \hat{S}^3$ denote a vector of military strengths for all countries.

When a country is at peace, its consumption reflects its resources and transfers arranged via bargaining. Let τ_{ij} denote net transfers from i to j.[4] In the absence of war, consumption in country i for a single period is $c_i = x_i - \sum_{j \in N \setminus \{i\}} \tau_{ij}$. Transfers are restricted by a bargaining protocol described below so that $c_i \geq 0$ for all i. The total payoff to a country that receives consumption stream $\{c_t\}_{t \geq 1}$ is $v_i = (1-\delta) \sum_{t=1}^{\infty} \delta^{t-1} c_t$. If only a single country remains, the game ends and the survivor receives the entire stream of available consumption $(1-\delta) \sum_{t=1}^{\infty} \delta^{t-1} X = X$.

We focus on one-time strength shifts where a single country's military power increases additively by $z \in \{1, 2, ..., K-1\}$ where the maximum achieved military strength cannot exceed K. Particular power shifts are denoted (i, z) where country i shifts up by z for a single period. A power shift can either be *unanticipated* in the sense that sticky alliances have already formed when the shift takes place or *anticipated*, which countries can take into account when forming alliances.[5] The SI discusses the case of persistent shifts that last longer than a single period.

Countries can form alliances which last for one-period and have the following traits: (1) Members make a binding commitment to not attack each other. (2) Members make a binding commitment to go to war with any country at war with an alliance member, regardless of who initiates war. Many de jure defensive alliances often behave this way. NATO, a defensive alliance, acted

[4]Transfers are flow payments rather than asset transfers, though the model is robust to asset transfers.

[5]If one objects to unanticipated strength shifts and insists that countries assign some probability to the possibility of unanticipated strength shifts, then it is always possible to define a probability $\epsilon > 0$ that is sufficiently small that no part of the analysis changes. The presence of probabilistic strength shifts makes this a stochastic game with complete information similar to Fearon (2004). Actors know for certain the relevant probabilities involved in the game and thus there is no requirement to specify how they form beliefs.

offensively when intervening in Bosnia and Kosovo. We consider purely defensive alliances in the SI. (3) Alliance members commit to sharing resources according to a known bargaining protocol, ϕ (described below). The credibility of alliance commitments is assumed as is standard in setups with cooperative game elements, but may be attributed to the existence of sufficient audience costs when signing alliance contracts.[6]

In every period, countries choose whether to form alliances upon observing the initial game state $(\mathbf{x}, \mathbf{s})$ and any anticipated strength shifts.[7] Once countries form alliances, $\mathbf{a}$ denotes the alliance state. $\mathbf{a}$ can either indicate the no alliance state, denoted 0, or any state in which alliances occur, denoted by specifying the allied countries, ij.

We assume that an alliance forms only if the alliance state is *stable*, defined as follows[8]:

Definition 1. Alliance state $\mathbf{a} \neq 0$ is stable if there is not a strictly positive deviation for any allied player and the outside player to another state $\mathbf{a}'$. The no alliance state, $\mathbf{a} = 0$, is stable if no two players have a weakly positive deviation to an alliance state.

Definition 1 implies that any two countries can object to the no alliance state and form an alliance. Furthermore, any alliance member can object to an alliance state and move to the no alliance state or form an alliance with the out country. The possibility of a grand coalition is excluded since any grand coalition would mimic the payoffs of one of these other coalitions, but deny countries the option value of fighting a war.

Wars are costly as countries involved receive zero consumption during conflict. Including

[6]The commitment reflects the sense in which formal alliances are distinct from informal alignments. This assumption allows alliances to credibly threaten a unified front during bargaining, even when one country may prefer to renege on the alliance and some do renege in practice.

[7]We suppress whether war occurred previously because the game is trivial after war.

[8]This is not Nash Equilibrium, but a cooperative game notion of stability closely related to the core solution concept in partition function games. The application of cooperative game theory avoids the necessity of exactly specifying the action space for the alliance formation stage. Given the large set of plausible actions, non-cooperative approaches tend to lead to a large multiplicity of equilibria (see Garfinkel, 2004). We mix cooperative and non-cooperative game elements in the tradition of Konishi and Ray (2003).

additional costs beyond forfeiting consumption would not qualitatively impact the analysis. For expositional purposes, we assume that any country is physically capable of attacking any other country. Our results are flexible to dropping this assumption as shown in the SI which formalizes a network structure specifying which countries can attack each other. In a two country war, i wins with probability $p_i = s_i/(s_i + s_j)$. In a three country war without a coordinated alliance, i wins with probability $p_i = s_i/(s_i + s_j + s_z)$. When fighting against a coordinated alliance, i faces a penalty for fighting on two fronts of $\chi \in \{1, ..., \min\{s_1, s_2, s_3\}\}$, so that i's probability of victory becomes $p_i(\chi) = (s_i - \chi)/(s_i - \chi + s_j + s_z)$.[9] If one country is defeated, the victor acquires the defeated country's consumption and military resources.[10] Defeated countries receive a continuation payoff of 0.

Let w_i represent i's war value for continuing to fight until all opponents are defeated. When fighting two uncoordinated opponents at once, $w_i = \delta\frac{s_i}{s_i+s_j+s_z}$. This value is discounted by δ since consumption is lost in the one period of fighting. Similarly, when fighting two coordinated opponents at once, $w_i = \delta\frac{s_i-\chi}{s_i-\chi+s_j+s_z}$. Due to our assumptions about war values, in equilibrium, any war will involve all countries. The SI describes off-the-equilibrium path wars where one country attacks the two other countries in sequence.

In the absence of power shifts, transfers are made according to some specific bargaining protocol where ρ represents the protocol between alliance groups and ϕ represents the protocol within alliances. The proofs go through for any ρ and ϕ that obey a set of assumptions (specified in the SI) that allow for a broad class of reasonable bargaining protocols. We give up some degree of

[9]These particular contest success forms are not necessary for the results but ease exposition. All that is necessary is that war functions for a non-allied country are strictly increasing in its own strength and strictly decreasing in opponent strength (and χ when facing an alliance), while an alliance total war value is strictly increasing in total alliance strength and χ and strictly decreasing in opponent strength. The SI provides general results without the assumption of having to fight on two fronts.

[10]If more than one country defeats a single country, they receive an equal share of the resources. Different sharing rules would impact the outcomes in particular cases, but any sharing rule that does not violate our bargaining assumptions would not invalidate our propositions.

sharpness in our predictions so that our results are robust to a wide variety of game specifications. However, for ease of exposition, examples use the following specification: ρ gives proposer power to the country or alliance group that has the largest war value or splits the bargaining surplus evenly between countries that all have the weakly largest war values; under ϕ alliance members evenly divide any surplus beyond their no-alliance values.

We refer to the full game as Γ. Each period players make moves in the stage game, labeled $\Gamma(t)$, before moving to the next stage game $\Gamma(t+1)$. The only distinction in timing from the standard preventive war model (Powell 2006), is that alliances must form every period. The one complicating factor is whether sticky alliances form before countries know of a power shift or form after countries know of the power shift. We focus on the latter case since our primary interest is how anticipated power shifts alter alliance choices. In each period, countries observe the power distribution, learn of any looming power shifts, form alliances and determine transfers, and decide whether to initiate war. The following describes the timing of the stage game, $\Gamma(t)$, beginning in period t:

1. Countries observe the state of the system, $(\mathbf{x}, \mathbf{s})$.
2. (Flexible alliances with anticipated power shifts case) Power shifts are announced.
3. Countries form alliances and all countries observe the alliance state $\mathbf{a}$.
4. (Sticky fixed alliances case) Power shifts are announced.
5. Countries determine transfers.
6. Countries (in random order) determine whether to engage in war.
7. War resolves.
8. Transfers are made by countries not engaged in war.
9. Power shifts take place for countries not engaged in war.
10. Move to period $t+1$.

