Commitment Problems in Alliance Formation *

Brett V. Benson[†] Bradley C. Smith[‡]

February 18, 2021

Abstract

If alliances cause significant shifts in the distribution of power, why does their creation or expansion rarely provoke preventive wars? We develop a theory to explain this puzzle, advancing three arguments about the connection between alliances, commitment problems, and war. First, we show that prospective allies can avoid provoking a common enemy by offering concessions to offset losses from an anticipated power shift from an alliance. Second, limits to an alliance's power or implementation speed are necessary to make such bargains possible. Allies manipulate these factors to set the terms of cooperation to avoid provoking a shared enemy. Finally, when such bargains are not possible, incentives for preventive war exist but the outbreak of such wars are rare. We show that while preventive war cannot be ruled out altogether, the conditions that make it most attractive also make it unlikely to be carried out.

*We are grateful to Andrew Coe, Casey Crisman-Cox, Hein Goemans, Brenton Kenkel, William Spaniel, Burcu Savun, audience members at the EPSA annual meeting 2019, and seminar participants at the Vanderbilt University Conflict Workshop and the University of Pittsburgh International Relations Workshop for helpful feedback.

†Associate Professor, Department of Political Science, Vanderbilt University, (brett.benson@vanderbilt.edu).

‡Assistant Professor, Department of Political Science, Vanderbilt University (bradley.carl.smith@gmail.com).

1 Introduction

Shifts in military power have a well-established theoretical connection to the outbreak of war.[1] Anticipating unfavorable shifts in the balance of power, states may launch preventive war to block them. Few power shifts in international relations are as sizable as the formation or expansion of a military alliance. For example, NATO provides members with the aggregate defensive military capability of 30 countries. States targeted by alliances often anticipate these shifts during the process of alliance formation and implementation. In spite of this, the anticipation of a new or expanded alliance rarely provokes preventive war.[2]

For example, in 1949 the US became aware that the USSR was likely to extend a treaty to the People's Republic of China. In spite of the substantial and well-anticipated power shift that would result from such a treaty, the Truman administration chose against military action to block the impending treaty.[3] Consequently, the Sino-Soviet Treaty of Friendship, Alliance and Mutual Assistance was signed peacefully on February 14, 1950. The 1999 expansion of NATO to include Hungary, Poland, and the Czech Republic provides another example. This expansion occurred even in the face of Russian opposition that can be traced back to 1990 negotiations on NATO's role in German unification.[4] Despite the near decade-long window of opportunity to block or forestall the impending expansion, Russia chose not to act aggressively, and NATO's eastward march proceeded peacefully.

This presents a puzzle: if alliances cause significant shifts in the distribution of power, why does their creation so rarely provoke preventive wars? We develop a theory to explain this puzzle, advancing three arguments about the connection between alliances, commitment problems, and war. First, we show that prospective allies can avoid provoking a common enemy by offering concessions to offset losses from an anticipated power shift from an alliance.

[1] See, e.g. Fearon (1995), Powell (2006) and Powell (2012).

[2] In a qualitative survey of wars between 1817-2007, the authors found little evidence of wars caused by preventive strikes in anticipation of an alliance.

[3] For more on US policy, see Chang (1990), especially chapters 1-2.

[4] For more on the process of unification and the historical debate over NATO expansion, see Sarotte (2010*a*) and Shifrinson (2016).

Second, limits to an alliance's power or implementation speed are necessary to make such bargains possible. Allies manipulate these factors to set the terms of cooperation to avoid provoking a shared enemy. Finally, when such bargains are not possible, incentives for preventive war exist but the outbreak of such wars are rare. We show that while preventive war cannot be ruled out altogether, the conditions that make it most attractive also make it unlikely to be carried out.

To develop this theory, we analyze a three-player, dynamic model of alliance formation, bargaining, and preventive war. Importantly, we treat security cooperation as a dynamic process. The benefits of a military alliance in our model do not arrive immediately. Rather, two prospective allies must exert effort to fully implement the terms of an alliance before realizing its military benefits. This creates a window of opportunity for the target to use military force to forestall the power shift that results once the alliance comes into force. Our analysis focuses on detailing the conditions under which this incentive for a preventive strike can be bargained away and when it persists.

Three main findings emerge from our equilibrium analysis. First, bargaining concessions play an important role in alleviating incentives for preventive strikes. In equilibrium, prospective allies may offer bargaining concessions prior to the implementation of an alliance to induce a targeted state to allow the alliance. These concessions serve to compensate the targeted state for the future power shift that occurs once an alliance is implemented. This presents a novel explanation for why alliances rarely provoke preventive war: allies may "buy off" shared enemies prior to the implementation of otherwise provocative alliances. We show that this logic was present in the 1990 negotiations over German unification, which resulted in peaceful inclusion of former East German territory in NATO. By offering concessions including substantial loans, negotiators won Soviet acquiescence to an otherwise provocative expansion of NATO.

Our second main finding details the conditions under which these bargains are possible. The size of the power shift resulting from an impending alliance and the expected speed of its implementation are key. If the anticipated power shift is too large and the speed of implementation too fast, then allies may not be able to offer concessions sufficient to secure the target's acquiescence. Consequently, preventive war results. This connects the logic of Powell (2006) to alliance politics by highlighting the size and speed of power shifts from alliance formation as key factors determining when peace occurs. In an equilibrium welfare analysis, we show that allies are often better off limiting the scope of an alliance's military benefits or the speed of its implementation so as to avoid provoking an enemy. The costs of such limits are justified by the benefit of avoiding war. Consistent with this logic, we provide evidence that limiting the number of NATO troops and the withdrawal timeline of Soviet troops were key in creating room for the bargain that won Soviet approval for unified Germany's 1990 inclusion in NATO.

Our third finding is that the conditions favorable for preventive war surprisingly also undermine the likelihood that it occurs. War provoked by alliance formation is rare even when we might most expect it. Preventive war is attractive when a large power shift from an alliance is expected to arrive rapidly and an attack would succeed at blocking its implementation. However, the speed of implementation is also the key factor that renders preventive war rare in equilibrium. Provocative alliances are often implemented before a preventive war would be effective. If an alliance is expected to be implemented rapidly, then there is a good chance that the window of opportunity for preventive war will close before an attack can be carried out in time. Hence, while preventive war cannot be ruled out altogether, the conditions that make it most attractive also make it unlikely to be carried out. Our equilibrium analysis demonstrates this formally.

Holistically, these findings highlight temporal dynamics as a key and novel factor connecting alliances and war. As alliances alter the future distribution of power, expectation of their future arrival alters behavior in the present. Bargaining and alteration of the terms of cooperation prior to an alliance's implementation may pave the way for peace. When such bargains are insufficient, targeted states race against time to complete preventive action before the alliance comes into force. These findings point to the importance of taking time seriously in the study of alliance politics and war.

As such, our main contribution is unifying the study of alliance formation with the logic of dynamic commitment problems. The logic of commitment problems in general is well-established in the theoretical literature (Fearon, 1995; Powell, 2006), and alliances are known to create power shifts as a result of capability aggregation and burden sharing (Olson and Zeckhauser, 1966). It is, therefore, reasonable to expect that military alliances might lead to conflict by inducing commitment problems. Our results formally establish the conditions when this problem may arise in a rationalist framework. As such, a major contribution of the present study is to expand the scope of interactions to which the logic of commitment problems applies to include alliance dynamics.

By establishing this connection, we also make important contributions to the study of security cooperation and alliance politics. Much of the existing work on alliances focuses on the credibility problem inherent in them: will an ally intervene on behalf of its partner in the event of war? Under incomplete information about an ally's willingness to intervene, formal alliances enhance a country's ability to communicate its willingness by providing a vehicle for sinking costs in peacetime (Morrow, 1991, 1994).

In contrast to studies focused on this credibility problem, we focus on the dynamic effect of security cooperation in a complete-information environment. Thus, our findings are related to existing studies of multilateral conflict in a repeated-game setting. In these studies, alliances are implemented immediately (Benson, Meirowitz and Ramsay, 2014) or are not modeled (Krainin and Wiseman, 2016). In contrast, our interest is in the implications of alliances when they are conceived and brought into force over time. Thus, we view our results as complementary to these existing studies.

The mechanism that we identify that connects alliance politics to conflict is the well-studied shifting-power commitment problem (Fearon, 1995; Powell, 1999, 2004, 2006). As Powell (2004) illustrates, this mechanism is present across a variety of substantive contexts, including political transitions (Acemoglu and Robinson, 2001), civil con ict (Fearon, 2004), and bureaucratic design (De Figueiredo, 2002). However, a systematic theoretical treatment of alliance formation as a cause of shifting-power commitment problems does not exist. Accordingly, the present study represents a \research bet" (Powell, 2012, 2017) that a focus on commitment problems will yield valuable and novel insights into the relationship between alliances and war.

Our theoretical framework is also related to models of arming, as developing arms and entering an alliance both represent costly attempts to alter the distribution of bargaining power. As such, our theoretical apparatus is related to those previously used to study arming. Meirowitz and Sartori (2008) show that partially observable arming creates incentives for strategic ambiguity that may lead to war. In Bas and Coe (2016)'s study of nuclear proliferation, they model the delay in nuclear development using a similar framework to the one studied here, in which a decision to implement an alliance may not immediately succeed. Coe (2018) studies the interaction between a powerful state that can arm with certainty and a less capable foe whose arming e orts may fail. As in our work, each of these studies explains con ict as the result of attempts to alter the balance of power.

Our model departs from these studies in two important ways. First, while implementation of an alliance may be delayed, the decision to attempt implementation is always observed by the target of the alliance. Second, the decision to implement an alliance is inherently multilateral. This contrasts with existing work on arming in which the decision to arm is made unilaterally. Modeling this multilateral process is crucial for capturing the strategic forces at play. In our model, formation of an alliance in equilibrium is subject to a participation constraint that requires all potential members to prefer to implement the alliance. The joint nature of the decision to implement an alliance also allows us to consider the possibility of divergent preferences among members over the speed of implementation and, consequently, the risk of war.

2 Model

Consider a world with three states. State 1 and State 2 repeatedly bargain over a disputed policy. State 3 has interests over the policy outcome that are aligned with those of State 1.[5] In every round of bargaining, State 1 and State 3 may attempt to implement a military alliance at some cost. Successful implementation of the alliance increases the payoff of war for both members and decreases the payoff of war for the shared enemy in all subsequent periods.

This interaction occurs over an infinite number of periods indexed by $t = 1, 2, \ldots$[6] In each period t, the interaction is characterized by a commonly observed state variable $s^t \in \{N, A\}$. The value of the state variable indicates that an alliance has been successfully implemented (A) in a previous period, or not (N). Both the sequence of actions in a period and the players' per-period payoffs are a function of s^t. As we are interested in how the process of alliance implementation influences bargaining and conflict, we assume that an alliance has not been implemented at the outset of the interaction, so $s^1 = N$.

The timing of the stage game in period t is as follows. In a period in which $s^t = N$, State 3 makes the first move of the stage game, choosing to extend an alliance commitment to State 1 or not. If State 3 extends an alliance, then State 1 chooses to join the alliance or not. If State 3 extends an alliance and State 1 joins, then the alliance is implemented with probability r and implementation fails with probability $1 - r$. If State 3 chooses not to extend or State 1 chooses not to join, the alliance is not implemented. In all cases, once States 1 and 3 have made their decisions, bargaining occurs, with State 1 proposing a settlement of the issue for the period, $x^t \in [0, 1]$. After observing this proposal, State 2 chooses to accept it or reject it. If State 2 rejects, players receive their war payoffs in all future periods. If State 2

[5] Following the alliance literature, we will sometimes refer to State 1 as the "protégé," State 2 as the "target," and State 3 as the "defender."

