Rising, Arming, and a Screening Effect of Alliances

Yuji Masumura*

Last Updated: February 26, 2025

I analyze a formal model under which a growing state decides to arm but this creates an incentive for preventive war due to the fear of hidden revisionism, and then, investigate how alliances alter this preventive war motive. The model reveals a *screening effect* of alliances. Alliances make arming more informative by changing the value and purpose of arming, and this screens the otherwise hidden intention of a rising protégé. In this screening effect, alliances solve a different information problem from other existing theories, and fully committed alliances can both constrain a protégé and deter an aggressor at the same time. Two implications are empirically tested: (a) alliance decreases the possibility of preventive war and (b) the costs of internal arming decrease military expenditure when coupled with alliances.

*University of Texas at Austin, email: yuji.masumura@utexas.edu.

1 Introduction

> *"First, Japan, a nation committed to peace, rejects the role of a military power" (Takeo Fukuda, Prime Minister of Japan, in Manila in 1979)*

Empirical records after the Second World War saw a drastic decline in hostile feelings toward Japan among Southeast and East Asian countries, as well as a similar decline in European hostilities toward West Germany. These two countries invaded their neighbors, caused devastating damage, and left strong hostility there. But, even though they grew at a historical rate in the post-war era, they did not suffer from preventive-war motives by their neighbors, which should have been fueled by the strong hostility and distrust they left in the past war. These "miracles" are famous among scholars, but why did preventive war motives not work in these very likely cases? Why were such miracles ever possible? Existing theories do not clearly answer this question.

This article focuses on how alliances change the incentive for preventive war and argues that military alliances can make peace by diminishing distrust among countries, not just by increasing the costs of aggression. Military alliances are often thought to bring peace by increasing the expected costs of aggression through capability aggregation or costly signaling (Walt, 1985; Morrow, 1994, 2000; Smith, 1995; Leeds, 2003a). This explanation, however, does not fully account for why some alliances seem to bring peace by genuinely weakening hostility among nations even in the context of power shifts, as happened in Japan and West Germany cases. If alliances can deter aggression just through the increased costs of attack, this does not reduce hostility itself because the hostility is a belief between a protégé and a potential aggressor, but the costs of attack are a function of a patron's military capability and resolve, which are irrelevant to such hostilities or distrust.

To investigate the effect of alliances on preventive war motives, I construct a formal model under which a growing state chooses arms buildups but this creates an incentive for a preventive war due to the fear of hidden revisionism of the rising country. Then, I analyze how introducing an alliance alters actors' strategic behavior. In the model, a growing country endogenously chooses to invest in its armaments, which causes a power shift in the future. The level of revisionism of the rising country is private information and is difficult to observe beforehand for an external actor. This creates uncertainty over the value of overlooking the power shift for a neighboring country, and thus, creates the fear of the miserable future and the incentive for a preventive attack.

The model reveals that, in such a strategic environment, alliances have a *screening effect*: Alliances make arming more informative by changing the value of arming at a different rate for the revisionists and status-quo-oriented (SQ-oriented) countries, and this screens otherwise hidden preference of the rising country. Without an alliance, both revisionists and the SQ-oriented countries have the incentive to arm to secure the division of their issue because the lack of the alliance creates the demand for internal buildups. This identical arming strategy makes a neighboring country unsure about the true intention of the arming country and the consequences of passing the arming decision (Jervis, 1978): it wants to prevent the arming if the arming country is a revisionist, but does not want to do so if it is the SQ-oriented type. This uncertainty creates the neighbors' incentive for preventive attacks even though it entails the risk of mistaken preventive wars against the SQ-oriented countries. Thus, a rising country suffers from preventive attack even when it is the SQ-oriented type, which is consistent with previous theories (Powell, 2006; Debs and Monteiro, 2014).

Alliances can solve this problem, not through capability aggregation or costly signaling but through revealing the hidden intentions of a rising country. Defensive alliances improve the status-quo payoffs, leading to a decrease in the relative benefit of arming mainly for the SQ-oriented protégés. Because alliances increase the country's deterrence if a protégé is attacked, extra arming would be a waste of resources. This, in turn, makes it clear that arming after alliance formation is a strong sign of revisionist intentions, which would be difficult to observe without alliances. Thus, revisionists cannot mimic the SQ-oriented type and avoid preventive attacks any longer, which deters even the revisionists from arming in the first place.

I also investigate the screening effect on the two main purposes of alliances: deterrence and constraining. I find that alliances are generally effective at deterrence even in the context of power shifts, and alliances can achieve both the deterrence effect and the constraining effect at the same time in a different way from existing theories (Fang et al., 2014; Benson, 2012). However, the model also shows that alliances sometimes encourage arming by the SQ-oriented types since they do not worry about mistaken preventive wars any longer thanks to the screening effect. This implies that the effect of alliances on military expenditure is difficult to find empirically, and this can explain why scholars find mixed and nuanced results on the association between alliances and military expenditure (Conybeare, 1992, 1994; Kimball, 2010; Sorokin, 1994; Plümper and Neumayer, 2015; Diehl, 1994; Horowitz et al., 2017; Alley, 2021; Digiuseppe and Poast, 2018).

I quantitatively test two implications of the model. First, by following Bell and Johnson (2015)'s empirical strategy to measure expected future power shifts, I analyze the effect of alliances on preventive war motives from 1818 to 2007. The result shows that, as consistent with previous studies, future power is positively associated with the possibility that a rising country is targeted at war. However, the effect of expected power shifts on war is substantively mitigated when the rising country has a defensive alliance, and this deterrence effect is stronger when expected power shifts are large.

Second, the model suggests that defensive alliances make a protégé more sensitive to the costs of internal arming. The empirical analysis of data from 1900 to 2015 shows that internal arming costs are negatively associated with military expenditure and this negative effect is amplified when countries have defensive alliances. These results reveal the uninvestigated role of military alliances in the context of power shifts and preventive war.

2 Literature

One of the core topics in the alliance literature is if and how alliances affect internal arms buildups[1]. Some argue that alliances and arms substitute with each other, as seen in Morrow (1993)' argument that alliances and arms have different domestic costs and states combine both to improve their security environment. According to this argument, alliances should decrease military spending, and some empirical studies support this argument (Conybeare, 1992, 1994; Kimball, 2010; Sorokin, 1994; Plümper and Neumayer, 2015; Kuokštytė and Kuokštis, 2024).

Others, however, posit that alliances and arms are complements, meaning that alliances

[1]There is a strand of studies about the interaction between alliances and arms transfers from a patron. See Yarhi-Milo et al. (2016), McManus and Yarhi-Milo (2017), and Kudo and Masumura (2022) for more on these studies.

and military spending are positively associated because alliance formation brings a new defense obligation (Diehl, 1994; Alley and Fuhrmann, 2021)[2] or because arm buildups work as a signal of a reliable alliance partner (Horowitz et al., 2017). The literature is divided on the assessment of the alliance effect on arming[3]. Recent studies try to capture a more nuanced relationship between alliance and arming: Digiuseppe and Poast (2018) focus on the regime type of a patron and show that democratic alliance partners decrease military spending of a protégé but autocratic partners increase it. Similarly, Alley (2021) argues that alliances with deep formal cooperation decrease but shallow alliances increase military spending.

These studies provide convincing arguments and evidence, but they have not paid attention to the idea that arming causes power shifts and power shifts make an incentive for preventive attacks. The idea that power shifts cause a preventive war, often called commitment problems, is one of the two reasons for war (Fearon, 1995; Powell, 2006), and this argument is empirically supported (Bell and Johnson, 2015; Bas and Schub, 2017). Even though the arming decision is endogenous in reality, this mechanism still works when coupled with some uncertainty (Debs and Monteiro, 2014). How do alliances affect the incentive of arms buildups and the incentive for preventive war? Benson and Smith (2023) investigate how alliance formation causes a power shift but still avoids preventive war. Although their argument is of full insight, alliances are a cause of power shifts in their model, whereas the focus of this chapter is the idea that alliances may solve a commitment problem. So far, the literature is silent about such a relationship. Because alliances, arming, and preventive war are directly connected, we need a theory that combines three

[2] For Alley and Fuhrmann (2021), see also Cooley et al. (2022)'s discussion.

[3] See Alley (2021) for more on these mixed results.

of them in a single model.

Looking at more broad literature on alliances, scholars often consider that alliances have mainly two distinct effects on international relations. One is a deterrence effect for external threats (Walt, 1985; Leeds, 2003b; Johnson and Leeds, 2011). Alliances deter aggressions through a costly signaling mechanism (Fearon, 1997; Morrow, 1994, 2000; Smith, 1995), so alliances are formed when external threats are high (Johnson, 2017). Recently, scholars apply the same mechanism to internal threats and argue that alliances, especially ones having consultation pacts, are formed when internal threats exist (Edry et al., 2021) and that allies are more likely to intervene in a civil war for the government side (Johnson et al., 2024) [4]. Although these studies help us understand the important role of alliances in world politics, one missing aspect is how alliances solve an incentive for a preventive war[5]. Existing studies investigate how alliances solve information asymmetry between a patron and a potential aggressor, but they are silent on the alliance effect on the uncertainty of a protégé's intention, which is especially relevant when a protégé is rising (Debs and Monteiro, 2014).

Another effect of alliances is to constrain a protégé (Fang et al., 2014; Benson, 2012). When a patron is status-quo-oriented, it can successfully constrain its protégé's aggression by obscuring its alliance commitment[6], as happened in the US-Taiwan relations in the 1950s (Benson, 2012), or by credibly threatening that such aggression damages the future

[4] McWard and Yoon (2024) suggest that autocracies with coop-proof devices are more likely to form an alliance to make up the lack of internal security, but they also imply that external threats have an opposite effect in such a sample of cases.

[5] One exception is Krainin and Schub (2021). They find that alliance formation and termination are likely when power shifts are expected because countries adjust their alliance relationship before power shifts happen so that preventive war is less likely to happen.

[6] Benson et al. (2014) show that unreliable alliances still have a deterrence effect under the assumption that peace payoffs can be divided in the same way as an insurance contract in domestic individuals.

patron-protégé relationship, including abandonment (Fang et al., 2014).

All of these studies are important and insightful. But some important examples show something different from these mechanisms: Japan and West Germany. As is often quoted, NATO was for "to keep the Soviet Union out, the Americans in, and the Germans down[7]". Constraining Germany's revisionism was one of the main goals for the US in the post-war era (Trachtenberg, 1999), and this was also the case for Japan. Unlike the existing theories, however, the US seemed to be fully committed to these alliances after WWII: unlike Taiwan, the US was clear in its military support for these countries and the threat of abandonment was not very credible, given the Cold War structure. Still, the US successfully constrained these countries' revisionism and deterred potential aggressors at the same time, even in the context where these countries grow dramatically. Why were these "historical miracles" possible? Why did not these rising countries suffer from preventive attacks? The literature does not give us good answers. To answer these questions, I constructed a formal model, which I introduce in the next section.

3 Model

I present a formal model under which a rising country decides to arm in the shadow of distrust from its neighboring state. I combine the incentive of preventive war due to power shifts and the uncertainty over the rising side (Fearon, 1995; Powell, 2006; Debs and Monteiro, 2014; Bas and Schub, 2017), and compare three versions of the model: one without an alliance, one with an exogenous alliance, and one with an endogenous alliance.

[7] Lord Ismay, the first Secretary General of NATO. (URL: `https://www.nato.int/cps/en/natohq/declassified_137930.htm`)

Assume that there is an (economically or politically[8], but not militarily) growing country, called A, that decides whether to invest its growing resources in its military capability. Its neighboring country, called B, faces the uncertainty over A's revisionism level, and it wants to prevent A's arming if A is a revisionist but it does not want to do so if A is not a revisionist. The game starts with Nature drawing the revisionism level of A, called r, and it takes either high ($r = 1$) with the probability of $\phi \in (0, 1)$ or low ($r = p \in (0, 1)$) with $1 - \phi$. I call the high revisionism type Revisionist and the low revisionism type SQ type. This information is revealed to A, but not to B or A's alliance partner, if it exists. Then, A decides to arm with some costs, $K = k > 0$, or not to arm without any costs $K = 0$. I denote the arming decision as a and it is either successful ($a = 1$) or not ($a = 0$), which I explain later. After observing the arming (or no arming) decision, B chooses to go to a preventive war against A or pass A's decision. If B preemptively attacks, the outcome is a preventive war in which the arming decision has not yet come to fruition ($a = 0$). This gives payoffs $rw_A(0)$ and $w_B(0)$ to A and B, respectively, where $w_i(\cdot)$ is a general war payoff function for i and it is an increasing function in the input $(\cdot)$ for A and a decreasing function in the input for B. I assume that war is more attractive for Revisionist than SQ type for various reasons, including lower war costs, higher value of the issue at stake, leaders' private benefits from war[9], or insulation from war costs[10]. The results of the model are not sensitive to the source of revisionism as long as it is difficult to observe for B beforehand and it makes wars more attractive or acceptable for Revisionist. A higher war payoff for

[8] Here, I suppose revolution, the end of civil war, or the change in the international political system, which may increase the government's resources.

[9] Examples are status or reputation concerns. See Renshon (2017), Dafoe et al. (2014), and Masumura and Tago (2023) for more on this.

[10] See Debs and Goemans (2010), Croco and Weeks (2016), Carter (2017), and Chiozza and Goemans (2004).

Revisionist is represented by the fact that $r \in \{1, p\}$ is multiplied by $w_A(\cdot)$ in A's war payoffs, where p is strictly lower than 1. To represent the inefficiency of war, I assume $w_A(\cdot) + w_B(\cdot) < 1$. Also, war does not have a negative value for both actors ($w_A(\cdot) > 0$ and $w_B(\cdot) > 0$).

If B passes, arming (if it's chosen) becomes completed ($a = 1$), and B can observe the realized type of A. This represents a common strategic trade-off in power shifts and uncertainty: passing reveals the state of the world but it is too late to go to preventive war (Debs and Monteiro, 2014; Bas and Schub, 2017; Wolford and Masumura, 2024). The level of revisionism is revealed after passing due to the longer time of interaction between leaders or more information from diplomatic or intelligence apparatus[11]. Assuming that uncertainty remains even after passing an arming decision unnecessarily complicates the analysis and causes the inability to differentiate the strategic environment of power shifts from the one of pure private information problems. After B decides to pass, the actors have a crisis bargaining in which B offers some concession denoted by $x \in (0, 1)$ to A out of a unit of the issue at stake. Then, A accepts or rejects the offer. Acceptance of the offer gives x to A and $1 - x$ to B. Rejecting the offer causes a war, and this gives $rw_A(a)$ to A and $w_B(a)$ to B. Note that it takes $w_i(0)$ when A does not arm at the earlier stage since $a = 0$, suggesting the same war payoffs as the preventive war. But, when B passes and A successfully arms ($a = 1$), the war payoff is $rw_A(1)$ for A and $w_B(1)$ for B. Remember that $w_i(\cdot)$ is an increasing function for A and a decreasing function for B, so, successful arming increases A's war payoffs, whereas it decreases B's war payoffs.

In versions of the model with an alliance, there is an external actor E which has or can

[11] Wu and Wolford (2018) show that years from leader turnover are associated with international conflict, implying that time is an important factor to solve an information problem.

have an alliance with A, and it can choose to intervene whenever war happens, if it has an alliance. E has a reference point about how much A should get from the distribution of the issue[12], and E's preference is constituted by this reference point. Let's denote E's reference point as $s \in (0, 1)$ and the distribution of the issue A actually gets as Δ_A. E gets a higher payoff as A secures something closer to the reference point, but gets a lower payoff when A's payoff is far from its reference point. Specifically, E's payoff is $1 - |\Delta_A - s|$. E and A are highly aligned when s is close to 1, and they are not aligned when s is close to 0. When A gets exactly what E wants it to get, then the payoff for E takes the highest value, 1. I do not assume any cost parameter for intervention because it does not change the fundamental incentives of actors, but substantively, it assumes a credible patron who does not hesitate to help if necessary. If E intervenes in a war, A and B's payoffs change to $w_i(a + \epsilon)$, where $\epsilon > 0$ is the additional power the alliance brings, and this increases A's war payoffs and decreases B's war payoffs in both the preventive and non-preventive war. A version of the model where A and E have choices to form an alliance has a slightly different structure, and I explain it later in the next section.

To summarize, the timing of the game is as follows: (1) Nature decides whether A is a revisionist or not, (2) A chooses to arm or not, (3) B chooses to go to a preventive war against A, (4) if B attacks, preventive war happens, (5) if B does not attack, A's types are revealed, arming is completed, and B offers a division of a pie, x, (6) A decides to accept the offer or not, and (7) if there is an alliance and war occurs, E decides to intervene to help A or not. The payoffs are also summarized below.

[12] This means that E cares about A's payoff from war or the crisis bargaining but not the costs of arming. So, this is different from A's utility itself.

$$u_i(\text{Settlement}) = \begin{cases} x - K \ \ (\text{if } i = A) \\ 1 - x \ \ (\text{if } i = B) \\ 1 - |x - s| \ \ (\text{if } i = E) \end{cases} \tag{1}$$

$$u_i(\text{War w/o Intervention}) = \begin{cases} rw_A(a) - K \ \ (\text{if } i = A) \\ w_B(a) \ \ (\text{if } i = B) \\ 1 - |rw_A(a) - s| \ \ (\text{if } i = E) \end{cases} \tag{2}$$

$$u_i(\text{War w/ Intervention}) = \begin{cases} rw_A(a+\epsilon) - K \ \ (\text{if } i = A) \\ w_B(a+\epsilon) \ \ (\text{if } i = B) \\ 1 - |rw_A(a+\epsilon) - s| \ \ (\text{if } i = E) \end{cases} \tag{3}$$

4 Key Features of the Model

The model has some features to note. First, preventive wars are caused by a commitment problem in a consistent way with the literature. As we will see later, the incentive of preventive war arises when a power shift is so large that the bargaining cannot absorb the loss of power. Other conventional models of commitment problems often have two or infinite stages where actors have a negotiation at each stage (Powell, 2006; Fearon, 1995). I simplify this environment by packing the multiple stages into a one-stage dynamic game and by dichotomizing B's decision into going to a preventive war or passing a power shift. This can be done without the loss of generality as long as a discount factor, often denoted by δ, is sufficiently large[13].