We restrict attention to Markov Perfect Equilibria (MPE). An MPE is a Subgame Perfect Equilibrium (SPE) such that players play Markov strategies. Markov strategies depend only on the payoff relevant state. Accordingly, past actions affect strategy only through their effect on the state

and players cannot explicitly condition on past actions.[11]

Analysis: Static Alliance Formation

To show how anticipated power shifts affect alliance dynamics, the analysis proceeds through three conditions. The first, covered in this section, analyzes the model without power shifts to provide a baseline against which to compare subsequent results. The second examines cases with (semi-)fixed alliances that countries cannot change if they learn of a power shift. These serve as counterfactuals that help solve the third condition in which countries form alliances accounting for anticipated power shifts. Showing the consequences of fixed alliances provides intuition for results with flexible alliances.

Begin with the model without power shifts. The analytical approach is straightforward: given a set of initial conditions, calculate the payoffs for each country in each possible alliance state. While intuitive, this process is intensive because it requires calculating 12 peace payoffs spanning across three countries and four possible alliance arrangements. Once completed, we construct an alliance-utility table that presents these payoffs and shows which alliance, if any, is optimal. These tables offer a visual means to present a large number of underlying calculations. Consider an example where $\mathbf{s} = (10,\ 6,\ 3)$, $\delta = 0.9$, and the penalty of fighting against coordinated opponents is $\chi = 2$. Since countries always have the outside option of war, it is necessary to first calculate each country or alliance group's war value in each possible circumstance. Start with the no alliance state where each country faces two uncoordinated opponents and less discounting occurs if a country faces both opponents immediately. War values in the no alliance state are given by $w_i = \delta s_i / (s_i + s_j + s_z)$. Using the war value function and initial strength parameters, one can find all war values in the no alliance state as $\mathbf{w}^0 \approx (0.474,\ 0.284,\ 0.142)$ where the superscript 0 represents the no alliance state. Since country 1 has the largest war value, it is able to extract the entire bargaining surplus, hence final values for this state are $\mathbf{v}^0 = (0.574,\ 0.284,\ 0.142)$. Transfers are made with the vector of net transfers $\tau \approx (0.241,\ -0.049,\ -0.191)$ since initial resources are $\mathbf{x} = (1/3,\ 1/3,\ 1/3)$.

[11]The SI argues why the MPE restriction is reasonable, though we acknowledge the relevance of historical relationships for alliance choices.

Having addressed the no alliance state, now consider alliance state $\{23\}$ where countries 2 and 3 are aligned. Country 1's war value is calculated as

$$w_1^{\{23\}} = \delta \frac{s_1 - \chi}{s_1 - \chi + s_2 + s_3} = 0.9 \frac{10-2}{10-2+6+3} \approx 0.423.$$

The alliance group's combined war value exceeds 1's war value and thus, according to ρ, the alliance group has proposer power and captures the bargaining surplus. According to ϕ, alliance members receive their no-alliance values plus an equal share of the surplus. Hence, 2 receives

$$v_2^{\{23\}} \;=\; v_2^0 + \frac{1}{2}\left[\left(1 - w_1^{\{23\}}\right) - \left(v_2^0 + v_3^0\right)\right].$$

One can similarly find 3's value and calculate the value vector for alliance state $\{23\}$ as $\mathbf{v}^{\{23\}} = (0.423,\ 0.360,\ 0.218)$. Repeating this procedure for each alliance, we construct the following alliance-utility table:

Table 1: Static power distribution with stable alliance state.

$\mathbf{s} = (10, 6, 3)$, $\chi = 2$, $\delta = 0.9$

	no alliance	$\{12\}$	$\{13\}$	**{23}**
v_1	0.574	**0.619**	0.611	0.423
v_2	0.284	0.329	0.211	**0.360**
v_3	0.142	0.053	0.179	**0.218**

Outcome: peace; balancing alliance $\{23\}$ forms

The bolded numbers represent a country's best outcome. We bold an alliance state if two countries most prefer that state and therefore form that alliance. We see that the balancing alliance, $\{23\}$, is the unique stable alliance state in this period. From Lemma 1, this is also the unique alliance that forms in the MPE of the repeated game.[12]

Lemma 1. *In the absence of strength shifts, there exists a stable alliance state, **a**, of the stage game $\Gamma(t)$. When **a** is the unique stable alliance state of the stage game, it forms every period in the MPE of the repeated game, Γ.*

[12]Lemma 1 extends the result in Krainin (2014) on static games to repeated games.

Analysis: Dynamic Alliance Formation

This section shows that alliance formation often depends on dynamic concerns. We demonstrate two propositions that characterize the novel implications of military strength shifts on war and consequently on alliance formation. The first subsection, with (semi-)fixed alliances, highlights how the presence of more than two actors can impact the conflict-causing potential of a strength shift. It also provides a necessary step toward understanding alliance dynamics when shifts are anticipated. It demonstrates how war may occur when alliances are initially immutable and non-anticipatory. The following subsection focuses on our main result. When alliances are flexible, anticipated strength shifts cause alliances to shift today which can allow states to avoid the losses of war. In such cases, power shifts can dominate the distribution of power in determining which alliances form.

Fixed Alliances and Shifting Power

In this subsection, states form sticky alliances that they cannot change if they learn a power shift will occur. Alternatively, we can think of power shifts as unanticipated. Laying out the consequences and shortcomings of sticky alliances illuminates the logic motivating states to alter alliances when they do anticipate power shifts (our central concern). Specifically, we show that fixed alliances can cause war in the presence of expected power shifts and that the same is true of alliances that can only change retrospectively.

Sticky alliances come in two varieties. First, states may form *fully-fixed* alliances that they cannot change, even in the next round after power shifts occur. Binding alliances are functionally equivalent to a single country. With these stipulations, the conditions for war are exactly analogous to Powell (2004). A sufficient condition for country or alliance group i to go to war is $w_i^{\mathbf{a}}(\mathbf{s}) > (1-\delta)\, c_i + \delta\left[v_i^{\mathbf{a}}(\mathbf{s}')\right]$. As in typical commitment problem models with shifting power, a country or alliance group i goes to war when its adversary's strength shifts sufficiently upward—i opts to obtain its current war value rather than any bargain today plus the discounted stream of its future bargains.

Second, states may form *semi-fixed* alliances that cannot change in anticipation of power shifts but can change after power shifts occur. A conception of alliances as semi-fixed animates much of the existing scholarship on alliance dynamics which evaluates how alliances respond after power changes (Morrow, 1991; Leeds and Savun, 2007). In the model, optimal alliances may change after a strength shift. A revised sufficient condition for war allows for strength shifts and subsequent alliance shifts:

$$w_i^{\mathbf{a}}(\mathbf{s}) > (1-\delta)\, c_i + \delta \left[v_i^{\mathbf{a}'}(\mathbf{s}') \right]. \tag{1}$$

The distinction between inequality (1), and the earlier condition for war is that (1) allows both the alliance state and the strength state to change. Alliance partners are not stuck together after the first period and may respond to strength shifts by altering alliances. Alliance changes in response to strength shifts create a potential secondary or indirect effect to any strength shift.