[6] We note that our results would also obtain in a two-period model in which the allies are able to make a single attempt at implementing an alliance between the first and second periods. However, as there is no natural empirical "end period" in which allies are unable to continue the process of alliance formation, we opt for the more realistic, infinite-horizon setup here.

accepts, play proceeds to the next period, with s^{t+1} determined by whether an alliance was implemented in the current period. In a period in which $s^t = A$, play proceeds with State 1 proposing $x^t \in [0,1]$ or attacking, and State 2 choosing to accept or reject. After an alliance has been implemented in period t, $s^{t'} = A$ for all periods $t' \geq t$.

States discount future period payoffs by $\delta \in (0,1)$, and we normalize states' dynamic payoffs by $1-\delta$. If State 2 accepts an offer of x^t in period t, then States 1 and 3 receive period payoffs of x^t and State 2 receives a period payoff of $-x^t$.

If a war occurs, then each player receives their (discounted) war payoff in all future periods. If $s^t = N$ and an alliance failed to be implemented in the current period, then State i's war payoff is w_i. If an alliance was implemented in the current period or in any previous period, then State i's war payoff is w_i'. To model the inefficiency of war, we assume that $w_2 + \max\{w_1; w_3\} \leq 1$ and $w_2' + \max\{w_1'; w_3'\} \leq 1$. To capture the effect of alliances on the players' war payoffs, we assume that $w_2' \geq w_2$, $w_3' \leq w_3$, and $w_1' \leq w_1$.[7] Finally, we assume that the process of implementing an alliance is costly. In any period in which the alliance has not yet been implemented, State 3 pays a cost of $a > 0$ if it chooses to extend an alliance. Similarly, in a period in which the alliance has not yet been implemented, State 2 pays a cost $a > 0$ if it accepts State 3's invitation. Throughout, we assume that the cost of alliance implementation is lower than the inefficiency generated by war, $a < 1 - w_2 - w_1$.[8]

2.1 Key Features of the Model

Before moving on to the analysis, we pause briefly to highlight some key features of the model. Consistent with existing work, in our model a military alliance alters the balance

[7] Note that we do not model the decision of State 3 to intervene explicitly. As this is a game of complete information, this is immaterial. Our results carry through unaltered if we extend the model to make this intervention a choice, assuming that State 3 is more likely to intervene after an alliance is implemented. As this only serves to make the presentation more cumbersome, we proceed with the simpler, reduced-form specification here.

[8] This assumption is not necessary to obtain the substantive results, but simplifies the proofs significantly. Further, the results are robust to relaxing the assumption that the players have the same cost of implementation. The results also persist if the cost of implementation persists even after the alliance is successfully implemented. In each case, the results are not altered, but the presentation of the results becomes more cumbersome. Accordingly, we opt for the simpler specification here.

of power, improving the war payoffs of the members (Smith, 1995; Morrow, 2000). Also in line with existing models, alliance members' payoffs for peaceful settlements are related, capturing the idea that the allies value the same policy goals (Fang, Johnson and Leeds, 2014; Benson, 2012; Benson, Meirowitz and Ramsay, 2014; Wolford, 2014), and alliance formation is subject to a participation constraint that requires all potential members to prefer to implement the alliance (Benson, 2012; Benson, Meirowitz and Ramsay, 2014). Finally, we follow the standard approach in models of multilateral conflict in which a third party may act to shape the outcome of war through intervention or alliance formation, but does not act to initiate conflict itself (Smith, 1995; Morrow, 2000; Fang, Johnson and Leeds, 2014; Benson, 2012; Benson, Meirowitz and Ramsay, 2014; Wolford, 2014). In sum, our theoretical setting incorporates standard features of existing models of alliance politics and multilateral conflict.

We build on this work by incorporating an important feature of security cooperation that has not been previously modeled; the military benefits of an alliance may not arrive immediately. Rather, exogenous factors often cause time to pass between the decision to form an alliance and the point at which the military benefits of the alliance are realized. Throughout, we refer to the time at which the military benefits of an alliance are realized as the point of *implementation*. This is a key distinguishing feature of our model. The possibility of such delay provides a window of opportunity for an adversary to act in an attempt to block the power shift resulting from an alliance's implementation. Thus, the possibility of implementation delay is an important feature connecting alliance politics to the logic of dynamic commitment problems. What might cause such delay in implementation?

Sources of exogenous delay in implementation are empirically prevalent. The difficulty of inter-operability is one important source. Disparate technologies and command structures must be brought into line before military cooperation can be effective in the event of war. Such concerns provided one roadblock to rapid NATO expansion in the 1990s (Goldgeier, 2010, p. 76). Exogenous domestic or international political problems may take priority and temporarily sideline progress on implementation, as with the French political scandal that sidelined negotiation and implementation of the Franco-Russian alliance. More generally, protracted negotiations may result from exogenous factors including divergent preferences of domestic actors, procedural rules for finalizing agreements, and even travel distances for diplomats between prospective allies.[9] In other cases, implementation stalls as a consequence of the terms of alliance membership itself. NATO's Membership Action Plan (MAP) provides an example; prospective members take costly measures in an attempt to meet MAP criteria, and these efforts may not immediately succeed.

To capture delay of this nature, we model alliance implementation as a stochastic process. Alliances in our model are subject to exogenous delay via the parameter r, which may be interpreted as the speed of implementation of an alliance due to factors beyond the control of prospective allies. Conceptualized as such, r impacts the amount of time between the moment when prospective allies decide to form an alliance together and the moment when the military benefits of the alliance are realized - the point of implementation. This modeling choice connects our work to existing models of bilateral conflict over arming, in which weapons production is subject to exogenous delay (e.g., Coe (2018)).

Another important feature of our model is that implementation may be blocked by a preventive war. This assumption is well-grounded empirically, as states are loathe to extend alliances to other states already embroiled in conflict. For example, a study on NATO enlargement commissioned at the 1994 Brussels Summit emphasized the importance of resolving ongoing territorial disputes as a prerequisite for receiving an invitation to join the alliance. However, our substantive results do not require that a preventive war entirely blocks alliance formation. All that is necessary is that a preventive strike diminishes the benefit of full implementation or prevents implementation with some probability.[10] For example,

[9] For a general treatment of barriers to alliance negotiations, see Poast (2019).

[10] More precisely, relaxing the assumption that preventive strikes block alliances in either of these ways does not eliminate any of the qualitative features of the equilibria we characterize. Preventive strikes still occur in equilibrium, though they may fail and are followed by successful implementation. We opt for the simpler specification of the model here.

many scholars agree that the attack by the People's Republic of China on Quemoy and other Chinese Nationalist-held islands in 1954 was a failed attempt to block an alliance between the Republic of China and the United States. Chinese Communist officials expressed their interest in trying to block the alliance and the future negative implications that would accrue if the alliance is allowed to be implemented. In the lead up to the attack in September 1954, the Chinese Communist Party Central Committee declared that if the United States and Jiang sign "a treaty, the relationship between us and the United States will be tense for a long period, and it will become more difficult [for the relationship] to turn around. Therefore, the central task of our struggle against the United States at present is to break up the U.S.-Jiang treaty of defense and the Southeast Asian treaty of defense."[11] (Christensen, 1996; Zhang, 1993; Tucker, 1989).

3 Analysis

As is standard in the analysis of dynamic games, we focus on stationary, Markov perfect equilibria (Maskin and Tirole, 2001).[12] For our model, this amounts to subgame perfect equilibrium with two additional restrictions. First, the players condition their behavior in period t only on the payoff-relevant state of the world, s^t. Second, for any t, t' such that $s^t = s^{t'}$, the players use the same strategies. Henceforth we refer to such a strategy profile as an "equilibrium." Proof of all propositions appears in the appendix.

Because of our interest in the relationship between alliances, power shifts, and conflict, we focus on conditions under which States 1 and 3 attempt to form an alliance in equilibrium. The implementation of an alliance entails substantial advantages in bargaining with a shared enemy. By contrast, attempting but failing to implement the alliance results either in a

[11] Telegram, CCP Central Committee to Zhou Enlai, Concerning Policies and Measures in the Struggle Against the United States and Jiang Jieshi after the Geneva Conference, July 27, 1954, CWIHPB no. 16, p. 83; quoted in (Christensen, 2011, p. 137).

[12] Our qualitative findings carry over to non-stationary equilibria as well. For some parameter values, it is possible to construct non-stationary equilibria in which alliance formation can be "bargained away" through the expectation of off-path punishment strategies. However, our main qualitative findings persist when considering such non-stationary behavior.

preventive attack by the target or a peaceful bargaining settlement. If implementation succeeds, the allies receive their bargaining benefit from the alliance, but if implementation fails they receive no benefit and must settle for either being subjected to a preventive strike or to an unfavorable settlement before continuing to attempt implementation again in the next period. Either way, they pay the cost associated with attempted implementation in the present.

Alliances only occur in equilibrium when the benefits justify the cost. The following result establishes that this trade-off is borne out in equilibrium.

Lemma 1. If $a \; r(w_2 \; w_2^0)=(1 \;)$, then in every equilibrium, State 3 extends an alliance in every period in which $s^t = N$ and State 1 always joins if State 3 extends.

Intuitively, if the cost of alliance formation is sufficiently low, States 1 and 3 will always attempt to implement an alliance in equilibrium. The cutoff value of a indicated in the proposition is also informative. Note that as either the martial benefit of the alliance, parameterized by $w_2 \; w_2^0$ or the likelihood of immediate implementation, parameterized by r, increases, the cost constraint becomes more permissive. Intuitively, this suggests that as the benefits of the alliance become larger or are expected to arrive more quickly, the allies become more tolerant of large costs of alliance formation.

3.1 Peaceful Alliances

With an understanding of when alliance formation is attempted in equilibrium, we now address our primary question: when do the power shifts from alliance formation provoke conflict, and when can such conflict be avoided? We first demonstrate that careful bargaining can allow alliances to come into force peacefully without provoking preventive action.

How do the allies achieve peace in spite of these power shifts, which create incentives for an adversary to block an alliance through preventive war? To avoid being attacked, States 1 and 3 become generous in all periods prior to successful implementation, making generous offers to State 2. These offers serve to compensate the target of the alliance for the expected future loss of bargaining power. If the size of the power shift resulting from alliance implementation is not too large, these generous offers suffice to render bargaining more attractive than launching a preventive attack for State 2. Under these conditions, the alliance is implemented peacefully along the equilibrium path of play. Proposition 1 demonstrates this formally.

Proposition 1. If $a \quad r(w_2 \quad w_2^0)=(1 \quad)$ and $w_2^0 \quad w_2 \quad (1 \quad)(1 \quad w_2)= r$ then the following strategy profile constitutes the unique equilibrium:

- State 1 joins if State 3 extends. If $s^t = N$ and an alliance has not been successfully implemented in the current period, State 1 offers $x^t = 1 \quad w_2 \quad r(w_2 \quad w_2^0)=(1 \quad) \quad x^N$. Otherwise, State 1 offers $x^t = 1 \quad w_2^0$.

- If $s^t = N$ and an alliance has not been implemented, State 2 accepts any x^t $1 \quad w_2 \quad r(w_2 \quad w_2^0)=(1 \quad)$ and rejects otherwise. If $s^t = A$ or if an alliance was implemented in the current period, State 2 accepts any $x^t \quad 1 \quad w_2^0$ and rejects otherwise.

- State 3 extends an alliance in every period in which $s^t = N$.