[13] See Wolford and Masumura (2024) as an example of papers that adopt the same modeling decision.

Second, the model combines commitment problems and information problems by assuming that the extent of A's revisionism is uncertain. As Chadefaux (2011) argue, in theory, power shifts do not cause a preventive war if the action that causes power shifts is endogenous and states can negotiate. Thus, private information is a necessary part for states to go to a preventive war when power shifts are endogenous (Debs and Monteiro, 2014). Other models also combine commitment problems and information problems in a similar way (Debs and Monteiro, 2014; Bas and Schub, 2017), but unlike Bas and Schub (2017)'s model where a rising country does not have a choice to rise and thus power shifts exogenously happen[14], a rising state endogenously chooses to arm in my model. Unlike Debs and Monteiro (2014)'s model, where the arming decision is only observable with an exogenously determined probability, the arming decision is perfectly observable in my model. Thus, although Debs and Monteiro (2014)'s model assumes that the incentive of a preventive war arises when a neighboring state does *not* observe arming, my model assumes such incentive arises when a neighboring country observes arming, which captures a different and more realistic strategic environment.

Third, unlike existing models, the model captures a credible alliance and a constraining effect in two ways. First, existing studies show that a patron can constrain its protégé's aggression by making its support probabilistic (Benson, 2012), but this article does not assume the uncertainty over intervention and still shows that a constraining effect emerges when coupled with an arming decision. Second, Fang et al. (2014)'s model demonstrates that the costs of damaging an alliance, a key parameter for their model, make a protégé follow its patron's recommendation, leading to a constraining effect. But, unlike their

[14] Krainin and Schub (2021)'s model on alliances and power shifts also assumes exogenous power shifts.

model, my model indicates that a constraining effect emerges without such a parameter, meaning a patron that could not credibly threaten to abandon or damage its protégé can still generate a constraining effect.

5 Analysis

The solution concept is Perfect Bayesian Equilibrium (PBE), in which strategies are sequentially rational and consistent with beliefs updated according to Bayes' Rule whenever possible.

I first analyze a version of the model without an alliance. This version offers a baseline about under which conditions a preventive war can happen when arming is endogenous. Then, I exogenously add an alliance to the model and examine how the alliance changes the incentive of a preventive war, the incentive of arming, and the belief that the rising country is Revisionist. Finally, I endogenize the alliance to discuss when alliance formation/continuation is likely.

5.1 Rising without an Alliance

Lemma 1 (Preventive War Incentive)**.** *After A's arming, B wants to go to a preventive war iff*

$$rw_A(1) + w_B(0) > 1 \tag{4}$$

$$\Leftrightarrow r(w_A(1) - w_A(0)) > 1 - rw_A(0) - w_B(0) \tag{5}$$

Proof is in the text.

In a version without an alliance, B has an incentive to go to preventive war after A's arming decision when the sum of A's war payoff after the power shift ($rw_A(1)$) and B's war payoff before the power shift ($w_B(0)$) is greater than the total value of the issue at stake, as Line 4 shows.

The logic is as follows. If A's arming is completed, A accepts x if and only if $x \geq rw_A(1)$. Because of no uncertainty after a power shift and war's inefficiency, B offers $x = rw_A(1)$ and gets $1 - rw_A(1)$. Given this, B prefers a preventive war if and only if the preventive war payoff, $w_B(0)$, is greater than $1 - rw_A(1)$, which can be transformed to Line 4. Substantively, this means that B goes to a preventive war when what it gets in a preventive war is greater than what it gets in the future concession. Another way of understanding this condition is Line 5. The left side is the extent of power shifts if arming is successful, and the right side is the bargaining surplus. Preventive war happens when power shifts are so large that the bargaining surplus cannot save the countries from war[15].

Note that B has no incentive for a preventive war when A does not arm because it does not change the power balance and, in such a situation, war's inefficiency makes sure B's passing anytime. Hereafter, I mainly focus on the following inequality.

$$w_A(1) + w_B(0) > 1 > pw_A(1) + w_B(0) \tag{6}$$

This means that, after A's arming, B wants to go to a preventive war when A is Revisionist, but not when A is SQ type. Revisionist credibly demands a larger future concession than SQ type thanks to its better war payoff. B prefers a preventive war to

[15]This result is consistent with Powell (2006)'s inefficiency condition shown in Line 1 in his article (Powell, 2006, p.182).

such a large concession, so it preemptively attacks A if A is Revisionist. But the future concession for SQ type is too small to justify a preventive war for B, leading to passing A's arming decision. This creates omnipresent uncertainty related to power shifts: countries do not know the result of passing. Countries want to prevent a power shift if passing such a power shift brings a miserable future, but how miserable the future is often unknown. Examples in the real world are plenty: The discourse about the alleged Weapons of Mass Destruction in the Iraq War (Debs and Monteiro, 2014), Chamberlain's appeasement against the Nazis in Munich in 1938, and Obama's "strategic patience" policy against China's military expansion in the 2010s are all centered around the inability to know the (un)desirability of the future concession. First of all, I analyze the conditions under which preventive war can happen.

Proposition 1 (Semi-Separating)**.** *When Line 6 is satisfied and*

$$p(w_A(1) - w_A(0)) > k \tag{7}$$

$$\phi < \phi^* = \frac{1 - pw_A(1) - w_B(0)}{w_A(1)(1-p)}, \tag{8}$$

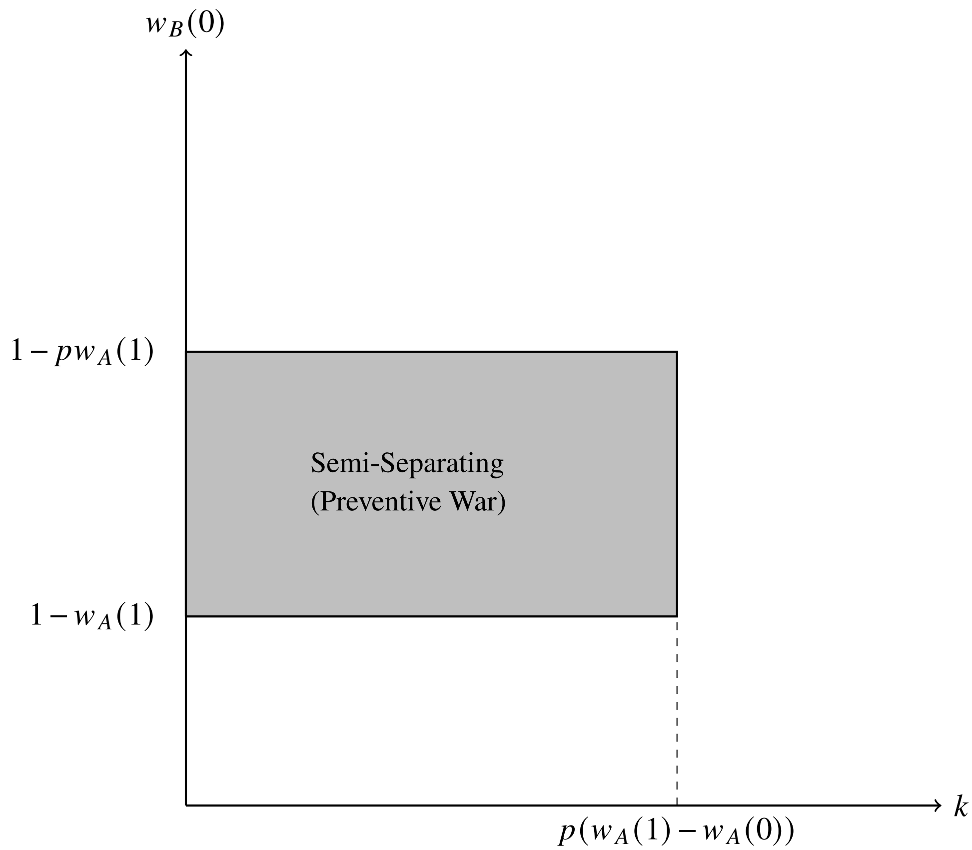
there exists a semi-separating PBE at which Revisionist A always arms, SQ type A arm with $l^* = \frac{\phi(w_A(1)+w_B(0)-1)}{(1-\phi)(1-pw_A(1)-w_B(0))}$*, B attack A with* $m^* = \frac{p(w_A(1)-w_A(0))-k}{p(w_A(1)-w_A(0)}$ *after arming, and does not attack after no arming. See Appendix for proof.*

Proposition 1 is a semi-separating equilibrium in which a preventive war can happen with a certain probability. Because of the low costs of arming (Line 7), both types of A want to arm if it can be done without prevention, which leads to B's inability to discern A's type. On the pass of play, Revisionist always arms, SQ type occasionally arms with the probability of l^*, and B prevents arming with the probability of m^*. Revisionist exploits B's inability to identify A's types and aims at successful arming and the avoidance of a preventive war at the same time. Given Revisionist's incentive, B partially, if not perfectly, deters Revisionist's arming by an occasional preventive war even though it entails the risk of a preventive war against SQ Type. This threat of mistaken preventive war against SQ Type makes SQ Type's arming decision probabilistic to balance out the costs of preventive war and the benefit of securing a better deal in the future.

This result is in contrast to other models of reassurance. Unlike conventional models of reassurance where types with Prisoner's-Dilemma preferences mimic types with Stag-Hunt preferences by cooperating (e.g. Kydd (2000)), the result here is based on the idea that even SQ Tyep does not cooperate to secure its share. In Proposition 1, both types of A want to arm if possible. B is fine with SQ type's arming, and thus, Revisionist has the incentive to mimic SQ type by arming to peacefully rise while hiding its true intention.

This is the only equilibrium where a preventive war can happen in this version of the model. If the condition of the costs of arming (Line 7) is not satisfied and only Revisionist type can profitably arms, or if B strongly believes that A is Revisionist unlike Line 8, then A does not arm in the first place since B would go to a preventive war with a conviction that A is Revisionist, which A wants to avoid beforehand. Similarly, in any separating equilibrium where Revisionist arms but SQ Type does not arm, B is sure that arming A is Revisionist and goes to a preventive war, which A avoids by not arming. Contrary, if B strongly believes that A is SQ Type, B always passes A's arming decision, and thus preventive war does not happen. Proposition 1, where arming is partially informative and Revisionist can (sometimes successfully) mimic SQ Type, is the only equilibrium characterized by a preventive war. Thus, I use Proposition 1 as a baseline to analyze the effect of alliances on preventive war motives. Figure 1 shows this equilibrium in terms of k for the vertical axis

Figure 1: Preventive War Equilibrium



and $w_B(0)$, B's war payoff in preventive war, for the horizontal axis.

5.2 Rising with an Alliance

Next, I introduce an (exogenous) alliance into the model to see how it can change the incentive of preventive war[16]. First, I analyze when the patron (E) intervenes in a war to help A.

[16] I make this exogenous to see the effect of alliances on the preventive war motives.

Lemma 2. *E intervenes in a war if and only if*

$$\frac{r(w_A(a+\epsilon)+w_A(a))}{2} < s < 1 \tag{9}$$

See Appendix for Proof.

E helps when its ideal point about A's payoff (s) is greater than the average of A's payoff with E's help ($rw_A(a+\epsilon)$) and its payoff without E's help ($rw_A(a)$). E's intervention improves A's war payoff, so E considers if intervention makes A's share of the issue closer to E's ideal point. The average of these two values is located in the center of them, so if s is greater than the average, then s is closer to $rw_A(a+\epsilon)$ than $rw_A(a)$, which drives E to choose the intervention. Hereafter, I assume the following inequality.

$$\frac{w_A(\epsilon)+w_A(0)}{2} < s < \frac{p(w_A(1+\epsilon)+w_A(1))}{2} \tag{10}$$

This ensures that E helps A when A does not arm or A is preemptively attacked, but does not intervene after A successfully arms regardless of its type[17]. Substantively, this represents a patron that is status-quo oriented: preventive war caused by B is not preferable for E and thus E intervenes, but successful arming by A is not favorable for E either regardless of A's type, because it makes B's concession larger in the future negotiation, leading to the revise of the status quo. Maintaining the status quo is often discussed in the literature as a policy goal that the US has aimed to achieve through alliances[18].

[17] This is a conservative case where E and A's ideal points are not closely aligned.

[18] Kinsella (1994) and Kudo and Masumura (2022) argue that the US maintained the regional status quo through its allies and arms transfers. Cha (2010) and Benson (2012) also discuss the US and its allies' role in sustaining the status quo in Asia in the Cold War period.

Moreover, to mirror Line 6, I mainly focus on the inequality below.

$$w_A(1) + w_B(\epsilon) > 1 > pw_A(1) + w_B(\epsilon)$$

$$\Leftrightarrow$$

$$\overline{w}_B = 1 - pw_A(1) + d > w_B(0) > \underline{w}_B = 1 - w_A(1) + d \tag{11}$$

, where $d = w_B(0) - w_B(\epsilon)$. Line 11 stipulates that B has the incentive of preventive war when Revisionist arms but not when SQ Type arms. This makes B's decision difficult due to the uncertainty of A's type in the same way above: a preventive war is a right strategy only when A is Revisionist. Line 6 and 11 have overlapping parameter space as long as $w_A(1) - pw_A(1) > w_B(0) - w_B(\epsilon)$, which means in a substantive sense that the difference of war payoffs for Revisionist and SQ Type is large enough. I start introducing several equilibria below, which are mapped in Figure 2. The horizontal and vertical axes are the same as Figure 1, and the area surrounded by a red-dotted line represents the preventive war equilibrium in a version of no alliance in Figure 1.

Proposition 2 (Pooling on $\neg$ arm)**.** *When Line 10 is satisfied and*

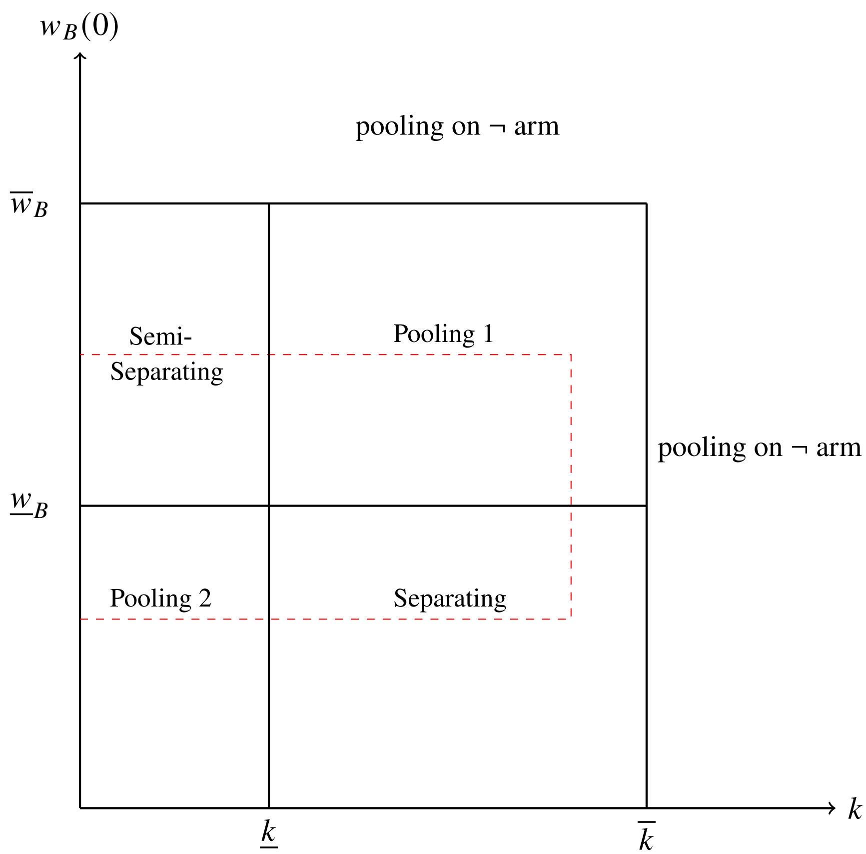
$$w_B(0) > \overline{w}_B \tag{12}$$

or

$$\overline{k} = w_A(1) - w_A(\epsilon) < k \tag{13}$$

there exist pooling PBEs at which neither type of A arms and B does not attack on the path of play. See Appendix for proof.

Figure 2: Equilibrium w/ Alliances



Proposition 2 is a pooling equilibrium where neither type of A arms and B does not go to a preventive war on the path of play. Two cases are falling into this equilibrium. First, the alliance decreases B's preventive war payoff by $d = w_B(0) - w_B(\epsilon)$, but when B's preventive war payoff is still too high as shown in Line 12, B prevents A's arming decision regardless of its type. This makes A's arming endeavor useless, so it pools on not arming. Second, when the cost of arming is too high (Line 13), both types of A do not have the incentive to arm in the first place even if A can avoid preventive war. Thus, A pools on not arming. In this equilibrium, B's off-equilibrium-path belief after A's arming can take any value since it does not affect actors' strategy.

Proposition 3 (Pooling 1)**.** *When Line 10 and 11 are satisfied and*

$$\underline{k} = p(w_A(1) - w_A(\epsilon)) < k < \overline{k} \tag{14}$$

there exists a pooling PBE at which neither type of A arms, B attack if A arms but does not attack if A does not arm. See Appendix for proof.

Proposition 3 is a pooling equilibrium in which neither type of A arms, and B does not go to preventive war on the path of play, which is the same as the previous pooling equilibrium. The difference from the previous one is in B's off-equilibrium-path belief. B's belief about the probability that A is Revisionist after A's arming is always 1 in Proposition 3. The reason is that the alliance changes the thresholds of the costs of arming, creating different incentives for SQ Type and Revisionist. The alliance improves the payoff of not arming for A, and thus, it relatively decreases the expected benefit from arming. The middle level of arming costs, as indicated in Line 14, completely eliminates SQ Type's incentive to arm, whereas it still preserves Revisionist's incentive to arm. Thus, after observing

A's arming, B rationally conjectures that A that has armed just now must be Revisionist. Technically, this satisfies the D_1-Criterion proposed by Banks and Sobel (1987) because B's off-equilibrium belief (the belief that A is Revisionist for sure after arming) is consistent with the fact that Revisionist is the only type who could profitably arm. Note that, as we can see in Figure 2, the area of this equilibrium overlaps with the one of the preventive war equilibrium in the no-alliance model (Proposition 1).