Having specified the war conditions for scenarios with fully- and semi-fixed alliances, we now illustrate by example how alliance changes *after* power shifts—that is, semi-fixed alliances—can alter outcomes versus what occurs when alliance switching is precluded. There are two cases to examine: the alliance change mitigates or exacerbates the risk of war. Evaluating these cases shows that the possibility of changing alliances after a strength shift can both restore peace when war would otherwise occur or prompt war when peace would otherwise prevail. The latter possibility is vital for our main result, discussed later, which shows that countries have incentives to change alliance arrangements *before* strength shifts occur.

First consider a scenario where alliance changes based on retrospective power shifts produce peace when fully-fixed alliances cause war. Alliance utility tables in Table 2 illustrate this scenario with a power shift of $(i, z) = (2, 3)$, with $\chi = 1$ and $\delta = 0.9$. Outcomes for each alliance scenario are summarized beneath.

Table 2: Demonstrating incentives against fully-fixed alliances.

Before shift: $\mathbf{s} = (4, 3, 2)$

	no alliance	$\{12\}$	$\{13\}$	**{23}**
v_1	0.500	**0.544**	0.538	0.338
v_2	0.300	0.344	0.225	**0.381**
v_3	0.200	0.113	0.238	**0.281**

After shift: $\mathbf{s} = (4, 6, 2)$

	no alliance	$\{12\}$	**{13}**	$\{23\}$
v_1	0.3	0.334	**0.371**	0.245
v_2	0.55	**0.584**	0.409	0.578
v_3	0.15	0.082	**0.221**	0.177

Outcomes

- *Fully-fixed alliances*: war; alliance $\{23\}$ forms and 1 initiates a preventive war against it
- *Semi-fixed alliances*: peace; alliance $\{23\}$ initially forms, 3 then balances 2's rise by forming alliance $\{13\}$ after the strength shift

As shown by the bolded column in the left utility table, countries 2 and 3 form an alliance given the initial conditions. If countries 2 and 3 are stuck together in a fixed permanent alliance, then war results. Country 1 prefers preventive war in this scenario since $0.338 > (1 - 0.9) + 0.9\,(0.245) \approx 0.321$. The 0.338 on the left hand side represents country 1's war value while the right hand side represents the maximum value country 1 could be guaranteed in peace—where the first term represents the present value of receiving the entire pie today and the second term represents 1's long-term peace payoff when confronting the $\{23\}$ alliance.

However, when allowing for alliance changes after the strength shift, 3 will switch to being allied with 1 in the second period. Intuitively, the $\{13\}$ alliance represents an effort to balance against the rising power. To see that a peaceful bargain exists (unlike the earlier case), calculate the consumption demands in the period before the shift:

$$\begin{aligned} c_1 &\geq \frac{1}{(1-0.9)}\left[0.338 - 0.9\,(0.371)\right] = 0.4 \\ c_2 &\geq \frac{1}{(1-0.9)}\left[0.331 - 0.9\,(0.409)\right] = -0.4 \\ c_3 &\geq \frac{1}{(1-0.9)}\left[0.231 - 0.9\,(0.221)\right] = 0.3. \end{aligned}$$

Consumption must be positive by assumption, so any transfer system $c_1 \geq 0.4$, $c_2 \geq 0$, and $c_3 \geq 0.3$ is peaceful. Moreover it is feasible as total resources (equal to 1) exceed total consumption. The alliance shift offsets the strength shift and restores feasible peaceful bargains. Unalterable alliances induce war while alliances that can change after strength shifts occur foster peace. The

implication is critical. The countries collectively prefer a system where alliances are not fixed. Absent strong counter forces that keep countries locked into existing alliances, strategic incentives make rigid alliance arrangements unsustainable.

Having shown a scenario demonstrating that fully fixed alliances are contrary to countries' interests, we now present a scenario highlighting deficiencies with arrangements that only allow states to alter alliances *after* strength shifts. In this case, allowing alliance shifts based on retrospective evaluation of power shifts prompts war when fully fixed alliances do not. The example in Table 3 and outcome summary below demonstrates a case of alliance switching exacerbating the risk of war with a strength shift $(i, z) = (2, 2)$ with $\chi = 50$ and $\delta = 0.9$.

Table 3: Demonstrating incentives against retrospective alliance changes and for anticipatory alliance changes.

Before shift: $\mathbf{s} = (100, 99, 98)$

	no alliance	$\{12\}$	$\{13\}$	**{23}**
v_1	0.403	**0.464**	0.4635	0.182
v_2	0.300	0.361	0.179	**0.411**
v_3	0.297	0.175	0.358	**0.408**

After shift: $\mathbf{s} = (100, 101, 98)$

	no alliance	$\{12\}$	**{13}**	$\{23\}$
v_1	0.301	0.362	**0.411**	0.181
v_2	0.404	**0.465**	0.184	0.464
v_3	0.295	0.173	**0.405**	0.355

Outcomes

- *Fully-fixed alliances*: peace; alliance $\{23\}$ forms
- *Semi-fixed alliances*: war; rising power 2 attacks before alliance $\{13\}$ can form after 2's rise
- *Flexible anticipatory alliances*: peace; alliance $\{13\}$ immediately forms to check 2's rise

There would be no war after the power shift if the initial alliance, $\{23\}$, were fixed. A bargain of the following form avoids war:

$$\begin{aligned} c_1 &\geq \frac{1}{(1-0.9)}\left[0.182 - 0.9\,(0.182)\right] = 0.18 \\ c_2 &\geq \frac{1}{(1-0.9)}\left[0.361 - 0.9\,(0.464)\right] = -0.57 \\ c_3 &\geq \frac{1}{(1-0.9)}\left[0.358 - 0.9\,(0.355)\right] = 0.38 \end{aligned}$$

which is feasible since the minimum positive consumption values sum to less than 1. Permanent alliances facilitate peace in this case.

However, when alliance shifts can occur based on retrospective evaluation of power dynamics,

from alliance $\{23\}$ to $\{13\}$, the best bargain 2 can receive is $(1-0.9)+0.9\,(0.184) \approx 0.266$ which is less than 2's war value before the shift of 0.361.

In words, 2's best bargain in the semi-fixed scenario would be receiving the present value of the entire pie in the first period plus its peace payoff when confronting alliance $\{13\}$ after the shift. This value is less than 2's war value before the alliance shift. Consequently, the rising power 2 initiates war because it recognizes that a balancing alliance will form against it after the power shift and thus limit its bargaining leverage. The example demonstrates that when peace would prevail with fully fixed alliances, a post-strength shift alteration in alliances can cause war. Whereas Table 2 showed deficiencies of fixed alliances, Table 3 shows deficiencies of alliances that change retrospectively. Together, it is clear that overwhelming institutional forces are necessary to sustain either type of alliance given countries' incentives to optimize decisions about war and peace.

Proposition 1 summarizes these results, comparing the effect of strength shifts when alliances are fully fixed to when alliances can change retrospectively.