Proposition 1 indicates two conditions for the peaceful implementation of an alliance. First, implementation costs cannot be too high; otherwise, allies are better off without the alliance. Second, the resulting power shift needs to be large enough to justify the costs of implementation but not so high that the allies cannot adequately compensate the target for its expected future bargaining loss.[13] In this case, States 1 and 3 always attempt to implement an alliance in equilibrium.

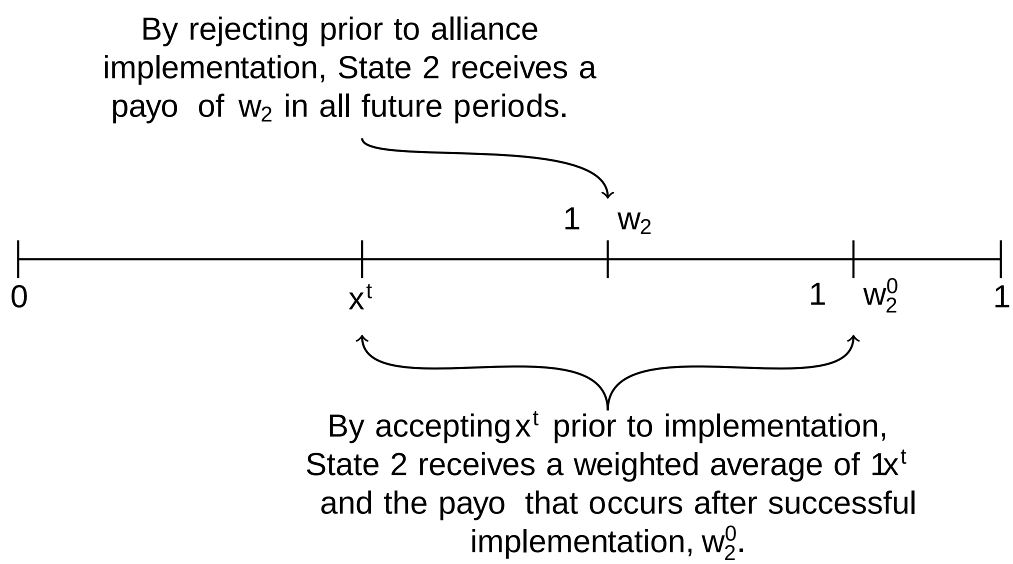
The equilibrium strategies detailed in Proposition 1 have two important features. First, the outcome is always peaceful. State 2 never chooses a preventive strike and instead is willing to accept the equilibrium settlements proposed by State 1 in all periods. Second, an

[13] Recall that higher values of w_2^0 correspond to smaller power shifts, as implementation lowers 2's war payoff to $w_2^0 \quad w_2$.

alliance is implemented with certainty in the long run because State 3 always extends an alliance, State 1 always joins that alliance, and State 2 never rejects a settlement proposal.

Peace obtains in this scenario because State 1 is willing to \buy o " State 2 from launching a preventive strike by making generous o ers in pre-implementation periods. In this case, the concessions State 1 is willing to make are su cient to compensate State 2 for the loss of bargaining power that occurs once the alliance is successfully implemented.

Figure 1: Pre-Implementation O er in Proposition 1



Note: Figure 1 illustrates how State 1 uses generous o ers prior to alliance implementation to pacify State 2. By rejecting, State 2 is able to lock in a favorable war payo , capturing all to the right of 1 w_2 in expectation. However, once implementation occurs in a future period State 2 knows that it has no choice but to accept less favorable o ers, only capturing all to the right of 1 w_2^0. By making a generous o er x^t, State 1 balances this trade-o , rendering State 2 indi erent in the present between ghting and allowing an alliance to form in the future.

Figure 1 illustrates this logic. The horizontal line in the gure represents the set of possible settlements. State 2 prefers outcomes closer to 0 while State 1 prefers outcomes closer to 1. By rejecting an o er in a period before successful implementation, State 2 receives its most favorable war payo in all future periods, receiving all to the right of the point 1 w_2. However, accepting a settlement o er in a pre-implementation period means that States 1 and 3 will have another attempt at implementing the alliance in the following period. If this attempt succeeds, the alliance leads to a shift in bargaining power and State 2 receives all to the right of $1 - w_2^0$ in all future periods.

Following this logic, accepting x^t prior to successful implementation of an alliance carries a significant drawback for State 2; it risks the realization of an alliance in the next period, and the subsequent loss of bargaining power that results. Accordingly, State 2's expected dynamic utility of accepting x^t in a pre-implementation period is a weighted sum of $1 - x^t$ and the payoff of waiting, represented by everything sitting to the right of $1 - w_2^0$.[14] As in the figure, this means that x^t must award State 2 more than its pre-implementation war payoff to render it willing to accept.

Thus, State 1's concessions in pre-implementation bargaining creates a disincentive for State 2 to launch a preemptive strike to block the alliance. If the size of the expected power shift is not too large, these concessions fully compensate State 2 for its expected post-implementation losses and undermines the incentive for a preventive strike. The outcome is peaceful implementation of the alliance.

Thus, the dynamic shifts in bargaining power that result from alliance implementation need not encourage aggression. However, this result rests on a critical condition: the size of the power shift caused by an alliance cannot not be too large. In the following section, we consider the case in which the alliance results in a more substantial shift in bargaining power.

3.2 Dangerous Alliances

So far, we have shown that alliances may arise peacefully in equilibrium. Can alliances cause wars? Our next result shows that if the power shift from an impending alliance is sufficiently large, the pacifying effect of bargaining may not be enough to stop State 1 from launching a preventive strike to block implementation. In spite of this, the alliance members will still choose to extend/join the alliance anyway. Hence, under some conditions, alliance members

[14] A more precise derivation of these continuation values appears in the proof of Proposition 1 in the appendix.

will knowingly form alliances that increase the probability of war.

While the risk of war lurks in the background in this case, war is far from guaranteed. The conditions indicated by our equilibrium analysis also suggest why wars provoked by the anticipation of an alliance are rare. By outlining the theoretical conditions under which alliances provoke war, our model ultimately sheds light on why this outcome is rarely observed empirically.

Proposition 2 formally states the conditions under which anticipation of an alliance provokes war.

Proposition 2. If a $r(w_2 \;\; w_2^0) = (1 \;\;)$ and $w_2^0 < w_2 \;\; (1 \;\;)(1 \;\; w_2) = r$ then all equilibria are equivalent in outcome distribution to the equilibrium in which players use the following strategies:

State 1 joins an alliance after State 3 extends in every period in which $s^t = N$. State 1 o ers $x^t = 0$ in every period in which $s^t = N$ and an alliance has not been implemented. Otherwise, State 1 o ers $x^t = 1 \;\; w_2^0$.

If $s^t = N$ and an alliance was not implemented in the current period, State 2 rejects all o ers x^t. Otherwise, State 2 accepts x^t if and only if $x^t \;\; 1 \;\; w_2^0$.

State 3 extends an alliance in every period in which $s^t = N$.

Why would States 1 and 3 attempt the implementation of an alliance that they know will risk war? A key feature of State 2's equilibrium strategy is that it only attacks in any period in which an alliance has not yet been implemented. Even if the conditions for a dynamic commitment problem are met, war is not guaranteed, because there will be no war if States 1 and 3 implement the alliance before State 2 can launch a preventive strike. Alliances can increase the risk of war, but the probability of war depends on the probability that an alliance comes into force before a preventive strike can block it. In this equilibrium, the formation of an alliance is a rational gamble, with the bene t of alliance weighed against the risk of provoking a preventive war.

How do the potential allies weigh these risks and bene ts of an alliance? Two factors are key. First, if the alliance can be implemented rapidly enough, then the probability of a preventive strike from State 1 is low. Second, the larger the power shift from the alliance, the more valuable the long-term bargaining bene ts from successfully implementing the alliance. The risk of trying to implement the dangerous alliance is worthwhile if it is likely to be successful and the payo s are high. In this case, States 1 and 3 gamble on the possibility of bringing the alliance into force before a preventive strike can block it. Understanding how potential allies evaluate this risk-reward tradeo is key to our explanation of the rarity of wars provoked by the anticipation of an alliance. As a prerequisite for this discussion in the section that follows, we complete our equilibrium analysis by outlining speci cally how the equilibrium probability of war reacts to changes in these two factors.

Propositions 1 and 2 indicate that whether an alliance entails a risk of provoking war depends crucially upon the value of w_2^0 relative to a key cuto . Recall that this cuto indicates that war will occur, as in Proposition 2 if

$$w_2^0 < w_2 \quad \frac{(1 \quad)(1 \quad w_2)}{r}; \tag{1}$$

and that otherwise the alliance will come into force peacefully as in Proposition 1. Inspection of inequality 1 reveals that whether it is satis ed depends on the size of the power shift, parameterized by w_2^0, and the speed with which the alliance will come into force, parameterized by r. To make the dependence upon clearer, note that the inequality can be stated equivalently as

$$r > \frac{(1 \quad)(1 \quad w_2)}{(w_2 \quad w_2^0)}: \tag{2}$$

Inequalities 1 and 2 respectively indicate the size and speed of the power shift resulting from an alliance as key to whether war occurs. These factors are familiar to scholars of con ict, operating to shape the prospects for peace across a variety of contexts (Powell, 2004).

The following result shows that the size of the power shift has a straightforward impact on the equilibrium probability of war.

Proposition 3. Suppose that a $r(w_2 \quad w_2^0) = (1 \quad)$. The equilibrium probability of war is

0 if $w_2^0 < w_2 \quad \frac{(1 \quad)(1 \quad w_2)}{r}$

$1 \quad r$ else.

All else equal, as the size of the power shift from alliance formation grows, war becomes more likely. This follows from a comparison of Propositions 1 and 2. If an alliance causes a relatively minor shift in the balance of power, bargaining may allow the alliance to come into force peacefully. However, as the alliance causes a larger shift, corresponding to lower values of w_2^0, its e ect on the future balance of power may become so large that State 2 is unwilling to allow an alliance to materialize in the present. In this case, peace only occurs if State 1 and State 2 are able to immediately bring the alliance into force, which occurs with probability r. Otherwise, war occurs with complementary probability.

Less straightforward is the relationship between the speed of the alliance's arrival and the equilibrium probability of war. Our equilibrium analysis indicates an important subtlety in the relationship between the speed with which an alliance can be brought into force and the risk of war. An increase in implementation speed, r, leads to two countervailing e ects in equilibrium. On the one hand, as inequality 2 suggests, increasing implementation speed may induce a commitment problem and lead to war where peace would have obtained otherwise. On the other hand, an increase in the speed of an alliance's arrival means that implementation is more likely to succeed before a preventive strike can occur. How do these competing e ects interact in the aggregate? The following result shows that they combine to generate a subtle nonmonotonicity in the relationship between the speed of alliance implementation and war.

Proposition 4. Suppose that a $r(w_2 \quad w_2^0) = (1 \quad)$. The equilibrium probability of war is nonmonotonic in r. In particular,

1. If $r \quad (1 \quad)(1 \quad w_2) = (w_2 \quad w_2^0) \quad r^w$, the equilibrium probability of war is 0.

2. If $r > r^w$, the equilibrium probability of war is $1 - r$.

First, note that whether war occurs at all depends upon the crucial value of indicated by inequality 2. Below this level, bargaining incentives allow an alliance to materialize peacefully in equilibrium. As r rises above it, the equilibrium probability of war jumps from 0 to a strictly positive value. But as r increases farther above this cutoff, the equilibrium probability of war begins to decrease, reaching 0 in the limit as $r \to 1$. Why is this?

