

Essays Relating to the Dynamics of Power and Policy Differences  
between States in the International System

By

Ryan L. Hill

Dissertation

Submitted to the Faculty of the  
Graduate School of Vanderbilt University  
in partial fulfillment of the requirements

for the degree of

DOCTOR OF PHILOSOPHY

in

Political Science

May 10, 2024

Nashville, TN

Approved:

Brett V. Benson, Ph.D.

Brenton Kenkel, Ph.D.

Bradley C. Smith, Ph.D.

Jesse C. Johnson, Ph.D.

Copyright © 2024 by Ryan L. Hill  
All Rights Reserved

To my wife, Rachel, and my daughter, Rylee

## ACKNOWLEDGEMENTS

I am extremely grateful to my dissertation committee, Dr. Brett Benson, Dr. Brenton Kenkel, Dr. Brad Smith, and Dr. Jesse Johnson, for their insight and guidance. I would also like to express my thanks to my professors in the Vanderbilt Political Science Department and to my amazing and supportive cohort. Most importantly, I wish to thank my wife, Rachel, and my daughter, Rylee, for their unwavering support.

## Contents

<b>Introduction</b>	<b>1</b>
<b>1 Restraining Moral Hazard in Asymmetric Alliances</b>	<b>4</b>
1.1 Introduction . . . . .	4
1.2 Literature Review . . . . .	8
1.3 Model Setup . . . . .	14
1.3.1 Key Features of the Model . . . . .	17
1.4 Equilibrium Results . . . . .	19
1.5 Discussion . . . . .	30
1.5.1 Case Study: US-ROK Alliance . . . . .	30
1.5.2 Private Information Extension . . . . .	39
1.6 Conclusion . . . . .	41
<b>2 Conceptualizing and Measuring Polar Tension in the International System</b>	<b>44</b>
2.1 Introduction . . . . .	44
2.2 Polarity and Polarization as Structural Aspects of the International System .	46
2.3 The Concept of Polar Tension . . . . .	52
2.4 Operationalizing and Measuring Polar Tension . . . . .	55
2.4.1 Defining the Actors . . . . .	55
2.4.2 Measuring Power and Parity . . . . .	58
2.4.3 Measuring Alignment . . . . .	64
2.4.4 Measuring Polar Tension . . . . .	71
2.5 Assessing Polar Tension . . . . .	74
2.6 Conclusion . . . . .	76
<b>3 The Effects of Polar Tension on Policy Alignment in Asymmetric Alliances</b>	<b>79</b>
3.1 Introduction . . . . .	79
3.2 Literature Review and Theoretical Underpinnings . . . . .	81
3.3 Formal Model . . . . .	86
3.4 Prioritization in Multilateral and Bilateral Alliances . . . . .	88
3.5 Research Design . . . . .	91
3.6 Results . . . . .	98
3.6.1 Testing the Mechanism . . . . .	100
3.7 Conclusion . . . . .	103
<b>4 References</b>	<b>105</b>
<b>5 Appendix A</b>	<b>116</b>

## List of Tables

1	Power Measures as Predictors of War and Dispute Outcomes, 1816-2010 (percentage of outcomes predicted correctly) . . . . .	61
2	Asymmetric Alliance Data Summary: 1946-2015 . . . . .	96
3	The Effects of Polar Tension on Policy Alignment of Asymmetric Alliances .	99
4	Bilateral versus Multilateral Asymmetric Alliance Agreements: 1946-2016 . .	102
5	Transitioned Asymmetric Alliance Agreements: 1946-2016 . . . . .	102
6	The Effects of Polar Tension on Policy Alignment of Asymmetric Alliances .	117

## List of Figures

1	Game Tree . . . . .	16
2	Relationship between the ‘alignment gap’ and the minimum deductible. . . .	21
3	Ally utility when $\underline{k} \leq k^* \leq \bar{k}$ . . . . .	25
4	Ally utility when $\bar{k}^* < \underline{k}$ . . . . .	27
5	Ally utility when $\bar{k} < k^*$ . . . . .	28
6	Ally utility when $\bar{k} < \underline{k} < k^*$ . . . . .	29
7	The Vertical and Horizontal Dimensions of Polar Tension . . . . .	55
8	CINC Index and Beckley Index as Measures of the Power of the Major Powers in the International System . . . . .	62
9	Power using the Hybrid (milexp, milper, and GDP x GDPPC) measure . . . .	63
10	Parity between the Major Powers (Greatest Power [US] = 1.00) . . . . .	65
11	UN voting similarity between the USSR/Russia and the US (Bailey et al. 2017, 439) . . . . .	68
12	Ideal Points for each of the Major Powers since World War II . . . . .	69
13	Ideal Point Distance between Major Powers (Greatest Power [US] = 0) . . . .	70
14	Hypothetical Polar Tension scores comparing typical Polarity categories . . . .	72
15	Polar Tension Variation in Bipolar System based on Spatial Distance . . . . .	73
16	Polarity Components (Parity and Alignment Gap) per year . . . . .	73
17	Comparing Polar Tension in ‘Bipolarity’ (1970) and ‘Unipolarity’ (2000) . . .	74
18	Polar Tension in the International System . . . . .	75
19	Comparing Concentration Scores to Polar Tension . . . . .	76
20	Mechanisms for Polar Tension’s Impact on Alignment in Asymmetric Alliances	90

## Introduction

This dissertation consists of three papers related to the differences in power (power gap) and policy preferences (alignment gap) between states in the international system. The alignment gap represents the difference in preferred international policy outcomes while the power gap represents the neorealist perspective of relative capabilities that enable nations to pursue their interests. Examining these dynamics between potential allies and potential enemies reveals incentives that drive individual state behaviors and helps explain the conditions that determine international system structure. The chapters herein explore these dynamics and generate findings that contribute to international relations literature on asymmetric alliances, international system structure, and the impact of system structure on policy alignment within asymmetric alliances. It also addresses potential international policy issues, such as alliance commitments and agreements.

The first chapter addresses how the power gap and alignment gap create potential risks for the stronger ally in an asymmetric alliance and proposes how the ally may incentivize its protégé to act with greater restraint. In an asymmetric alliance, where the stronger ally provides a security guarantee and the weaker ally agrees to surrender autonomy in the way of policy concessions, the weaker state is inherently incentivized to act with greater boldness in pursuing its own interests, creating a potential moral hazard problem. This paper examines how the stronger ally can manage such an alliance to restrain its protégé's behavior. I propose that the stronger ally can induce a cost conditional on conflict initiation, comparable to a deductible in insurance terms. I use a formal model to demonstrate that the stronger ally, by limiting its initial protective measures, can incentivize the protégé to act with restraint. Therefore, the stronger state reduces the risk of moral hazard by avoiding over-insuring its protégé. The model also suggests that as the states grow in alignment regarding policy preferences, the stronger ally can lower the deductible and provide greater protection, enhancing the protégé's bargaining power and deterring the challenger from making excessive demands.

I use the alliance between the US and South Korea as a case study to demonstrate how the theory explains changes in alliance behaviors over time. This theory provides a novel mechanism for restraining moral hazard and demonstrates how a stronger power can vary the protection of its ally to pursue its own policy objectives.

Chapter 2 examines how a narrow power gap (power parity) and a large alignment gap between major powers in the international system can define the system structure. The term polarity is predominately used to describe a single dimension of the international system, the distribution of power, with no regard for the second dimension, the degree of opposition between the major powers, limiting its theoretical utility. I propose a new measure that captures both the power distribution and degree of opposition between the world's major powers to define the international system structure. The measure, "polar tension", is defined as the degree to which international policy interests are contested. I provide face validity tests to demonstrate the contribution of both power parity and an alignment gap to overall polar tension and to show its superiority to a measure of power concentration. The measure allows for a more nuanced and continuous assessment of international system structure and provides a new mechanism to explain how the structure may affect state behavior.

Chapter 3 examines the impact of polar tension on the international policy preference alignment between nations in asymmetric alliances. While the stronger protectorate state in these alliances has broad and strategic interests, weaker protégé states are likely to have particular interests that are localized and proximate. When polar tension increases, it represents a threat to the general interests of the major powers and increases the relevance of the alliance to its weaker partners, who respond by increasing their alignment. This response is moderated by the degree to which the protégé state prioritizes the shared general interests of its major power ally. The greater the priority it places on the general interests, the more relevant the alliance becomes when changes in the international system structure threaten those interests. High and low prioritization levels also correspond with multilateral and bilateral alliances, respectively, which means that protégés in multilateral alliances are more

likely to align with the major power in times of increased polar tension. I use OLS regression on alliance data from 1946 to 2015 to test these hypotheses and find that as polar tension rises, protégés increase alignment. I also find that the degree of alignment is moderated by whether the protégé is in a multilateral or bilateral alliance with the major power.

The dynamics associated with the power and alignment gap between nations drive behaviors at the state-to-state and system levels. By assessing these forces, this dissertation contributes to the rich literature on asymmetric alliances and international system structure.

# 1 Restraining Moral Hazard in Asymmetric Alliances

## 1.1 Introduction

Throughout the first thirty-five years of the US alliance with the Republic of Korea (ROK), the US maintained a relatively low number of troops on the Korean peninsula, limited the range and payload of South Korean rockets, and did little to protect its densely populated capital city, Seoul—all measures that would make the initial stages of a potential conflict with its revisionist enemy, the Democratic People’s Republic of Korea (DPRK), extremely costly. This raises the question: Why might a strong state in an asymmetric alliance intentionally keep its ally vulnerable? While such a strategy seems counterintuitive, this paper demonstrates that by employing such an approach, the stronger state, or “ally,” can control the incentives of its weaker partner, or the “protégé,” to act in accordance with its own international policy preferences.

The need for such a strategy arises when the protégé is incentivized to act with greater risk than the ally would prefer, a phenomenon known as *moral hazard*. Two incentives to act with such risk are inherent in asymmetric alliances. The first is a *desire* for the protégé to pursue its own interests. Typically, when states ally, they both exchange a degree of sovereignty for the security benefits they can realize by combining or coordinating their defenses with the other state. However, in asymmetric alliances, the stronger ally provides the predominance of military power while its weaker protégé provides very little. The protégé, in turn, forfeits a potentially large amount of its sovereignty to the ally (Morrow 1994). Thus, the ally receives policy concessions from its partner that allow it to further its own international interests, and the protégé receives a security guarantee. The more divergence there is between the interests of the states (i.e. the greater the “alignment gap”), the more the protégé will be incentivized to seek its own policy goals.