Proposition 1. *(War with Sticky Alliances) Consider an exogenous, one period, one country,* ***unanticipated****, positive strength shift where alliances can shift after one period compared to the case where alliance groups are permanently treated as a single country.*

1. *(Direct effect of a strength shift) For a given strength shift: (a) If non-allied country i initiates war when alliance shifting is possible, then i also initiates war when alliance shifting is not possible. (b) The converse does not hold; allowing alliance shifting could avoid war.*

2. *(Indirect effect of a strength shift) For a given strength shift: (a) War may occur when alliances can shift, but not occur when alliances are fixed. (b) There exists a nonempty set of parameters where a rising power initiates war when alliance shifts are possible.*

Proposition 1 illustrates one reason why simplifying international systems to a two country setting can mislead. Proposition 1.1 demonstrates that two country models (or three country models where alliance partners are permanently fixed) may predict war when no war would occur when

accounting for alliance shifts (as shown in Table 2). Proposition 1.2 demonstrates the opposite: simpler models may fail to predict war when it would occur in a model where retrospective alliance shifts are possible (as shown in Table 3). The latter case is essential for our main result. Countries have strong incentives to alter alliances *before* strength shifts when those strength shifts would otherwise later induce a country to change alliances which increases the risk of preventive war.

Flexible Alliances and Anticipated Power Shifts

We now consider flexible alliances that can change *before* power shifts and show by example that the distribution of power is not a sufficient statistic to determine alliance formation with anticipated power shifts. Dynamic effects can dominate static effects in determining alliance choices.

Consider the situation analyzed in Table 3 which produced war when alliances could only change after power shifts (see Proposition 1.2). We now examine when countries anticipate the strength shift and can change alliances *before* the shift. Forming alliance $\{13\}$ before the power shift occurs salvages peace. Country 3's choice is pivotal. In a setting with semi-fixed alliances we know from above that 3 joins alliance $\{23\}$, goes to war, and secures a war value of 0.358. If 3 instead opts to initially form $\{13\}$, it produces peace values before the shift of

$$\begin{aligned} v_1 &= (1-0.9)\,0.4635 + 0.9\,(0.411) \approx 0.416 \\ v_2 &= (1-0.9)\,0.179 + 0.9\,(0.184) \approx 0.184 \\ v_3 &= (1-0.9)\,0.358 + 0.9\,(0.405) \approx 0.400 \end{aligned}$$

All countries prefer peace under alliance $\{13\}$ to their war values: $w_1 = 0.414$, $w_2 = 0.179$, and $w_3 = 0.308$. Critically, 3 prefers peace under $\{13\}$ to war under $\{23\}$ ($0.400 > 0.358$). Substantively, the case illustrates that countries can preserve peace by rapidly balancing against a rising power before that rising country attacks isolated members of the balancing coalition. An anticipated power shift alters optimal alliances. When able to form alliances in anticipation of looming strength shifts, alliance arrangements exist such that all countries prefer peace to war. This example demonstrates the following proposition.

Proposition 2. *(Dynamic Alliances) When countries anticipate strength shifts at the time of alliance formation, there exists a nonempty set of parameters where the optimal static alliance does not form.*

Proposition 2 indicates that optimal alliances when strength shifts were absent, minimal, or unanticipated may not be optimal when strength shifts are looming and known. Proposition 1.2 establishes conditions when the ability to switch alliances after a strength shift produces war, whereas Proposition 2 establishes that cases exist where those countries can avoid war if they anticipate the strength shift and alter alliances beforehand.

One implication is that the likelihood of alliance changes increases when moving from a world with fixed power to one with anticipated changes. Corollary 1 formally establishes this point. We can interpret the likelihood of alliance changes as the size of the parameter set for which an alliance shift occurs.

Corollary 1. There exists a $Z > 0$ such that the likelihood of an alliance shift is weakly increasing and sometimes strictly increasing in the absolute value of an anticipated strength shift for the range $[0, Z]$.

Alliance shifts becomes more probable (or at least no less probable) as the size of an anticipated strength shift increases. The precise value of Z depends on multiple parameters, but is sufficiently large as to be essentially irrelevant for empirical purposes (e.g., requiring a country's strength to more than double). Barring extreme values, greater anticipated strength shifts increase the likelihood of alliance shifts. Corollary 1 provides a falsifiable expectation which motivates the empirical analysis.

Dynamic Alliances in Practice

The empirical analysis emphasizes one model implication: greater anticipated strength shifts increase the likelihood of alliance changes. This section specifies the testable implication before operationalizing concepts. Both the power distribution *and* power dynamics affect alliance formation and termination. For a given power distribution, states form optimal alliances and maintain

them so long as power remains static (Lemma 1). If states anticipate that at least one state's power will change, then the optimal alliance also may change (Proposition 2). Corollary 1 clarifies the relationship between these points. Moving from a static system to one with looming strength shifts can induce alliance changes. Intuitively, the size of the anticipated strength shift matters. Minuscule changes in strength, akin to static environments, are unlikely to sufficiently alter war payoffs and induce alliance change. Larger strength shifts produce more dramatic changes in war payoffs. When states anticipate sizable changes, the probability that optimal alliance arrangements change increases, causing states to terminate alliances or form new ones. A testable hypothesis follows: *An increase in expected strength shifts increases the probability of alliance changes*.[13]

Because it is new and operationalization is feasible, we emphasize the relationship between forward-looking power dynamics and alliance changes as opposed to other implications from the model. We focus on alliances rather than war for several reasons. First, as in many theories of alliances, the threat of war motivates alliance behavior because states seek to avoid inefficient conflict. While the possibility of war looms over the interaction, alliance behavior rather than war is the outcome of interest. Second, the theoretical model does not yield a simple monotonic relationship predicting war. Amid space constraints, testing this complex relationship requires diluting attention from the alliance results.

By operating within the power-centric conceptions of alliances, our approach predicts alliance changes when states anticipate strength changes even if states have opposing interests. Factors outside the theoretical model, such as shared interests, of course are likely to affect alliances in practice. Unaligned preferences may limit—though not eliminate—enthusiasm for forging alliances which could dampen the effect of strength shifts.

Operationalizing Concepts

Mapping the hypothesis to data involves three steps: specify the relevant states, measure anticipated power shifts, and capture alliance changes. First, who are the players? We operationalize a set as a geographic region—the Americas, Europe, Sub-Saharan Africa, Middle East and North

[13]The analysis discusses the empirical relevance of Z from Corollary 1.

Africa, and Asia-Pacific. The analysis uses two approaches for region membership. A first includes all states geographically within the region plus all major powers (Correlates of War Project, 2017). Asia-Pacific during the modern era includes the US and other major powers on the premise that these actors influence regional dynamics. A second approach, restricts membership based on geography and excludes external major powers. Regions provide an intuitive grouping of states that monitor one another's power dynamics and consider one another as potential allies. The appropriateness of a non-global framework likely varies with time, which we address in the analysis.

A regional approach offers advantages over a strict three-player approach, which poses several problems. Many three-state configurations are politically irrelevant with no plausible alliance-formation possibilities. Moreover, alliances rarely reflect strictly three-player considerations because they are multilateral (26% of alliances) or target more than one state (68% of alliances).[14] Regions, in contrast, often contain all alliance members and their targets.

Despite each region consisting of more than the modeled three players, the hypothesis likely remains appropriate. While difficult to show analytically, the intuition is straightforward. Alliance-utility tables rapidly grow in complexity as the number of players increases. Finding analytical solutions after scaling to a 40-state setting is beyond this paper's scope. Nonetheless, we expect the basic insight holds. With no expected power shifts, some set of optimal alliances exists. Provided the power distribution remains stable, this optimal set likely holds. In contrast, when states anticipate meaningful strength shifts, the optimal alliance set likely changes. As expected strength shifts increase, so does the probability of alliance changes.