This nonmonotonic relationship arises precisely because alliance implementation in the case when r is high represents a rational gamble. If implementation fails, State 2 launches a preventive strike to block States 1 and 3 from reaping the benefits of the alliance coming into force in a future period. If implementation succeeds, State 2 has no choice but to adjust to the new reality that its bargaining position has been undermined. Thus, State 2 begrudgingly bargains, granting concessions that acknowledge the newfound bargaining position of the allies.

Because of this, under the strategies in Proposition 2, the probability of delay in the implementation of an alliance, $1 - r$, mirrors the equilibrium probability of war. Equilibrium behavior introduces an endogenous correlation between fast implementation of an alliance and peace. If the allies succeed at bringing the alliance into force quickly, then peace obtains. Delay in implementation represents a window of opportunity for State 2 to attack and block a future alliance. In line with Proposition 2, if the resulting shift in bargaining power is large, seizing this opportunity is irresistible. As a result, war follows failed implementation in this case.

This analysis shows that, as is standard in models of shifting power, the size and speed of the power shift determine whether peace obtains. But this leaves an important question unanswered. Do States 1 and 3 prefer to risk war, or would they prefer to alter either the size of the power shift or the speed of the alliance's arrival prior to play of the game to preserve peace? Below we address this question with a welfare analysis, analyzing the equilibrium welfare of States 1 and 3 as a function of w_2^0 and r. With this analysis, we show that the allies will often prefer to take action to alter the terms of an alliance, manipulating the size of the power shift it causes or the speed with which it comes into force in order to avoid war.

3.3 How Allies Avoid Provocation

So far, our results have established a theoretical connection between the logic of commitment problems and the power shifts caused by impending alliance formation. Allies may choose to bring into force an alliance, knowing full well that it may provoke war. However, war in response to anticipated alliance formation is empirically rare. What accounts for this discrepancy between theory and empirics? In highlighting the conditions under which provocation occurs theoretically, our formal analysis suggests an answer; by manipulating the size and speed of the power shift caused by an impending alliance, allies may avoid the provocative e ects outlined in Proposition 2.

In this section, we perform a welfare analysis to highlight how changes in these factors relate to (1) the allies' equilibrium payo and (2) the equilibrium probability of war. We demonstrate that in many cases, reducing the risk of war is in the allies' best interest even if it comes at a cost of reducing the speed of the alliance's arrival or limiting the size of the power shift resulting from the alliance. In fact, our welfare results show that the conditions under which the allies prefer an alliance that entails a positive risk of war to an alliance that is not provocative are quite narrow. These theoretical results have important empirical implications, highlighting tools of alliance design and maintenance that allies may use to avoid provocation while providing a theoretical frame to understand the empirical rarity of wars provoked by alliance formation.

Our rst welfare result highlights an important feature of our model: the equilibrium utilities of the allies are not monotonic in the size and speed of the power shift resulting from an alliance.

Proposition 5. The equilibrium utilities of State 1 and State 3 are:

nonmonotonic in w_2^0

nonmonotonic in r

This result suggests that allies do not unambiguously bene t from increases in the speed with which an alliance is brought into force or the size of the power shift from an alliance. This is despite clear bene ts of increases in these factors. If an alliance comes into force quickly, the allies bene t by more quickly enjoying the bene ts of military cooperation in bargaining with a shared enemy. If the size of the power shift from such cooperation grows, the bargaining bene ts grow further. However, Proposition 5 indicates that changes in these factors are not unambiguously bene cial. What is the drawback?

For each factor, the drawback is the same: an increased risk of provocation and war. Because war is costly, provocative alliances entail a signi cant drawback, in spite of the bene ts outlined above. Indeed, in many cases allies may wish to alter the details of an alliance in order to avoid provoking a shared enemy. How do they do so?

One tool is in limiting the scope of military cooperation resulting from an alliance. This serves to reduce the size of the anticipated power shift, making the alliance less threatening to a shared enemy. In many cases, the allies may prefer to accept a lower level of military cooperation in order to reduce the risk of war. The following corollary formalizes this.

Corollary 1. Let $w_2 \quad (1 \quad)(1 \quad w_2) = r \quad \overline{w}_2$ be the highest value of w_2^0 for which peace occurs in equilibrium with probability 1. There exists a unique $w_P < \overline{w}_2$ such that State 1 and State 3's equilibrium utilities are both higher under $\overline{w}_2$ than under any $w_2^0 \in (w_P; \overline{w}_2)$.

According to corollary 1, there always exists a range of values of w_2^0 such that the bene t of the power shift is not worth the risk of provocation. Under this range, the allies are better o limiting the scope of military cooperation, even though doing so results in a weaker alliance.

The techniques used by prospective allies to mute such power shifts are varied. In many cases, the content of an alliance contract is manipulated to limit the nature of the e ective power shift under which the alliance applies. Previous work has suggested that a primary driver for these content-based manipulations is aimed at establishing deterrence (Leeds, 2003) without increasing the risk of moral hazard (Benson, 2012). This previous work points to a desire to manipulate the content of an alliance to reign in the ambitions of one's alliance partner. The present analysis suggests that these manipulations to content have another important feature - they avoid provoking an attack from a shared enemy. By limiting the scope of an alliance during the period of formation, it becomes less threatening, possibly avoiding provocation.

In addition to manipulating the size of the power shift resulting from alliance formation, allies may avoid provocation by limiting the speed with which an alliance is put into force. The following result formalizes that, under some conditions, the allies may prefer to delay or slow the implementation of an alliance in order to avoid a risk of war.

Corollary 2. Let r (1)(1 w_2)= (w_2 w_2^0) r^w be the highest value of r for which peace occurs in equilibrium with probability 1. There exists a unique $\underline{r}$ with $r^w < \underline{r} < 1$ such that State 1 and State 3's equilibrium utilities are both higher under r^w than under any r 2 $(r^w; \underline{r})$.

Though all else equal allies would prefer to enjoy the bargaining benefits of a new alliance more quickly, our model suggests that all else is not equal. If an alliance is expected to arrive quickly, the resulting threat may provoke aggression from an enemy. This creates the nonmonotonicity in equilibrium utility outlined in proposition 5 and corollary 2. If an alliance cannot be brought into force quickly enough to outpace a preventive strike from the enemy, it may be in the allies' best interest to intentionally put off the implementation of the alliance, delaying its arrival to avoid provoking aggression from an enemy concerned with the alliance's otherwise rapid and threatening arrival.

How do states achieve a deliberately drawn-out implementation of a security commitment? NATO's membership action plan is one example. This plan provides for a drawn-out, staged entry of new members. By imposing a number of requirements on prospective members, the entry process unfolds over a relatively long period of time rather than occurring rapidly.

Eliminating the window of time in which an enemy might react is another viable strategy to avoid provocation. The logic of Proposition 2 showed that the target of an alliance will only launch a preventive war if it is afforded a window of opportunity to attack before the alliance comes into force. This suggests that, by reducing or eliminating this window of opportunity, allies may be able to reduce the risk of preventive war, even when the alliance would result in a substantial power shift once it is implemented. How do allies eliminate this incentive?

One tool is secrecy. By keeping knowledge of an impending alliance confidential until the alliance is already implemented, allies can prevent a shared enemy from responding until it is too late and the alliance has already come into force. In terms of the formal model, this corresponds to a very high value of, with the alliance coming into force almost immediately before the target has a chance to take preventive action. The 1894 Franco-Russian alliance is an example. After nearly three years of negotiations, France and Russia came together to form a secret alliance targeting Germany. France and Russia wanted to form a power two front alliance to counter Germany without provoking it. Although it took time to negotiate the specifics of the alliance, its implementation was done in secrecy (Snyder, 1997, pp. 120–121).

The following proposition shows that, if the allies are able to bring the alliance into force before an ally can react, then welfare can be high even though there is a small risk of war.

Proposition 6. Fix some value $r < 1$. There exists an $\bar{r}$ with $r < \bar{r} < 1$ such that both State 1 and State 3's equilibrium utilities are higher under every $r' \geq \bar{r}$ than under r.

This finding builds on prior work by indicating a novel role of secrecy in multilateral conflict processes. Previous studies of secrecy among allies point to either its role in generating asymmetric information (Bas and Schub, 2016) or in facilitating communication among military partners (Smith, Forthcoming). While these studies focus on the important role of secrecy in an environment of incomplete information, the present analysis highlights the dynamic e ects of secrecy. In this context, secrecy can allow allies to \surprise" a shared enemy with the existence of a new alliance. This serves to bring the alliance into force before the shared enemy can mount an attack to block the alliance. Accordingly, extremely fast implementation of an alliance due to secrecy can allow the allies to avoid war with high probability.

4 Discussion

Our theoretical results have important implications for our understanding of observed patterns of alliances and con ict. Below, we identify how the incentives highlighted by the model shaped the prospects for peace or con ict in two important cases from the historical record. Throughout, we aim to show that the causal mechanisms identi ed by the formal model operated in these cases. We contrast two important cases, tracing the motivation and strategic reasoning of key decision-makers when possible (Goemans and Spaniel, 2016; Lorentzen, Fravel and Paine, 2017). First, we argue that bargaining concessions alongside limits to the timeline and scope of expansion won Soviet approval for the 1990 reuni cation of Germany within NATO. Second, we argue that the spectre of Georgia's accession to NATO, which accelerated after a 2008 summit meeting, provoked Russia to invade Georgia in an attempt to block its membership in the alliance.

In contrasting these cases, we aim to identify and illustrate the formal model's logic. We argue that the strategic logic advanced by the formal model helps to explain the di erent outcomes in each case. In the case of German uni cation, U.S. and German negotiators were able to manipulate the terms of German NATO membership, rendering it less provocative by slowing the withdrawal of Soviet troops and placing limits on the number of western troops that would replace them. Consistent with the logic of Proposition 1, these manipulations created room for peaceful bargaining, with Germany compensating the Soviets with loans in exchange for their admission into NATO. In contrast, the 2008 war in Georgia illustrates the logic of the effect of an increase in implementation speed on preventive war as described in Proposition 2. The acceleration of Georgia's path to membership following a crucial NATO summit meeting in 2008 raised the spectre of threat to Russia from an anticipated power shift due to eventual Georgian NATO membership. Russia invaded in an attempt to forestall the otherwise inevitable accession. Below we describe each case in more detail, tying them to the logic of the formal model.

4.1 German Reunification and the NATO Question

The 1990 reunification of Germany, and subsequent inclusion of the newly unified state into NATO represented a fundamental upending of the Cold War security order. Though West Germany had been a NATO member since 1955, unification posed a provocative new question. Would East German territory, formerly covered under the Soviet-led Warsaw pact, come under the security umbrella of NATO? This question weighed heavily in the minds of diplomats during talks on German reunification. In February 1990, Gorbachev indicated that the idea of a unified Germany in NATO was incompatible with Soviet aims (Newnham, 1999, 428-429). Nevertheless, Germany reunified and peacefully joined NATO without provoking Soviet action. The strategic logic of our formal model explains how troop limits, withdrawal timelines, and German loans were instrumental in ensuring peaceful reunificaton under NATO.

The case bears a striking resemblance to the structure of the formal model. NATO allies, led by the U.S.,[15] sought the inclusion of unified Germany in NATO. As a result, German diplomats worked to find solutions that would allow this to occur. These roles are analogous to those of states 3 (U.S.) and 1 (Germany) in our model, respectively. The Soviets played a role akin to that of State 2, bargaining with the prospective allies while maintaining the threat to use military action to block the alliance from coming into force.