This equilibrium suggests a screening effect of alliances: alliances make arming more informative by reducing the benefit of arming uniquely for SQ Type. Thus, B can correctly infer that the decision to arm despite the existence of an alliance is a sign of high revisionism, and thus B can prevent a power shift with a conviction. However, because alliances screen a protégé's hidden type, a protégé with high revisionism gives up causing a power shift to avoid such preventive war. This means that screening happens on a off-equilibrium path. Successful screening removes observable preventive wars caused by revealed revisionism through alliances from empirical records, but this is consistent with the strategic consequences of the screening effect of alliances. Because of the screening effect, Revisionist cannot militarily rise even though it wants, and this reduces the possibility of preventive war. In Proposition 3, both the probability of preventive war and the probability of A's arming are 0.

Proposition 4 (Separating)**.** *When Line and 10 is satisfied and*

$$w_B(0) < \underline{w}_B \tag{15}$$

$$\underline{k} < k < \overline{k} \tag{16}$$

$$\tag{17}$$

There exists a separating PBE at which Revisionist arms, SQ Type does not arms, and B *always pass. See Appendix for proof.*

This is a separating equilibrium where Revisionist arms, SQ Type does not arm, and B passes A's arming decision. As happens in the previous pooling equilibrium (Proposition 3), the middle level of arming costs (Line 16) stop arming by SQ Type but does not stop arming by Revisionist. This is because the alliance improves the payoff of preventive war for A, and this, in turn, makes the threshold of the arming costs that justify arming lower and stricter. This alters SQ Type's incentive and it does not think arming is beneficial, even if it can do so without preventive war. Moreover, the alliance reduces the preventive war payoff for B, requiring a higher value of $w_B(0)$ to justify preventive war. Thus, when Line 15 is satisfied, B does not find preventive war profitable even against Revisionist. In such a situation, only Revisionist, which has an incentive to arm, can profitably and peacefully improve its military capability, leading to a separating equilibrium. In this separating equilibrium, The probability of preventive war is 0, but the probability of Revisionist's arming is 1, whereas the probability of arming by SQ Type is 0.

This equilibrium still shows a screening effect in a different way: After arming, B correctly updates its belief and thinks that A is Revisionist. But at the same time, the alliance deters B's preventive war so much that it leads to the failure to constrain Revisionist's arming. The alliance discourages SQ Type's arming but encourages Revisionist's arming even though the alliance reveals the hidden nature of A. This suggests a heterogeneous effect of alliances on protégés' arming. Also, preventive wars against revisionists are sometimes deleted from the data, even though the screening effect reveals the true nature of protégés.

Proposition 5 (Pooling 2)**.** *When Line 10 is satisfied and*

$$k < \underline{k} \tag{18}$$
$$w_B(0) < \underline{w}_B \tag{19}$$

there exists a pooling PBE at which both types of A arms, B does not attack. See Appendix for proof.

Proposition 5 is a pooling equilibrium where both Revisionist and SQ Type arm, but B does not want to prevent it. The very low costs of arming in Line 18 incentivize both SQ Type and Revisionist to arm if possible, even after the alliance makes the thresholds of arming more strict. Line 19 shows that B's very low payoff of preventive war due to the alliance eliminates any incentive to prevent A's arming regardless of its type. Thus, both types of A arm in equilibrium without any worry about preventive war. Thus, the probability of preventive war in equilibrium is 0, but the probability of A's arming is 1 in both types. In addition, unlike Proposition 4, Proposition 5 does not have a screening effect since the low costs of arming preserve both types of A's incentive of arming, which leads to the inability to discern the type of A based on its arming decision.

Proposition 6 (Semi-Separating)**.** *When Line and 10 and 11 are satisfied and*

$$k < \underline{k} \tag{20}$$
$$\phi < \phi^{\dagger} = \frac{1 - pw_A(1) - w_B(\epsilon)}{w_A(1)(1-p)} \tag{21}$$

there exists a semi-separating PBE at which Revisionist A arms with $l^{\dagger} = \frac{\phi(w_A(1)+w_B(\epsilon)-1)}{(1-\phi)(1-pw_A(1)-w_B(\epsilon))}$*, B attack A after arming with* $m^{\dagger} = \frac{p(w_A(1)-w_A(\epsilon))-k}{p(w_A(1)-w_A(\epsilon)}$*, and does not attack after no arming. See Appendix for proof.*

Proposition 6 is a mirror equilibrium of Proposition 1, but with more restrictive conditions as the alliance alters the thresholds of arming costs and B's preventive war payoff.

Revisionist always arms with the hope that it can successfully mimic SQ Type, whereas SQ Type arms with $l^{\dagger}$ since the threat of mistaken preventive war is large enough to discourage its arming to some extent. Given Revisionist's incentive to mimic, B tries to deter it with an occasional preventive war with a probability of $m^{\dagger}$ after arming, even though it entails the risk of mistaken preventive war against SQ Type. In this equilibrium, the alliance does not change the incentive to arm itself: both Revisionist and SQ Type want to arm if not prevented by B due to the low costs of arming (Line 20). However, as Figure 2 shows, thanks to the alliance, the area of preventive war decreases from the without-alliance model, and as will be shown in the next section, the probability of preventive war and the probability of A's arming are smaller even within this equilibrium.

5.3 Effects of an Alliance

Next, I take a closer look at how an alliance changes actors' behavior in equilibrium. Specifically, I examine the deterrence and constraining effects of an alliance as well as its influence on B's belief.

5.3.1 Deterrence Effect

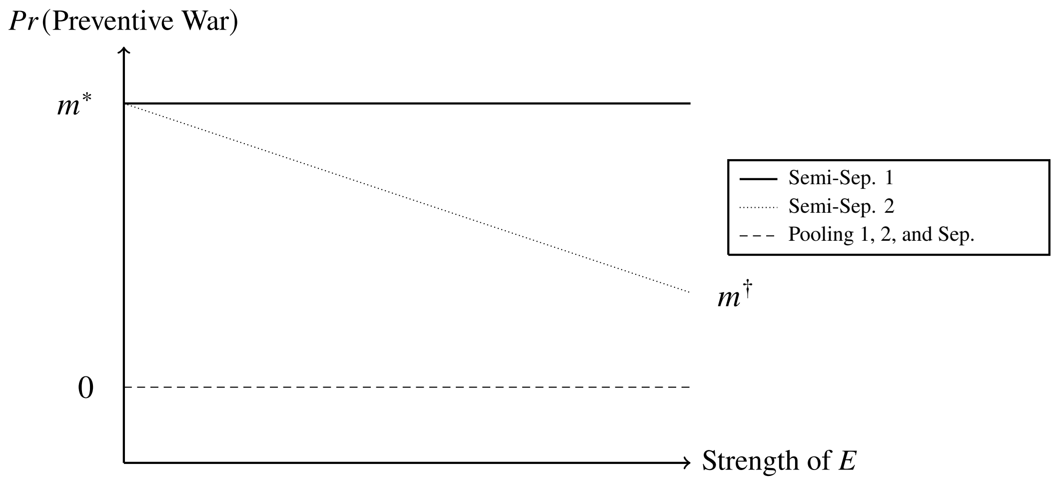
First, I analyze the deterrence effect of alliances. Figure 3 shows the probability of preventive war in each equilibrium on the vertical axis and the strength of E on the horizontal axis. The probability of preventive war in the first semi-separating equilibrium, the one without an alliance (Proposition 1), is m^*, and this is represented by a solid line in Figure 3. When assessing the deterrence effect of an alliance, this line is used as a reference point. The probability of preventive war in the second semi-separating equilibrium, the one with an alliance (Proposition 6), is $m^\dagger$. $m^\dagger$ is strictly decreasing in the strength of E, suggesting that a strong patron can generate a stronger deterrence effect even when the probability of preventive war cannot be eliminated. This is shown in a dotted line in Figure 3[19]. In the two pooling equilibria and the separating equilibrium (Proposition 3, 5, and 4), the probability of preventive war is 0 thanks to the screening effect, which is shown in the dashed line in the figure.

Let's see how much an alliance decreases the probability of preventive war in Figure 3. As m^* is the highest regardless of the strength of E, meaning that the deterrence effect of an alliance is always positive. When compared to Proposition 3, 5, and 4, the probability of preventive war decreases by $m^* - 0 = m^*$. When compared to Proposition 6, the probability of war decreases by $m^* - m^\dagger = k(w_A(\epsilon) - w_A(0))$, which is, again, strictly positive and increasing in the strength of E.

Conventional alliance models often argue that alliances solve a private-information problem on E's intervention through costly signaling (Morrow, 1994, 2000; Smith, 1995). The model in this chapter, however, does not assume any private information on E's intervention. Instead, the model assumes the uncertainty over a protégé's intention, which is a different information problem from existing models. The model shows us that alliances can still bring a deterrence effect by screening the intention of a protégé, which is an important factor in wars caused by endogenous commitment problems (Fearon, 1995; Powell, 2006; Chadefaux, 2011; Debs and Monteiro, 2014). This deterrence effect is effective even when the alliance cannot completely eliminate the possibility of preventive wars because of very low arming costs, as shown by the fact that $m^\dagger$ is lower than m^*.

[19] Note that the war payoffs, $w_i(\cdot)$ is a general function, so the slope of the line for Proposition 6 is not exact.

Figure 3: Deterrence Effect



Existing theories of alliances have not focused on these effects of alliances[20].

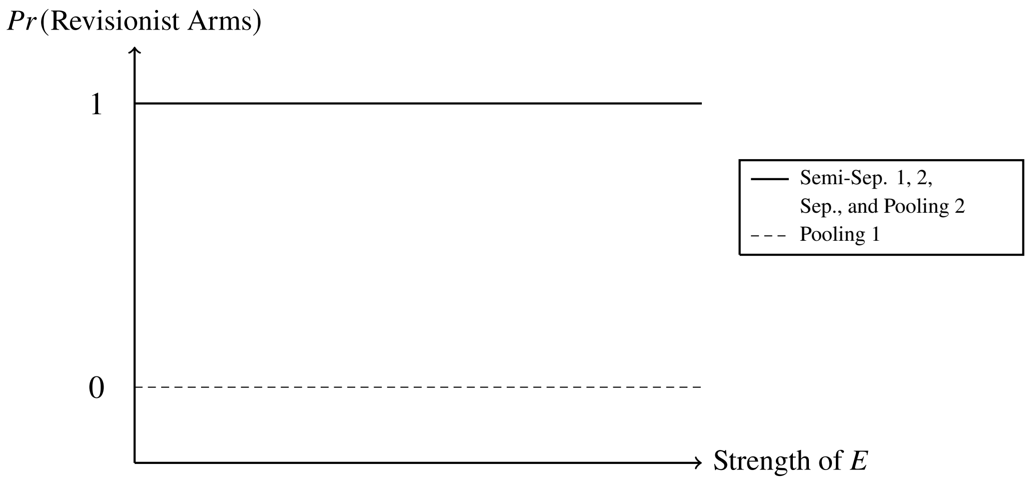
5.3.2 Constraining Effect

Next, I analyze how an alliance decreases a protégé's arming decision for both Revisionist and SQ Type[21]. Figure 4 shows the equilibrium probability of Revisionist's arming. Revisionist arms for sure in the first separating equilibrium, where preventive war happens without an alliance. Even after introducing an alliance, in the second semi-separating equilibrium (Proposition 6), the separating equilibrium (Proposition 4), and the second pooling equilibrium (Proposition 5), Revisionist still always arms. The probability of arming is 1 in these equilibria, and the alliance does not have a constraining effect in these

[20]An exception is Bas and Schub (2017)

[21]In this article, I regard a constraining effect as the decrease in arming by A. This is because a decrease in the arming level leads to a decrease in demand in the later crisis bargaining, and this is exactly what Fang et al. (2014) regard as a constraining effect.

Figure 4: Constraining Effect for Revisionist



cases. This is represented in the solid line in Figure 4.

In the first pooling equilibrium (Proposition 3), however, arming is a clear sign of high revisionism thanks to the screening effect of alliances, and B wants to prevent arming by Revisionist. Thus, to avoid preventive war beforehand, Revisionist does not arm, and the probability of arming is 0, which is shown in the dotted line in Figure 4. The effect size is $1 - 0 = 1$ in Proposition 3. Thus, on average, we should see a constraining effect of alliances in empirical data *if* we see only revisionists, as the size of the constraining effect is either 0 or 1, and the average effect should be positive.

The constraining effect on SQ Type is more complicated, as shown in Figure 5. In the baseline semi-separating equilibrium without an alliance (Proposition 1), SQ Type arms with l^*, and the solid line in the figure shows this value. In the second semi-separating equilibrium, where preventive war can happen even with an alliance (Proposition 6), SQ Type arms with the probability of $l^\dagger$, which is shown as the dotted line. Compared with the baseline value, the size of the constraining effect is $l^* - l^\dagger = (1-p)(w_B(0) - w_B(\epsilon))$, which is strictly positive and increasing in the strength of E. In the first pooling equilibrium (Proposition 3) and the separating equilibrium (Proposition 4), SQ Type does not arm, and thus, the probability of arming is 0. The dashed line in the figure shows this value. In these cases, an alliance has a constraining effect on a protégé's arming level.

However, the second pooling equilibrium (Proposition 5) complicates the constraining effect. In this equilibrium, the alliance deters B from going to a preventive war, which incentivizes SQ Type to always arm. SQ Type's arming probability is 1, which is shown in the dash-dotted line in Figure 5. Given that SQ Type arms with $l^* < 0$ without an alliance (Proposition 1), I can say that an alliance increases the probability of arming by SQ Type because the screening effect of alliances removes the concern of mistaken preventive wars.

This implies that introducing an alliance may have a mixed result: It sometimes deters a protégé's arming, but in other times it rather encourages arming, and thus, the empirical relationship between alliances and protégés' arming is difficult to find. Alliances remove the worry of mistaken preventive war, which pushes SQ Type to arm *more* than before. Thus, even if we could collect a good sample of status-quo-oriented alliance partners, empirical assessment of the average relationship between alliances and military expenditure is not easy. The endeavor of searching for an average effect of alliances mixes many cases, where some of them have positive and others have negative effects. The model tells us that we need to differentiate the second pooling equilibrium from other equilibria since this is the cause of the countervailing effects of alliances. In fact, scholars find mixed results on the effect of alliances on a protégé's military expenditure (Conybeare, 1992, 1994; Kimball, 2010; Morrow, 1993; Sorokin, 1994; Diehl, 1994; Horowitz et al., 2017), and the model can explain why[22].

Even though the constraining effect is difficult to find empirically, this does not mean that such an effect does not exist. Remember that in the first pooling equilibrium (Proposition 3), Revisionists as well as SQ Type do not arm. Given that the deterrence effect is also positive in that equilibrium, this suggests that an alliance can achieve both the deterrence effect and the constraining effect at the same time. This is because alliances have a screening effect: an alliance improves the payoff of the status quo, and this changes Revisionist and SQ Type's arming calculations at a different rate. Alliances push SQ Type to find arming unprofitable, whereas Revisionist still desires arming. This makes arming more informative and thus B can successfully infer A's hidden type.

This is an important insight: reliable alliances can generate both deterrence and constraining effects, without assuming probabilistic support in a war (Benson, 2012) or the costs of disobeying a patron for a protégé (Fang et al., 2014). Also, unlike other alliance models, this model does not assume that a patron knows its protégé's type, meaning the deterrence and constraining effects can be generated without any knowledge about a protégé's intention.

5.3.3 B's Belief

Finally, I analyze B's subjective belief that A is Revisionist after arming in order to see if B can conjecture A's type more correctly thanks to an alliance. This is a core part of the screening effect since it proposes that alliances make arming more informative. Figure 6 shows, on the vertical axis, B's belief about the probability that A is Revisionist

[22]Note that this argument is different from a selection effect, which says that a protégé's threat perception drives alliance formation and a protégé's arming.

after arming and, on the horizontal axis, the strength of the alliance. In the first semi-separating equilibrium without an alliance (Proposition 1), B's belief takes, by using the Bayes Rule, $Pr(\bar{r}|\text{arm}) = \frac{\phi}{\phi+l^*(1-\phi)} = \frac{1}{w_A(1)+w_B(0)}$. The solid line in the figure represents this value. Similarly, B's belief in the second semi-separating equilibrium with an alliance (Proposition 6) is $\frac{1}{w_A(1)+w_B(\epsilon)}$, which is increasing in the strength of E since $w_B(\epsilon)$ becomes smaller as ϵ gets larger. This is shown in a dotted line[23]. B in the separating equilibrium (Proposition 4) and in the first pooling equilibrium (Proposition 3) similarly think that A must be Revisionist after observing A's arming decision. This is because the alliance removes SQ Type's incentive to arm but does not do so for Revisionist. Thus, $Pr(r = 1|\text{arm})$ takes 1, and this is represented in the dash-dotted line in Figure 6.

The second pooling equilibrium (Proposition 5), where both Revisionist and SQ Type successfully arm, can be supported by any value of ϕ, and since both types adopt the same strategy, actors' equilibrium behavior preserves B's prior belief after the arming stage without any updates. Thus, B's belief after arming is $Pr(r = 1|\text{arm}) = \phi$, which can be situated anywhere in Figure 6. So, I do not show this in the figure.

As we can see in Figure 6, the alliance makes arming more informative because B's belief about A's type becomes 1 or gets closer to 1 once an alliance is introduced. This effect is stronger when the strength of E is large, especially in the second separating equilibrium, where preventive war motives still exist. This means that the alliance makes arming more informative than before, and this effect becomes larger as the patron gets stronger when the alliance cannot guarantee peace. Overall, the result shows that the alliance screens A's hidden intention, and thanks to this, B correctly updates its belief after A's arming, even in

[23] The slope is not exact since war payoff functions, $w_i(\cdot)$, is general.

a context where preventive war could happen.