Region-year is the unit of analysis with year representing a period from the theory. Regions vary in their temporal spans based on data availability and alliance behavior. The sample starts from the year in which the first regional alliance formed.[15] Across the five regions there are 655 observations with Europe contributing the most (183) and Sub-Saharan Africa the least (43). When

[14]Percentages based on alliances with an offensive or defensive component (Leeds et al., 2002).

[15]Robustness tests in the SI include years preceding the first regional alliance and find similar results.

excluding major powers from regions, the number of observations drops (526) because the first alliance featured an external major power.

The second step measures the explanatory variable: anticipated strength shifts. We adopt the approach from Bell and Johnson (2015), which provides the methodological details. It models the relationship between observable state characteristics from time $t-1$ and that state's military strength at time t. It then uses coefficients from the models and plugs in values of observable characteristics in t to generate predicted military strength for $t+1$. The outcome variable takes the average of a state's share of global military personnel and expenditures, which are two components of CINC scores (Singer, Bremer and Stuckey, 1972). The predictors include economic potential (represented by the other CINC components), the log of that year's battle deaths (Sarkees and Wayman, 2010), regime type (Marshall and Jaggers, 2002), an indicator for whether a state has alliance obligations that would demand military support if triggering conditions were met (Leeds et al., 2002)[16], and an indicator for strategic rivalries. As Bell and Johnson (2015) describes, rivalry captures whether a state is under threat, which can stem from clashing interests or the rival's perceived strength. Theoretically, states may expand their capabilities to address the threat that rivals present. Consistent with expectations, states with strategic rivals tend to increase their military strength. Bell and Johnson (2015) posits that states limit military spending if they have alliance partners. However, the indicator for alliance obligations has limited association with changes in power. The generative model employs OLS with country-year as the unit of analysis and country fixed effects.

We take the expected change in each state's military power: country i's estimated capabilities in $t+1$ minus actual capabilities in t, or $\widehat{Mil_{it+1}} - Mil_{it}$. ***Expected Power Shift*** is the mean absolute value expected change of states in a region, or $\frac{1}{n}\sum_{i=1}^{n}|\widehat{Mil_{it+1}} - Mil_{it}|$, for a region with n states. Because power dynamics alter war payoffs and thus optimal alliance configurations, we are interested in power increases *and* decreases. The measure grants more weight to powerful actors as they produce larger shifts in war payoffs and thus optimal alliances. Using the region

[16]States without alliances are coded 0.

mean captures the general level of anticipated strength shifts. Robustness tests using the *maximum* anticipated shift yields similar results. When including major powers in regions, *Expected Power Shift* has a mean of 0.33 with higher values indicating more extreme anticipated shifts.

The third step constructs a measure for regional alliance changes—the outcome variable. We use only alliances with offensive or defensive provisions because the model does not offer clear implications for alternative agreements (Leeds et al., 2002). We identify all alliances for a given region. When the region includes external major powers, alliances must have at least one country within the geographic region. An alliance strictly between European major powers does not count as an Asia-Pacific alliance. To assure a match between the unit of analysis and the outcome variable, we exclude inter-regional alliances if the outside actor is not a major power.[17] We calculate the number of regional alliances at the start of year t, formed during t, and terminated during t. *Alliance Changes* represents the sum of regional alliances formed and terminated because both reflect alliance dynamics.[18] On average, 0.70 alliances change per region-year. Figure 1 shows alliance shifts across time.

The models include a set of control variables that may affect *Alliance Changes* or be confounding variables for *Expected Power Shift* and *Alliance Changes*. *Alliance Start Count* measures the number of alliances at the year's start. More existing alliances increases the number of potential alliance terminations. We calculate the variable as described above. *Number of States* indicates the number of states in a given region-year. Alliance changes may increase with the number of potential partners. Having more *Major Powers* within a region may increase *Expected Power Shift* as their expected strength changes are larger than those for small states. It may also affect the frequency of alliance changes if multipolar systems are associated with alliance jockeying. *Year* accounts for possible time trends. Empirically, alliance changes become increasingly frequent in

[17]NATO does not qualify for inclusion by this criteria because it includes non-major powers from several regions, such as Canada, Belgium, and Turkey. Given the importance of NATO, a robustness test alters the qualification criteria to include NATO and results are unchanged.

[18]Formation and termination pertain to the entire alliance's start and end dates as opposed to changes in alliance phases or membership (Leeds et al., 2002).

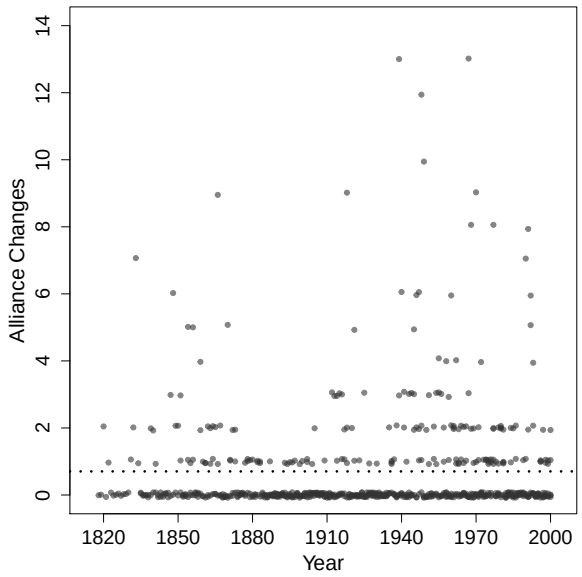

Figure 1: *Alliance Changes* by region-year when region includes major powers. Jitter added for readability; dashed line indicates mean.

later years when conditioning on the region. Results remain similar when including a quadratic term for time trends. *Regional Uncertainty* offers a summary measure of a region's power distribution (Bas and Schub, 2016). Its value increases with greater power parity and decreases with a more strict hierarchy among states. When lower, indicating a relatively unbalanced power distribution, states may balance powerful actors through internal arms acquisition and external alliance formation, thus increasing *Expected Power Shift* and *Alliance Changes*. The SI provides all summary statistics.

Results: Strength Shifts Produce Alliance Shifts

Figure 2 plots the explanatory (*Expected Power Shift*) and outcome (*Alliance Changes*) variables with a fitted line and 95% confidence intervals from an OLS specification with control variables. As hypothesized, greater expected strength shifts are associated with more alliance changes. When the mean anticipated strength shift for states within a region is small, few alliance changes occur. States have formed the optimal alliances and the optimal set does not change. In contrast, when power dynamics become more pronounced, the optimal alliances changes.