Consistent with our theoretical analysis, concerns over what the Soviets might do in an-

[15] It is widely acknowledged that the U.S. led the way in reunification talks, with NATO allies largely following the agenda set by U.S. negotiators (e.g., see Moens (1991) and (Sarotte, 2010b, p. 130)).

ticipation of unification and NATO expansion figured prominently in the 1990 negotiations. Gorbachev's approval of a unification deal was far from guaranteed at the outset of negotiations. At the time, roughly 400,000 Soviet troops were stationed in East Germany. The possibility of military resistance to an unfavorable reunification arrangement was real, and Gorbachev repeatedly resisted pushes from Soviet military officials for action in the period leading up to reunification (Newnham, 1999, p. 428). A simple refusal to remove Soviet troops from the east could have had disastrous consequences. Accordingly, U.S. diplomats recognized early on that Soviet approval was key, and worked throughout to gain Gorbachev's assent for a unification process that would result in unified Germany's receipt of full NATO membership.

How was this approval gained and provocation avoided? The logic of our equilibrium analysis highlights some important factors that are clearly present in the historical record. Proposition 1 describes the conditions under which an alliance may come into force peacefully without provoking a shared enemy. Two especially important factors are the size of the power shift resulting from an alliance and the speed with which the alliance is implemented. For an alliance to come into force peacefully in equilibrium, the resulting shift in power cannot be too high. In addition, it cannot be too fast or the alliance target will perceive the realization of the power shift to be imminent and inevitable. If the power shift is too high or its implementation too speedy, the condition for preventive war in Proposition 2 is satisfied, and the target will initiate a strike to block the implementation of the alliance.

Consistent with this logic, provisions such as troop limits and withdrawal timelines were key to winning Soviet assent for unified Germany's membership. In particular, the U.S. pushed for \nine assurances" that were aimed to make the transition less threatening. Key among these provisions were limits on NATO troops in future Soviet territory, as well as a relaxed withdrawal timeline and limits on the placement of nuclear weapons in German territory. A particularly important restriction was Helmut Kohl's agreement to limit the number of troops in the Bundeswehr to 370,000 (Sarotte, 2010, p.125-130). These provisions served to moderate the size and speed of the power shift resulting from membership, rendering the change less provocative. Indeed, Gorbachev's own memoirs indicate these limits to the power and timeline of the transition to unification and full NATO membership were key in opening up the possibility of a negotiated settlement (Gorbachev, 1996, p. 527-529).

With an opening for peaceful bargaining over unification and NATO expansion, how was the final agreement constructed to achieve Gorbachev's approval? Consistent with the logic provided in Proposition 1 of our model, Germany made concessions to pacify the Soviets. As the proposition demonstrates, when the power shift from a prospective alliance is not too high and implementation speed not too fast, then bargaining concessions from prospective allies to a shared enemy are key to inducing the alliance target to allow peaceful implementation of the alliance.

The historical record provides evidence consistent with the model's logic. In a meeting early in the negotiation process, U.S. President George H.W. Bush suggested that West German Chancellor Helmut Kohl's \deep pockets" might prove advantageous (Rice and Zelikow, 1995, p. 215) in facilitating a deal. Germany then followed through by providing substantial loans to the Soviets. Even public observers noted the importance of these loans in securing Soviet agreement on the NATO issue. Illustrating this, the 1990 July issue of *The Economist* ran a political cartoon depicting Gorbachev holding open door for Helmut Kohl labeled \NATO," while Gorbachev's other hand clutched a bag of Deutsche Marks.[16] Subsequent historical analysis has reiterated the widespread perception that these loans were pivotal in ensuring the success of the negotiations that resulted in German unification and inclusion in NATO (Newnham, 1999; Sarotte, 2010b).

4.2 2008 Russo-Georgian War

Our theory also demonstrates that preventive war to block an alliance might occur if the alliance will result in a very large shift of power and its implementation speed is too fast.

[16] This cartoon is reproduced in the appendix.

We find evidence for this logic in the attempt to hasten Georgian entry to NATO that resulted in the 2008 Russo-Georgian conflict. We provide evidence that efforts to accelerate Georgia's path to NATO membership provoked Russia to act aggressively to block Georgia's membership bid.

Our equilibrium analysis, as demonstrated in Proposition 2, suggests that alliances may provoke conflict if implementation occurs within a relatively short time horizon. The outcome of the April 2008 NATO summit represents such an acceleration of Georgia's path to NATO membership. Georgia had been on a steady but slow path to membership since the 2003 Rose Revolution, when Georgian President Mikheil Saakashvili made Georgian accession to NATO a foreign policy priority. Georgia increased its ties with NATO through active steps such as contributing forces to the ISAF in Afghanistan. Efforts continued in October 2004 with the initiation of a NATO Individual Partnership Action Plan for Georgia. This plan required that the separatist claims over the Abkhazia and South Ossetia regions be resolved. While these actions all represented steps toward Georgian accession, membership was still a distant event on the horizon.

The speed of accession efforts suddenly accelerated at the April 2008 NATO summit in Bucharest, where US President George W. Bush lobbied for the extension of a Membership Action Plan for Georgia. His proposal was met with reluctance from German and French leaders because of the status of Georgia's internal politics, including separatist movements. In spite of this, NATO Secretary General Jaap de Hoop Scheffer confirmed for the first time the inevitably of Georgian membership, stating that the issue would be taken up again at a summit in December 2008.[17] This represented a significant change in NATO's orientation towards Georgia. NATO membership was no longer a hypothetical. It was now an inevitability. In our model, alliances provoke only when they are expected to be implemented relatively quickly but experience short-term delay. NATO's explicit commitment to granting Georgia membership, along with the promise to revisit the membership discussion in less

[17] http://news.bbc.co.uk/2/hi/europe/7328276.stm

than a year, represented such a situation.

This acceleration of Georgia's NATO membership bid appears to have provoked a preventive strike from Russia, NATO's main target. Foreshadowing what was to come, Vladimir Putin warned on the nal day of the Bucharest meeting that NATO's promise to Georgia did not contribute to trust and predictability in NATO-Russia relations and would be destabilizing.[18] Russian ambassador to NATO Dmitry Rogozin echoed this, saying that \The attempt to push Georgia into NATO is a provocation."[19]

In August 2008, within just a few months of the April Bucharest meeting, war broke out. Consistent with our model, evidence indicates that Russia was motivated to forestall Georgia's NATO quest. Internal con ict with separatists in Abkhazia and South Ossetia had long been a barrier to Georgia's NATO membership aspirations. So Russia choose to aggravate that con ict. Mere weeks after the Bucharest summit, Russia lifted sanctions against Abkhazia, which emboldened the separatists. When Georgia sent troops to South Ossetia to end separatist attacks, Russia responded by sending their troops into Georgia on a purported \peace mission." The ghting lasted for 9 days and resulted in over 300 deaths and more than 1,000 wounded. In the wake of the 2008 con ict, Russia granted o cial recognition to Abkhazia and South Ossetia, creating an additional hurdle on Georgia's path to membership. As Mearsheimer (2014) argues, this is consistent with the idea that Russia's goal in 2008 was to stoke internal divisions in Georgia in order to keep them out of NATO.

Statements of Russian decision makers after the war make it clear that, at least in the minds of Russian decision-makers, Georgian accession to NATO was provocative. For example, in September 2008, just as the con ict had died down, Russian President Dmitry Medvedev stated that \NATO only provoked the con ict, and not more than that."[20] In 2011 Dmitry Medvedev addressed Russian troops by saying that Georgia would have already

[18] https://www.thetimes.co.uk/article/vladimir-putin-tells-summit-he-wants-security-and-friendship-96655h3k9nf

[19] https://www.spiegel.de/international/world/interview-with-russia-s-ambassador-to-nato-the-attempt-to-push-georgia-into-nato-is-a-provocation-a-540426.html

[20] https://www.reuters.com/article/us-russia-medvedev-west/russia-says-nato-provoked-georgia-conflict-idUSLJ45058520080919

been a NATO member if Russia had failed to act.[21]

Thus, our theoretical model provides a useful lens for interpreting the conflict between Russia and Georgia. Provoked by the acceleration in Georgia's path to NATO membership, Russia took military action to block implementation of an alliance commitment. As of November 2020, Georgia was still not a full NATO member.

5 Conclusion

We began by asking why military alliances rarely provoke preventive war. We argued that temporal dynamics provide the missing theoretical link between alliances and the logic of commitment problems. This provides a key conceptual innovation: to understand this connection, security cooperation must be treated as a dynamic process that unfolds over time.

We found that alliances rarely provoke commitment problems because prospective allies have a variety of tools to avoid provocation. In designing alliances, allies can manipulate the speed of implementation and the size of the power shift to reduce incentives for preventive attack. This creates conditions in which concessions to the target of the alliance can induce their peaceful assent to the formation of the alliance.

Even though allies have options to form alliances that avoid commitment problems and war, we cannot rule out war altogether. Some alliances are indeed dangerous and might induce preventive attacks. If the speed of implementation is sufficiently high and the resulting power shift is sufficiently large, alliances do indeed sometimes lead to war in equilibrium. Importantly, when the power shift is sufficiently high and the implementation speed sufficiently fast, allies do not have an incentive to take actions to avoid war. Forming the alliance if the face of a high risk of preventive war is the result of an intentional decision on behalf of the allies. The choice to implement an alliance represents a costly gamble in which the allies race to implement the alliance before a preventive strike can occur.

Our main contribution consists in highlighting temporal dynamics as a key and novel

[21] https://in.reuters.com/article/idINIndia-60645720111121

factor connecting alliances and war. As alliances alter the future distribution of power, expectations of their future arrival alter behavior in the present. These findings point to the importance of taking time seriously in the study of alliance politics and war. Our theory does precisely this, with implications expanding beyond the present analysis. Our novel framework speaks to two important topics in the study of alliances and security cooperation including research on alliance entrapment, as well as research on the relationship between alliances and war.

First, the logic of preventive war in our model adds an equilibrium explanation to extant explanations of alliance entrapment (Snyder, 1984; Benson, 2012). In our model, allies' preferences about a war might diverge because allies have different payoffs from war. As a result, the ally with the higher war payoff has an incentive to bargain aggressively with the target or, if it has control over the speed of alliance implementation, choose a provocative implementation speed so as to cause war to break out. The other ally will fight reluctantly when so obligated by the conditions of the alliance. Why does the reluctant ally join an alliance that entails an increased risk of war? Such states will join when the risk of war breaking out is low because of high implementation speed and the increased payoff from a peaceful bargaining settlement is high. Given the expected benefits of this lottery, allies, even reluctant ones, will willingly join alliances that they know might result in war. By painting alliance implementation as a rational gamble, our analysis provides a new frame to understand entrapment. Under some conditions, allies may enter contracts that they know entail a small, but significant risk of entrapment in an undesirable war. Our analysis suggests they do so in the hopes that an alliance will be implemented before such a strike can occur.

Second, an implication of our theory is that alliances are effective in deterring war after implementation. If war happens as a result of alliance formation and commitment problems, then the timing of conflict will be around the time of formation. In this sense, our study contributes to the debate about conflicting findings from statistical analyses of the effect of alliances on conflict. While some studies suggest alliances deter conflict (Leeds, 2003; Benson, 2012; Johnson and Leeds, 2011; Leeds and Johnson, 2017), others argue that alliances increase the risk of war in the years immediately preceding and following alliance formation (Kenwick, Vasquez and Powers, 2015; Kenwick and Vasquez, 2017).

Morrow (2017) reconciles these findings using a formal model that incorporates the possibility of both provocation and deterrence. In Morrow's model, uncertainty about the conflict of interest between the target of an alliance and the recipient of the alliance commitment in the main mechanism that might lead to conflict. Our model offers a complementary qualitative prediction; under different circumstances, alliances may deter or provoke. However, the causal mechanism that generates these behaviors is distinct from the informational logic of Morrow's analysis.