The second is the protégé’s perception of the added *potential* the alliance gives it to achieve its interests, given the support of the ally. The greater the ally’s contribution to the

probability of victory (i.e. the greater the “power gap”), the more confidence the protégé has that its policy goals can be achieved. The relationship can be mutually beneficial; however, such a relationship may also introduce incentives for the protégé to act in a way that creates a moral hazard. The increased *desire* and *potential* to attain its interests motivate the protégé in an asymmetric alliance to rationally act in ways that put its stronger ally partner’s interests at risk.

Given that the nature of asymmetric alliances creates the incentives for a protégé to act with risk, creating a moral hazard problem, the ally must leverage other incentives to manage the alliance relationship and mitigate this risk. To this end, there are two additional risk-incentivizing factors that the ally may manipulate. First, the ally may vary its *assurance* (i.e. the clarity and credibility of the signal that the ally sends the protégé regarding its intent to intervene in a conflict). The second measure the ally may take is to vary the amount of *insurance* it extends to the protégé. This is the portion of war costs the protégé perceives the ally will internalize in the event of a conflict. If the protégé believes that the ally will indeed join the fight (assuring the protégé) and that it will bear the majority of the costs for its actions (over-insuring the protégé), then it may be further incentivized to take unacceptable risks. On the other hand, if the protégé is uncertain about its ally’s reliability, or if it perceives it will bear a large cost on its own, then it may be disincentivized to behave with risk.

Extant literature offers only one explanation for how a state can successfully manage these incentives to reduce the risk of moral hazard. Benson (2012) demonstrated that the ally may be ambiguous, or probabilistic, about its commitment in order to restrain its protégé. While such a method may be effective in reducing moral hazard, ambiguity may also reduce the deterrent effect of the alliance, a prospect that may be strategically unjustified in some situations. In such a case, an ally may choose to credibly assure its ally, but vary the level of insurance it provides. No research has addressed using this mechanism to disincentivize a protégé from risk-taking behavior that can result from moral hazard. This paper bridges

that gap.

The theory presented in this paper posits that the stronger ally may restrain the protégé's risky behavior by redistributing the costs of conflict. That is, it may shift a portion of its up-front potential costs (e.g. troop deployment levels or defensive measures) to conditional up-front loss for its protégé. Instead of providing a full but ambiguous security guarantee, the ally can be clear about its intentions but induce a conditional cost for initiating conflict, much like charging a deductible in insurance terms. Using this strategy, the ally provides less protection during an initial attack. The deductible ensures that the protégé bears the initial burden of its risky behavior if the conflict is initiated. At the same time, the stronger ally assures its weaker partner that it will bring forces to bear which greatly enhances the probability of victory. By using a deductible, the stronger ally leaves its protégé to shoulder the majority of the costly effects of an initial attack, weakening its bargaining position and reducing its incentives to act in a way that will initiate conflict, but assures the protégé of its commitment.

Assuming that moral hazard is a potential problem for the ally, I necessarily accept the scope condition that the ally's ideal policy position lies between that of the protégé and the challenger. This position allows the ally to set the deductible to a level that both reduces the risk of over-insuring the protégé while simultaneously deterring the enemy from making excessive demands. In other words, the ally can provide a level of protection that corresponds to its own level of interest, emboldening its protégé to resist a challenger's demands that exceed those interests. So, while the deductible allows the ally to avoid over-insuring the protégé and incentivizing overly risky behavior, placing limits on the deductible also empowers the protégé to accept from the challenger only those demands to which it is willing to concede, resulting in its ideal position. This allows the ally to reap the benefits of the alliance while minimizing its risk of entrapment.

To state the theory clearly and derive precise empirical predictions, I use a formal model, examining the effects of a deductible on alliance formation and behavior. The model includes

three states, two of which are in conflict over an issue, and a third stronger state that joins in an asymmetric alliance with one of the other states. I find that when the allied states can agree upon the size of the deductible, it is an effective mechanism for not only restraining an ally from creating a moral hazard problem but also for deterring the enemy from making excessive demands. The model also suggests that when the alignment gap decreases between the allied states, the optimal deductible decreases as well. This makes intuitive sense; as the protégé becomes more aligned with the interests of its stronger ally, its desire to act on its own desires decreases, lowering the risk of moral hazard. Likewise, the ally is more comfortable granting the protégé more bargaining power and emboldening it to resist excessive demands from the challenger.

I follow the model with a case study of the United States-South Korean alliance to serve as an “existence proof” that validates the theory’s logic (Goertz 2017, Goemans & Spaniel 2016, Lorentzen, Fravel & Paine 2017). I find that the United States has consistently levied a deductible on South Korea, deploying a relatively low number of troops, limiting the ROK’s missile systems, and neglecting the protection of Seoul. However, as the alignment between the states’ interests has increased over time, the deductible has been reduced. Furthermore, this case is one in which the US has consistently and credibly signaled its intent to intervene on the ROK’s behalf if it is attacked, making the deductible mechanism a better explanation for South Korea’s restraint than the strategy of ambiguity.

This novel theory introduces a new mechanism for how stronger nations successfully manage asymmetric alliances. Using a perfect information model, this mechanism provides an alternative to the ambiguity theory, demonstrating how the risks of a moral hazard problem may be reduced in cases where the stronger state is unable or unwilling to sacrifice a level of deterrence. Additionally, it contributes to the rich literature on alliances and extends the use of insurance principles into the realm of alliance management. The theory may also have international policy implications as it proposes a viable strategy for cost redistribution within alliances and for steering ally behavior in the pursuit of policy preferences.

## 1.2 Literature Review

Alliances between states are typically thought of as mechanisms in which both states may increase their security vis-à-vis another state or alliance. Examples of such security arrangements date back centuries. In the biblical account of an alliance formed between Israel and the city of Gibeon in the thirteenth century BCE, the king of Jerusalem responded to the resulting threat by forming an alliance with four other kings in the region (English Standard Version Bible 2001, Josh: 10). Likewise, Thucydides' account of the Peloponnesian War in the fifth century BCE describes the development of the Delian League by the Athenians and the Peloponnesian League by the Spartans as measures to counter the threat that each perceived from the other (Strassler 1996).

The long-standing prevalence of alliances formed for the purpose of aggregating power has generated several to theorize that the *raison d'être* that nations ally is to balance power (Altfeld 1984, Walt 1985, Morgenthau 1973, Kaplan 1957). According to this view, alliances form because each nation within the alliance desires—perhaps requires—greater power to protect itself from the potential threat posed by aspiring hegemonic powers with revisionist ambitions. This concept implies that states ally regardless of whether their views align more with the potential ally or the potential adversary, so long as the partnership protects the status quo from disruption (McGowan and Rood 1975, Walt 1985, Morgenthau 1973).

Morrow challenges this view, maintaining that the status quo should not be viewed as a “public good” (Morrow 2000, 74). He argues that even nations and alliances that uphold the status quo do so, not out of satisfaction with it, but out of an understanding that the cost of changing it would be greater than the cost of upholding it. These “status quo states” choose to constrain their ideal points (i.e., their preferred international policy outcomes, or interests) for a number of reasons. In general, states may not find that the cost of pursuing their unconstrained interests is worth the reward. Weaker states may simply be unable to revise the status quo due to a lack of military strength, while stronger states, which have broad international interests and place a high premium on international stability, may

constrain their specific interests to protect their broader interests (Gilpin 1981). Morrow illustrates the concept of having simultaneous revisionist and status quo desires, pointing out that the US desired a peaceful status quo during the Cold War but did not attempt to restore the Soviet Union when it fell. Rather, the US desired its fall but constrained its ideal point during the Cold War due to the high price of pursuing that goal (Morrow 2000, 74).

Understanding that states within an alliance may have both status quo and revisionist interests implies that alignment within the alliance is important. Morrow (2000) asserts that when a state joins an alliance, it must surrender a portion of its autonomy, which he describes as “the degree to which it pursues desired changes in the status quo,” or the “ability to determine its own policies” (Morrow 1991, 908-909). The greater the “alignment gap” between the interests of allies, the more autonomy one or more of the nations must surrender. This is of particular concern in asymmetric alliances. In these partnerships, the weaker protégé state in the alliance has little to offer in the way of security and therefore depends on a security guarantee from the stronger ally state. In exchange, the protégé agrees to forfeit a potentially large amount of its sovereignty by constraining its own ideal positions to align with those of the ally (Morrow 1991).

While the stronger ally is likely to constrain the particular interests that it shares with the protégé to preserve its general interests in broader stability, the protégé is likely to have narrower interests and be constrained only by a lack of military strength or its alliance commitments (Snyder 1984, Rothstein 1968). If the difference between the ally’s constrained interests and the protégé’s unconstrained interests, or the *alignment gap*, is significant, then the protégé may feel overly encumbered and desire to pursue its own ideal policies. This divergence in preferences has been identified in other studies as a significant obstacle to cooperation between nations. Smith (2021) finds that differences in preferences were a primary factor in credible communication regarding the reliability of partners, while Wolford (2014) demonstrates that divergence in nations’ valuation of costs affects their willingness to join a coalition. These studies demonstrate that a significant alignment gap in policy

preferences, or interests, can induce behaviors that are antithetical to those outlined in the alliance agreement.

If the ally is significantly more powerful than the protégé, creating a large *power gap*, then it increases the possibility that, provided it is joined by the stronger ally, the protégé can successfully pursue its misaligned interests. With increased desire and with amplified potential to attain its own interests, both of which are inherent dynamics in the asymmetric alliance, the protégé may be incentivized to act with greater boldness toward challengers and risk pulling its ally into a conflict. This incentivized behavior to act with risk to the ally is known as *moral hazard*.

The concept of moral hazard comes from economics research concerned with optimizing insurance policies (Rowell & Connelly 2012). The term is often described as the lack of “care” given to avoiding loss when the insured realizes that the insurance company will bear the costs (Vaughan 1997, Pauly 1974, Shavell 1979). Pauly (1974) uses formal modeling to demonstrate that “over-insurance” creates the incentives that lead to moral hazard, showing that this surplus of costs paid by the insurer alleviates cost-averse incentives for the insured to act with care. The rational behavior of the insured, motivated by the cost coverage of the insurer, is analogous to the protégé’s incentives to behave boldly due to the security guarantee from its stronger ally. The protégé, even if it is void of any exogenous risk-acceptant predispositions (i.e., even in the absence of an adverse selection problem), may be rationally incentivized to act in ways that pose a risk to its ally (Pauly 1968, Pauly 1974, Rowell & Connelly 2012). Therefore, despite the exogenous incentives of *desire* and *potential* that are inherent in asymmetric alliances, the ally has two factors that it can manipulate to reduce the incentives that embolden the protégé: the *assurance* that the ally is committed, and *over-insurance* that the ally will shoulder an excessive amount of the cost of war.