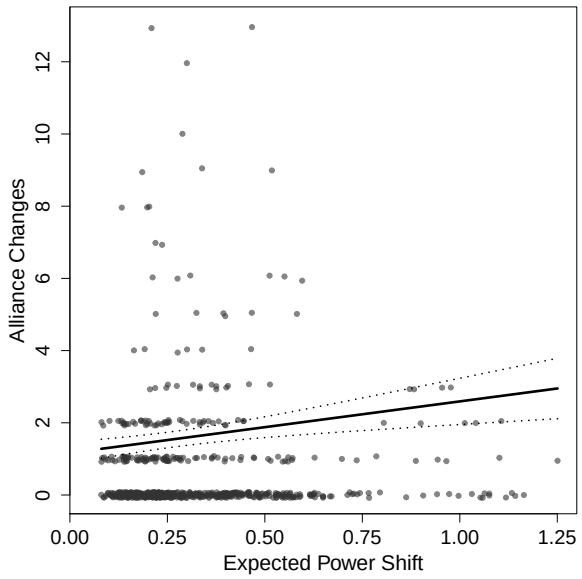

Figure 2: *Expected Power Shift* and *Alliance Changes* with fitted line and 95% confidence intervals from Table 4 Model 3. Jitter added for readability.

An historically familiar region and era help illustrate the relationships. Consider European power dynamics and alliance behavior from the Congress of Vienna until the Crimean War. This period of stability saw few dramatic power changes or intra-European wars. The data follow this pattern. CINC scores for the premier European powers (UK, France, Prussia, Austria-Hungary, Italy, and Russia) exhibit minimal flux. Appropriately, the *Expected Power Shift* values for the region are relatively low (mean=0.22). As expected, there are few alliance changes during this period (mean=0.9; 66% with zero changes). In contrast, from the Crimean War to the Franco-Prussian war is popularly known for its volatility. Stark power shifts are evident in CINC scores with Prussia and Italy rising while Austria-Hungary and Russia fall. *Expected Power Shift* captures this volatility (mean=0.32). With larger anticipated strength shifts, we expect and find that the number of alliance changes (mean=2.1; 39% with zero changes) significantly exceeds the earlier period's number.

Regression specifications ensure that the relationship holds after accounting for other potential causes of alliance changes. Table 4 presents results from models using OLS with the region-year as the unit of analysis and number of *Alliance Changes* as the outcome. Models 1-5 employ regions

that include external major powers while Models 6-8 exclude them. Models 2-5 and 7-8 include region fixed effects to account for unobserved time-invariant factors that may cause the frequency of alliance changes to differ given differences in the international politics across these regions (Kang, 2003).

Greater anticipated changes in capabilities increase the expected number of alliance changes within a region. This holds with and without control variables and with and without fixed effects with only one exception. When states expect the regional balance of power to substantially change in the following year, the likelihood of alliance termination and/or formation increases. The sole exception is Model 1's bivariate specification that excludes fixed effects. Absent fixed effects, the pooled sample makes inter-regional comparisons which might be inappropriate. As the SI details, the null result primarily stems from the inclusion of 120 Middle East and North Africa observations during a period in which external major powers outnumber local states and no fully local alliance exists.[19] A high ratio of external major powers to local states increases expected power shifts and decreases alliance changes because there are fewer potential local allies. Our main result holds when excluding these observations from the sample and even holds *within* the subsample of these 120 observations. Unique elements of this region period highlight problems with inter-regional comparisons. Once accounting for unmeasured regional differences, the results corroborate the theoretical expectations. Models 4 and 5 separately examine alliances formed and terminated. Higher anticipated strength shifts are associated with both more alliance formation and termination.

Returning to the main analysis (Model 3), consider the effect within the European region of increasing *Expected Power Shift* from its 5th to 95th percentile values (0.12 and 0.77) when holding other variables at their means. At the lower value, when states anticipate stasis, the model predicts 1.7 alliance changes. At the higher value, when states anticipate sizable alterations to the regional power distribution, we expect 2.6 alliance changes. The marginal effect of a 0.9 increase

[19]The first alliance between an external major power and local actor formed in the 1830s. The first local alliance formed in the 1950s.

Table 4: Anticipated Strength Shifts and Alliance Dynamics

	Region + Major Powers					Region		
	Total Changed			Formed	Terminated	Total Changed		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Expected Power Shift	0.15	0.76**	1.43***	0.69***	0.74***	1.87***	1.19**	2.61***
	(0.31)	(0.32)	(0.42)	(0.26)	(0.26)	(0.50)	(0.54)	(0.67)
Year			0.004*	0.003**	0.001			0.002
			(0.002)	(0.001)	(0.001)			(0.003)
Alliance Start Count			0.03	−0.02	0.05***			0.02
			(0.02)	(0.01)	(0.01)			(0.02)
Number Major Powers			−0.12	−0.16**	0.03			−0.20
			(0.11)	(0.07)	(0.07)			(0.12)
Number States			−0.02	−0.02	−0.002			−0.02
			(0.02)	(0.01)	(0.01)			(0.02)
Regional Uncertainty			−2.46**	−1.91***	−0.55			−3.51**
			(1.02)	(0.64)	(0.63)			(1.70)
Constant	0.65***	1.34***	−4.81	−3.60	−1.21	0.37***	1.18***	0.31
	(0.12)	(0.14)	(3.92)	(2.45)	(2.42)	(0.11)	(0.17)	(4.85)
N	655	655	655	655	655	526	526	526
Region FEs	N	Y	Y	Y	Y	N	Y	Y
Outcome Mean	0.70	0.70	0.70	0.37	0.33	0.68	0.68	0.68

$^{*}p < 0.1$ $^{**}p < 0.05$ $^{***}p < 0.01$

Notes: OLS with region-year as unit of analysis; region includes all major powers in M1-5 and excludes them in M6-8. Standard errors in parentheses. Outcome is total number of alliance changes (M1-3; M6-8), alliances formed (M4), and alliances terminated (M5). Region fixed effects not shown.

in alliance changes ($\pm$ 0.5 at the 95% confidence interval) represents a roughly 55% relative increase. Expected strength shifts and actual alliance jockeying in 1921 Europe typify the empirical pattern. The aggregate anticipated power shift that year was relatively high (75th percentile) due to large expected changes in Russian strength after the Polish-Soviet War and in France due to rising economic output. Using the observed values from the region-year, Model 3 predicts nearly two alliance changes, which is a high number within the data set (91st percentile). In reality, states formed four new intra-regional alliances that year.

Robustness: Era, Retrospection, and War

Tests in the SI demonstrate the robustness of the finding across additional specifications.[20] These include models using robust standard errors[21] and negative binomial models. Results are similar when evaluating anticipated *persistent* shifts as opposed to single year changes (Krainin, 2017). We reconstruct the explanatory variable by estimating each state's military capabilities three years out rather than one year out. Figure 2 shows many observations clustered around low values for the explanatory and outcome variables which could contribute to the positive relationship between the two. The joint occurrence of low values on the main variables is itself evidence supporting the hypothesis, but we nonetheless wish to show that these observations are not driving the results. SI specifications dropping observations with expected power shift values in the lowest quartile show the core relationship holds. Finally, results hold when accounting for possibly very large anticipated strength shifts that might exceed parameter Z in Corollary 1. Extreme strength shifts can swamp the effect of alliance changes at which point the hypothesis no longer applies. While we cannot specify Z across observations, it is apparent in Figure 2 that large strength shifts produce fewer alliance changes than the model expects. The model's over prediction of alliance changes is consistent with extreme anticipated strength shifts negating the benefits of alliance changes.

Table 5 shows results are robust to several substantive considerations. Limiting the sample to pre-WW1 (Models 1 and 5), when intra-regional concerns likely superseded inter-regional concerns, produces similar findings. Accounting for retrospective power shifts, which are related to alliance termination and might also cause states to anticipate future power shifts, does not alter the substantive finding (Models 2 and 6). Nor does controlling for the number of regional actors—restricted to countries geographically within the region, excluding external major

[20]The SI presents results using regions that include major powers. Results are generally even stronger when excluding major powers.