The dynamic conceptual framework of our analysis outlines the equilibrium conditions and timing of both provocation and deterrence.[22] Our theory finds that deterrence succeeds after an alliance is fully implemented: the target accepts the new distribution of power and is effectively deterred from making aggressive demands or initiating a preventive war. In this way, our model is consistent with the notion that alliances have long-term deterrent effects (Leeds and Johnson, 2017). Indeed, our model suggests that alliances achieve their maximal deterrent benefits in the long run, once they have been fully implemented and no longer represent an expected future power shift.

However, in a world where alliance implementation is not guaranteed to be instantaneous, war is still possible. Importantly, and consistent with Kenwick, Vasquez and Powers (2015), this logic identifies the time around alliance formation as the most dangerous.[23] Any delay in implementation represents a window of opportunity for the target and may lead to war if the

[22] Interestingly, the commitment-problem logic speaks directly to the importance of the timing of an alliance. As Leeds and Johnson point out, "The consequential difference between Johnson and Leeds (2011) and Kenwick, Vasquez and Powers (2015) is whether the effect of alliances should be observed throughout their existence or only for a short time after gaining a new ally." (2017, p.373)

[23] A possible objection to this logic is that the target of the alliance should either adjust its demands (Kenwick and Vasquez, 2017, p.330) or abandon them altogether in response to this anticipated power shift (Johnson and Leeds, 2011). This is precisely what occurs in equilibrium in Proposition 1, as the target of the alliance begrudgingly accepts that an alliance will occur and adjusts its demands once implementation has succeeded.

anticipated power shift is large enough. However, our analysis suggests that alliances are not always provocative. Thus, our model provides an additional framework for understanding the con icting ndings in the empirical literature.

In conclusion, our theoretical framework suggests some potential avenues for future empirical and theoretical work. An avenue for future research is to identify which cases of alliance formation and con ict t the informational logic of Morrow (2017) or the complementary commitment problem dynamics that we have presented here. While we have provided evidence that this mechanism exists in the context of speci c cases, quantitative evaluation of this mechanism across a larger body of cases is a worthwhile endeavor. Such a targeted test of our model should begin by collecting data on the timing of alliance implementation, as our model identi es the period during which implementation has not yet occurred as the most vulnerable. Additionally, future studies might also evaluate the implications of dynamic alliance formation for alliance design and bargaining with the enemy.

Finally, our ndings pave the way for further theoretical inquiry. Future work might look to augment our framework by incorporating exogenous power shifts into the model of alliance implementation. Indeed, prominent explanations of alliance formation, and alignments more generally, paint alliances as responses to shifts in either the actual or perceived balance of power (Waltz, 1979; Walt, 1990). Incorporating exogenous power shifts alongside the endogenous, alliance-driven power shifts we have modeled would allow for a systematic analysis of the interaction of these forces as well as a game-theoretic analysis of bandwagoning and balancing, expanding on the analysis of (Powell, 1999).

References

Acemoglu, Daron and James A Robinson. 2001. \A Theory of Political Transitions." American Economic Review 91(4):938{963.

Bas, Muhammet A and Andrew J Coe. 2016. \A Dynamic Theory of Nuclear Proliferation and Preventive War." International Organization 70(4):655{685.

Bas, Muhammet and Robert Schub. 2016. \Mutual optimism as a cause of con ict: secret alliances and con ict onset." International Studies Quarterly 60(3):552{564.

Benson, Brett V. 2012. Constructing International Security: Alliances, Deterrence, and Moral Hazard. Cambridge University Press.

Benson, Brett V, Adam Meirowitz and Kristopher W Ramsay. 2014. \Inducing Deterrence Through Moral Hazard in Alliance Contracts." Journal of Con ict Resolution 58(2):307{335.

Chang, Gordon H. 1990. Friends and Enemies: The United States, China, and the Soviet Union, 1948-1972. Stanford University Press.

Christensen, Thomas J. 1996. Useful Adversaries: Grand Strategy, Domestic Mobilization, and Sino-American Con ict, 1947-1958. Princeton University Press.

Christensen, Thomas J. 2011. Worse than a Monolith: Alliance Politics and Problems of Coercive Diplomacy in Asia. Princeton University Press.

Coe, Andrew J. 2018. \Containing Rogues: A Theory of Asymmetric Arming." The Journal of Politics 80(4):1197{1210.

De Figueiredo, Rui JP. 2002. \Electoral Competition, Political Uncertainty, and Policy Insulation." American Political Science Review 96(2):321{333.

Fang, Songying, Jesse C Johnson and Brett Ashley Leeds. 2014. \To concede or to resist? The restraining e ect of military alliances." International Organization 68(4):775{809.

Fearon, James D. 1995. \Rationalist Explanations for War."International organization 49(3):379{414.

Fearon, James D. 2004. \Why Do Some Civil Wars Last So Much Longer than Others?" Journal of peace research 41(3):275{301.

Goemans, Hein and William Spaniel. 2016. \Multimethod research: A case for formal theory." Security Studies 25(1):25{33.

Goldgeier, James M. 2010. Not whether but when: The US decision to enlarge NATO. Brookings Institution Press.

Gorbachev, Mikhail Sergeevich. 1996. Mikhail Gorbachev: Memoirs. Doubleday.

Johnson, Jesse C and Brett Ashley Leeds. 2011. \Defense Pacts: A Prescription for Peace?" Foreign Policy Analysis 7(1):45{65.

Kenwick, Michael R and John A Vasquez. 2017. \Defense Pacts and Deterrence: Caveat Emptor." The Journal of Politics 79(1):329{334.

Kenwick, Michael R, John A Vasquez and Matthew A Powers. 2015. \Do Alliances Really Deter?" The Journal of Politics 77(4):943{954.

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

Leeds, Brett Ashley. 2003. \Do Alliances Deter Aggression? The In uence of Military Alliances on the Initiation of Militarized Interstate Disputes." American Journal of Political Science 47(3):427{439.

Leeds, Brett Ashley and Jesse C Johnson. 2017. "Theory, Data, and Deterrence: A Response to Kenwick, Vasquez, and Powers." The Journal of Politics 79(1):335–340.

Lorentzen, Peter, M Taylor Fravel and Jack Paine. 2017. "Qualitative investigation of theoretical models: the value of process tracing." Journal of Theoretical Politics 29(3):467–491.

Maskin, Eric and Jean Tirole. 2001. "Markov Perfect Equilibrium: I. Observable Actions." Journal of Economic Theory 100(2):191–219.

Mearsheimer, John J. 2014. "Why the Ukraine Crisis is the West's Fault: The Liberal Delusions that Provoked Putin." Foreign Aff. 93:77.

Meirowitz, Adam and Anne E Sartori. 2008. "Strategic Uncertainty as a Cause of War." Quarterly Journal of Political Science 3(4):327–352.

Moens, Alexander. 1991. "American diplomacy and German unification." Survival 33(6):531–545.

Morrow, James D. 1991. "Alliances and Asymmetry: An Alternative to the Capability Aggregation Model of Alliances." American journal of political science pp. 904–933.

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

Morrow, James D. 2000. "Alliances: Why Write Them Down?" Annual Review of Political Science 3(1):63–83.

Morrow, James D. 2017. "When Do Defensive Alliances Provoke Rather than Deter?" The Journal of Politics 79(1):341–345.

Newnham, Randall. 1999. "The price of German unity: The role of economic aid in the German-Soviet negotiations." German Studies Review 22(3):421–446.

Olson, Mancur and Richard Zeckhauser. 1966. \An Economic Theory of Alliances." The review of economics and statistics pp. 266{279.

Poast, Paul. 2019. Arguing About Alliances: The Art of Agreement in Military-Pact Negotiations. Cornell University Press.

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

Powell, Robert. 2004. \The Ine cient Use of Power: Costly Con ict with Complete Information." American Political science review 98(2):231{241.

Powell, Robert. 2006. \War as a Commitment Problem." International organization 60(1):169{203.

Powell, Robert. 2012. Commitment Problems and Shifting Powers a Cause of Con ict. In The Oxford Handbook of the Economics of Peace and Con ict.

Powell, Robert. 2017. \Research Bets and Behavioral IR." International Organization 71(S1):S265{S277.

Rice, Condoleezza and Philip Zelikow. 1995. \Germany Uni ed and Europe Transformed." A Study in .

Sarotte, Mary Elise. 2010a. \Not one inch eastward? Bush, Baker, Kohl, Genscher, Gorbachev, and the origin of Russian resentment toward NATO enlargement in February 1990." Diplomatic History 34(1):119{140.

Sarotte, Mary Elise. 2010b. \Perpetuating US Preeminence: the 1990 deals to \Bribe the Soviets out" and move NATO in." International Security 35(1):110{137.

Shifrinson, Joshua R Itzkowitz. 2016. \Deal or no deal? The end of the Cold War and the US o er to limit NATO expansion." International Security 40(4):7{44.

Smith, Alastair. 1995. \Alliance Formation and War." International Studies Quarterly 39(4):405{425.

Smith, Bradley C. Forthcoming. \Military Coalitions and the Politics of Information." Journal of Politics .

Snyder, Glenn H. 1984. \The Security Dilemma in Alliance Politics."World Politics 36(4):461{495.

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

Tucker, Nancy Burnkopf. 1989. Cold War Contacts: America and China, 1952-1956. In Sino-American Relations, 1945-1955, ed. Harry Harding and Yuan Ming. Wilmington, DE: Scholarly Resources pp. 238{66.

Walt, Stephen M. 1990. The origins of alliance. Cornell University Press.

Waltz, Kenneth N. 1979. Theory of international politics. McGraw-Hill.

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

Zhang, Shu Guang. 1993. Deterrence and Strategic Culture: Chinese-American Confrontations, 1949-1958. Cornell University Press.

Online Appendix for \Why Alliances Rarely Provoke War"

Contents

A Proof of Propositions in Main Text

Lemma 1. *If $a \leq r(w_2 - w_2^0)=(1- \ )$, then in every equilibrium, State 3 extends an alliance in every period in which $s^t = N$ and State 1 always joins if State 3 extends.*

Proof. We proceed in two steps. First, we show that if $a < a$, then State 1 must join if State 3 extends an alliance. Second, we show that State 3 must extend an alliance.

We prove the first step by contradiction. Suppose that   is a stationary MPE in which State 1 does not join if State 3 extends an alliance and that $a < a$. Under , the payoff of war in every period for State 1 is w_1. Standard arguments establish that, under such a , in every period in which $s^t = N$, State 1 offers $1 - w_2$ and State 2 accepts x^t if and only if $x^t \leq 1 - w_2$. Now, consider State 1's reaction to State 3's (potentially off-path) extension of an alliance in a period in which $s^t = N$. As   is an equilibrium, State 2 cannot profitably deviate to joining. Therefore, it must be that

$$r(1 - w_2^0) + (1 - r)(1 - w_2) - (1- \ )a \geq 1 - w_2 \implies a > r(w_2 - w_2^0)=(1- \ ) \equiv a \ ;$$

a contradiction. Therefore, in equilibrium State 1 must join after State 3 extends an alliance.