The first of these protégé emboldening factors, assurance, is the degree to which a partner can be depended upon to uphold its alliance commitments. Benson (2012) demonstrates how a stronger ally can reduce the protégé’s incentives by being ambiguous about assurances in

the contract design. He explains that if a contract between allies is unambiguous, meaning that the obligation of support is unconditional in its trigger and complete in its support, the ally is more likely to intervene in a time of crisis. This reliability creates an assurance that promotes moral hazard. On the other hand, if the contract is ambiguous as to the conditions that invoke the treaty and probabilistic about the amount of obligated support, the ally has more flexibility to stay out of the conflict and the protégé, being less confident that the ally will join it, acts with greater restraint.

While contract ambiguity can be an effective method of restraining a protégé, it has some weaknesses. The lack of clarity that is inherent in such a contract has the potential to bring reliability into question and therefore diminish the deterrent effects of the alliance (Fearon 1997, Morrow 2000, Mattes 2012). Smith (1995) finds that a potential aggressor is more likely to attack an alliance that appears less reliable than one it believes to be strong. Though not specifically pertaining to alliances, Sartori (2005) has a similar finding. She demonstrates that a state with credible deterrence is less likely to even receive threatening demands from a challenger. This implies that sending a strong signal of reliability to the protégé serves two purposes for the ally. First, it assures the protégé such that it is not willing to concede too much to an enemy, and second, it demonstrates resolve to potential challengers, preventing them from not only attacking but from demanding more than the alliance signals it is willing to concede. Providing assurance creates a predicament that Snyder (1984) sees as part of the “alliance security dilemma” (Snyder 1984, 462-463). On one hand, reliability can promote the conflict-inducing problem of moral hazard. On the other hand, if a partner is known to be reliable, they present an increased deterrent effect on an enemy. So, while muddying the signal of reliability can incentivize a protégé to act with greater restraint, it may also reduce the deterrent effect, inviting potential challengers to make excessive demands.

Confidence in the reliability of a partner is a necessary condition for a weaker ally to act with risk, creating a moral hazard problem; however, assurance does not have to be

diminished in order to reduce the likelihood of moral hazard. Instead, the stronger ally can focus on reducing the impact of the second protégé emboldening factor: *over-insurance*. This term indicates that the costs are disproportionately distributed, and the stronger ally is carrying more than required to meet its objectives. Fang, Johnson, and Leeds (2014) use a formal model to demonstrate that the cost distribution, measured relative to the partners' respective values of the issue at stake, motivates each partner's willingness to maintain the alliance. The similarity or dissimilarity of these values is comparable to the degree of alignment between the states, which affects the willingness of each to pay the costs associated with attaining its objective. The authors moderated this willingness by ascribing a value to maintaining the alliance such that if a state highly values the alliance, it might be willing to pay a higher price than its valuation of the issue warranted.

Another perspective on cost distribution within military partnerships has been presented by Wolford (2014). He demonstrates that a leading coalition member may have to lower the costs incurred by its partners in order to keep them in the coalition. In this way, the coalition is held together because the price that each attributes to the issue is proportionate to the cost each pays.

When costs are distributed such that the ally covers the costs that the protégé is not willing to pay, then the protégé is over-insured, altering its valuation for action. The proposed solution of redistributing costs to control the risks of moral hazard due to over-insurance is prevalent in economics literature. Shavell (1979) employs a formal model exploring moral hazard and found that the optimal choice of insurance is to provide incomplete coverage. The findings from Pauly's (1974) formal analysis agree; he assessed that the optimal solution to moral hazard situations is "for the insured to retain some part of his losses" (Pauly 1974, 45). Benson, Meiorowitz, and Ramsay (2014) capitalize on this idea in dealing with moral hazard in alliances but focus on deterring the aggressor rather than on restraining the protégé. They use a three-player game and show that in cases of asymmetric alliances, the stronger ally, by accepting more indemnity (i.e. providing more coverage), can motivate the protégé to

behave in a way that creates greater moral hazard, therefore deterring would-be aggressors. This illustrates how, by redistributing the costs, the stronger state controls the incentivizes of its weaker protégé.

While altering the amount of coverage is an effective way of redistributing the costs and modifying a partner's level of aggressive behavior, economists have found that simply providing partial coverage is most effective in cases where the principal is interested in reducing losses when incidents occur (Lee, Lee, & Hong 2021). This insurance is ideal for cases in which the incidents cannot be prevented, such as natural disasters, when the insured can take measures to minimize loss if the incident does occur. If an insurer is interested in reducing the probability of the incident itself, on the other hand, the most effective means is to employ a deductible (Arrow 1971, Lee, Lee, & Hong 2021). Rees and Wambach (2008) explore the difference between loss-prevention and loss-reduction and find that "If the agent can only influence the probability of an accident, then the optimal insurance contract has a deductible" (Rees & Wambach 2008, 130). Introducing a deductible places the conditional losses up-front, motivating the insured to avoid an incident altogether. Whereas if the stronger ally were to simply limit the overall coverage, then the losses would be distributed over the duration of the conflict and the insured would be more inclined to act in ways that reduce loss rather than taking measures to prevent it. This assessment confirms Arrow's (1971) theorem, which states that a deductible is the optimal form of insurance against moral hazard. Additionally, it is welfare improving to both parties in that it incentivizes the insured to reduce the risk of an incident occurring, and it provides the insured with greater protection against larger degrees of loss (Pauly 1968, Zeckhauser 1970). In terms of alliances, a deductible incentivizes the protégé to act in ways that avoid initiating a conflict while preserving their confidence that the ally will honor its security guarantee.

My theory applies this insurance mechanism to alliances by thinking about the application of a deductible primarily as a shifting of costs from the ally to the protégé. More specifically, it is a shift in the potential up-front cost of the ally to a conditional up-front loss of the

protégé. An example of this would be if an ally, rather than using its own forces right away, purposefully delays its military response to ensure that the brunt of an initial attack is absorbed by the weaker ally. Another way the ally could apply the deductible is by not only neglecting to defend certain of the protégé’s potential targets but hindering the protégé’s ability to protect them as well. If the major power ally were to allow the protégé to defend these assets, then the transferred cost is not conditional. By preventing them from shoring up defenses in certain areas, the ally is ensuring that if war is to occur, the protégé pays the conditional costs via the losses it incurs.

Though we conceptually think about the deductible as an alliance agreement, in practical terms, it may be continually adjusted so long as both parties agree. In the following model, I refer to the agreement over the deductible as part of the alliance formation, but it is to be understood, not only as part of the initial agreement, but also as a continual updating and maintenance of the alliance. The malleability of the agreement is an important feature; not only can the deductible be adjusted to keep the protégé from acting with risk, incentivizing it to concede to enemy demands, but it can also be reduced, such that the protégé does not concede more than the stronger partner desires. It may be adjusted as the dynamics of the relationship change. Unlike using contract ambiguity to restrain a partner, employing an optimal deductible allows the ally to reap the benefits of the alliance while preserving its ideal international policy positions. The following section formalizes this theory.

### 1.3 Model Setup

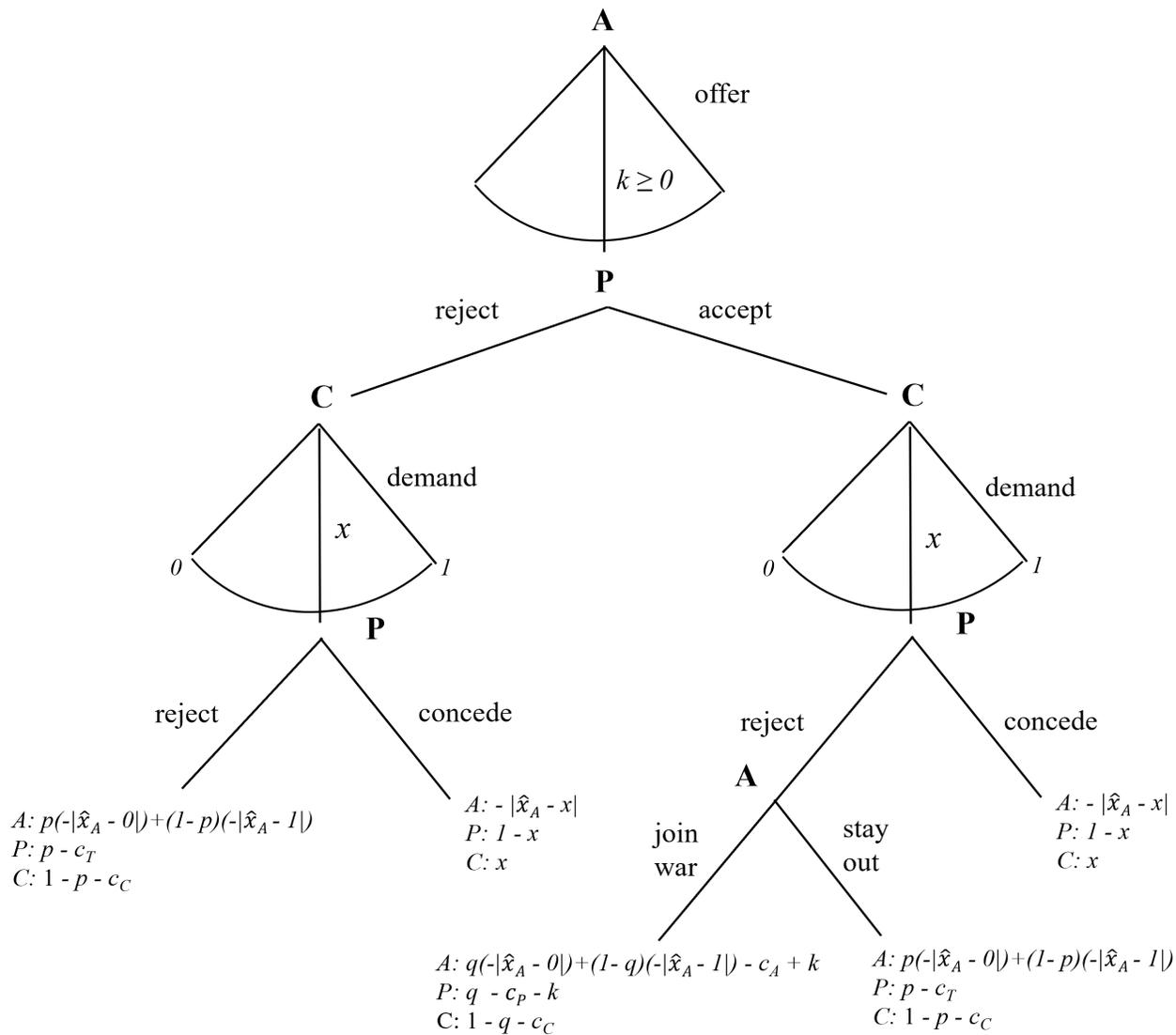
To formalize the theory, I use a three-player sequential game including ultimatum bargaining and perfect information. Consider a world in which an asymmetric alliance may be formed between a powerful Ally, A, and its Protégé, P. A Challenger, C, is in a dispute with the Protégé over an international issue,  $x$ , where  $x \in [0, 1]$ . C has an ideal point normalized to 1,  $\hat{x}_C = 1$ , while P has an unconstrained ideal point of 0,  $\hat{x}_P = 0$ . State A has a single-peaked utility with an interior ideal point at,  $\hat{x}_A \in [0, 1]$ .