5.4 Endogenous Alliance

Finally, I endogenize the alliance formation decision to study if the screening effect can exist in a more realistic setting. Specifically, I allow E to offer an alliance after Nature's move and then allow A to accept or reject the offer before its arming decision. Note that E does not know A's type, thus its strategy is based on its (common) prior belief, ϕ. I do not assume any costs for alliance formation to remove the mechanism that alliances are costly signaling (Morrow, 1994; Smith, 1995)[24].

Remember that E's payoffs are constituted based on its ideal point, s. When Line 10 is satisfied, E prefers A getting a division of the issue with its support but without additional arming ($rw_A(\epsilon)$) to A getting the issue with arming ($rw_A(1)$) or without E's support ($rw_A(0)$). This is consistent with the idea that E is a status-quo power: E does not like A not being able to get a minimum share of the issue but it does not like A getting too much as a result of arming, either.

Proposition 7 (Pooling 3)**.** *When*

$$\frac{w_A(0)+w_\epsilon}{2} < s < \frac{p(w_A(1)+w_\epsilon)}{2} \tag{22}$$

$$\frac{w_A(0)+w_A(\epsilon)}{w_A(1)+w_A(\epsilon)} < p \tag{23}$$

$$\underline{k} < k < \min\{\overline{k}, p(w_A(1)-w_A(0))\} \tag{24}$$

$$\underline{w}_B < w_B(0) < 1-pw_A(1) \tag{25}$$

$$\phi < \phi^* \tag{26}$$

[24] Even if we assume some costs of alliance formation, we would not get a new insightful conclusion because the results will be always something like "states form an alliance when it is cheap enough".

Figure 5: Constraining Effect on SQ Type

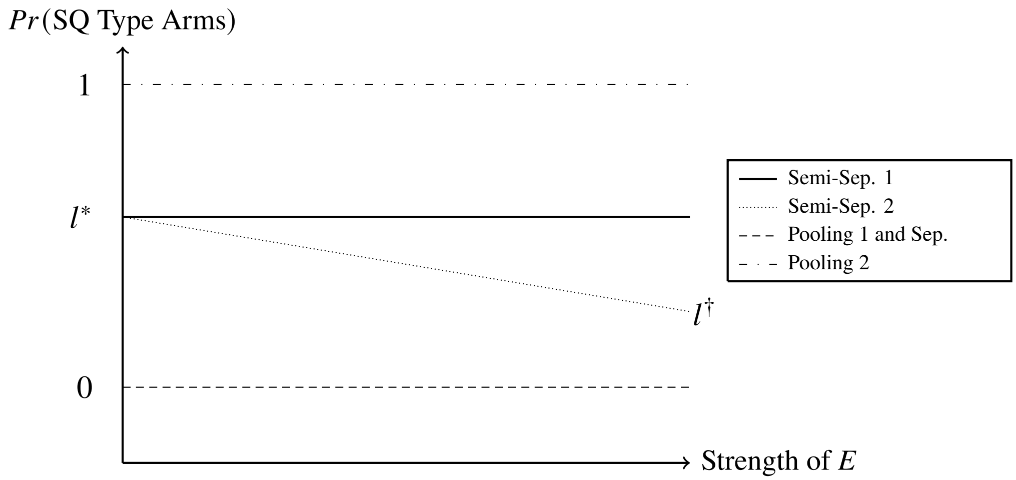
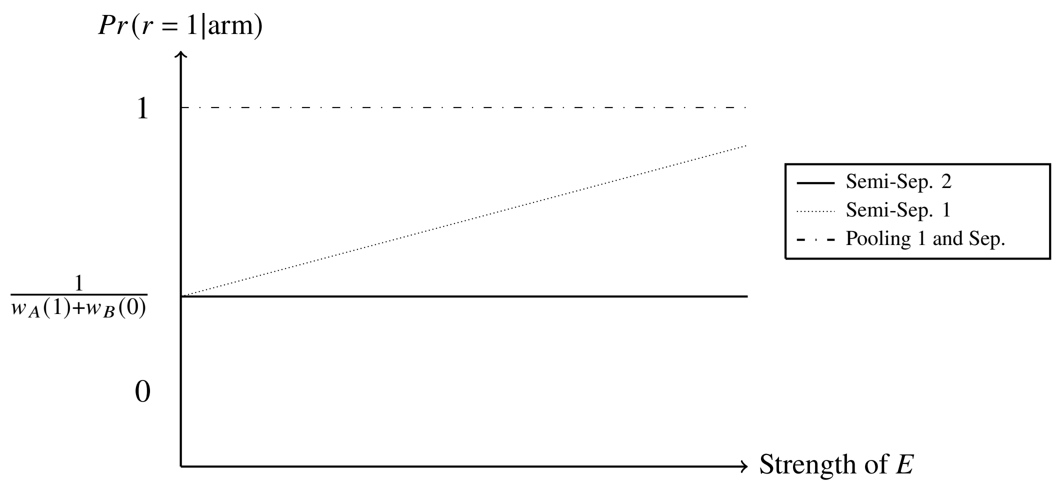


Figure 6: B's Belief after arming



under the parameters where preventive war can happen without a choice of an alliance, there exists a pooling PBE at which E offers an alliance, both types of A accept the offer and do not arm, and B does not attack on the path of play. See Appendix for proof.

Proposition 7 confirms that the screening effect of alliances exists even when alliance formation is endogenous, and this reduces the probability of preventive war and the probability of A's arming. As we discussed in Proposition 1, preventative war can happen without an alliance decision. In the parameters of such an equilibrium, allowing a choice of alliance formation changes the actors' strategy. A pools on accepting the alliance offer and giving up arming. This is because the alliance improves the payoff of the status quo (the payoff of not arming), and this makes Revisionist the only type who could profitably arm. Given this change in actors' preferences, B rationally conjectures that A must be Revisionist after arming. Thus, R cannot mimic SQ Type any longer and gives up arming to avoid preventive war. Since the alliance strictly improves the payoff of not arming and the alliance also makes arming less attractive, A chooses the alliance offer over no alliance, regardless of its type. In other words, Revisionist is well constrained, and causing a power shift is not an option for it, so it improves its payoff by choosing an alliance. Given this screening effect of alliances, E offers an alliance simply because it can achieve a better payoff. Without the alliance offer, the outcome is preventive war or A's successful arming, either of which E does not prefer. But if E offers an alliance, the outcome is surely the improved payoff for A without arming, thanks to both the deterrence and constraining effects.

6 Hypothesis

One main implication of the formal model is that alliances alleviate the incentive for preventive war. A protégé with high revisionism cannot mimic the status-quo-oriented type any longer thanks to the screening effect, which removes the incentive to arm to avoid preventive attack. This screening effect of alliances should decrease the probability of preventive war. So I hypothesize that defensive alliances decrease the probability of preventive war.

Some considerations about alternative explanations are worth doing here. If alliances have no deterrence effect on preventive war motives but have some effect on information asymmetry (Morrow, 1994; Smith, 1995), alliances should have the same level of deterrence effects regardless of the extent of expected power shifts. This is because successful signaling deters a potential enemy at a certain rate based on the current, not future, power balance, and thus, future power shifts are irrelevant to the deterrence effect. Put differently, if alliances can effectively mitigate the preventive war motives, the deterrence effect should be larger when expected power shifts are large, simply because the screening effect is salient when expected power shifts are large.

Similarly, if alliances can deter preventive war because of the capability aggregation mechanism, the deterrence effect of alliances should wear out in expected power shifts. Alliances may have a deterrence effect on preventive war not through the screening effect but through simple power aggregation. Aggregation of power decreases the impact of the protégé's power shifts on the total power balance by introducing the patron's power in the calculation. This makes the opponent less sensitive to the protégé's power shifts, which leads to a lower probability of preventive war. However, such a deterrence effect should be large when expected power shifts are small enough because the (fixed) power of the patron has a larger influence on the total power balance. Conversely, the power of the patron becomes less influential as the future power shifts get larger. Thus, if alliances can deter preventive war through the capability aggregation mechanism, such deterrence effect should wear out as the expected power shifts become large. In other words, if alliances decrease the impact of power shifts at a larger rate as the future expected power shifts get larger, this is good evidence that alliances can deter preventive war through the screening effect, not capability aggregation. Based on this, I derive the first hypothesis.

H1: *Defensive alliances decrease the probability of preventive attack, especially when expected power shifts are large.*

The second main implication of the model is that the average effect of alliances on military expenditure is not easy to find empirically. The screening effect of alliances has a dual effect on the arming decision of a protégé. On the one hand, alliances reduce the necessity of arming by increasing the payoff of staying at the status quo, and this constrains the growth of military expenditure (Proposition 3 and 4). On the other hand, alliances also remove the fear of preventive attack by deterring the potential opponent, which incentivizes a protégé to arm more (Proposition 5). Figure 5 in the previous chapter shows this dual effect of alliances on a status-quo-oriented type.

To derive the model's implication for the arming decision of a protégé, I focus on the costs of internal arming, which is represented by a parameter called k. Suppose that there is an alliance and we move k from low to high in Figure 2. This clearly decreases a protégé's arming. When k is low, a protégé may arm more often because it does not fear the (mistaken) preventive wars any longer thanks to the deterrence effect of alliances. But when k is high, the protégé loses the incentive to arm because of the improved payoffs of staying in the status quo, meaning the protégé arms less likely. Thus, by changing the cost of internal arming from low to high, we should see a clear decline in a protégé's arming.

When a protégé does not have an alliance, however, this clear decline should not be observed. Remember that the first semi-separating equilibrium, where states go to a preventive war when they do not have an alliance (Proposition 1), occupies a larger area in Figure 2. In this equilibrium, Revisionist arms for sure and Pacifist arms with the probability of l^*. Since l^* does not have the arming cost parameter, the probability of a protégé's arming does not change in k in this equilibrium. Thus, when states do not have an alliance, the internal costs of arming should have a smaller effect on a protégé's military expenditure. In sum, the model shows that countries become more sensitive to the internal cost of arming uniquely when they have an alliance. Alliances incentivize a protégé to arm more when the costs are low, but they also discourage a protégé from arming when the costs are high. This effect should not appear when states do not have an alliance. Thus, I derive the following two hypotheses.

H2a: *The larger the internal arming costs are, the less a protégé spends on military expenditure.*

H2b: *Defensive alliances strengthen the effect of the costs of internal arming on military expenditure.*

These expectations are not derived from existing theories of alliances and military expenditure. These existing theories do not explicitly consider the interaction between internal arming and alliances in the first place. For example, those who consider that alliances decrease members' military expenditure often assume that alliances are a substitution for internal arming (Morrow, 1993; Conybeare, 1992, 1994; Kimball, 2010; Sorokin, 1994; Plümper and Neumayer, 2015). And those who support that alliances increase their military expenditure often rely on the idea that new alliances bring new military obligations (Diehl, 1994; Alley and Fuhrmann, 2021; Horowitz et al., 2017). In both cases, it is not clear how the effect of internal arming costs is conditioned on the alliance status. This is the same for the studies that look at the conditional effects of alliances. Digiuseppe and Poast (2018) argue that the regime type of a patron conditions the effect of alliances and Alley (2021) claims that alliances with deep cooperation reduce military expenditure. Both of the studies focus on the characteristics of alliances themselves or alliance partners, and do not touch on the domestic factor of protégé's. Thus, the interaction of the costs of internal arming and alliances is not derived from these existing theories.

7 Empirical Design

7.1 Measuring Expected Power Shifts

A key variable in the first hypothesis is expected power shifts. To construct this variable, I follow (Bell and Johnson, 2015)'s procedures. Bell and Johnson (2015) measured expected power shifts by predicting the future military capability based on the results of regression analysis of past military capability[25] on several observable variables such as economic potential[26], defensive and offensive alliances from the Alliance Treaty Obligation Project (ATOP) (Leeds et al., 2002), rival relationships (Thompson, 2001), logged battle deaths from the COW project (Sarkees and Wayman, 2010), and the level of democracy from the Polity project (Marshall et al., 2017). Specifically, they measure the future power shift in a directed dyad between Target (T) and Challenger (C) in the following way.

$$\text{Expected PowerShift}_{TC,t} = \frac{\hat{\text{Power}}_{T,t+1}}{\hat{\text{Power}}_{T,t+1} + \text{Power}_{C,t+1}} - \frac{\text{Power}_{T,t}}{\text{Power}_{T,t} + \text{Power}_{C,t}}$$

, where $\hat{\text{Power}}_{T,t+1}$ is a predicted military capability of T in year $t+1$ and $\text{Power}_{i,t}$ is observed military capability of country i at year t. This value takes from -1 to 1. This value captures the difference between the predicted future power balance in year $t+1$ and the actual power balance in year t.

Since their measurement includes alliances as a predictor of future military capability, using their replication data directly will bias the estimation of the alliance effect as the effect of alliances will appear through both predicted power shifts and the coefficient of alliance. Instead, I recalculate future military capability by removing alliances from the predictor variables. Also, I extend their data from 1816-2000 to 1816-2007 and, to make up for the lack of prediction power, I newly include some predictors[27] such as the difference in the

[25]This is an average of the share of military expenditure and the share of military personnel in the world in a given year, which is the same as a military part of the Composite Index of National Capability (CINC) score (Singer, 1988).

[26]This is an average of the share of urban and total population, steel production, and primary energy consumption. This is the same as the economic part of the CINC score (Singer, 1988).

[27]The purpose of this analysis is to predict future power based on observable characteristics of states. Thus, adding many variables does not cause a concern of post-treatment bias (Montgomery et al., 2018).

Polity index from the previous year (Marshall et al., 2017), estimated logged GDP, GDP growth, estimated logged population, population growth, estimated GDP per capita, GDP per capita growth (Anders et al., 2020), COW's major power status, and nuclear powers, with country-fixed effects[28]. These data are gathered and analyzed by using an R package `peacesciencer` developed by Miller (2022)[29]. The distribution of the newly predicted power shifts, as well as the OLS regressions of military capability on these predicters, are shown in C.1 in Appendix. In the result section, I report the solo effect of power shifts on war, in addition to the effect of alliances, to validate the new measure of power shifts.

7.2 Alliance and Preventive War

Next, I explain the empirical model to test the first hypothesis. As a dependent variable, I code a dummy of the onset of war initiated by one country against another in a directed dyad (Sarkees and Wayman, 2010). My independent variable is the interaction of expected power shifts and defensive alliances. The former variable is explained above, and the latter variable is from ATOP (Leeds et al., 2002), where I record if a target in a directed dyad has a defensive alliance[30]. Following Bell and Johnson (2015), I include the same control variables such as mutual democracy measured by Polity (Marshall et al., 2017), foreign policy similarity called S-score (Chiba et al., 2015), a dummy variable for contingency, the logged distance of capitals, and up-to-cubic polynomial of peace years to control temporal dependencies (Carter and Signorino, 2010). The unit of analysis is directed dyad year, and the year coverage is from 1818 to 2007. The empirical model I estimate is a logistic

[28] I did not include year-fixed effects because this makes predictions using new data difficult as there is no information of the fixed effects of the years in the new data. Bell and Johnson (2015) adopt the same strategy.
[29] I use `peacesciencer` in the later analyses, as well.
[30] I did not count defensive alliances that are signed by both countries in a dyad.

regression with directed-dyad- and year-fixed effects, which is shown below.

$$\text{Logit}(\mathbb{E}(Y_{it})) = \beta_1 \text{Power shift}_{it} + \beta_2 \text{Alliance}_{it} + \beta_3 \text{Power shift}_{i,t} \times \text{Alliance}_{i,t} + \mathbf{X}'\beta + \lambda_i + \eta_t + \epsilon$$

, where Y_{it} is a dummy variable of war in year t and in directed dyad i, βs are coefficients, $\mathbf{X}'$ is a matrix of control variables, λ_i and η_t are directed-dyad- and year- fixed effects, and ϵ is an error term.

7.3 Arming Costs

The key variable to test H2a and H2b is the cost of arming, especially the cost of internal arming[31]. The context of the formal model is whether states successfully convert their (internal) resources into military capability. States increase their military capability in various ways, including direct taxation on citizens, forced labor, conscription, and voluntary military recruitment (Zielinski, 2016). In so doing, leaders have to get political "consent" from citizens to implement armament. Essentially, citizens pay the costs of arming with their own blood, money, and labor, so this consent is important even in autocracies. If leaders fail to get such consent, they face punishment at ballot boxes (Kreps, 2018, See), domestic unrest, or revolution, which shortens the political life of the leaders (DiGiuseppe, 2015). One example is Bloody Sunday in Russia in 1905, which was caused by the social

[31] Another way of arming is external arming, namely sovereign borrowing. States can avoid the "butter versus gun" trade-offs by sovereign borrowing (Poast, 2019), so countries may prefer it to internal arming. In fact, Krainin et al. (2022) and Landry (2023) operationalize the costs of sovereign borrowing by using the UK's (global) risk-free interest rate. This variable takes the same value for all countries in a given year, and thus, the effect of the costs of external arming is absorbed in the year-fixed effect in the analysis of this paper.

costs of the Russo-Japanese War. The Russian government faced a domestic opposition movement claiming a better working environment, the introduction of a general election, and the end of the war, and its oppressive reaction towards the movement triggered the later Russian Resolution. Imposing the costs of armament on domestic people risks the tenure of leaders. Thus, how leaders manage to get domestic consent is a key consideration in internal arming. In fact, existing studies often study war financing through the perspective of leaders' survival and citizens' dissatisfaction with internal arming (Levi, 1988; Zielinski, 2016; DiGiuseppe, 2015).

The operationalization of the difficulty of getting consent from citizens is a challenging task. One may think of the level of democracy as a proxy for the need for leaders to listen to domestic actors. Although this captures one aspect of democracy, democratic institutions are also correlated with many other things such as the efficiency of the government (Adam et al., 2011; Charron and Lapuente, 2010), the ability to collect tax effectively (Wang and Xu, 2018; Besley and Persson, 2009), and nationalism and patriotism (Tudor and Slater, 2021; Smith, 2021; Mylonas and Tudor, 2021), all of which make internal arming easier and cheaper. Given that autocracies can generate audience costs as well (Weeks, 2008; Li and Chen, 2021), the assumption that democracies listen to people's voices more needs some caution. In fact, DiGiuseppe and Shea (2022) shows that US support for a country fosters local property rights and this improves its extractive capacity, implying democratic institutions decrease the costs of arming. Regime types encompass various institutional consequences, and it is difficult to disentangle the costs of internal arming from other factors. Thus, democracy is a limited proxy for the costs of internal arming.