Our second step is to show that State 3 must extend an alliance. Again, we proceed by contradiction. Suppose that   is a stationary MPE in which State 3 never extends in a period in which $s^t = N$ and that $a < a$. As State 1 must join, State 3 does not have a profitable deviation if

$$r(1 - w_2^0) + (1 - r)(1 - w_2) - (1- \ )a \geq 1 - w_2 \implies a > r(w_2 - w_2^0)=(1- \ ) \equiv a \ ;$$

a contradiction. This completes the proof. □

Lemma 2. *If   is a stationary MPE, then in every period in which $s^t = A$, State 1 and State 2 use the following strategies*

- *State 1 offers* $x^t = 1 - w_2'$
- *State 2 accepts* x^t *if and only if* $x^t \leq 1 - w_2'$.

Proof. First, we show that State 1 may profitably deviate from any value x^t other than $x^t = 1 - w_2'$. To begin, suppose that State 1 is offering some value $x^t > 1 - w_2'$ in every period in which $s^t = A$. If State 1 is using such a strategy, then State 2 must be rejecting x^t, as

$$w_2' > 1 - x^t.$$

Now, consider a deviation to offering $1 - w_2' - \epsilon$ for some small $\epsilon > 0$. State 2 will accept this in period t, as

$$(1-\delta)(w_2' + \epsilon) + \delta w_2' > w_2.$$

This deviation is profitable for State 1 if

$$(1-\delta)(1 - w_2' - \epsilon) + \delta w_1' > w_1',$$

which holds for sufficiently small $\epsilon > 0$.

Next, suppose that State 1 is offering some $x^t < 1 - w_2'$ in equilibrium. State 2 must accept this offer, as it yields a strictly higher payoff than war. Now, consider a one-shot deviation from x^t to $x' = (x^t + 1 - w_2')/2$. State 2 will accept, as

$$(1-\delta)(1 - x') + \delta(1 - x^t) > 1 - x^t > w_2'.$$

Because State 2 accepts and $x' > x^t$, this deviation is profitable for State 1. Therefore, State 1 cannot offer $x^t < 1 - w_2'$ in equilibrium. As State 1 cannot offer $x^t < 1 - w_2'$ or $x^t > 1 - w_2'$, it follows that State 1 must be offering $x^t = 1 - w_2'$.

□

Lemma 3. *Suppose that* $w_2^0 \geq w_2 - (1 - \ )(1 - w_2) = \ r$*. If* *is a stationary MPE in which State 3 extends and State 1 joins in every period in which* $s^t = N$*, then in every period in which* $s^t = N$ *and an alliance was not successfully implemented, State 1 o ers* $x^t = \max\{0; 1 - w_2 - \ r(w_2 - w_2^0) = (1 - \ )\}$*.*

Proof. First, suppose that $w_2^0 \geq w_2 - (1 - \ )(1 - w_2) = \ r$.

Let be a stationary MPE in which State 3 extends an alliance in every period in which $s^t = N$ and State 1 always joins in response.

We show that State 1 must be offering $x^t = \max\{0; 1 - w_2 - \ r(w_2 - w_2^0) = (1 - \ )\} \equiv x^N$. To do this, we show that if State 1 uses any other stationary strategy in such a period, offering $x^t < x^N$ or $x^t > x^N$, there exists a profitable deviation. As a preliminary step, we form State 2's continuation value under of accepting x^t if State 1 offers x^t in every period in which $s^t = N$. Denoting this continuation value V^2, we have

$$V^2 = (1 - \ )(1 - x^t) + \ [r w_2^0 + (1 - r) V^2]:$$

Rearranging this expression yields

$$V^2 = \frac{(1 - \ )(1 - x^t) + \ r w_2^0}{1 - \ + \ r}:$$

It is optimal for State 2 to accept x^t in a period in which $s^t = N$ under if $V^2 \geq w_2$. Substituting and solving for x^t, we find that State 2 will accept if

$$x^t \leq 1 - w_2 - \ r(w_2 - w_2^0) = (1 - \ ):$$

Note that our assumption that $w_2^0 \geq w_2 - (1 - \ )(1 - w_2) = \ r$ implies that $1 - w_2 - \ r(w_2 - w_2^0) = (1 - \ ) \geq 0$.

Now suppose that State 1 offers some $x^t < x^N$ in a period in which an alliance has not been implemented and $s^t = N$. By the argument above, State 2 will accept x^t. As before, let V^2 be State 2's continuation value of accepting x^t under . Consider a deviation to $x^t +$ for some arbitrarily small value > 0. State 2 must accept this offer if

$$(1 - \ )(1 - x^t - \ ) + \ [rw_2^0 + (1 - r)V^2] > w_2:$$

Recall that $x^t < x^N \implies V^2 > w_2$. Therefore, it suffices to show that

$$(1 - \ )(1 - x^t - \ ) + \ [rw_2^0 + (1 - r)w_2] > w_2:$$

This holds if

$$x^t + \ < 1 - w_2 - \ r(w_2 - w_2^0)=(1 - \ );$$

which holds for sufficiently small > 0 as $x^t < x^N$. This contradicts the assumption that is an equilibrium, and so State 1 cannot offer $x^t < x^N$ in equilibrium.

Next, suppose that in every period in which $s^t = N$, State 1 offers some $x^t > x^N$ after an alliance fails to be implemented under . Consider a one-shot deviation from x^t to $x^N -$ for some arbitrarily small > 0. As we have assumed that is an equilibrium, State 2 must react optimally to this deviation, accepting if

$$(1 - \ )(1 - x^N - \ ) + \ [rw_2^0 + (1 - r)w_2] > w_2:$$

Rearranging this, we find that State 2 must accept after such a deviation if

$$x^N - \ < 1 - w_2 - \ r(w_2 - w_2^0)=(1 - \ );$$

which holds by definition of x^N. This deviation is profitable for State 1 if

$$w_1 - (1 - \ )a < (1 - \ )(x^N - \ - a) + \ [r(1 - w_2^0) + (1 - r)w_1 - (1 - \ )a];$$

which holds as long as

$$a < (1 - \ )(1 - w_2 - w_1 - \ ) + \ r(1 - w_2 - w_1):$$

Recall that we have assumed that $a < 1 - w_2 - w_1$, and that the limit of the right hand side of the above inequality as $\ \to 0$ is equal to $1 - w_2 - w_1$. Therefore, by continuity, there exists sufficiently small $\ > 0$ such that the above inequality is satisfied. As this deviation is profitable, State 1 cannot be using $x^t > x^N$ under . This completes the proof.

□

Proposition 1. *If $a \leq r(w_2 - w_2^0)$=$(1 - \ )$ and $w_2^0 \geq w_2 - (1 - \ )(1 - w_2)$= $\ r$ then the following strategy pro le constitutes the unique equilibrium:*

- *State 1 joins if State 3 extends. If $s^t = N$ and an alliance has not been successfully implemented in the current period, State 1 o ers $x^t = 1 - w_2 - \ r(w_2 - w_2^0)$=$(1 - \ ) \equiv x^N$. Otherwise, State 1 o ers $x^t = 1 - w_2^0$.*
- *If $s^t = N$ and an alliance has not been implemented, State 2 accepts any $x^t \leq 1 - w_2 - \ r(w_2 - w_2^0)$=$(1 - \ )$ and rejects otherwise. If $s^t = A$ or if an alliance was implemented in the current period, State 2 accepts any $x^t \leq 1 - w_2^0$ and rejects otherwise.*
- *State 3 extends an alliance in every period in which $s^t = N$.*

Proof. Uniqueness of strategies is implied by lemmas 1, 2, and 3.

Next, we proceed to show that the strategy profile is an equilibrium by demonstrating that no player has a profitable one-shot deviation.

To begin, consider State 3's choice to extend an alliance in every period in which $s^t = N$. To do this, we first compute State 3's continuation value of extending under . First, note that under , with probability $1 - r$, the alliance is not implemented after 3 extends and 1 joins. If this occurs, then State 3's continuation payoff once implementation fails, denoted

V^3, is given by

$$V^3 = (1 - \ )(1 - w_2 - \ r(w_2 - w_2')/(1 - \ )) + \ [r(1 - w_2' - a(1 - \ )) + (1 - r)V^3;$$

which implies

$$V^3 = 1 - w_2 - a\frac{(1 - \ )(1 + \ r)}{1 - \ + \ r}.$$

Continuing to form 3's continuation payoff of extending under , note that with probability r the alliance is implemented after State 1 joins, leading to a continuation payoff of $1 - w_2' - a(1 - \ )$. Pulling these together, the continuation value of extending under   for player 3 is

$$(1 - r) \ \ 1 - w_2 - a\frac{(1 - \ )(1 + \ r)}{1 - \ + \ r} \ \ + r(1 - w_2' - a(1 - \ ));$$

which simplifies to

$$1 - w_2 + r(w_2 - w_2') - \frac{a(1 - \ )}{1 - \ + \ r}.$$

With this, under , a deviation to not extending is not profitable if

$$1 - w_2 + r(w_2 - w_2') - \frac{a(1 - \ )}{1 - \ + \ r} \geq (1 - \ ) \Big[1 - w_2 - \frac{r(w_2 - w_2')}{1 - \ } \Big] + \ \Big[1 - w_2 + r(w_2 - w_2') - \frac{a(1 - \ )}{1 - \ + \ r} \Big].$$

This inequality is satisfied if

$$\frac{1 - \ + \ r}{1 - \ } \quad \frac{a(w_2 - w_2')}{1 - \ } \ \ ;$$

which holds as we have assumed that $a \leq r(w_2 - w_2')/(1 - \ )$. Therefore, State 3 cannot profitably deviate to not extend. Next, note that the same argument above establishes that State 1 cannot profitably deviate from joining after 3 extends under .

Next, we show that State 1 cannot profitably deviate from offering $x^t = x^N$ in a period in which $s^t = N$ and an alliance was not successfully implemented in the current period. First, consider a deviation to some $x' > x^N$. In response to such a deviation, State 2 will reject. Comparing this to the continuation value of offering x^N derived above, we find that this deviation is not profitable if

$$1 - w_2 + r(w_2 - w_2') - \frac{a(1 - \ )}{1 - \ + \ r}: \geq w_1;$$

which holds if

$$a \leq \frac{1 - \ + \ r}{1 - } (1 - w_2 - w_1 + r(w_2 - w_2')):$$

The above inequality is satisfied as we have assumed that $a < 1 - w_2 - w_1$. Therefore, 1 cannot profitably deviate to offer some x^N.

Next, consider a deviation by 1 to offer some $x' > x^N$. Denoting the value of the game beginning in an arbitrary period under   in which $s^t = N$ as V^1, such a deviation is not profitable as long as

$$(1 - \ )(x' - a(1 - \ ) + \ V^1 \leq (1 - \ )(x^N - a(1 - \ ) + \ V^1;$$

which holds as $x' < x^N$ and $\ \in (0;1)$.

Finally, note that by construction of State 1's offer, State 2 weakly prefers to accept x^N in a period in which $s^t = N$ and an alliance has not been implemented. Similarly, in a period in which $s^t = A$ or an alliance has been implemented, State 2's payoff of rejecting is w_2', and so State 2 cannot profitably deviate from the proposed strategy in response to an offer by State 1.

□

Lemma 4. *If $a \leq r(w_2 - w_2')=(1 - \ )$ and $w_2' < w_2 - (1 - \ )(1 - w_2)= \ r$, then in every equilibrium, in a period in which $s^t = N$ and an alliance was not implemented in the current period, State 2 rejects all o ers x^t.*

Proof. We proceed by proving that there does not exist any offer x^t that will be accepted in equilibrium under the conditions indicated in the proposition.