Figure 1 diagrams the game tree, displaying the sequence of play and potential payoffs for each player. The game begins with the Ally and the Protégé negotiating the terms and structure of the alliance partnership, in which an alliance forms if A and P agree on the price that P must pay, conditional on the initiation of conflict. This price becomes the deductible,  $k$ , which may be explicitly stated or implied by the structural details of A's posturing and response plan. The deductible is a shifting of costs from A to P in the form of a vulnerability that A could have potentially defended but chose not to. The measures that constitute the deductible are intended to ensure that if a war is initiated, P suffers  $k \geq 0$  cost before A enters the conflict in force or begins sharing a larger portion of the cost.

If the Protégé rejects the Ally's offer of  $k$ , then the alliance does not form (or breaks down) and P must face the demands of the Challenger without any commitment from A to assist. This scenario is displayed on the left-hand side of the overall game tree in Figure 1. If P does accept the offer,  $k$ , as depicted on the right-hand side of Figure 1, then P and A form (or maintain) an alliance in which A commits to provide military assistance to P if it is attacked by C.

In the next step of the game, C makes an ultimatum demand,  $x \in [0, 1]$ , of P, which P can concede or reject. This demand is made in the absence or presence of an alliance, as displayed on the left-hand-side and right-hand-side of Figure 1, respectively. In either case, if P concedes, the game ends and the players receive payoffs. The policy is set to the accepted value  $x$ , which results in a payoff of  $x$  for C and  $1 - x$  for P. A receives a single-peaked settlement payoff of  $-|\hat{x}_A - x|$ , such that it is dissatisfied with a shift from  $\hat{x}_A$  in either direction.

If P rejects the demand, then conflict is initiated. In the absence of an alliance, each state, P and C, receives its war payoffs (displayed in Figure 1) and the game ends. In the presence of the alliance, when conflict begins, A chooses to stay out of the fight or to join in. If A stays out, P wins the war with probability  $p \in [0, 1]$ , and if A joins the fight, then the alliance (P and A together) wins the war with probability  $q \in [0, 1]$ . We can assume



**Figure 1.** Game Tree

that the aggregate military capabilities of the alliance make the probability of the alliance winning together greater than if P fights alone, so we assume  $q > p$ .

War is modeled as a costly lottery where the winner imposes its ideal point as the policy outcome. In the case that A does not join P in fighting, either because the alliance was not formed or because A chose not to fight, A is forced to accept the difference between P's ideal point,  $\hat{x}_p = 0$ , and its own,  $\hat{x}_a$ , with probability  $p$ , or to accept the difference between C's ideal point,  $\hat{x}_C = 1$ , and  $\hat{x}_A$  with probability  $1 - p$ . Since A does not fight in these cases, it does not pay a war cost, so its expected utility for war is  $p(-|\hat{x}_A - 0|) + (1 - p)(-|\hat{x}_A - 1|)$ . P, with no help from A, pays the *total* cost of fighting,  $-c_T$ , making its war payoff  $p - c_T$ . C pays a war cost of  $-c_C$ , receiving a war payoff of  $1 - p - c_C$ . I assume that in the case of indifference between fighting outside the alliance and fighting inside the alliance when A does not join, P chooses to stay out of the alliance.<sup>1</sup>

When fighting together in an alliance, A pays a cost for fighting,  $-c_A$ , but it shifts a portion of its costs back to P via the deductible, so its costs are  $-c_A + k$ . Therefore, A's payoff is  $q(-|\hat{x}_A - 0|) + (1 - q)(-|\hat{x}_A - 1|) - c_A + k$ . P's cost for fighting is  $-c_P$ , which is less cost than  $-c_T$ . This difference represents P's anticipation that it will not internalize all the costs of war and that A will cover a portion of its costs. However, since they are fighting together and A has induced a deductible, P's cost for fighting when joined by A is  $-c_P - k$ . This makes P's payoff  $q - c_P - k$ . C's war payoff in this case is  $1 - q - c_C$ .

### 1.3.1 Key Features of the Model

Before examining the results, there are four key features of the model worth considering in more detail. First, my treatment of the spatial preferences of the allies is similar to that of Smith (2021). I assume that A's ideal policy is less extreme than P's and that its preference is interior to the preferences of P and C, which are 0 and 1, respectively. Furthermore, the

---

<sup>1</sup>I make this assumption because it would emerge as equilibrium behavior if P were to incur any costs from an alliance breakdown in wartime. This is plausible if A's defection creates additional vulnerabilities in its abandoned areas of responsibility, causing P to incur greater costs due to the alliance breakdown than if the alliance had not formed.

single-peaked utility for A's policy preference opens the possibility that marginal changes in the settlement value benefit one ally's interests while harming the other's; in the case that the policy,  $x$ , moves away from A's preference,  $\hat{x}_A$ , and toward P's preference of 0, there is a loss of utility for A and a gain for P. These assumptions create the potential disparity between the preferences of each party and the increased desire to pursue independent goals, which accounts for A's desire to restrain P's behavior.

Second, I make the plausible assumption that if A joins P in fighting against C, then A will accept P's ideal position ex-post. This assumption follows Morrow's (2000) logic that even a strong state may desire a different ideal point than the status quo but be compelled to constrain that desire in order to maintain peace and avoid the costly lottery of war. Therefore, in this model, I assume that A's constrained ideal position,  $\hat{x}_A$ , is based on its value for maintaining its general interests of a peaceful status quo stability of the international environment, but that the decision to go to war and incur war costs would likely change A's valuation of the constrained ideal point. Furthermore, it is a reasonable inference, based on the formation of the alliance, that the nations have similar unconstrained particular interests. For the purpose of this model, I assume that A's unconstrained ideal point is aligned with P's.

Third, since the deductible theory does not rely on manipulating uncertainty, as does the ambiguity mechanism, I assume perfect information in this model. This allows us to isolate the effect of the deductible and examine the important strategic tradeoffs faced by the players.<sup>2</sup>

Finally, I assume that the deductible does not lower the probability of victory. State A has strategic options that increase costs to P without sacrificing any military advantages. For example, an initial attack, whether it involves A's forces or P's, will cost lives and resources. If A initially withholds troops and resources, it is not reducing capabilities; rather, it is

---

<sup>2</sup>In a later section, I examine a private information extension to this model in which war could occur. The extension reveals that adding a deductible has a deterrent effect, decreasing the probability of war and demonstrating that the outcomes of this model do not depend on the assumption of perfect information.

shifting the initial losses that it would have absorbed, had it been in the initial attack, to P. The loss or damage of non-military assets, especially those that are valuable to society but have little to no military use, is arguably also a price that does not translate to a lower probability of victory.

#### 1.4 Equilibrium Results

I solve for the subgame perfect equilibrium, which I will henceforth simply refer to as an “equilibrium,” by backward induction. Therefore, I begin with A’s decision to stay out or join P in fighting and then work my way up to assessing A’s decision regarding the amount of  $k$  to offer P in order to get its best possible outcome. I begin with A’s decision to join the war or to stay out.

The Ally will choose to join in the fight if its expected utility for such an action is greater than or equal to its expected utility for staying out of the war. This is captured in the inequality  $q(-|\hat{x}_A - 0|) + (1 - q)(-|\hat{x}_A - 1|) - c_A + k \geq p(-|\hat{x}_A - 0|) + (1 - p)(-|\hat{x}_A - 1|)$ . Solving this inequality gives us  $k \geq (q - p)(2\hat{x}_A - 1) + c_A$ , which establishes the lower boundary of what A is willing to offer. Therefore, we can define the lower boundary of the deductible as

$$\underline{k} = (q - p)(2\hat{x}_A - 1) + c_A, \tag{1}$$

which provides the equilibrium behavior for Lemma 1.

**Lemma 1.** *In equilibrium, A joins P in the war if  $k \geq \underline{k}$ , and stays out otherwise.*

Intuitively, this lower boundary,  $\underline{k}$ , is the minimum cost that A can shift to P such that P is not over-insured. We can view this in terms of A’s resulting costs after it transfers  $k$ , (i.e.  $-c_A + \underline{k}$ ), which avoids over-paying for the expected outcome. Formally speaking, if A were to impose less cost, then the expected utility from fighting would not surpass the utility of staying out; therefore, A would choose the latter.

Understanding this decision criteria for A allows us to assess the two primary factors that

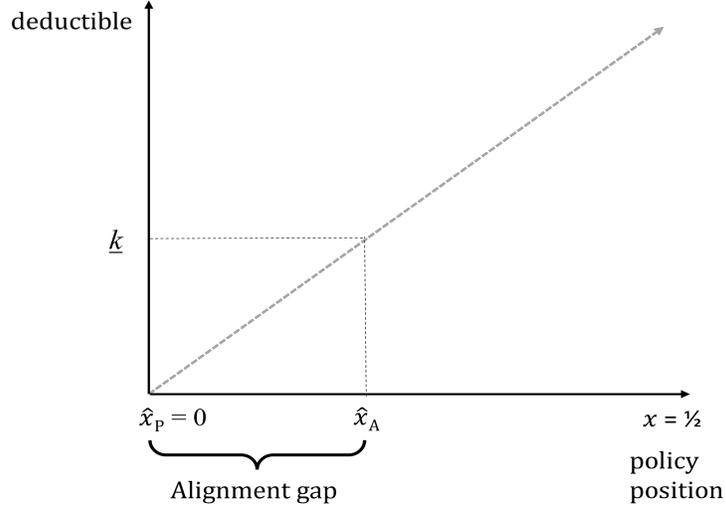
increase the risk of moral hazard: (1) P is confident that A will join it in war, and (2) A will shoulder an excessive amount of the war costs, over-insuring P. In this perfect information model, Equation 1 reveals that P can be confident that A will join it in war when A's ideal point is sufficiently close to P's and A shifts at least a minimally acceptable amount of its costs to P via the deductible,  $\underline{k}$ . This indicates how important the issue is to A and therefore the credibility of A's commitment. A lower value of  $\underline{k}$  indicates that the issue is important to A, while the higher the value of  $\underline{k}$ , the less important the issue and the greater incentive for abandonment by A. We can also see from Equation 1 that the amount of  $k$  is dependent on two primary factors: the "alignment gap" and the "power gap".