[21]We do not cluster standard errors on the region because there are only five clusters. Rejection rates substantially exceed 0.05 when using few unbalanced clusters (Cameron and Miller, 2015, p. 341).

Table 5: Robustness Tests: Era, Retrospection, and War

	Region + Major Powers				Region			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Expected Power Shift	1.63***	1.35***	0.74*	1.02**	2.74***	2.06***	1.71**	1.83***
	(0.53)	(0.51)	(0.42)	(0.41)	(0.95)	(0.73)	(0.68)	(0.67)
Retrospective Power Shift		0.03				0.36*		
		(0.10)				(0.19)		
War			0.19***				0.15***	
			(0.03)				(0.03)	
War Termination				0.30***				0.25***
				(0.04)				(0.05)
Constant	5.77	−4.76	−4.07	−3.87	3.27	1.00	2.94	2.45
	(7.38)	(3.92)	(3.80)	(3.78)	(15.18)	(4.86)	(4.80)	(4.76)
N	264	655	655	655	179	526	526	526
Region FEs	Y	Y	Y	Y	Y	Y	Y	Y
Controls	Y	Y	Y	Y	Y	Y	Y	Y
Sample	Pre-1914	All	All	All	Pre-1914	All	All	All
Outcome Mean	0.46	0.70	0.70	0.70	0.61	0.68	0.68	0.68

$^{*}p < 0.1$ $^{**}p < 0.05$ $^{***}p < 0.01$

Notes: OLS with region-year as unit of analysis; region-year includes major powers in M1-4 and excludes them in M5-8. Standard errors in parentheses. Outcome is number of alliance changes. Region fixed effects and controls not shown.

powers—involved in any war as coded in Reiter, Stam and Horowitz (2014). War could confound the relationship of interest as belligerents increase their military strength and may form new alliances or break old ones. Even after controlling for war participation, anticipated strength shifts still have a significant effect (Models 3 and 7). War termination is also likely a common cause of anticipated power shifts (defeated parties are disarmed) and alliance dynamics (war-winning alliances dissipate) (Bennett, 1997). Results remain similar when including a variable counting the number of states participating in a war ending that year (Models 4 and 8). Other theories attribute alliance choices to states' foreign policy interests. To account for this possibility, SI robustness tests show that the results hold using static and dynamic measures of state interests. The static measure captures the variance of interests within a region while the dynamic measure reflects the retrospective change in those regional interests over the preceding year. The analysis uses United National General Assembly ideal point scores and is thus limited to after World War II (Bailey, Strezhnev and Voeten, 2017). A range of quantitative tests consistently suggest that anticipated

strength shifts alter the optimal set of alliances. States frequently end existing alliances or seek new ones to adapt to the changing landscape of power.

Conclusion

States look to the future, not only the past and present, when forming new alliances and exiting old ones. In addition to explaining bargaining failures, anticipated shifts in military strength play a vital role in explaining alliance dynamics. Rather than wait for strength shifts to unfold, states anticipate shifts when possible and prospectively alter alliance configurations to maximize their bargaining position and minimize the threat of war. A novel theoretical model that incorporates power dynamics, three players, bargaining, and flexible roles for all actors generates the hypothesized relationship. Whereas prior work privileges a single state's alliance choices—such as a hegemon (Gilpin, 1981)—all states make choices and no configurations are precluded in our model. We find that future power dynamics can supersede existing power distributions in swaying states' alliance choices. A testable implication follows: as the size of anticipated shifts in military strength increases, the number of expected alliance changes grows. Empirical specifications using the region-year as the unit of analysis corroborate the implication. When states anticipate few power shifts among the region's actors, alliance arrangements rarely change. As expected strength shifts in a region increase, the number of alliance changes grows.

Our results suggest two ways that alliance dynamics might contribute to peaceful relations. First, anticipated power shifts mitigate risks of balancing coalitions forming in the short term. Consider the general lack of hard balancing against the US after the Cold War. While it is possible a coalition including China could effectively balance the US and provide bargaining gains for those chafing under US hegemony, this study suggests that China's rise relative to the US could disincline countries against such balancing. Aligning with the US and dynamically balancing China may improve the chances of peace compared to static balancing against the US. This operates by limiting US incentives to preventively attack a balancing coalition before it coalesces or preventively attack an unchecked rising China. While we believe this logic applies, dynamic balancing is one force among many related to this phenomenon (Ikenberry, Mastanduno, and Wohlforth, 2009;

Gowa and Ramsay, 2017). Second, the findings suggest that alliance flexibility can foster peace. Alliance changes lower concerns about future power shifts and dampen the appeal of preventive conflict. Deeply institutionalized alliances with binding provisions—such as commitments to fight without consultation (Chiba, Johnson and Leeds, 2015)—help states signal to adversaries. However, these signaling gains come at the expense of alliance flexibility which allows states to alter alliance arrangements in ways that can promote peace.

This paper opens several pathways for future research. Its lessons may travel to other settings with potential for violence beyond interstate behavior, such as revolutions or civil wars. For theory, the formal model provides a framework for addressing other alliance topics. The model accommodates different network structures which could represent geographic linkages or force projection capabilities. War and alliance changes could emerge due to network shifts, even absent power shifts. Empirically, subsequent work can incorporate *anticipated* changes in other attributes, such as interests, that may drive alliance shifts.

The findings show the importance of looking beyond dyads and accounting for expected *future* dynamics when studying states' choices. On the first, states frequently consider multiple actors when approaching strategic interactions. Scholars can productively follow this lead. Alliance outcomes reflect multilateral concerns given that alliances include minimally two parties plus an external threat that the alliance aims to check. On the second, states are typically forward-looking and optimize choices based on their expectations for the future which might diverge from the present conditions. When beliefs about the future are in flux, alliance arrangements may prove particularly unstable. States' decisions to accept or discard formal international obligations depend upon not only how matters stand today, but also how matters might stand tomorrow.

References

Bailey, Michael, Anton Strezhnev and Erik Voeten. 2017. "Estimating Dynamic State Preferences from United Nations Voting Data." *Journal of Conflict Resolution* 61(2):430–456.

Bas, Muhammet and Robert Schub. 2016. "How Uncertainty about War Outcomes Affects War Onset." *Journal of Conflict Resolution* 60(6):1099–1128.

Bell, Sam and Jesse Johnson. 2015. "Shifting Power, Commitment Problems, and Preventive War." *International Studies Quarterly* 59(1):124–132.

Bennett, Scott. 1997. "Testing Alternative Models of Alliance Duration, 1816-1984." *American Journal of Political Science* 41(3):846–878.

Cameron, A. and Douglas Miller. 2015. "A Practitioner's Guide to Cluster-Robust Inference." *Journal of Human Resources* 50(2):317–372.

Castillo, Jasen and Alexander Downes. 2020. "Loyalty, Hedging, or Exit: How Weaker Alliance Partners Respond to the Rise of New Threats." *Journal of Strategic Studies* pp. 1–42.

Chiba, Daina, Jesse Johnson and Brett Ashley Leeds. 2015. "Careful Commitments: Democratic States and Alliance Design." *The Journal of Politics* 77(4):968–982.

Correlates of War Project. 2017. "State System Membership List, v2016." http://correlatesofwar.org.