For a contradiction, suppose that   is an equilibrium in which State 2 accepts x^t on the path of play in a period in which $s^t = N$ and an alliance has not been implemented in the current period. Let V^2 denote State 2's continuation value of accepting this offer x^t under   in such a period. As in the proof of lemma 3, forming this continuation value and manipulating it reveals that

$$V^2 = \frac{(1- \ )(1-x^t) + \ r w_2'}{1- \ + \ r}.$$

State 2 cannot profitably deviate to rejecting this offer x^t as long as $V^2 \geq w_2$, which implies

$$x^t \leq 1 - w_2 - \ r(w_2 - w_2')/(1- \ ).$$

Note that

$$w_2' < w_2 - (1- \ )(1-w_2)/ \ r \implies 1 - w_2 - \ r(w_2 - w_2')/(1- \ ) < 0.$$

Because $x^t \in [0,1]$, this contradicts inequality A above. Therefore, there does not exist a value x^t such that State 2 is willing to accept it along the equilibrium path of play. □

Proposition 2. *If $a \leq r(w_2 - w_2')/(1- \ )$ and $w_2' < w_2 - (1- \ )(1-w_2)/ \ r$ then all equilibria are equivalent in outcome distribution to the equilibrium in which players use the following strategies:*

- *State 1 joins an alliance after State 3 extends in every period in which $s^t = N$. State 1 o ers $x^t = 0$ in every period in which $s^t = N$ and an alliance has not been implemented. Otherwise, State 1 o ers $x^t = 1 - w_2'$.*
- *If $s^t = N$ and an alliance was not implemented in the current period, State 2 rejects all o ers x^t. Otherwise, State 2 accepts x^t if and only if $x^t \leq 1 - w_2'$.*
- *State 3 extends an alliance in every period in which $s^t = N$.*

Proof. To begin, note that lemma 1 implies that, under the conditions of the proposition, in every equilibrium, in period 1 State 3 extends and State 1 joins, and with probability r and alliance is implemented and play proceeds according to lemma 2, and with probability $1 - r$ State 2 rejects any offer and the game concludes with war in period 1. This establishes that, under the parameter values indicated in the proposition, every equilibrium is equivalent in outcome distribution. Finally, note that the equilibrium strategies outlined in the proposition produce precisely this outcome distribution.

Next, we establish that this strategy profile is indeed an equilibrium by demonstrating that no player has a profitable one-shot deviation.

First consider a deviation by State 3 to not extend an alliance in a period in which $s^t = N$. Such a deviation is not profitable if

$$w_3 \leq r(1 - w_2^{\ell}) + (1 - r)w_3 - a(1 - \quad);$$

which holds if

$$a \leq \frac{r(1 - w_2^{\ell} - w_3)}{1 - \quad}:$$

This holds, as we have assumed that $a \leq r(w_2 - w_2^{\ell})=(1 - \quad)$ and

$$a \leq \frac{a(w_2 - w_2^{\ell})}{1 - \quad} < \frac{r(1 - w_2^{\ell} - w_3)}{1 - \quad}:$$

Therefore, State 3 cannot profitably deviate from extending an alliance in such a period. Next, we demonstrate that State 1 cannot profitably deviate from joining after State 3 extends in a period in which $s^t = N$. Such a deviation is not profitable if

$$w_1 \leq r(1 - w_2^{\ell}) + (1 - r)w_1 - a(1 - \quad);$$

which holds if

$$a \leq \frac{r(1 - w_2^{\ell} - w_1)}{1 - \quad}:$$

This holds, as we have assumed that $a \leq r(w_2 - w_2^0)=(1 - \ )$ and

$$a \leq \frac{a(w_2 - w_2^0)}{1 - \ } < \frac{r(1 - w_2^0 - w_1)}{1 - \ }:$$

Next, consider a deviation by State 1 to offer some $x^t \neq 0$ in a period in which $s^t = N$ and an alliance was not implemented in the current period. Given State 2's strategy, any such offer will be rejected, yielding a payoff of $w_1 - a(1 - \ )$. This is equivalent to the payoff of offering $x^t = 0$ in such a period, therefore no such deviation is profitable.

Next, consider a deviation from State 2's acceptance strategy in response to an offer x^t in a period in which $s^t = N$ and an alliance has not been implemented in the current period. Such a deviation is profitable if $(1 - \ )(1 - x^t) + \ [rw_2^0 + (1 - r)w_2]$. Note that if this inequality holds for $x^t = 0$, it will also hold for all $x^t \in (0; 1]$. Substituting $x^t = 0$ and rearranging reveals that the inequality is satisfied if

$$w_2^0 \leq w_2 - \frac{(1 - \ )(w_2 - 1)}{r};$$

which holds as we have assumed that $w_2^0 \leq w_2 - (1 - \ )(1 - w_2)= \ r$, and

$$w_2^0 \leq w_2 - \frac{(1 - \ )(1 - w_2)}{r} < w_2 - \frac{(1 - \ )(w_2 - 1)}{r}:$$

Therefore, State 2 cannot profitably deviate from rejecting all offers x^t in a period in which $s^t = N$ and an alliance was not implemented in the current period.

Finally, the same argument employed in the proof of proposition 1 establishes that no player has a profitable deviation from the proposed strategies in a period in which an alliance has been successfully implemented or $s^t = A$. □

Proposition 4. *Suppose that $a \leq r(w_2 - w_2^0)=(1 - \ )$. The equilibrium probability of war is nonmonotonic in r. In particular,*

1. *If $r \leq (1 - \ )(1 - w_2)= \ (w_2 - w_2^0) \equiv r^w$, the equilibrium probability of war is 0.*

2. *If $r > r^w$, the equilibrium probability of war is $1 - r$.*

Proof. Follows from the equilibrium strategies outlined in Propositions 1 and 2. □

Proposition 5. *The equilibrium utilities of State 1 and State 3 are:*

- *nonmonotonic in $w_2^\emptyset$*
- *nonmonotonic in r*

Proof. We prove each component of the proposition in turn. First, note that the highest value of $w_2^\emptyset$ for which peace occurs with probability 1 in equilibrium is $w_2^\emptyset = w_2 - (1- \quad)(1-w_2) = \quad r \equiv \overline{w}_2$. Note that for $w_2^\emptyset > \overline{w}_2$, the equilibrium utility of States 1 and 3 is decreasing in $w_2^\emptyset$. This follows from the proof of Proposition 1, which demonstrated that the equilibrium utilities of State 1 and State 3 under these conditions are equal to

$$1 - w_2 + r(w_2 - w_2^\emptyset) - \frac{a(1- \quad)}{1- \quad + \quad r};$$

which is decreasing in $w_2^\emptyset$.

Next, we show that, for sufficiently small , State 1 and State 3's equilibrium utilities are higher under $w_2^\emptyset = \overline{w}_2$ than under $w_2^\emptyset = \overline{w}_2 - \quad$. This is true for State 1 if

$$r(1 - (\overline{w}_2 - \quad)) + (1-r)w_1 - a(1- \quad) < 1 - w_2 + r(w_2 - \overline{w}_2) - \frac{a(1- \quad)}{1- \quad + \quad r}:$$

Algebra yields that this inequality holds if

$$< \frac{r}{1-r}\left[1 - w_1 - w_2 - \frac{a(1- \quad)}{1- \quad + \quad r}\right]:$$

Note that our assumption that $a < 1 - w_2 - w_1$ implies that the right hand side of this inequality is strictly positive. Therefore, the inequality holds for sufficiently small $\quad > 0$, as required. A similar argument, replacing w_1 with w_3 in the previous inequality establishes the result for State 3's utility as well.

Finally, note that if $w_2^0 < \overline{w}_2$, the equilibrium utility of State 1 is

$$r(1 - w_2^0) + (1 - r)w_1 - a(1 - \ );$$

which is decreasing in w_2^0. Similarly, if $w_2^0 < \overline{w}_2$, the equilibrium utility of State 3 is

$$r(1 - w_2^0) + (1 - r)w_3 - a(1 - \ );$$

which is decreasing in w_2^0.

Finally, note that the preceding analysis implies that, for both State 1 and State 3, equilibrium utility as a function of w_2^0 is decreasing for $w_2^0 < \overline{w}_2$. Then, equilibrium utility for both State 1 and State 3 jumps up discontinuously at $w_2^0 = \overline{w}_2$, and decreases for $w_2^0 > \overline{w}_2$. Therefore, the equilibrium utility of both State 1 and State 3 is nonmonotonic in w_2^0.

To prove the second component of the proposition, note that the highest value of r for which peace occurs with probability 1 in equilibrium is $r = (1 - \ )(1 - w_2) = \ (w_2 - w_2^0) \equiv r^w$. Note that for $r > r^w$, the equilibrium utility of States 1 and 3 is increasing in r. This follows from the equilibrium strategies outlined in Proposition 2.

Next, we show that for sufficiently small , State 1 and State 3's equilibrium utilities are lower under $r = r^w + \ $ than under $r = r^w$. This is true for State 1 if

$$(r^w + \ )(1 - w_2^0) + (1 - r^w - \ )w_1 - a(1 - \ ) < 1 - w_2 + r(w_2 - w_2^0) - \frac{a(1 - \ )}{1 - \ + \ r}:$$

Algebra yields that this inequality holds if

$$\frac{1 - w_2^0 - w_1}{1 - r} \ < 1 - w_1 - w_2 - \frac{a(1 - \ )}{1 - \ + \ r}:$$

Note that our assumption that $a < 1 - w_2 - w_1$ implies that the right hand side of this inequality is strictly positive. Therefore, the inequality holds for sufficiently small $\ > 0$, as required. A similar argument, replacing w_1 with w_3 in the previous inequality establishes the result for State 3's utility as well.

Finally, note that if $r < r^w$, the equilibrium utilities of State 1 and 3 are

$$1 - w_2 + r(w_2 - w_2^0) - \frac{a(1 - \ )}{1 - \ + \ r};$$

which is increasing in r. □

Finally, note that the preceding analysis implies that for both State 1 and State 3, equilibrium utility as a funciton of r is increasing for $r < r^w$. Then, equilibrium utility for both State 1 and State 3 jumps discontinuously down at $r = r^w$ and again increases for $r > r^w$. Therefore, the equilibrium utility of both States 1 and 3 are nonmonotonic in r.

Corollary 1. *Let $w_2 - (1 - \ )(1 - w_2)= \ r \equiv \overline{w}_2$ be the highest value of w_2^0 for which peace occurs in equilibrium with probability 1. There exists a unique $w_P < \overline{w}_2$ such that State 1 and State 3's equilibrium utilities are both higher under $\overline{w}_2$ than under any $w_2^0 \in (w_P; \overline{w}_2)$.*

Proof. This result follows from the proof of the previous proposition combined with the knowledge that State 1 and State 3's equilibrium utilities are continuous in w_2^0 for all $w_2^0 < \overline{w}_2$). □

Corollary 2. *Let $r \leq (1 - \ )(1 - w_2)= \ (w_2 - w_2^0) \equiv r^w$ be the highest value of r for which peace occurs in equilibrium with probability 1. There exists a unique $\underline{r}$ with $r^w < \underline{r} < 1$ such that State 1 and State 3's equilibrium utilities are both higher under r^w than under any $r \in (r^w; \underline{r})$.*

Proof. This result follows immediately from the proofs of the previous two results. □

Proposition 6. *Fix some value $r < 1$. There exists an $\overline{r}$ with $r < \overline{r} < 1$ such that both State 1 and State 3's equilibrium utilities are higher under every $r^0 \geq \overline{r}$ than under r.*

Proof. This follows immediately from the equilibrium strategies described in propositions 1 and 2. □

B Political Comic Mentioned in Text

This comic ran in the July 1990 issue of *The Economist* magazine. It depicts Helmut Kohl striding through a door labeled NATO, while Gorbachev is depicted holding the door open, holding a bag labeled DMs.