The first factor, the "alignment gap," is the difference between A's ideal point,  $\hat{x}_A$ , and P's ideal point, which we normalized to zero,  $\hat{x}_P = 0$ . The gap, then, as it equates to the distance from  $\hat{x}_A$  to 0, is equivalent to the absolute value of  $\hat{x}_A$ . We can assess this gap formally by examining Equation 1. We see that as the alignment gap increases, represented here by  $\hat{x}_A$ ,  $\underline{k}$  increases. Intuitively, this means that if the two alliance members were further apart in their ideal points, then A would be less inclined to fight and would therefore raise the deductible on P, encouraging it to concede to larger demands from C. In fact, we can see from Equation 1 that if the alignment gap grew to  $1/2$ , then A would want to shift all of its costs into the deductible before fighting, and for an alignment gap greater than  $1/2$ , P would have to compensate A to get them to fight.

Next, we can assess the "power gap," which is the difference in the probabilities of victory between fighting as an alliance or for P fighting alone. In Equation 1, this gap is represented by  $(q - p)$ . We see that  $\underline{k}$  is increasing with this power gap; as the power gap grows (i.e. as  $q$  increases or  $p$  decreases), A desires to impose a higher deductible.<sup>3</sup> The relationship between the power gap and the deductible is similar to the relationship between the alignment gap and the deductible in that an increase in either factor creates an increase in the deductible. As a conceptual representation of this relationship, Figure 2 displays the general correlation

---

<sup>3</sup> $k$  is increasing with the power gap so long as  $\hat{x}_A < 1/2$ .



**Figure 2.** Relationship between the ‘alignment gap’ and the minimum deductible.

between the alignment gap and the deductible, where the x-axis represents the difference in ideal points and the y-axis represents the corresponding deductible. The power gap represents the slope or the impact that the alignment gap has at any given value of  $\hat{x}_A$ . The same graph could be drawn to represent the correlation between the power gap and the deductible, with the alignment gap determining the slope; however, for the sake of simplicity, I have chosen to display one graph.

We next examine how A’s war criteria affect P’s acceptance criteria. Since P and C know whether or not A would join in fighting, P is able to calculate, given the probabilities of victory and the level of alignment, its criteria for conceding or rejecting a demand from C, and it allows C to calculate its demand, given P’s strategy. The strategies are represented in Lemma 2.

**Lemma 2.** *In the case that P has an alliance with A:*

- *If  $k \geq \underline{k}$ , then P will concede to any demand from C such that  $x \leq 1 - q + c_P + k$ .*
- *If  $k < \underline{k}$ , then P will concede to any demand from C such that  $x \leq 1 - p + c_T$ .*

*In the case that P does not have an alliance with A:*

- *P will concede to any demand from C such that  $x \leq 1 - p + c_T$ .*

*In all cases, C will demand the highest value to which P will concede.*

Having already established the lower boundary for A's willingness to fight based on  $\underline{k}$ , we backward induct from each alliance scenario. First, in the case that  $k \geq \underline{k}$ , P would accept a demand from C if and only if P's expected utility for accepting the offer,  $x$ , is greater than or equal to its expected utility for fighting alongside its ally, which is  $q - c_P - k$ . This calculation reveals that P accepts if  $x \leq 1 - q + c_P + k$ . In a perfect information environment, C knows P's acceptance level and demands  $x = 1 - q + c_P + k$  to maximize its own expected utility without going to war. P accepts, following the criteria just explained.

Next, we examine the case in which  $k < \underline{k}$ . This indicates that, in the case that an alliance has been formed, A stays out of the war, and P fights alone. In this scenario, P's willingness to concede a demand from C depends on whether its expected utility for conceding  $x$  is greater than or equal to its expected utility for going to war alone, which is  $p - c_T$ . Therefore, P accepts a demand from C if and only if  $x \leq 1 - p + c_T$ . C demands exactly  $x = 1 - p + c_T$  to maximize its own utility, and since this demand meets P's criteria, P concedes. Finally, in the case in which P and A do not ally, the resulting payoffs are the same as those above; P willingly concedes  $x \leq p - c_T$ , therefore C's demands  $x = p - c_T$ .

In the case that P and A agree on a value  $k$  and are allied, we can see the restraining effect of the deductible in action. Higher values of  $k$  place a higher burden on P for initiating a conflict, lowering its incentive to fight and its bargaining power with C. This lowers the probability of A being dragged into a war due to P's aggressiveness. The deductible shoulders P with a higher cost for war, giving pause to P's aggressiveness and raising its willingness to concede to demands from C. The obvious risk for A is that a high  $k$  could force P to concede too much, failing to deter C's greater demands and pushing A away from its ideal position in the opposite direction, undermining its willingness to fight. The question now is whether a deductible acts as a one-way street or whether A can set an ideal  $k$  such that it maintains its own ideal point,  $\hat{x}_A$ . That is, can A prevent P from acting with too much risk, but also leave it sufficiently empowered such that it does not become overly passive? This

is answered in the following lemma:

**Lemma 3.** *In any equilibrium where the final policy is A's ideal, the deductible offer is  $k = k^* \equiv \hat{x}_A - 1 + q - c_P$ .*

The Ally's ideal point is realized when there is no difference between the demand made by C and A's ideal position, formally written  $x = \hat{x}_A$ . This condition indicates that  $-\hat{x}_A - x = 0$ , which occurs when C knows that P will not concede more than  $x$ , so C demands  $x$ , which is equivalent to A's ideal policy. The question for A is how to set  $k$  such that this condition is met. If we make  $x$  a function of  $k$  to examine how a change in  $k$  affects  $x$ , then we have A's optimal expected utility function as  $U_A = -|\hat{x}_A - x(k)| = 0$ . So the question is, how does A set  $k$  such that  $x(k) = \hat{x}_A$ ?

We can assume that  $x(k)$  represents the demand by C that is conceded by P, since, per Lemma 2, demanding a value of  $x$  that would be rejected by P would give C less utility. Therefore, we can set the value of  $x(k)$  to the maximum value of  $x$  that C would offer, knowing that P would concede:  $x(k) = 1 - q + c_P + k$ . So, A's optimal utility becomes  $U_A = -|\hat{x}_A - 1 + q - c_P - k| = 0$ , which, when solved for  $k$ , provides us with A's optimal value of  $k$ ,

$$k^* = \hat{x}_A - 1 + q - c_P. \quad (2)$$

By imposing this value of  $k = k^*$  on P, A sets a precise limit on how much P is willing to concede to C such that C does not offer any more than A is willing to concede. Thus, by setting  $k = k^*$ , not only does A avoid over-insuring P and creating an incentive for P to act with excessive risk, but it also limits what P is willing to concede. This ensures that A is not, as a consequence of implementing a risk-reducing deductible, incentivizing P to become overly-conciliatory and concede an issue for which A is willing to fight. This concept is similar to one presented by Benson et al. (2014), where reducing the costs to an ally partner can purposefully induce a level of moral hazard, signaling resolve to the enemy and limiting its demands. The model presented here, however, demonstrates that, by transferring costs

as a deductible rather than distributing it over the course of a war, the Ally is able to set a specific cost, not only empowering it to limit the challenger's demands but simultaneously limiting the incentives for the Protégé to act in a way that risks conflict initiation.

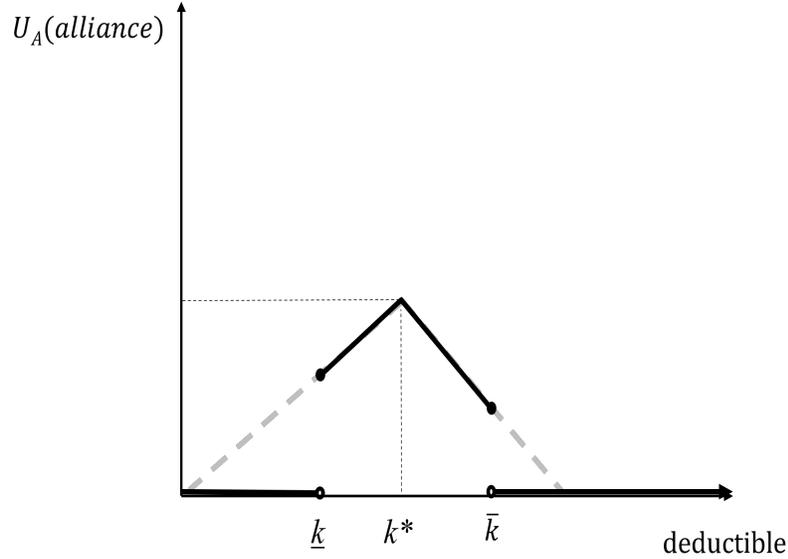
Having calculated A's optimal value for the deductible, we must now determine whether A is willing and able to enforce it. To accomplish this, we must examine whether  $k^*$  falls within the lower and upper boundaries of  $k$ . We have already determined that  $\underline{k}$  is the lower boundary, such that the value of  $k < \underline{k}$  will not be offered because the high cost is unacceptable to A. So, we now look to the upper boundary. If  $k^*$  is too costly to the Protégé, then it will not accept the offer and the alliance will not form. To determine the upper limit, we find the highest value of  $k$  that P is willing to accept. Continuing with backward induction, we can assess this value by examining when P's concessions within the alliance are less than or equal to the demands it would have to concede outside the alliance.<sup>4</sup> So, P would accept  $k$  such that  $1 - q + c_P + k \leq 1 - p + c_T$ . This means that P would accept if and only if  $k \leq q - p + c_T - c_P$ , which we can label as  $\bar{k}$ .

$$\bar{k} = q - p + c_T - c_P \tag{3}$$

The upper boundary decision criteria for the Protégé is intuitive; Equation 3 consists of two criteria that are essential to understanding if an alliance is advantageous. The first is  $q - p$ , which represents the difference in the probability of victory between being in the alliance and fighting outside it. The second factor,  $c_T - c_P$ , represents the difference between the costs involved in fighting alone or together. If the costs of  $k$  were greater than the advantages gained by the increased probability of victory and the reduction in costs, then P would have no reason to ally. The upper boundary on  $k$  also allows us to state the next lemma:

---

<sup>4</sup>I assume that P would accept a value of  $k$  that makes it indifferent toward joining the alliance because in most cases a higher value of  $k$  gives P more confidence that A will join. Additionally, it would emerge as equilibrium behavior if P received even the slightest increase in  $q$  from forming the alliance.