Crescenzi, Mark, Jacob Kathman, Katja Kleinberg and Reed Wood. 2012. "Reliability, Reputation, and Alliance Formation." *International Studies Quarterly* 56(2):259–274.

Fang, Songying, Jesse Johnson and Brett Ashley Leeds. 2014. "To Concede or To Resist? The Restraining Effect of Military Alliances." *International Organization* 68(4):775–809.

Fearon, James. 1995. "Rationalist Explanations for War." *International Organization* 49(3):379–414.

Fearon, James. 2004. "Why Do Some Civil Wars Last So Much Longer Than Others?" *Journal of Peace Research* 41(3):275–301.

Fordham, Benjamin and Paul Poast. 2016. "All Alliances Are Multilateral Rethinking Alliance Formation." *Journal of Conflict Resolution* 60(5):840–865.

Garfinkel, Michelle. 2004. "Stable Alliance Formation in Distributional Conflict." *European Journal of Political Economy* 20:829–852.

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

Gowa, Joanne and Kristopher Ramsay. 2017. "Gulliver Untied: Entry Deterrence Under Unipolarity." *International Organization* 71(3):459–490.

Ikenberry, G. John, Michael Mastanduno and William C. Wohlforth. 2009. "Unipolarity, State

Behavior, and Systemic Consequences." *World Politics* 61(1):1–27.

Ikenberry, John. 1998. "Institutions, Strategic Restraint, and the Persistence of American Postwar Order." *International Security* 23(3):43–78.

Jackson, Matthew and Stephen Nei. 2015. "Networks of Military Alliances, Wars, and International Trade." *Proceedings of the National Academy of Sciences* 112(50):15277–15284.

Johnson, Jesse. 2017. "External Threat and Alliance Formation." *International Studies Quarterly* 61:736–745.

Kang, David. 2003. "Getting Asia Wrong: The Need for New Analytical Frameworks." *International Security* 27(4):57–85.

Keohane, Robert. 1984. *After Hegemony: Cooperation and Discord in the World Political Economy*. Princeton University Press.

Konishi, Hideo and Debraj Ray. 2003. "Coalition Formation as a Dynamic Process." *Journal of Economic Theory* 110:1–41.

Krainin, Colin. 2017. "Preventive War as a Result of Long Term Shifts in Power." *Political Science Research and Methods* 5(1):103–121.

Krainin, Colin and Thomas Wiseman. 2016. "War and Stability in Dynamic International Systems." *Journal of Politics* 78(4):1139–1152.

Kydd, Andrew. 2020. "Switching Sides: Changing Power, Alliance Choices and US–China–Russia Relations." *International Politics* pp. 1–30.

Lai, Brian and Dan Reiter. 2000. "Democracy, Political Similarity, and International Alliances, 1816-1992." *Journal of Conflict Resolution* 44(2):203–227.

Leeds, Brett Ashley. 2003. "Alliance Reliability in Times of War: Explaining State Decisions to Violate Treaties." *International Organization* 57(4):801–827.

Leeds, Brett Ashley and Burcu Savun. 2007. "Terminating Alliances: Why Do States Abrogate Agreements?" *Journal of Politics* 69(4):1118–1132.

Leeds, Brett Ashley, Jeffrey Ritter, Sara McLaughlin Mitchell and Andrew Long. 2002. "Alliance Treaty Obligations and Provisions." *International Interactions* 28(2):237–260.

Levy, Jack. 1987. "Declining Power and the Preventive Motivation for War." *World Politics*

40(1):82–107.

MacDonald, Paul and Joseph Parent. 2011. "Graceful Decline? The Surprising Success of Great Power Retrenchment." *International Security* 35(4):7–44.

Marshall, Monty and Keith Jaggers. 2002. *Polity IV Dataset [Computer file]*. College Park, MD: University of Maryland.

Morrow, James. 1991. "Alliances and Asymmetry: An Alternative to the Capability Aggregation Model of Alliances." *American Journal of Political Science* 35(4):904–33.

Morrow, James. 1994. "Alliances, Credibility, and Peacetime costs." *Journal of Conflict Resolution* 38(2):270–297.

Niou, Emerson, Peter Ordeshook and Gregory Rose. 1989. *The Balance of Power: Stability in International Systems*. New York, NY: Cambridge University Press.

Powell, Robert. 1999. *In the Shadow of Power: States and Strategies in International Politics*. Princeton, N.J.: Princeton University Press.

Powell, Robert. 2004. "The Inefficient Use of Power: Costly Conflict with Complete Information." *American Political Science Review* 98(2):231–241.

Powell, Robert. 2006. "War as a Commitment Problem." *International Organization* 60(1):169–203.

Powell, Robert. 2017. "Taking Sides in Wars of Attrition." *American Political Science Review* 111(2):219–336.

Reiter, Dan, Allan Stam and Michael Horowitz. 2014. "A Revised Look at Interstate Wars, 1816-2007." *Journal of Conflict Resolution* 60(5):956–976.

Sarkees, Meredith and Frank Wayman. 2010. *Resort to War: 1816-2007*. Washington DC: Congressional Quarterly Press.

Schroeder, Paul. 1994. "Historical Reality vs. Neo-realist Theory." *International Security* 19(1):108–148.

Schroeder, Paul. 2004. Alliances, 1815–1945: Weapons of Power and Tools of Management. In *Systems, Stability, and Statecraft: Essays on the International History of Modern Europe*. Springer pp. 195–222.

Schweller, Randall. 1992. "Domestic Structure and Preventive War: Are Democracies More Pacific?" *World Politics* 44(2):235–269.

Schweller, Randall. 1993. "Tripolarity and the Second World War." *International Studies Quarterly* 37(1):73–103.

Schweller, Randall. 1994. "Bandwagoning for Profit: Brining the Revisionist State Back In." *International Security* 19(2):72–107.

Singer, J. David, Stuart Bremer and John Stuckey. 1972. Capability Distribution, Uncertainty, and Major Power War, 1820-1965. In *Peace, War, and Numbers*, ed. Bruce Russett. Sage: Beverly Hills.

Smith, Alastair. 1998. "Extended Deterrence and Alliance Formation." *International Interactions* 24(4):315–343.

Smith, Bradley. 2016. "Military Coalitions and the Politics of Information." *Presented at the 2016 Formal Models in International Relations Conference* .

Snyder, Glenn. 1997. *Alliance Politics*. Cornell University Press.

Trager, Robert. 2015. "Diplomatic signaling among multiple states." *The Journal of Politics* 77(3):635–647.

Tullock, Gordon. 1980. Efficient rent seeking. In *Toward a Theory of the Rent Seeking Society*. College Station, TX: Texas A&M University Press pp. 3–15.

Walt, Stephen. 1987. *The Origins of Alliance*. Ithaca, NY: Cornell University Press.

Waltz, Kenneth. 1979. *Theory of International Politics*. Long Grove, IL: Waveland Press.

Werner, Suzanne. 2000. "Deterring Intervention: The Stakes of War and Third-Party Involvement." *American Journal of Political Science* 44(4):720–732.

Wolford, Scott. 2014. "Showing Restraint, Signaling Resolve: Coalitions, Cooperation, and Crisis Bargaining." *American Journal of Political Science* 58(1):144–156.

Yuen, Amy. 2009. "Target Concessions in the Shadow of Intervention." *Journal of Conflict Resolution* 53(5):745–773.