Instead, following Chapman et al. (2015)'s idea that the level of military capability is a result of domestic bargaining between the government and the opposition, I present the size of potential opposition groups as a proxy of arming costs. The size of potential opposition groups captures the degree to which leaders are able to extract recourse from domestic actors without strong resistance. When the size of a potential opposition group is large, the price of getting citizens' consent is high because the latent threat of overturning the government is large, and thus, even smaller social costs of arming can trigger an actual punishment regardless of the form of resource extract, including taxation, conscription, or forced labor. This makes leaders unable to increase military capability or make them spend their own political or economic resources to carry out the armament. In other words, the size of the potential opposition group measures the need for leaders to pay careful attention to domestic actors due to the fear of losing office when they extract resources from the private sector.

Also, the size of potential opposition groups would correlate with the (in)effectiveness of resource extraction: The size of potential opposition groups increases when citizens' utility of the status quo is small due to factors such as bad economic performance, political corruption, or the unequal distribution of wealth. These factors worsen both the efficiency and the extent of possible resource extraction, leading to higher costs of arming. The limited ability to maintain public security and to provide public service increases the size of opposition as well, and these factors worsen the government's ability to collect revenue at the same time because of the incapability of exercising authority. Thus, the size of potential opposition groups is highly related to the conversion rate of domestic resources into military capability.

I measure this variable using the index of regime opposition groups size, which is contained in the V-dem dataset (Coppedge et al., 2020). This index measures the share of the domestic population that are opposition actors to the current political regime. This index can include a wide range of opposition groups from rural/urban working classes such as peasants and labor unions, civil servants, and business elites to ethnic, racial, or religious groups. They can be counted as opposition groups as far as they oppose the current political regime. To be counted in this variable, "groups need not be actively mobilized or explicitly engaged in high-level opposition activities" (Coppedge et al., 2020), and thus, this is a good measure of *potential* political costs of imposing the burden of internal arming. This index covers years from 1900 to 2023, and the minimum and maximum values of this variable are -3.870 and 3.161, respectively, with a mean of -0.321. The distribution of this variable is shown in Figure 7. To check the robustness of my empirical results, in the later analysis, I use State Capacity data developed by Hanson and Sigman (2021), which covers years from 1960 to 2015, as a proxy of the *easiness* of internal armament (as opposed to arming costs), and report the results later.

7.4 Alliance and Arming

To test the hypotheses about alliances and military expenditure, I construct panel data from 1900 to 2016 based on the COW's state membership (Correlates of War, 2011) with state-year as a unit of analysis. Following Digiuseppe and Poast (2018), the dependent variable is the natural log of military expenditure from the National Material Capability data Singer (1988), and I include the one-year lagged military expenditure in the right-hand side to address potential concerns of serial correlation. The independent variable is the interaction of defensive alliances from the ATOP dataset (Leeds et al., 2002) and the cost of internal

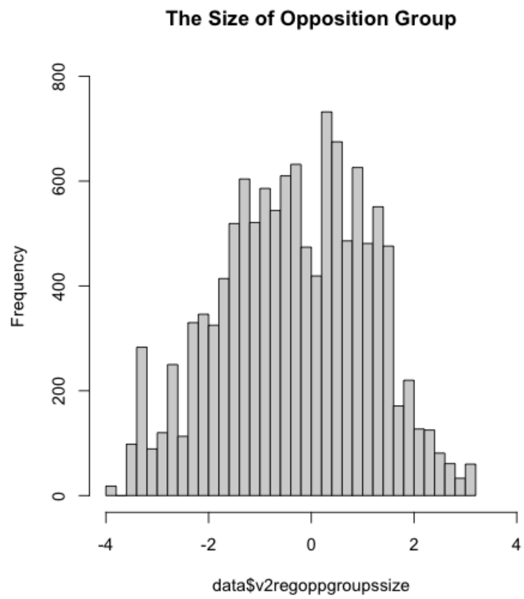


Figure 7: The Distribution of Arming Costs

arming measured by the size of opposition groups. I also add several control variables. I control international war and civil war (Sarkees and Wayman, 2010), nuclear weapons, rivalry (Thompson, 2001), the democracy level by using V-dem's polyarchy index, lagged GDP and population, as well as their growth from the previous year (Anders et al., 2020). I estimated the empirical model by using OLS with both country- and year-fixed effects.

8 Results

8.1 Effects on Preventive War

I report the results of preventive war and alliances in this section. Table 1 shows the results of the regression analysis. Standard errors are displayed in the parentheses and they are clustered by the directed dyad. Model 1 is a simple regression analysis on expected power shifts, and Model 2 has full control variables, but does not include the interaction term of alliances and expected power shifts. Models 3, 4, 5, and 6 have the interaction term. Model 3 contains only the main independent variables. Models 4 and 5 have the full control variables but have only directed-dyad-fixed effect or year-fixed effect, respectively. Model 6 has all control variables and both fixed effects.

First, I confirm that expected power shifts cause a preventive war. In all models, expected power shifts are associated with the higher possibility that a growing country is attacked by another country. This relationship is positive and significant at the $\alpha = 5\%$ level. The effect size is very large. Let's take the coefficient in Model 5 as an example since it has the smallest coefficient. The coefficient of expected power shifts is 3.515, and, substantively, this means that a one-unit increase in power shift makes the odds of war occurrence $e^{3.515} \simeq 33$ times more. In other words, a 5% increase in the expected balance of power in the next year is associated with 1.68 times more chances of war. This effect size is stronger if we use different models. These results are consistent with Bell and Johnson (2015)'s main finding: large and rapid power shifts cause a preventive war (Fearon, 1995; Powell, 2006). The results also confirm the good performance of the new measure of power shifts.

Table 1: Alliance and Preventive War

Dependent Variable:	War					
Model:	(1)	(2)	(3)	(4)	(5)	(6)
Variables						
Expected Power Shift	10.39***	21.66***	10.92***	16.72***	3.515***	47.08***
	(2.586)	(5.960)	(3.573)	(3.797)	(0.8440)	(18.27)
Alliance		1.660	-0.4550	-0.4003	0.1485	2.273
		(4.162)	(0.4944)	(1.036)	(0.6292)	(4.182)
Alliance × Expected Power Shift			-1.751	-7.809**	-1.185	-33.75**
			(3.768)	(3.501)	(2.013)	(16.95)
Mutual Democracy		-3.842*		-4.169***	-1.705**	-6.581
		(2.169)		(1.556)	(0.8537)	(5.700)
Foreign Policy Similarity		-0.2545		-0.0245	-2.369**	-0.0470
		(3.443)		(1.056)	(0.9905)	(3.690)
Contingency		18.54***		4.218***	1.816***	21.25***
		(2.805)		(1.406)	(0.4424)	(4.119)
Distance		-100.3		-33.78	-0.6536***	-264.1
		(745.3)		(145.1)	(0.1419)	(683.5)
Peace Year		0.0831		-0.0101	-0.0973*	0.0929
		(0.1091)		(0.0749)	(0.0499)	(0.0832)
Peace Year^2		0.0027		0.0017	0.0025**	0.0031
		(0.0047)		(0.0018)	(0.0012)	(0.0027)
Peace Year^3		-2.52×10^{-5}		-9.82×10^{-6}	-1.58×10^{-5}**	-2.76×10^{-5}
		(3.48×10^{-5})		(1.07×10^{-5})	(7.84×10^{-6})	(2.07×10^{-5})
Fixed-effects						
Directed Dyad	Yes	Yes	Yes	Yes		Yes
Year	Yes	Yes	Yes		Yes	Yes
Fit statistics						
Observations	1,933	476	1,933	2,370	141,624	476
Squared Correlation	0.16798	0.48619	0.16950	0.12089	0.00162	0.52575
Pseudo R^2	0.24276	0.46864	0.24533	0.22449	0.22970	0.51640
BIC	1,574.0	576.38	1,587.1	669.19	1,029.2	568.98

*Signif. Codes: ***: 0.01, **: 0.05, *: 0.1*

Next, let's examine how defensive alliances alleviate this tragic relationship. Defensive alliances alone do not seem to have a strong effect on the probability of war. Overall, the alliance variable does not have significant results even in Model 2, where there is no interaction term, and the signs of the coefficients are not stable.

Note that the first hypothesis posits that the probability of preventive war decreases when a rising country has a defensive alliance and this relationship is strong when expected power shifts are large. To investigate the interaction between alliances and expected power shifts, I plot the marginal coefficients of expected power shifts in Figure 8 based on Model 6[32]. Figure 8 shows the alliance status on the x-axis and the coefficient of preventive war on the y-axis. Error bars are 95 confidence intervals. As we can see, the alliance status effectively decreases the effect of power shifts on war. When a rising state does not have an alliance, the coefficient of expected power shifts is about 47. When a rising state has an alliance, the coefficient of power shifts decreases to 15.6 and becomes not significant. Thus, it is clear that alliances mitigate the effect of power shifts. As I discussed earlier, this deterrence effect does not solely come from alliances, but the interaction of alliances and power shifts as the alliance term in Model 6 does not have a negative impact on the occurrence of war. The negative and significant interaction term in Model 6 implies that the deterrence effect of alliances becomes stronger as the expected power shifts get larger. This is consistent with my expectation in **H1**. Defensive alliances generate a deterrence effect when power shifts are expected, and thus, they successfully mitigate the risk of preventive war.

To further investigate the relationship between alliance and preventive war, I plot the predicted probability of war depending on the alliance status in Figure 9. Calculating the

[32]The same plot based on Model 4 and 5 are shown in C.2.1 in Appendix. The main finding holds in these models.

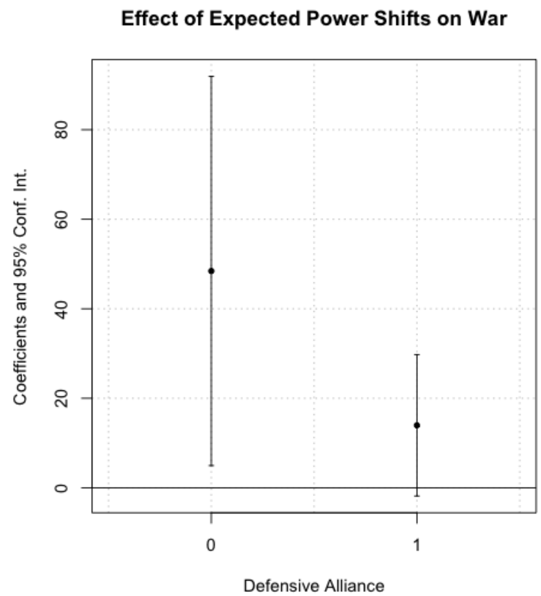


Figure 8: Marginal Coefficient of Expected Power Shifts

predicted probability of fixed effect models with a binary dependent variable is challenging, so I report results based on a logistic regression without any fixed effects here. The figure shows expected power shifts on the x-axis and the probability of war on the y-axis. It is indicated that the predicted probability of war is lower when states have an alliance, and this relationship becomes stronger as the expected power shifts get larger, suggesting that the deterrence effect of alliances becomes stronger as the expected power shifts get larger. This is consistent with the implication that alliances can deter preventive wars in the context of large power shifts.

I checked the robustness of this finding in two ways. First, while the main analysis used expected power shifts in year $t+1$, states may look at the longer future. I confirmed that the main findings still hold even when expected power shifts in the year $t+3$ and $t+5$

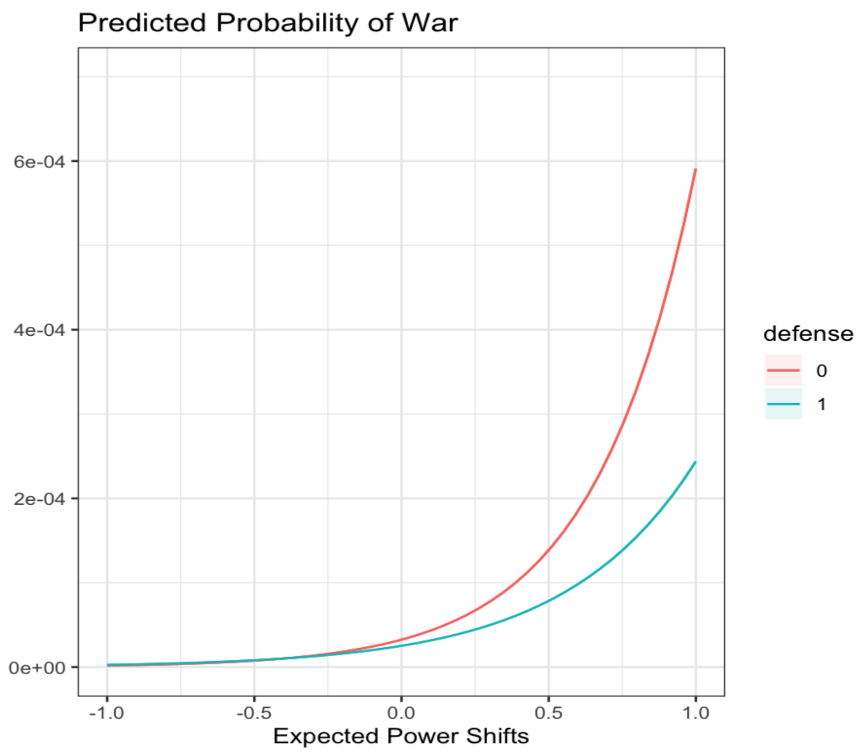


Figure 9: Alliance and the Probability of Preventive War

(See C.2.3 in Appendix). Second, to keep consistency with Bell and Johnson (2015) and to cover as many years as possible, I use the S-score as foreign policy similarity. But the S-score is constructed based on alliances, which may bias the effect of alliances. To address this concern, I conduct robustness checks using the S-score calculated UN voting data, covering after 1946 (Häge, 2011). The result shows the same pattern: alliances reduce preventive wars. This is reported in C.2.2 in Appendix.

8.2 Effects on Military Expenditure

Next, I report the result of the analysis of military expenditure. Table 2 shows the results of the regressions. Model 7 is a simple OLS regression of military expenditure on arming costs (the size of potential opposition groups). Model 8 is an OLS regression with full control variables but the alliance variable and the interaction term between arming costs and alliances. I include the alliance variable in Models 9 and 10 and the interaction term in Model 10.

Models 7, 8, and 9 show a negative correlation between the costs of arming and military expenditure. The coefficients are all negative and significant at the $\alpha = .05$ level. This supports the hypothesis that arming costs decrease military expenditure. Notably, The arming costs (the size of potential opposition) have a stronger correlation with military expenditure than the democracy index (Polity). Remember that the measure of the size of potential opposition groups takes values from -3.87 to 3.16, whereas Polity takes values from -10 to 10. Given that the coefficients of these variables are different by 4-5 times (For example, -0.0267 for the size of potential opposition and -0.0056 for Polity in Model 8), the effect of a one-unit increase in the potential opposition size variable is equivalent to a four-

Table 2: Military Expenditure, Alliance, and Arming Costs

Dependent Variable:	ln(Mil. Exp.)			
Model:	(7)	(8)	(9)	(10)
Variables				
Arming Costs	-0.0271	-0.0267***	-0.0268***	-0.0196**
	(0.0453)	(0.0078)	(0.0081)	(0.0094)
Alliance			0.0737***	0.0702***
			(0.0175)	(0.0172)
Alliance × Arming Costs				-0.0116
				(0.0096)
International War		0.2237***	0.2216***	0.2203***
		(0.0333)	(0.0324)	(0.0324)
Civil War		0.1308***	0.1334***	0.1335***
		(0.0311)	(0.0315)	(0.0315)
ln(Mil. Exp.)$_{t-1}$		0.8052***	0.8027***	0.8021***
		(0.0178)	(0.0177)	(0.0178)
Nuclear		0.0383	0.0472	0.0428
		(0.0367)	(0.0336)	(0.0330)
Rivalry		-0.0043	-0.0067	-0.0065
		(0.0174)	(0.0178)	(0.0180)
Democracy		-0.0056***	-0.0057***	-0.0057***
		(0.0015)	(0.0015)	(0.0015)
ln(GDP)		0.1571***	0.1660***	0.1646***
		(0.0259)	(0.0269)	(0.0269)
ln(Population)		0.0319	0.0029	0.0041
		(0.0283)	(0.0279)	(0.0280)
GDP Growth		5.301*	5.192*	5.151*
		(3.054)	(3.091)	(3.095)
Population Growth		23.14	21.84	22.01
		(16.38)	(16.49)	(16.47)
Fixed-effects				
Country	Yes	Yes	Yes	Yes
Year	Yes	Yes	Yes	Yes
Fit statistics				
Observations	10,877	8,983	8,983	8,983
R^2	0.93233	0.98448	0.98454	0.98455
Within R^2	0.00076	0.76021	0.76122	0.76130

Clustered (Country) standard-errors in parentheses
*Signif. Codes: ***: 0.01, **: 0.05, *: 0.1*

or five-point increase in the Polity index. A four- or five-point increase in Polity is a large jump, so the effect size of the potential opposition size variable is substantively meaningful. The opposition size is a better predictor of military expenditure than the democracy index, and thus, it shows a feature of a good proxy.

Finally, I test the last hypothesis. Remember that H2b posits that the negative effect of arming costs on military expenditure is stronger when states have a defensive alliance. This is because alliances make a state sensitive to its internal arming costs: with low arming costs, alliances encourage arming by removing the fear of preventive attack thanks to the screening effect, and with high arming costs, alliances discourage arming by improving the utility of the status quo. In Figure 10, I display the marginal coefficients of the arming costs when states have a defensive alliance and when they do not based on Model 10. The error bars show the 95% confidence intervals. The result is clear. Alliances make the effect of arming costs stronger. The negative coefficient of arming costs on military expenditure is large and significant when states have an alliance. However, the coefficient of arming costs when states do not have an alliance is still negative but the effect size is smaller and not significant. Thus, my hypothesis is supported. To check the robustness of this result, I use a different measure of arming costs: state capacity. State capacity is the ability to implement policy, and this captures the "easiness" of internal arming. This measure is from Hanson and Sigman (2021)'s data. The result is shown in Figure 11. Note that the sing of the coefficients should be flipped as the arming easiness is the opposite concept. The result is consistent with the main finding: the easiness of arming has a larger effect on military expenditure uniquely when states have an alliance.