**Figure 3.** Ally utility when  $\underline{k} \leq k^* \leq \bar{k}$

**Lemma 4.** *In equilibrium, if A offers  $k^*$  such that  $\underline{k} \leq k^* \leq \bar{k}$ , then P accepts, the alliance is formed, and A receives its ideal policy,  $\hat{x}_A$ .*

Figure 3 is a conceptual graph that displays A's increase in utility for forming the alliance over the utility it would receive when not forming the alliance. The deductible is displayed on the x-axis and the resulting utility is displayed on the y-axis. The thick dashed line represents A's single-peaked utility for its policy preference, which, per Figure 2, corresponds to its optimal deductible,  $k^*$ . The thick solid black line between the lower and upper boundaries of  $k$  represents A's utility for any acceptable offer of  $k$ , with the peak at  $k^*$ . The thick black lines to the left of  $\underline{k}$  and to the right of  $\bar{k}$ , which run horizontally along the x-axis, represent the Ally's lack of utility for offering  $k$  outside the protégé's and its own acceptable limits. The thin dotted vertical line represents A's offer of  $k$ . This figure shows, given that it offers  $k^*$  when  $k^*$  falls within the criteria outlined in Lemma 4,  $k^*$  will be accepted, the alliance will form, and A will obtain its ideal policy preference.

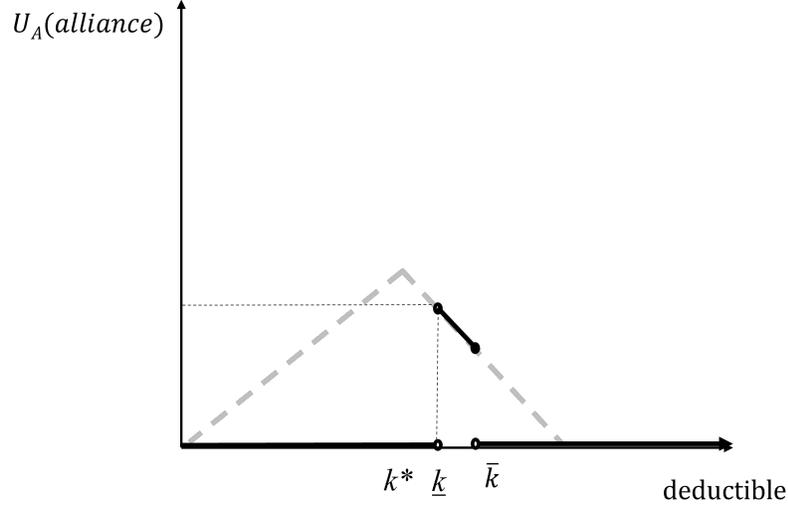
Per Lemma 4, the Protégé would accept an offer of  $k^*$  such that  $\underline{k} \leq k^* \leq \bar{k}$ , but what if  $k^*$  falls outside these acceptable boundaries? To answer this question, we first look at the case where  $k^* < \underline{k}$ . Practically speaking, this means that the deductible that A would

need to induce to get to its ideal point,  $k^*$ , is less than the deductible that A would require to be indifferent about fighting,  $\underline{k}$ . Inducing  $k^*$  would not restrain P from fighting, so if A were to join the fight while only inducing  $k^*$ , it would be worse off than if it joined while inducing  $\underline{k}$ . Therefore, if A intended to join, it would offer  $\underline{k}$ . A would join in fighting if inducing  $\underline{k}$  would give it greater utility than if it were not to join and P were to fight alone. To see if this is the case, we determine which policy outcome found in Lemma 2 (joining or staying out) is closest to A's ideal point. Formally, A would make the offer  $\underline{k}$  and join if  $|\hat{x}_A - (1 - q + c_P + \underline{k})| \leq |\hat{x}_A - (1 - p + c_T)|$ . We assume here that  $k^* < \underline{k}$ , so we know that A's utility for going to war when inducing  $k^*$  is less than its utility for going to war when it induces  $\underline{k}$ ; therefore, we can remove the absolute value from that portion of the inequality and rewrite it as  $\hat{x}_A - (1 - q + c_P + \underline{k}) \leq |\hat{x}_A - (1 - p + c_T)|$ . Solving this inequality depends on whether, on the right side of the inequality, the utility that A gets from not going to war is greater or less than its ideal utility. This could vary depending on the value of  $p$  or  $c_T$ . So, we must calculate A's criteria under three circumstances: when its value for staying out of war is less than A's ideal point, when they are equal, and when it gets greater utility for staying out of war. Doing so, we find that, even if  $k^* < \underline{k}$ , A offers  $\underline{k}$  and joins if:

$$\underline{k} \geq \begin{cases} q - p + c_T - c_P & \text{if } 1 - p + c_T < \hat{x}_A \\ 1 + q - \hat{x}_A - c_P & \text{if } 1 - p + c_T = \hat{x}_A \\ q - p + c_T - c_P - 2\hat{x}_A & \text{if } 1 - p + c_T > \hat{x}_A \end{cases} \quad (4)$$

This concept is displayed graphically in Figure 4, where the thick black line illustrates the available options for  $k$  that would yield additional utility;  $\underline{k}$  represents the maximum amount of utility available.

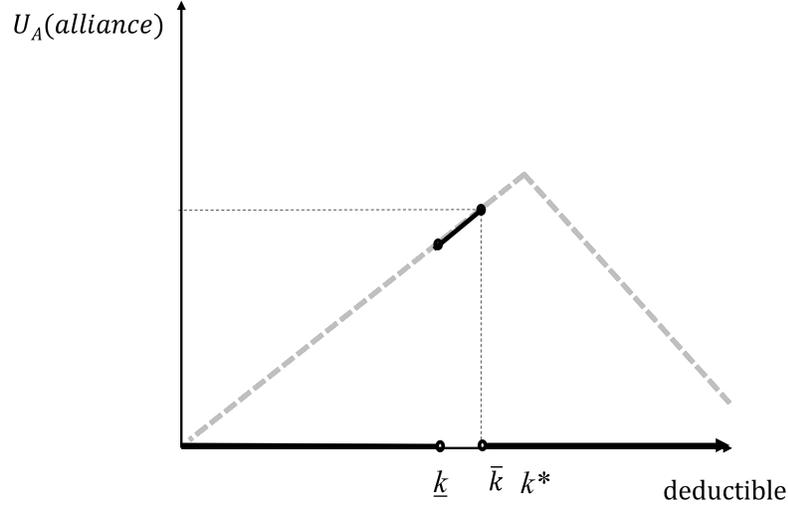
Next, we examine the case in which  $k^* > \bar{k}$ . What offer, if any, would the Ally make if faced with this situation? Lemma 4 demonstrates that P would reject an offer of  $k > \bar{k}$ , so offering  $k^* > \bar{k}$  would provide no additional utility and would be the equivalent of making no



**Figure 4.** Ally utility when  $k^* < \underline{k}$

offer. If foregoing the alliance would get it closer to its ideal point than would forming the alliance and imposing  $k^*$ , then A would make no offer; however, offering  $\bar{k}$ , even if it is less than  $k^*$ , may achieve a more acceptable outcome for A than making no offer. This would be the case as long as A's utility for offering  $\bar{k}$  is greater than or equal to its value for staying out of the fight,  $\underline{k}$ . Formally, if  $k^* > \bar{k}$  and  $\bar{k} \geq \underline{k}$ , which means that  $\hat{x}_A \leq \frac{1}{2} - \frac{c_T - c_P - c_A}{2(q-p)}$ , then A would offer  $\bar{k}$ . To examine when this would be the case, we use a similar inequality to the one examining the condition of  $k^* < \underline{k}$ , only, in this case, we substitute in  $\bar{k}$  for  $\underline{k}$ , which tells us that A offers  $\bar{k}$  if  $|\hat{x}_A - (1 - q + c_P + \bar{k})| \leq |\hat{x}_A - (1 - p + c_T)|$ . Here, we assume that  $k^* > \bar{k}$ , thus  $\hat{x}_A > (1 - q + c_P + \bar{k})$  and we can rewrite the inequality to read  $\hat{x}_A - (1 - q + c_P + \bar{k}) \leq |\hat{x}_A - (1 - p + c_T)|$ . Again, the right-hand side of the inequality is dependent on the policy value for A not joining the war,  $x = 1 - p + c_T$ . Therefore, even if  $k^* > \bar{k}$ , A would offer  $\bar{k}$  in the following cases.

$$\bar{k} \geq \begin{cases} q - p + c_T - c_P & \text{if } 1 - p + c_T < \hat{x}_A \\ 1 + q - \hat{x}_A - c_P & \text{if } 1 - p + c_T = \hat{x}_A \\ q - p + c_T - c_P - 2\hat{x}_A & \text{if } 1 - p + c_T > \hat{x}_A \end{cases} \quad (5)$$



**Figure 5.** Ally utility when  $\bar{k} < k^*$

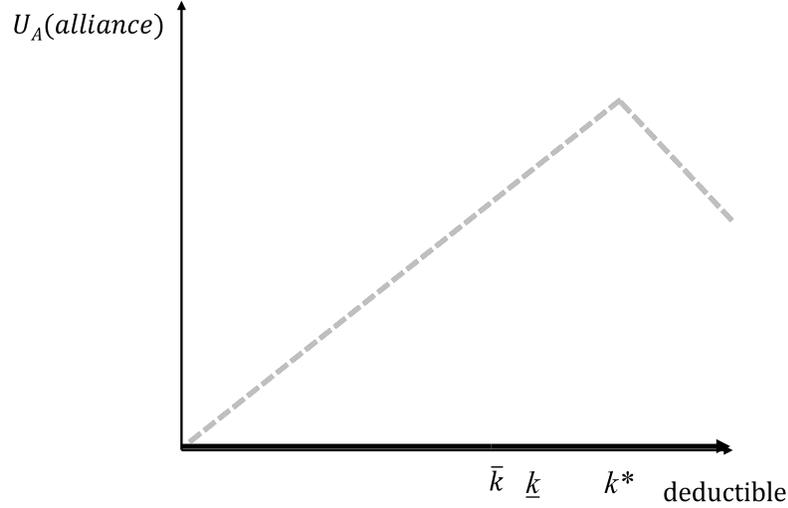
Figure 5 provides a graphical representation of this case, where the thick black line again illustrates the available options for  $k$ ;  $\bar{k}$  represents the maximum amount of utility available.

However, if  $k^* > \bar{k}$  and  $\underline{k} > \bar{k}$ , then A would receive no additional utility for forming the alliance and would, therefore, not make an offer, and an alliance would not form. This relationship is displayed in Figure 6, where the thick black line remaining steady along the x-axis illustrates that there is no acceptable offer of  $k$  that would raise A's utility, and the lack of a vertical dotted line shows that A would make no offer.

This allows us to characterize the equilibrium of this model in Proposition 1.

**Proposition 1.** *The game has a unique equilibrium that depends on the relationship between  $k^*$ ,  $\underline{k}$ , and  $\bar{k}$ .*

- *If  $\underline{k} \leq k^* \leq \bar{k}$ , then the Ally offers its Protégé a deductible value of  $k^*$ . P accepts the offer and an alliance is formed between P and A. The Challenger demands  $x = 1 - q + c_P + k^*$ , which P concedes, resulting in peace and giving A a payoff of  $\hat{x}_A$ . In the off-path case in which P rejected the offer, A joins P in war.*
- *If  $k^* < \underline{k} \leq \bar{k}$ , then A offers P a deductible value of  $\underline{k}$ . P accepts the offer and an alliance is formed between P and A. C demands  $x = 1 - q + c_P + \underline{k}$ , which P concedes,*



**Figure 6.** Ally utility when  $\bar{k} < \underline{k} < k^*$

resulting in peace and giving A a payoff of  $-|\hat{x}_A - 1 + q - c_P| < \hat{x}_A$ . In the off-path case in which P rejected the offer, A joins P in war.

- If  $\underline{k} \leq \bar{k} < k^*$ , then A offers P a deductible value of  $\bar{k}$ . P accepts the offer and an alliance is formed between P and A. C demands  $x = 1 - q + c_P + \bar{k}$ , which P concedes, resulting in peace and giving A a payoff of  $-|\hat{x}_A - 1 + q - c_P| < \hat{x}_A$ . In the off path case in which P rejected the offer, A joins P in war.
- If  $\bar{k} < k^*$  and  $\bar{k} < \underline{k}$ , then A makes no offer and does not attempt to form an alliance with P. C demands  $x = 1 - p + c_T$ , which P concedes, resulting in peace and giving A a payoff of  $-|\hat{x}_A - 1 - p + c_T|$ .

Provided that the optimal deductible,  $k^*$ , is not less than the Ally's minimum acceptable nor higher than the Protégé's maximum acceptable deductible, then A will be able to enforce a deductible that allows it to achieve its ideal policy position. If its optimal deductible is less than its minimum for joining the war, then A will desire to induce the minimum and join, provided it will produce a greater payoff than not joining. Similarly, if A's optimal deductible is greater than P's tolerance, A will desire to enforce P's maximum acceptable value, provided it is greater than its own minimum and P's maximum will produce an outcome greater than

not joining the alliance at all.

The incentives that drive each state in this equilibrium allow us to imply what the model predicts regarding the states' behaviors. In the following section, I will examine the comparative statics of Equation 2 to understand how the exogenous factors of the power gap and alignment gap affect the deductible in the equilibrium condition. I will then discuss the intuition behind this dynamic and will use a case study to demonstrate the existence of the resulting observable implications, or expected outcomes and behaviors.

## **1.5 Discussion**

In this section, I will examine the comparative statics found in Equation 2, which describe A's optimal value of the deductible. Using the US-ROK alliance as a case study I will determine how each of the relevant comparative statics translate into behaviors we should expect to see from the Ally, given the exogenous factor of the alignment gap, and will then illustrate the existence of these observable implications and corroborate the logic of the model.

### **1.5.1 Case Study: US-ROK Alliance**

The alliance between the United States and the Republic of Korea is an ideal case study to display the observable implications of this theoretical model. It is an asymmetric alliance in that the ROK has forfeited a great deal of its sovereignty to the US. Additionally, the ROK has expressed desires that are out of alignment with the US, indicating that the difference between South Korea's unconstrained interests and those imposed on it by its alliance with the US is significant. This indicates an alignment gap. Additionally, there is a substantial power gap, as the US provides a significant security guarantee to the ROK, supplying the preponderance of military power and providing the ROK with a much greater capability of achieving its policy preferences than if it were to fight alone. So, the implicit incentives to act in a way that creates a risk of moral hazard, which are associated with an asymmetric

alliance, are present in this alliance.

The US-ROK alliance also represents a case that is not well explained by other models. Benson (2012) demonstrated that contract ambiguity can be an effective way to constrain moral hazard; however, in this alliance, the US has intentionally and credibly signaled that it will intervene on behalf of the ROK if it is attacked. There are at least three clear signals that it would do so. First, the US signed a Mutual Defense Treaty with the ROK in 1953, obligating itself to intervene to “meet the common danger” in the case of external attack (Mutual Defense Treaty, 1953). Second, even before the treaty was established, the US demonstrated the importance of South Korea to its overall strategy of containment when it came to its aid and even pushed for reunification in 1950. Finally, the US pays a high peacetime price to keep the alliance strong, including the continued presence of troops and equipment on the peninsula and annual combined exercises. These signals reveal that the US is a reliable partner (Fearon 1997, Morrow 2000, Mattes 2012, Morrow 1994), thus providing the ROK with assurance.

The US has also employed other, more nuanced assurances that will be highlighted in the case study. By providing such guarantees to the ROK, the US further incentivizes the ROK to act with risk. The existence of incentives inherent in an asymmetric alliance and those induced by assurance provide an opportunity to assess the ability of an ally to constrain its partner by altering the insurance it provides via a deductible, thereby verifying the model’s logic. Furthermore, the alliance has, in its nearly 70-year existence, undergone changes in the power gap and the alignment gap, allowing us to assess if these dynamics induce the changes in behaviors we expect to see in the model as well.

## **Comparative Statics**

Given the merits of using this alliance as a case study, we begin by examining the comparative statics found in Equation 2, which demonstrate that  $k^*$  increases with  $\hat{x}_A$ . Intuitively, this means that the greater the difference between the policy preferences of the Ally and

the Protégé, the more the Ally will want to levy a deductible, limiting the bargaining power of the Protégé. Conversely, as the policy preferences of the two nations come in line with one another, the less deductible the Ally will impose to ensure that the Protégé has the bargaining power to reject over-zealous demands from the Challenger. If this is true, then we should observe that when the policy preferences of the Ally and Protégé are misaligned, the Ally will seek to limit the initial protection of the Protégé. This may be accomplished through such measures as reducing defensive posturing, opting not to protect valuable assets with limited military value, or limiting the Protégé's independent and immediate retaliatory capabilities. When policy alignment becomes more synchronized, causing the Ally to be less concerned about the ramifications of moral hazard, the Ally will seek to provide greater protection by reversing or easing these measures. In the case of the US and South Korean alliance, we have indeed observed these anticipated behaviors corresponding to changes in alignment.

### **A Large Alignment Gap**

Despite South Korea's agreement to support the US's objectives, in the years following the Korean War, the US perceived the policy preferences of the South Korean government to be significantly different than its own. Declassified documents reveal that the US feared that South Korea's young authoritarian government under President Syngman Rhee, who was staunchly anti-communist and advocated strongly for the forceful reunification of the peninsula under his regime, would drag it back into conflict (Oberdorfer 2001). A conversation among President Eisenhower's National Security Council in 1957 highlights this concern. General Lemnitzer, commander of the Eighth Army and Far East Command, expressed fear that if the South Koreans were in charge of all the military force, President Rhee may attack north (Memorandum 1957). President Eisenhower expressed his own concern over the South's ambitions in a National Security Council meeting in June of 1959 (Memorandum 1959). The strong unconstrained desire by the South Koreans to attack north and reunify

the Korean peninsula stood in contrast to the constrained interests of the US, which deemed such an effort too costly for the benefit it would receive from the outcome. The allies' policy preferences were misaligned.

As the model predicts, the US responded to the growing alignment gap by limiting the amount of protection against the damage the South would sustain in the initial stages of any conflict with the North. One way the US did this was by limiting its defensive posturing, keeping a limited number of troops on the peninsula. The distance between the two nations allowed the US to credibly delay a response to join the ROK in a conflict. Had the US left the peninsula entirely, the time it would take the US to return would certainly have ensured the South Koreans paid an upfront cost for war; however, it would also have failed to send a credible signal of assurance to the ROK or of resolve to the DPRK. So, while having no presence would place a limit on South Korea's incentives to attack north, it would also have generated risks concerning the level of deterrence and defense the US was providing.

On the other hand, too much presence on the peninsula, though it would provide greater assurance, deterrence, and defense if needed, would have over-insured the ROK, increasing the incentive to reject reasonable demands from the North and inducing conflict. Therefore, the US sought to deploy the optimum number of troops to deter the North while limiting the risk of moral hazard from the South. In the first year after the Korean War ended, the US only marginally reduced its troop levels on the peninsula; it left six of its eight deployed divisions in place, maintaining the ability to repel an attack from the North and reestablish the 38th parallel (Cavendish 2004). After the first year, however, the troop level was reduced to only two US divisions, along with an Army missile command, a tactical missile unit, and a fighter-bomber wing (Memorandum 1959). The US forces, combined with those of the ROK, were no longer enough to repel the attack but were instead determined to be sufficient to limit North Korean gains until the US could deploy significant follow-on assets (Kim 2002, Work 2020). The projected follow-on forces, including over 600,000 troops plus equipment, were estimated to be sufficient to repel an attack, but the induced delayed response time ensured

that the ROK would pay a cost for conflict initiation (Heo & Roehrig 2018). While the US intended for the initial troop level to be enough to hold the North (Memorandum 1959), the number was under constant scrutiny, as the US feared that too high of a level would incentivize the ROK to act with risk. This was illustrated in a declassified memorandum between members of the National Security Council in 1962 when Robert Komer argued to Deputy National Security Advisor Carl Kaysen that the US military was “over-insured” on the peninsula and that even Secretary of Defense McNamara felt the conventional forces were more than enough to defeat the North Koreans (Memorandum 1962). These comments illustrate the the US was steadfast in its efforts to achieve the optimum deductible.