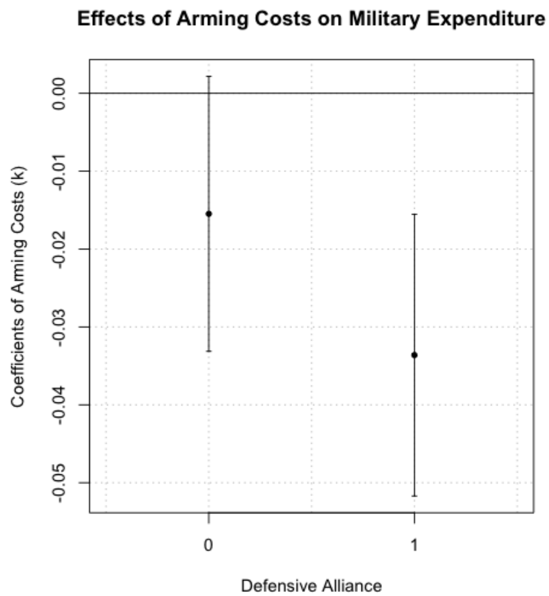


Figure 10: Marginal Coefficients of Arming Costs

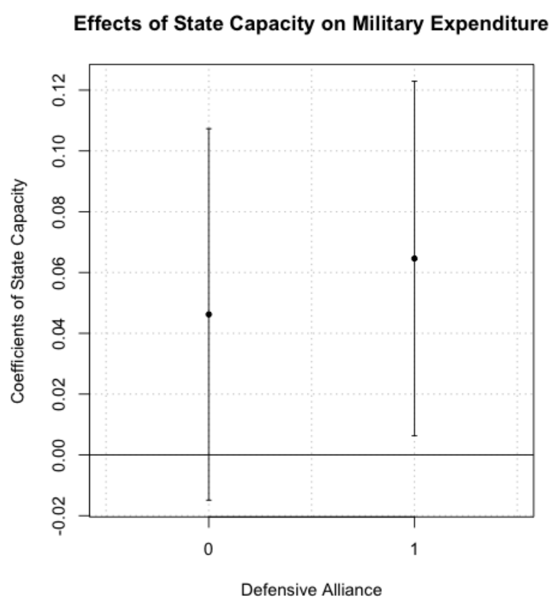


Figure 11: Marginal Coefficients of Arming "Easiness"

9 Conclusion

In this chapter, I investigated a formal model in which a rising protégé chooses arms buildups but this decision could make it suffer from a preventive war by another country facing the fear of the revisionism of the rising country. Through the comparison of versions of the model with and without an alliance, as well as a model with an endogenous alliance, this chapter proposed a *screening effect* of alliances. Alliances make arming more informative by changing the thresholds of arming at a different rate for revisionists and status-quo-oriented types, and this reveals an otherwise hidden intention of a rising state. This screening effect eliminates or reduces the incentive of preventive war. Existing theories often argue that alliances bring peace by revealing uncertainty of a patron's resolve (Morrow, 1994; Smith, 1995), which is often referred to as an information problem (Fearon, 1995). In this chapter, however, alliances solve a different information problem: alliances can reveal the uncertainty on the protégé side, and this reduces wars caused by commitment problems(Fearon, 1995; Powell, 2006; Debs and Monteiro, 2014; Bas and Schub, 2017).

Specifically, I examined how the screening effect affects the two main goals of alliances: the deterrence effect and the constraining effect. I find that alliances always reduce the probability of preventive war and discourage arms buildups by a revisionist at the same time. But the model also identifies that sometimes alliances rather encourage arms buildups by a protégé because the screening effect removes the fear of mistaken preventive wars. This suggests that the deterrence effect is easier to find empirically, but the constraining effects must exist only in a proper sample of cases, and this can explain why scholars find mixed results on the effect of alliances on military expenditures.

More broadly, the screening effect suggests that alliances can bring peace by genuinely removing distrust among countries, not just by increasing the costs of aggression as existing theories posit. This insight provides a new understanding of why US alliances in the post-WWII era were so successful: they constrained the US's past enemies, deterred new communist competitors at the same time and, more importantly, reduced hostile feelings among countries, which consolidated the US hegemony by bringing the stability of the international system.

References

Adam, A., Delis, M. D., and Kammas, P. (2011). Are democratic governments more efficient? *European Journal of Political Economy*, 27(1):75–86.

Alley, J. (2021). Alliance participation, treaty depth, and military spending. *International Studies Quarterly*, 65(4):929–943.

Alley, J. and Fuhrmann, M. (2021). Budget breaker? the financial cost of us military alliances. *Security Studies*, 30(5):661–690.

Anders, T., Fariss, C. J., and Markowitz, J. N. (2020). Bread before guns or butter: Introducing surplus domestic product (SDP). *International studies quarterly: a publication of the International Studies Association*, 64(2):392–405.

Banks, J. S. and Sobel, J. (1987). Equilibrium selection in signaling games. *Econometrica: Journal of the Econometric Society*, pages 647–661.

Bas, M. A. and Schub, R. (2017). Peaceful uncertainty: When power shocks do not create commitment problems. *International studies quarterly: a publication of the International Studies Association*, 61(4):850–866.

Bell, S. R. and Johnson, J. C. (2015). Shifting power, commitment problems, and preventive war. *International studies quarterly: a publication of the International Studies Association*.

Benson, B. V. (2012). *Constructing international security: Alliances, deterrence, and moral hazard*. Cambridge University Press.

Benson, B. V., Meirowitz, A., and Ramsay, K. W. (2014). Inducing deterrence through moral hazard in alliance contracts. *Journal of Conflict Resolution*, 58(2):307–335.

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

Besley, T. and Persson, T. (2009). The origins of state capacity: Property rights, taxation, and politics. *American economic review*, 99(4):1218–1244.

Carter, D. B. and Signorino, C. S. (2010). Back to the future: Modeling time dependence in binary data. *Political analysis: an annual publication of the Methodology Section of the American Political Science Association*, 18(3):271–292.

Carter, J. (2017). The political cost of war mobilization in democracies and dictatorships. *Journal of Conflict Resolution*, 61(8):1768–1794.

Cha, V. D. (2010). Powerplay: Origins of the US Alliance System in Asia. *International Security*, 34(3):158–196.

Chadefaux, T. (2011). Bargaining over power: when do shifts in power lead to war? *International Theory*, 3(2):228–253.

Chapman, T. L., McDonald, P. J., and Moser, S. (2015). The domestic politics of strategic retrenchment, power shifts, and preventive war. *International Studies Quarterly*, 59(1):133–144.

Charron, N. and Lapuente, V. (2010). Does democracy produce quality of government? *European journal of political research*, 49(4):443–470.

Chiba, D., Johnson, J., and Leeds, B. A. (2015). Careful commitments: Democratic states and alliance design. *The Journal of Politics*.

Chiozza, G. and Goemans, H. E. (2004). International conflict and the tenure of leaders: Is war still ex post inefficient? *American Journal of Political Science*, 48(3):604–619.

Conybeare, J. A. (1992). A portfolio diversification model of alliances: The triple alliance and triple entente, 1879-1914. *Journal of Conflict Resolution*, 36(1):53–85.

Conybeare, J. A. (1994). Arms versus alliances: The capital structure of military enterprise. *Journal of Conflict Resolution*, 38(2):215–235.

Cooley, A., Nexon, D. H., Poast, P., Alley, J., and Fuhrmann, M. (2022). Estimating alliance costs: An exchange. *Security Studies*, 31(3):510–532.

Coppedge, M., Gerring, J., Knutsen, C. H., Lindberg, S. I., Teorell, J., Altman, D., Bernhard, M., Fish, M. S., Glynn, A., Hicken, A., Luhrmann, A., Marquardt, K. L., McMann, K., Paxton, P., Pemstein, D., Seim, B., Sigman, R., Skaaning, S.-E., Staton, J., Cornell, A., Gastaldi, L., Gjerløw, H., Mechkova, V., von Römer, J., Sundtröm, A., Tzelgov, E., Uberti, L., ting Wang, Y., Wig, T., and Ziblatt, D. (2020). V-dem codebook v10. Varieties of Democracy (V-Dem) Project.

Correlates of War (2011). State system membership list, v2016.

Croco, S. E. and Weeks, J. L. (2016). War outcomes and leader tenure. *World Politics*, 68(4):577–607.

Dafoe, A., Renshon, J., and Huth, P. (2014). Reputation and status as motives for war. *Annual Review of Political Science*, 17(1):371–393.

Debs, A. and Goemans, H. E. (2010). Regime type, the fate of leaders, and war. *American Political Science Review*, 104(3):430–445.

Debs, A. and Monteiro, N. P. (2014). Known unknowns: Power shifts, uncertainty, and war. *International organization*, 68(1):1–31.

Diehl, P. F. (1994). Substitutes or complements?: The effects of alliances on military spending in major power rivalries. *International Interactions*, 19(3):159–176.

DiGiuseppe, M. (2015). The fiscal autonomy of deciders: Creditworthiness and conflict initiation. *Foreign Policy Analysis*, 11(3):317–338.

Digiuseppe, M. and Poast, P. (2018). Arms versus democratic allies. *British Journal of Political Science*, 48(4):981–1003.

DiGiuseppe, M. and Shea, P. E. (2022). Us patronage, state capacity, and civil conflict. *The Journal of Politics*, 84(2):767–782.

Edry, J., Johnson, J. C., and Leeds, B. A. (2021). Threats at home and abroad: Interstate war, civil war, and alliance formation. *International Organization*, 75(3):837–857.

Fang, S., Johnson, J. C., and Leeds, B. A. (2014). To concede or to resist? the restraining effect of military alliances. *International Organization*, 68(4):775–809.

Fearon, J. D. (1995). Rationalist explanations for war. *International organization*, 49(3):379–414.

Fearon, J. D. (1997). Signaling foreign policy interests: Tying hands versus sinking costs. *Journal of conflict resolution*, 41(1):68–90.

Häge, F. M. (2011). Choice or circumstance? adjusting measures of foreign policy similarity for chance agreement. *Political Analysis*, 19(3):287–305.

Hanson, J. K. and Sigman, R. (2021). Leviathan's latent dimensions: Measuring state capacity for comparative political research. *The Journal of Politics*, 83(4):1495–1510.

Horowitz, M. C., Poast, P., and Stam, A. C. (2017). Domestic signaling of commitment credibility: Military recruitment and alliance formation. *Journal of Conflict Resolution*, 61(8):1682–1710.

Jervis, R. (1978). Cooperation under the security dilemma. *World politics*, 30(2):167–214.

Johnson, J. C. (2017). External threat and alliance formation. *International Studies Quarterly*, 61(3):736–745.

Johnson, J. C. and Leeds, B. A. (2011). Defense pacts: A prescription for peace? *Foreign Policy Analysis*, 7(1):45–65.

Johnson, J. C., Leeds, B. A., and Savun, B. (2024). Alliances and civil war intervention. *International Studies Quarterly*, 68(4):sqae128.

Kimball, A. L. (2010). Political survival, policy distribution, and alliance formation. *Journal of Peace Research*, 47(4):407–419.

Kinsella, D. (1994). Conflict in context: Arms transfers and third world rivalries during the cold war. *American Journal of Political Science*, pages 557–581.

Krainin, C., Ramsay, K. W., Wang, B., and Ruggiero, J. J. (2022). Preventive war and sovereign debt. *Conflict Management and Peace Science*, 39(5):487–519.

Krainin, C. and Schub, R. (2021). Alliance dynamics in the shadow of shifting power. *International Studies Quarterly*, 65(4):905–918.

Kreps, S. (2018). *Taxing wars: The American way of war finance and the decline of democracy*. Oxford University Press.

Kudo, Y. and Masumura, Y. (2022). A constraining effect of arms transfers: Why do arms transfers deter a recipient's aggression? *Midwest Political Science Association. Unpublished Manuscript*.

Kuokštytė, R. and Kuokštis, V. (2024). A spatiotemporal analysis of nato member states' defense spending: how much do allies actually free ride? *Political Science Research and Methods*, pages 1–8.

Kydd, A. (2000). Trust, reassurance, and cooperation. *International Organization*, 54(2):325–357.

Landry, J. (2023). *The politics of arming, postwar power, and war outcomes*. PhD thesis, The University of Texas at Austin.

Leeds, B., Ritter, J., Mitchell, S., and Long, A. (2002). Alliance treaty obligations and provisions, 1815-1944. *International Interactions*, 28(3):237–260.

Leeds, B. A. (2003a). Alliance reliability in times of war: Explaining state decisions to violate treaties. *International Organization*, 57(4):801–827.

Leeds, B. A. (2003b). Do alliances deter aggression? the influence of military alliances on the initiation of militarized interstate disputes. *American Journal of Political Science*, 47(3):427–439.

Levi, M. (1988). *Of rule and revenue*. University of California Press.

Li, X. and Chen, D. (2021). Public opinion, international reputation, and audience costs in an authoritarian regime. *Conflict Management and Peace Science*, 38(5):543–560.

Marshall, M. G., Gurr, T. R., and Jaggers, K. (2017). Polity IV project: Political regime characteristics and transitions, 1800-2016. University of Maryland, Center for International Development and Conflict Management.

Masumura, Y. and Tago, A. (2023). Micro-foundations of the quest for status: Testing self-status perception and the multilateral use of force. *Foreign Policy Analysis*, 19(4):orad027.

McManus, R. W. and Yarhi-Milo, K. (2017). The logic of "offstage" signaling: Domestic politics, regime type, and major power-protégé relations. *International Organization*, 71(4):701–733.

McWard, A. C. and Yoon, H. (2024). Preventing coups and seeking allies: The demand and supply of alliances for coup-proofing regimes. *Journal of Conflict Resolution*, 68(4):730–755.

Miller, S. V. (2022). peacesciencer: An r package for quantitative peace science research. *Conflict Management and Peace Science*.

Montgomery, J. M., Nyhan, B., and Torres, M. (2018). How conditioning on posttreatment variables can ruin your experiment and what to do about it. *American Journal of Political Science*, 62(3):760–775.

Morrow, J. D. (1993). Arms versus allies: Trade-offs in the search for security. *International Organization*, 47(2):207–233.

Morrow, J. D. (1994). Alliances, credibility, and peacetime costs. *The Journal of Conflict Resolution*.

Morrow, J. D. (2000). Alliances: Why write them down? *Annual Review of Political Science*.

Mylonas, H. and Tudor, M. (2021). Nationalism: What we know and what we still need to know. *Annual Review of Political Science*, 24(1):109–132.

Plümper, T. and Neumayer, E. (2015). Free-riding in alliances: Testing an old theory with a new method. *Conflict Management and Peace Science*, 32(3):247–268.

Poast, P. (2019). *Arguing about alliances: The art of agreement in military-pact negotiations*. Cornell University Press.

Powell, R. (2006). War as a commitment problem. *International organization*, 60(1):169–203.

Renshon, J. (2017). *Fighting for Status: Hierarchy and Conflict in World Politics*. Princeton University Press.

Sarkees, M. R. and Wayman, F. (2010). *Resort to War: 1816 - 2007.* Washington DC: CQ Press.

Singer, J. D. (1988). Reconstructing the correlates of war dataset on material capabilities of states, 1816–1985. *International Interactions*, 14(2):115–132.

Smith, A. (1995). Alliance formation and war. *International Studies Quarterly*, 39(4):405–425.

Smith, S. B. (2021). *Reclaiming patriotism in an age of extremes*. Yale University Press.

Sorokin, G. L. (1994). Arms, alliances, and security tradeoffs in enduring rivalries. *International Studies Quarterly*, 38(3):421–446.

Thompson, W. (2001). Identifying rivals and rivalries in world politics. *International Studies Quarterly*, 45(4):557–586.

Trachtenberg, M. (1999). *A constructed peace: the making of the European settlement, 1945-1963*, volume 79. Princeton University Press.

Tudor, M. and Slater, D. (2021). Nationalism, authoritarianism, and democracy: Historical lessons from south and southeast asia. *Perspectives on Politics*, 19(3):706–722.

Walt, S. M. (1985). Alliance formation and the balance of world power. *International security*, 9(4):3–43.

Wang, E. H. and Xu, Y. (2018). Awakening leviathan: The effect of democracy on state capacity. *Research & Politics*, 5(2):2053168018772398.

Weeks, J. L. (2008). Autocratic audience costs: Regime type and signaling resolve. *International Organization*, 62(1):35–64.

Wolford, S. and Masumura, Y. (2024). Preventive war and domestic politics. *Unpublished Manuscript*.

Wu, C. X. and Wolford, S. (2018). Leaders, states, and reputations. *Journal of Conflict Resolution*, 62(10):2087–2117.

Yarhi-Milo, K., Lanoszka, A., and Cooper, Z. (2016). To arm or to ally? the patron's dilemma and the strategic logic of arms transfers and alliances. *International Security*, 41(2):90–139.

Zielinski, R. C. (2016). *How states pay for wars*. Cornell University Press.

Appendix

A Proof for Proposition 1

Proposition 1 is a semi-separating equilibrium where the SQ type arms with l^*, Revisionist always arms, and B attacks with m^*. With this equilibrium behavior in mind, I first analyze the SQ type's incentive to arm.

For the SQ type,

$$u_A(\text{arm}) = mpw_A(0) + (1-m)pw_A(1) - k$$
$$u_A(\neg\text{arm}) = pw_A(0)$$

Since B's strategy makes these two utilities indifferent for the SQ type, we get

$$\begin{aligned} &u_A(\text{arm}) = u_A(\neg\text{arm}) \\ &\Leftrightarrow m = m^* = \frac{p(w_A(1) - w_A(0)) - k}{p(w_A(1) - w_A(0))} \end{aligned} \tag{27}$$

Since $0 < m < 1$, we have

$$p(w_A(1) - w_A(0)) > k \tag{28}$$

For Revisionist,

$$
\begin{aligned}
&u_A(\text{arm}) = mw_A(0) + (1-m)w_A(1) - k \\
&u_A(\neg\text{arm}) = w_A(0) \\
&u_A(\text{arm}) > u_A(\neg\text{arm}) \\
&\Leftrightarrow m < \frac{w_A(1) - w_A(0) - k}{w_A(1) - w_A(0)}
\end{aligned}
$$

Thus, we should check if the following condition is true.

$$
\begin{aligned}
&m^* < \frac{w_A(1) - w_A(0) - k}{w_A(1) - w_A(0)} \\
&\Leftrightarrow p < 1
\end{aligned}
$$

, which is sure to be true.

For B, by using Baye's Rule,

$$
Pr(\text{Revisionist}|\text{arm}) = \frac{\phi}{\phi + (1-\phi)l}
$$

Thus, the payoffs for B are

$$
\begin{aligned}
&u_B(\text{prevent}) = w_B(0) \\
&u_B(\text{pass}) = \frac{\phi}{\phi + (1-\phi)l}(1 - w_A(1)) + \frac{(1-\phi)l}{\phi + (1-\phi)l}(1 - pw_A(1))
\end{aligned}
$$

The SQ type's strategy of arming makes these payoffs are indifferent, thus,

$$\begin{aligned}
&u_B(\text{prevent}) = u_B(\text{pass}) \\
&\Leftrightarrow l = l^* = \frac{\phi(w_B(0) + w_A(1) - 1)}{(1-\phi)(1 - pw_A(1) - w_B(0))}
\end{aligned} \tag{29}$$

Let's check this value is with the range of $0 < l < 1$. Thanks to Line 6 and $0 < \phi < 1$, both the numerator and denominator of l^* are positive, so $0 < l^*$ is true.

$$\begin{aligned}
&l^* < 1 \\
&\phi < \frac{1 - pw_A(1) - w_B(0)}{w_A(1)(1-p)} = \phi^*
\end{aligned} \tag{30}$$

Thus, $l^* < 1$ is satisfied when Line 30. Also, $0 < \phi^* < 1$ is satisfied because of Line 6. Therefore, Proposition 1 exists when Line 6, 28, and 30 are satisfied.

B Proof for Lemma 2

E intervenes iff

$$\begin{aligned}
&u_S(\text{War w/ Intervention}) > u_A(\text{War w/o Intervention}) \\
&\Leftrightarrow 1 - |rw_A(a+\epsilon) - s| > 1 - |rw_A(a) - s| \\
&\Leftrightarrow |rw_A(a) - s| > |rw_A(a+\epsilon) - s|
\end{aligned} \tag{31}$$

When $rw_A(a) > s$, Line 31 is

$$rw_A(a) - s > rw_A(a+\epsilon) - s$$
$$rw_A(a) > rw_A(a+\epsilon)$$

This is not true for sure. Thus, E does not help when $rw_A(a) > s$.

When $rw_A(a) < s < rw_A(a+\epsilon)$, Line 31 is

$$s - rw_A(a) > rw_A(a+\epsilon) - s$$
$$s > \frac{r(w_A(a+\epsilon) + w_A(a))}{2}$$

Thus, E intervenes when $rw_A(a+\epsilon) > s > \frac{r(w_A(a+\epsilon)+w_A(a))}{2}$ and does not intervene when $rw_A(a) < s < \frac{r(w_A(a+\epsilon)+w_A(a))}{2}$

When $rw_A(a+\epsilon) < s$, Line 31 is

$$s - rw_A(a) > s - rw_A(a+\epsilon)$$
$$rw_A(a) < rw_A(a+\epsilon)$$

This is sure to be true. Thus, E intervenes for sure when $rw_A(a+\epsilon) < s$.

In sum, E intervenes iff

$$\frac{r(w_A(a+\epsilon)+w_A(a))}{2} < s \tag{32}$$

Therefore, Lemma 2 is proved.

C Proof for Proposition 2

This is a pooling equilibrium where both types of A do not arm. Tow cases fall down in this equilibrium.

The first case is when B's payoff of preventive war is so large that A can't profitably arm.

Let's consider B's strategy. Line 10 makes sure that E intervenes only when A does not arm or is preemptively attacked by B.

After observing A's arming, B's payoffs of preventing and passing are

$$u_B(\text{prevent}) = w_B(\epsilon)$$
$$u_B(\textbf{pass}) = \begin{cases} 1 - w_A(1) & (A \text{ is Revisionist}) \\ 1 - p w_A(1) & (A \text{ is the SQ type}) \end{cases}$$

B attacks A's arming regardless of its type when

$$
\begin{aligned}
& w_B(\epsilon) > 1 - p w_A(1) \\
& \Leftrightarrow w_B(0) > 1 - p w_A(1) + d
\end{aligned} \tag{33}
$$

, where $d = w_B(0) - w_B(\epsilon)$

On that situation, both types of A does not have any incentive to arm because

$$
\begin{aligned}
u_A(\text{arm}) &= r w_A(\epsilon) - k \\
u_A(\neg\text{arm}) &= r w_A(\epsilon)
\end{aligned}
$$

So ¬arm is a dominant strategy. Thus, A pools on ¬arm when Line 33 and 10 is satisfied.

The second case is when the costs of arming is too expensive for both types of A to profitably arm.

For both Revisionist and the SQ-type,

$$
\begin{aligned}
u_A(arm) &= \begin{cases} r w_A(1) - k & \text{(if } B \text{ passes)} \\ r w_A(\epsilon) - k & (if B \text{ attacks)} \end{cases} \\
u_A(\neg\text{arm}) &= r w_A(\epsilon)
\end{aligned}
$$

Regardless of B's reaction towards arming, both types of A does not arm when

$$
\begin{aligned}
& r(w_A(1) - w_A(\epsilon)) < k \\
& \Leftrightarrow w_A(1) - w_A(\epsilon < k
\end{aligned} \tag{34}
$$

Finally, I check that B does not attack after A's arming.

$$
\begin{aligned}
u_B(\text{attack}|\neg\text{arm}) &= w_B(\epsilon) \\
u_B(\text{pass}|\neg\text{arm}) &= 1 - r w_B(\epsilon) \\
u_B(\text{pass}|\neg\text{arm}) &> u_B(\text{attack}|\neg\text{arm}) \\
\Leftrightarrow 1 &> r w_A(\epsilon) + w_B(\epsilon)
\end{aligned}
$$

This is sure to be true because of the inefficiency of war ($1 > w_A(\epsilon) + w_B(\epsilon)$).

Thus, when Line 34 and 10 satisfied, A pools on not arming regardless of B's strategy. Proposition 2 is proved.

D Proof for Proposition 3

Proposition 3 is a pooling equilibrium where both types of A do not arm because arming invokes B's prevention for sure. B is sure that A that arms is Revisionist because this is the only type of A that could profitably arm.

Because of Lemma 1, E intervenes only when A does not arm or is preemptively attacked when Line 10 is satisfied.

The payoffs for Revisionist are

$$u_A(\text{arm}) = \begin{cases} w_A(1) - k & (\text{if } B \text{ passes}) \\ w_A(\epsilon) - k & (if B \text{ prevents}) \end{cases}$$

$$u_A(\neg\text{arm}) = w_A(\epsilon)$$

Thus, Revisionist has an incentive to arm if B passes when

$$k < w_A(1) - w_A(\epsilon) \tag{35}$$

But it does not arm because B prevents on the path of play ($w_A(\epsilon) - k < w_A(\epsilon)$).

For the SQ type,

$$u_A(\text{arm}) = \begin{cases} pw_A(1) - k & (\text{if } B \text{ passes}) \\ pw_A(\epsilon) - k & (if B \text{ prevents}) \end{cases}$$

$$u_A(\neg\text{arm}) = pw_A(\epsilon)$$

Thus, the SQ type does not have an incentive to arm eve if B passes when

$$k > p(w_A(1) - w_A(\epsilon)) \tag{36}$$

Thus, when Line 35 and 36 are satisfied, or

$$p(w_A(1) - w_A(\epsilon)) < k < w_A(1) - w_A(\epsilon) \tag{37}$$

, Revisionist type has the incentive to arm but the SQ type does not have such an incentive. This leads to B's off-equilibrium-path belief that A that arms is Revisionist for sure. (Thus, this satisfies the D_1 criterion.)

Next, let q be B's belief that $Pr(A \text{ is Revisionist}|\text{arming})$.

For B,

$$u_B(\text{pass}) = q(1 - w_A(1)) + (1-q)(1 - pw_A(1))$$
$$u_B(\text{prevent}) = w_B(\epsilon)$$

B prevents

$$\begin{aligned} & u_B(\text{prevent}) > u_B(\text{pass}) \\ \Leftrightarrow\ & \frac{1 - pw_A(1) - w_B(\epsilon)}{(1-p)w_A(1)} < d \end{aligned}$$

This satisfies $0 < d < 1$ when

$$pw_A(1) + w_B(\epsilon) < 1 < w_A(1) + w_B(\epsilon)$$

In other words, when this condition is satisfied, B wants to prevent Revisionist's arming but not the SQ type's arming. Thus, Line 11 is shown.

Here, Revisionist is the only A that could profitably arm as Line 37 is satisfied, So, $d = 1$. In this context, B prevents A's arming for sure.

Thus, Proposition 3 exists when Line 37, 11, and 10 are satisfied.

E Proof for Proposition 4

Proposition 4 is a separating equilibrium where Revisionist arms but the SQ type does not arm, and B does not prevent A's arming.

For B, it is sure that A that arms is Revisionist. Thus, its payoffs are

$$
\begin{aligned}
&u_B(\text{pass}) = 1 - w_A(1) \\
&u_B(\text{prevent}) = w_B(\epsilon) \\
&u_B(\text{pass}) > u_B(\text{prevent}) \\
&\Leftrightarrow \overline{w_B} = 1 - w_A(1) + d > w_B(0) \qquad (38)
\end{aligned}
$$

Thus, B does not prevent A's arming when Line 38 is satisfied.

Next, the payoffs for Revisionist are

$$
\begin{aligned}
&u_A(\text{arm}) = w_A(1) - k \\
&u_A(\neg\text{arm}) = w_A(\epsilon)
\end{aligned}
$$

Revisionist arms iff

$$
\begin{aligned}
&u_A(\text{arm}) > u_A(\neg\text{arm}) \\
&\Leftrightarrow \overline{k} = w_A(1) - w_A(\epsilon) > k \qquad (39)
\end{aligned}
$$

For the SQ type,

$$u_A(\text{arm}) = pw_A(1) - k$$
$$u_A(\neg\text{arm}) = pw_A(\epsilon)$$

The SQ type does not arm iff

$$u_A(\text{arm}) < u_A(\neg\text{arm})$$
$$\Leftrightarrow \underline{k} = p(w_A(1) - w_A(\epsilon)) < k \tag{40}$$

Thus, when Line 38, 39, and 40 are satisfied, Revisionist arms, Pacifist does not arm, and B does not prevent. Thus, Proposition 4 is shown.

F Proof for Proposition 5

Proposition 5 is a pooling equilibrium where both types of A arms and B does not prevent.

For both types of A, the payoffs are

$$u_A(\text{arm}) = rw_A(1) - k$$
$$u_A(\neg\text{arm}) = rw_A(\epsilon)$$

A arms iff

$$\begin{aligned}
&u_A(\text{arm}) > u_A(\neg\text{arm}) \\
&\Leftrightarrow r(w_A(1) - w_A(\epsilon)) > k \\
&\Leftrightarrow \underline{k} = p(w_A(1) - w_A(\epsilon)) > k
\end{aligned} \tag{41}$$

For B,

$$\begin{aligned}
&u_B(\text{pass}) = \phi(1 - w_A(1)) + (1 - \phi)(1 - pw_A(1)) \\
&u_B(\text{prevent}) = w_B(\epsilon)
\end{aligned}$$

B passes

$$\begin{aligned}
&u_B(\text{prevent}) < u_B(\text{pass}) \\
&\Leftrightarrow \frac{1 - pw_A(1) - w_B(\epsilon)}{(1-p)w_A(1)} > \phi
\end{aligned}$$

Since $\phi < 1$, the condition above is sure to be satisfied when

$$\begin{aligned}
\frac{1 - pw_A(1) - w_B(\epsilon)}{(1-p)w_A(1)} &> 1 \\
1 &< w_A(1) + w_B(\epsilon) \\
\underline{w_B} = 1w_A(1) + d &> w_B(0)
\end{aligned} \tag{42}$$

Therefore, when Line 41 and 42 are satisfied, Proposition 5 exists.

G Proof for Proposition 6

Proposition 6 is a semi-separating equilibrium where the SQ type arms with $l^{\dagger}$, Revisionist always arms, and B attacks with $m^{\dagger}$. With this equilibrium behavior in mind, I first analyze the SQ type's incentive to arm.

For the SQ type,

$$u_A(\text{arm}) = mpw_A(\epsilon) + (1-m)pw_A(1) - k$$
$$u_A(\neg\text{arm}) = pw_A(\epsilon)$$

Since B's strategy makes these two utilities indifferent for the SQ type, we get

$$\begin{aligned} & u_A(\text{arm}) = u_A(\neg\text{arm}) \\ \Leftrightarrow\ & m = m^{\dagger} = \frac{p(w_A(1) - w_A(\epsilon)) - k}{p(w_A(1) - w_A(\epsilon))} \end{aligned} \tag{43}$$

Since $0 < m < 1$, we have

$$p(w_A(1) - w_A(\epsilon)) > k \tag{44}$$

For Revisionist,

$$
\begin{aligned}
&u_A(\text{arm}) = mw_A(\epsilon) + (1-m)w_A(1) - k \\
&u_A(\neg\text{arm}) = w_A(\epsilon) \\
&u_A(\text{arm}) > u_A(\neg\text{arm}) \\
&\Leftrightarrow m < \frac{w_A(1) - w_A(\epsilon) - k}{w_A(1) - w_A(\epsilon)}
\end{aligned}
$$

Thus, we should check if the following condition is true.

$$
\begin{aligned}
&m^{\dagger} < \frac{w_A(1) - w_A(\epsilon) - k}{w_A(1) - w_A(\epsilon)} \\
&\Leftrightarrow p < 1
\end{aligned}
$$

, which is sure to be true.

For B, by using Baye's Rule,

$$
Pr(\text{Revisionist}|\text{arm}) = \frac{\phi}{\phi + (1-\phi)l}
$$

Thus, the payoffs for B are

$$
\begin{aligned}
&u_B(\text{prevent}) = w_B(\epsilon) \\
&u_B(\text{pass}) = \frac{\phi}{\phi + (1-\phi)l}(1 - w_A(1)) + \frac{(1-\phi)l}{\phi + (1-\phi)l}(1 - pw_A(1))
\end{aligned}
$$

The SQ type's strategy of arming makes these payoffs are indifferent, thus,

$$
\begin{aligned}
& u_B(\text{prevent}) = u_B(\text{pass}) \\
& \Leftrightarrow l = l^{\dagger} = \frac{\phi(w_B(\epsilon) + w_A(1) - 1)}{(1-\phi)(1 - p w_A(1) - w_B(\epsilon))}
\end{aligned}
\tag{45}
$$

Let's check this value is within the range of $0 < l < 1$. Thanks to Line 11 and $0 < \phi < 1$, both the numerator and denominator of l^* are positive, so $0 < l^{\dagger}$ is true.

$$
\begin{aligned}
& l^{\dagger} < 1 \\
& \phi < \frac{1 - p w_A(1) - w_B(\epsilon)}{w_A(1)(1-p)} = \phi^{\dagger}
\end{aligned}
\tag{46}
$$

Thus, $l^{\dagger} < 1$ is satisfied when Line 46. Also, $0 < \phi^{\dagger} < 1$ is satisfied because of Line 11. Therefore, Proposition 6 exists when Line 11, 44, and 46 are satisfied.

H Proof for Proposition 7

Proposition 7 is a pooling equilibrium where E offers an alliance, A accepts it, and both types of A do not arm because B correctly think that A that arms is Revisionist.

First of all, if E does not offer an alliance, the subgame is exactly the same as the model without an alliance analyzed at the beginning of the Analysis section in the main text. Thus, as Proposition 1 shows, preventive war can happen when

$$w_A(1) + w_B(0) > 1 > pw_A(1) + w_B(0)$$

$$p(w_A(1) - w_A(0)) > k$$

$$\phi < \phi^*$$

In such an occasion, E's payoff is

$$\begin{aligned} u_E(\neg\text{offer an alliance}) =& \phi m^*(1 - |w_A(0) - s|) \\ &+ \phi(1 - m^*)(1 - |w_A(1) - s|) \\ &+ (1 - \phi)l^* m^*(1 - |pw_A(0) - s|) \\ &+ (1 - \phi)l^*(1 - m^*)(1 - |pw_A(1) - s|) \\ &+ (1 - \phi)(1 - l^*)(1 - |pw_A(0) - s|) \end{aligned} \tag{47}$$

Next, I consider A and B's strategy after E offers an alliance.

B's payoffs of preventing A's arming is.

$$u_B(\text{prevent}) = \begin{cases} w_B(\epsilon) & (\text{if } A \text{ accepts an alliance}) \\ w_B(0) & (\text{if } A \text{ does not accept an alliance}) \end{cases}$$

B's payoffs of passing do not depend on the existence of the alliance because E does not help A after arming (See Line 10). Thus,

$$u_B(\text{pass}) = \begin{cases} 1 - w_A(1) & \text{(if Revisionist)} \\ 1 - pw_A(1) & \text{(if the SQ type)} \end{cases}$$

Thus, B has an incentive to prevent Revisionist but not the SQ type when

$$w_A(1) + w_B(0) > 1 > pw_A(1) + w_B(0)$$

if A has an alliance and

$$w_A(1) + w_B(\epsilon) > 1 > pw_A(1) + w_B(\epsilon)$$

if A does not an alliance.

By combining these two conditions, B wants to prevent Revisionist's arming but not the SQ type's arming regardless of the alliance when

$$\begin{aligned} & w_A(1) + w_B(\epsilon) > 1 > pw_A(1) + w_B(0) \\ \Leftrightarrow\ & 1 - pw_A(1) > w_B(0) > \underline{w_B} \end{aligned} \tag{48}$$

Next, A's payoffs are analyzed. Since A can combine the alliance choice and the arming choice, it has four choices: arming with alliance, arming without alliance, no arming with alliance, and no arming without alliance.

A's payoffs are

$$u_A(\text{arming with alliance}) = \begin{cases} rw_A(\epsilon) - k & \text{(if } B \text{ prevents)} \\ rw_A(1) - k & \text{(if } B \text{ passes)} \end{cases}$$

$$u_A(\text{arming without alliance}) = \begin{cases} rw_A(o) - k & \text{(if } B \text{ prevents)} \\ rw_A(1) - k & \text{(if } B \text{ passes)} \end{cases}$$

$$u_A(\text{not arming with alliance}) = rw_A(\epsilon)$$

$$u_A(\text{not arming without alliance}) = rw_A(0)$$

Note that B wants to prevent Revisionist's arming but not the SQ type's arming regardless of the alliance under Line 48.

If, as happens in equilibrium, B prevents A's arming, "not arming with alliance" is obviously a dominant strategy since $rw_A(\epsilon) > rw_A(0)$ and $k > 0$.

If, hypothetically, B does not prevent A's arming, Revisionist has an incentive to arm when

$$\begin{aligned} & w_A(1) - k > w_A(\epsilon) \\ & \Leftrightarrow w_A(1) - w_A(\epsilon) > k \end{aligned} \tag{49}$$

But the SQ type does not have an incentive to ar when

$$\begin{aligned} & pw_A(1) - k < pw_A(\epsilon) \\ & \Leftrightarrow p(w_A(1) - w_A(\epsilon)) > k \end{aligned} \tag{50}$$

Thus, when Line 49 and 50 are satisfied, or

$$p(w_A(1) - w_A(\epsilon)) = \underline{k} < k < \overline{k} = w_A(1) - w_A(\epsilon) \tag{51}$$

, B rationally conjectures that A that arms must be Revisionist and prevents A's arming on the off-equilibrium path.

Given R's strategy, "not arming with alliance" is a dominant strategy for both types of A, as discussed above. Thus, after E's alliance offer, A pools on "not arming with alliance" and B does not attack A.

On such an occasion, E's payoff of offering an alliance is

$$\begin{aligned} u_E(\text{offer an alliance}) =& \phi(1 - |w_A(\epsilon) - s|) \\ &+ (1-\phi)(1 - |pw_A(\epsilon) - s|) \end{aligned} \tag{52}$$

Remember that, as shown Line 47, E's payoff of not offering an alliance is

$$\begin{aligned} u_E(\neg\text{offer an alliance}) =& \phi m^*(1 - |w_A(0) - s|) \\ &+ \phi(1-m^*)(1 - |w_A(1) - s|) \\ &+ (1-\phi)l^* m^*(1 - |pw_A(0) - s|) \\ &+ (1-\phi)l^*(1-m^*)(1 - |pw_A(1) - s|) \\ &+ (1-\phi)(1-l^*)(1 - |pw_A(0) - s|) \end{aligned}$$

Because of Line 10, $w_A(\epsilon)$ is closer to s than $w_A(0)$ and $pw_A(\epsilon)$ is closer to s than $pw_A(0)$, meaning $|w_A(1) - s| > |w_A(\epsilon) - s|$, and $|pw_A(1) - s| > |pw_A(\epsilon) - s|$.

Also, let's assume that

$$
\begin{aligned}
& |pw_A(1) - s| > |pw_A(\epsilon) - s| \\
& \Leftrightarrow \frac{p(w_A(1) + w_\epsilon)}{2} > s
\end{aligned} \tag{53}
$$

, meaning that the combination of E's intervention and A' no arming brings a closer outcome to E's ideal point than the combination of A's arming without E's intervention. This is consistent with the idea that E is a status-quo power.

When Line 53 is satisfied, $u_E(\text{offer an alliance})$ is larger than $u_E(\neg\text{offer an alliance})$ for sure because $rw_A(\epsilon)$ is closer to s than $rw_A(0)$ and $rw_A(1)$. For this reason, E chooses to offer an alliance.

Therefore, when Line 48, 51, 53, and the conditions for Proposition 1 are satisfied, Proposition 7 exists. The conditions are summarized as follows.

$$
\begin{aligned}
& \frac{w_A(0) + w_\epsilon}{2} < s < \frac{p(w_A(1) + w_\epsilon)}{2} \\
& \frac{w_A(0) + w_A(\epsilon)}{w_A(1) + w_A(\epsilon)} < p \\
& \underline{k} < k < \min\{\overline{k}, p(w_A(1) - w_A(0))\} \\
& \underline{w}_B < w_B(0) < 1 - pw_A(1) \\
& \phi < \phi^*
\end{aligned}
$$

I Newly Estimated Power Shifts

Figure 12 shows the distribution of the expected power shifts.

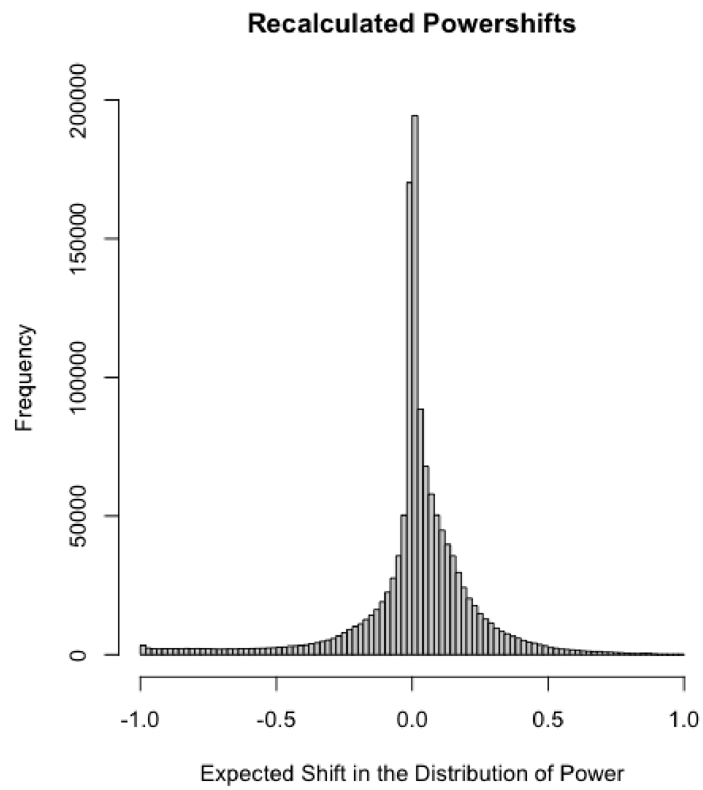


Figure 12: Histogram of Newly Estimated Power Shifts

J Alliance and Preventive War

J.1 Results of Model 4 and 5

Figure 13 and 14 are the marginal coefficients of expected power shifts in Model 4 and Model 5, respectively. Both of them show that the effect of power shifts on war decreases when states have an alliance. As the interaction terms in the main text are negative, the deterrence effect of alliances become stronger when the expected power shifts are large.

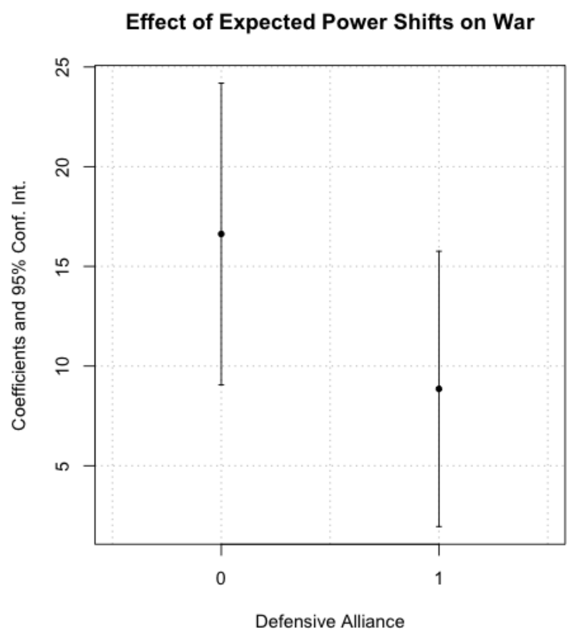


Figure 13: Marginal Coefficient of Expected Power Shifts in Model 4

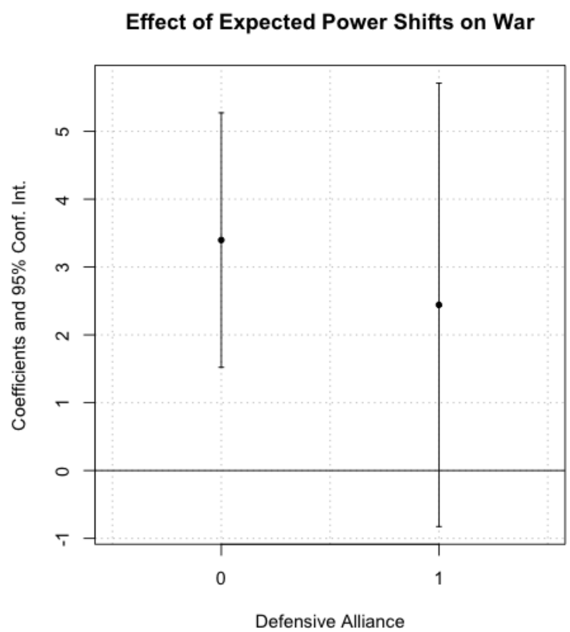


Figure 14: Marginal Coefficient of Expected Power Shifts in Model 5

J.2 Robustness to Different Similarlity Measure

For robustness checks, I use the S score based on the UNGA voting data, which are offered by Häge (2011) and contained in Miller (2022).

Table 3 shows the regression table of the result. Also, Figure 15 shows the marginal coefficient of expected power shifts. The results show the same tendency as the main model.

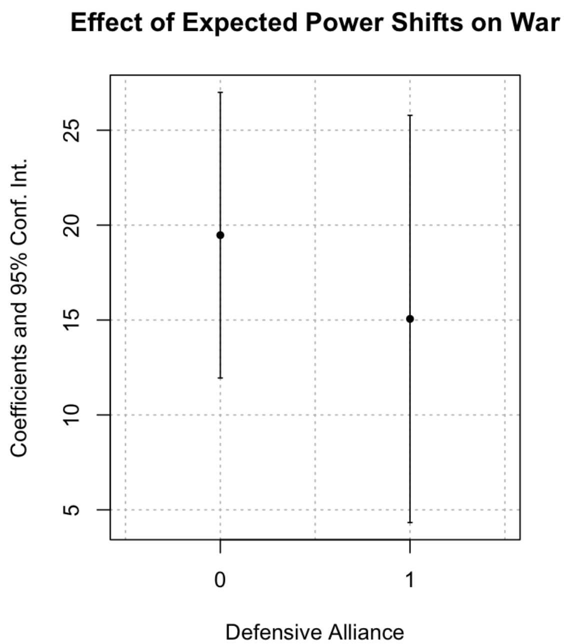


Figure 15: Marginal Coefficients of Power Shifts using Different Similarlity Score

Table 3: Alliance and Preventive War (Different Similarlity Score)

Dependent Variable:	waronset_target
Model:	(1)
Variables	
Alliance	-1.851
	(1.713)
Expected Power Shifts	18.26***
	(4.010)
Alliance × Expected Power Shifts	-0.3488
	(5.483)
Mutual Democracy	-2.280***
	(0.6954)
S-Score based on UN Voting	5.867**
	(2.826)
Contingency	8.263*
	(4.353)
Distance	-399.1***
	(90.26)
Peace Year	0.5825**
	(0.2608)
Peace Year^2	-0.0283**
	(0.0125)
Peace Year^3	0.0004**
	(0.0002)
Fixed-effects	
Directed Dyad	Yes
Fit statistics	
Observations	1,452
Squared Correlation	0.42606
Pseudo R^2	0.50919
BIC	464.06

Clustered (Directed Dyad) standard-errors in parentheses
*Signif. Codes: ***: 0.01, **: 0.05, *: 0.1*

J.3 Expected Power Shifts at Year $t+3$

Table 4 shows the regression table of the result. Also, Figure 16 shows the marginal coefficient of expected power shifts.

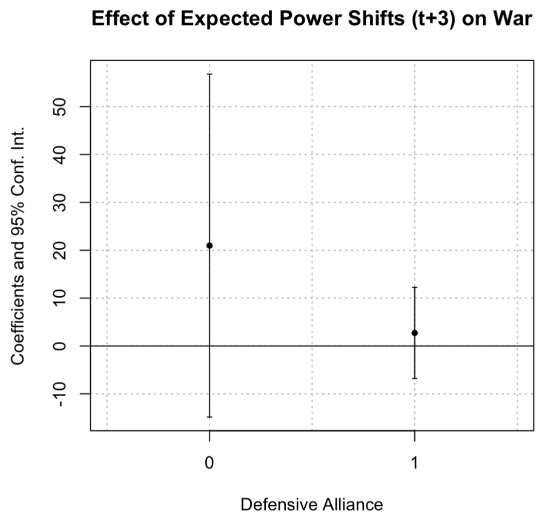


Figure 16: Marginal Coefficients of Power Shifts using Expected Power Shifts at Year $t+3$

Table 4: Alliance and Preventive War (Year $t+3$)

Dependent Variable:	waronset_target
Model:	(1)
Variables	
Alliance	2.994
	(4.181)
Expected Power Shift ($t+3$)	21.17
	(16.40)
Alliance × Expected Power Shift ($t+3$)	-18.23
	(14.39)
Mutual Democracy	-1.869
	(1.428)
Foreign Policy Smilarlity	-1.453
	(3.756)
Contingency	19.52***
	(4.139)
Distance	-67.02
	(415.6)
Peace Year	0.0969
	(0.1323)
Peace Year^2	0.0037
	(0.0043)
Peace Year^3	-3.41×10^{-5}
	(3.4×10^{-5})
Fixed-effects	
Directed Dyad	Yes
Year	Yes
Fit statistics	
Observations	435
Squared Correlation	0.42118
Pseudo R^2	0.41011
BIC	558.54

Clustered (Directed Dyad) standard-errors in parentheses
*Signif. Codes: ***: 0.01, **: 0.05, *: 0.1*

J.4 Expected Power Shifts at Year $t+5$

Table 5 shows the regression table of the result. Also, Figure 17 shows the marginal coefficient of expected power shifts.

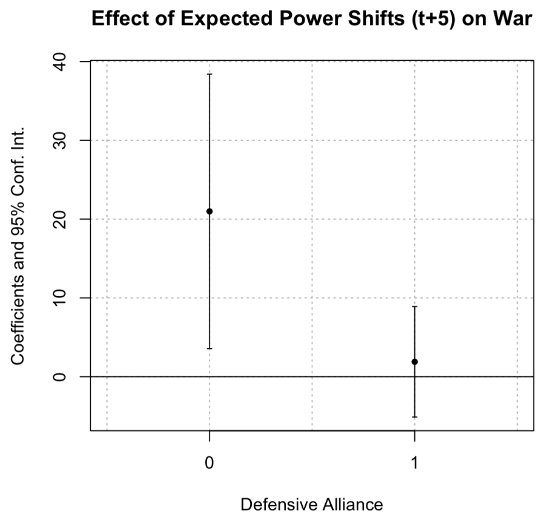


Figure 17: Marginal Coefficients of Power Shifts using Expected Power Shifts at Year $t+5$

Table 5: Alliance and Preventive War (Year $t+3$)

Dependent Variable:	waronset_target
Model:	(1)
Variables	
Alliance	4.718
	(6.580)
Expected Power Shift ($t+5$)	24.58**
	(12.16)
Alliance × Expected Power Shift ($t+5$)	-23.09**
	(11.39)
Mutual Democracy	-2.446*
	(1.302)
Foreign Policy Similarlity	0.4155
	(7.361)
Contingency	4.978*
	(2.565)
Distance	-131.6
	(261.1)
Peace Year	0.2517
	(0.2006)
Peace Year^2	0.0024
	(0.0039)
Peace Year^3	-2.1×10^{-5}
	(3.49×10^{-5})
Fixed-effects	
Directed Dyad	Yes
Year	Yes
Fit statistics	
Observations	387
Squared Correlation	0.47891
Pseudo R^2	0.45983
BIC	498.81

Clustered (Directed Dyad) standard-errors in parentheses
*Signif. Codes: ***: 0.01, **: 0.05, *: 0.1*

K Alliance and Military Expenditure

Table 6 shows a regression table of the robustness check using State Capacity Measure from Hanson and Sigman (2021).

Table 6: Military Expenditure and State Capacity

Dependent Variable:	ln(Mil. Exp.)	
Model:	(1)	(2)
Variables		
Alliance	0.0626**	
	(0.0270)	
Capacity	0.0585**	
	(0.0280)	
International War	0.1002***	0.1008***
	(0.0355)	(0.0355)
Civil War	0.1247***	0.1258***
	(0.0356)	(0.0352)
ln(Mil. Exp.)$_{t-1}$	0.7652***	0.7664***
	(0.0288)	(0.0291)
Nuclear	-0.0148	-0.0220
	(0.0535)	(0.0509)
Rivalry	0.0278	0.0271
	(0.0206)	(0.0210)
Democracy	-0.0070***	-0.0073***
	(0.0020)	(0.0021)
ln(GDP)	0.1875***	0.1755***
	(0.0371)	(0.0353)
ln(Population)	-0.0165	0.0009
	(0.0459)	(0.0466)
GDP Growth	1.068	1.415
	(3.137)	(3.119)
Population Growth	-5.502	-4.024
	(9.207)	(9.292)
Capacity × defense = 0		0.0462
		(0.0309)
Capacity × defense = 1		0.0646**
		(0.0295)
Fixed-effects		
ccode	Yes	Yes
year	Yes	Yes
Fit statistics		
Observations	6,184	6,184
R^2	0.98261	0.98257
Within R^2	0.71158	0.71095

Clustered (ccode) standard-errors in parentheses
*Signif. Codes: ***: 0.01, **: 0.05, *: 0.1*