The US’s concern for limiting the bargaining power of the South Koreans was also seen in the restrictions it placed on South Korea’s immediate retaliatory capabilities. The US restricted the range and payload of South Korean missiles. The missile guidelines, which were instituted in 1979, were the result of South Korea attempting to build its own system beginning in 1971. After facing setbacks in preventing South Korea from taking these steps, the US intervened and provided missiles primarily intended for air defense, with a limited capability for surface-to-surface employment, in exchange for South Korea agreeing to the restrictive missile guidelines (Pinkston 2012). The move by the US was intended to reduce ROK options by placing significant limits on South Korea’s ability to not only initiate conflict but on their ability to respond quickly to North Korean provocations (Kim 2008). This limited South Korea’s ability to defend itself, creating more conditional costs.

The US also limited the protection it provided for the city of Seoul. Two factors indicate that it is a part of the US strategy. First, Seoul is inherently and obviously vulnerable to attack; nearly 50% of the South Korean population lives in the Greater Seoul Metropolitan Area, which is 40 km from the North Korean border. Even in 1957, the Chairman of the Joint Chiefs, Admiral Radford, acknowledged that the city was constantly under threat and expressed “considerable sympathy” for those “sitting under the guns” (Memorandum 1957). Through the years, Seoul has come under an increasing threat from artillery bombardment.

The North Koreans now have nearly 6,000 artillery systems located within range of Seoul, causing RAND experts to conclude that “because so much harm could be done so quickly, the United States and South Korea should try to avoid military provocation cycles that could lead to these attacks” (Barnett et al, 2020).

Second, despite Seoul’s vulnerability and North Korea’s growing missile and artillery threat, the US has done very little to protect it. While the US has contributed over \$1.6 billion to Israel for the development of a counter-artillery system called the Iron Dome (Horton 2021), which has proven to be 85-90% effective against incoming artillery rounds (Grudo 2016), it has not supplied or developed any such system for South Korea. According to former Deputy Assistant Secretary of Defense Brad Roberts, providing the ROK with such deterrent capabilities would embolden them, making the US “concerned about entrapment” (Bowers & Hiim 2020, 34). The vulnerability of such an important city is a major concern for the South Koreans and acts as a key constraint, disincentivizing actions that may lead to conflict.

One possible alternative to the theory that the optimum deductible drove these behaviors by the US is that these measures are instead driven by an effort to avoid a security dilemma. If the latter were the case, we would expect the US to avoid building capabilities that may be perceived as offensive while maintaining or increasing defensive capabilities. One may view the limited troop levels in theater, which kept the overall capability below suitable for offensive operations, as an attempt to meliorate security dilemma concerns; however, the paucity of troops goes beyond what an effort to avoid a security dilemma would require. The troop levels were, in fact, even lower than one would expect for purely defensive measures as they would not be enough to repel the attack, but only to protect South Korea against an invasion to the point of holding off an enemy and enabling the deployment of follow-on forces. Meanwhile, the rocket restrictions, which made the South Korean rockets primarily defensive, did not prevent their offensive use. Finally, the overall lack of protection for Seoul simply cannot be explained by a theory that the US was attempting to avoid a security

dilemma. These behaviors can, however, be explained through the lens of the US enforcing an optimum deductible.

## **A Reduced Alignment Gap**

In addition to predicting that a misalignment between the allies would lead to the US enforcing a deductible on the South Koreans, the theoretical model also predicts that greater alignment between the states should reduce the deductible. Indeed, we see this result as well. Both South Korea and the international system underwent significant changes in the 1980s, bringing the US and the ROK into greater alignment. South Korea, which fully democratized in 1988, changed its policies toward its common enemies with the US. Due to its extraordinary economic development and its desire to open trade with Asia and the Soviet Bloc, the ROK found it best to take a more conciliatory approach to North Korea and its allies. South Korean policy was no longer to isolate the North Koreans or to reunify the peninsula via force. Instead, they pursued a path of nonaggression. The policy change was concurrent with Gorbachev's perestroika in the Soviet Union, which softened the relationship between the US and USSR (Oberdorfer 2001). Consequently, the US became less concerned about the DPRK threat being backed by the Soviets, changing its hard-line anti-communist policies. The policy changes by no means pacified the relationship between North and South Korea; however, the shift significantly increased alignment between the US and the ROK and reduced American fears that the South would create a moral hazard problem, making it less interested in restricting the ROK's bargaining abilities.

With this change in alignment, the model predicts that the US would take measures to reduce the deductible. One way in which this has manifested itself is that the US began maintaining more troops on the peninsula than were required to hold off a North Korean attack. US discussions about troop reductions on the peninsula have been a recurring theme throughout the history of the alliance. However, in each of the cases prior to the policy changes that led to greater alignment between the US and the ROK, the US sought to avoid

over-insuring the ROK by committing to minimum troop levels, given the level of threat posed by the North and the perceived ability of the South to withstand an attack. This was illustrated by the earlier example of Kennedy's staff being concerned the troop level was over-insuring the ROK. We see this again in 1970 as the Nixon administration reduced troops after painstakingly evaluating the threat and assessing the ROK military ability to confront it (Minutes 1970). Additionally, when President Carter sought to reduce troops in the late 1970's, the administration found it wiser to leave the majority in place due to updated intelligence about a stronger-than-expected North Korean force (Oberdorfer 2001). After the alignment, however, the nature of discussions pertaining to troop reductions changed; the US began showing less concern about providing a minimum force to avoid over-insuring and instead provided more forces than needed to assure South Korea and keep it from caving to North Korean demands.

The South Korean military grew significantly stronger throughout the 1980s and early 90s, shrinking the power gap and making the ROK less dependent on the US for security in general. Additionally, the threat posed by the North diminished significantly when its backing from the Soviet Union fell through and China normalized relations with South Korea. At the same time, the South Korean economy skyrocketed from the 1960s through the 1980s, especially compared to that of the North. The actual and potential growth of the South Korean military forces precluded a need for the US to be able to hold the North Koreans until follow-on forces arrived, so any presence on the peninsula was more than required. Under Secretary of Defense for Policy, Paul Wolfowitz acknowledged before Congress that "Seoul is economically capable of matching Pyongyang's military buildup," yet he defended the strategy of maintaining a strong troop presence as insurance against potential North Korean aggression (Taylor et al. 1990, 6-7). The US maintained about 40,000 troops throughout the 1980s and only reduced that number to an average of about 35,000 in the 1990s (Kane 2006). After the nations became more aligned, the US became less concerned about over-insuring its ally and instead, ensured that it provided enough

assurance that the ROK did not make excessive concessions.

A second indication that the US reduced the deductible it placed on South Korea is that it eased restrictions on the ROK's immediate retaliatory capabilities. After nearly three decades of limiting South Korea's ability to respond against North Korea's growing threat from short- and medium-range ballistic missiles with its own missiles, the US extended and expanded the missile guidelines in 2001, 2012, 2017, and 2020, raising the payload and operational ceiling each time. The limitations were finally lifted completely in 2021 (Kim 2008, Wright 2021). The liberalization and ultimate dismantling of the restrictions returned a degree of sovereignty and negotiating power to the South Koreans. Boo Seung-chan, a spokesman for the ROK's Defense Ministry, stated that the termination of the restrictions was a result of "Washington's trust in South Korea" (Kim 2021). His view is corroborated by the fact that the move came only after the US had greater confidence that the ROK's interests aligned more closely with its own.

The third outcome validating the theoretical model is that after the nations came into greater alignment, the US moved to provide some protection for Seoul. In 2016, the US deployed a Terminal High Altitude Area Defense (THAAD) missile defense system into the Korean theater. The significance of the additional protection it provided against a possible North Korean attack is highlighted by the fact that, even after China protested the move and enforced an economic penalty on the ROK, the South Korean government is moving forward with accepting a second system. A senior ROK official in President Yoon's administration described the issue as a "matter of South Korea's self-defense" (Lee 2022). This increase in defense, which further lowered the deductible by making Seoul less vulnerable, came as a result of growing alignment between the nations.

An alternative explanation for why the US would lift restrictions and add protection over time is to avoid a commitment problem. Benson and Smith (2022) examine the commitment problem that could result from alliance formation, focusing on the speed with which the alliance is formed. They found that slowing the speed and offering concessions to the

opposing side reduced the enemy’s concerns about the ensuing power shift and ameliorated the risk of conflict. Similarly, it could be argued that if the US-ROK alliance increased its capabilities too much and too quickly on the peninsula, it would incentivize North Korea to launch a preventive war (Powell 2006). If the US-ROK alliance were attempting to avoid such an eventuality, they would take similar measures to those proposed by Benson and Smith. Indeed, the power shift has been gradual and incremental, much like we would expect to see. However, when the allied nations had divergent policy preferences, there were only decreases in capabilities, not the gradual increase one would expect. The increase in relative capabilities on the peninsula came only after the change in policy alignment between the US and the ROK, giving greater credibility to the deductible theory. These additional protective measures demonstrate, as the model predicted, that the US’s optimal deductible decreased only as the gap in alignment decreased. Overall, the dynamics of the US-ROK alliance demonstrate the merit of the deductible theory.

### 1.5.2 Private Information Extension

As demonstrated above, the perfect information model suggests that the ally can induce a deductible,  $k$ , to get its desired policy outcomes,  $\hat{x}_A$ , by restraining the ally and deterring the challenger from making excessive demands. However, the model does not include war as a potential equilibrium outcome. As long as the challenger knows what will be conceded, it offers exactly that, and there is no possibility of war. Do the expectations hold if we relax the assumption of perfect information? To answer this question, I introduce private information into the model using the common approach of assuming that the challenger is uncertain of the protégé’s costs. This makes it more difficult for the challenger to make a demand that maximizes its own utility (i.e. extracting the most it can without inducing war). By examining the private, or incomplete information model, we can determine whether the deductible affects the probability that the challenger will demand more than the protégé will accept (i.e. does the deductible affect the probability of war?).

In the private information game, I assume that nature chooses a high or a low cost of war for P, which I label  $c_P^H$  and  $c_P^L$  respectively if an alliance is formed and  $c_T^H$  and  $c_T^L$  respectively if an alliance is not formed. I also assume that C does not know the costs but has some belief,  $\mu$ , about whether P's costs are high,  $\mu(H)$ , or low,  $1 - \mu(H)$ . Using this notation, we can examine C's options and determine what belief it requires in order to risk going to war.

**Lemma 5.** *If an alliance has not formed between A and P, then making the more demanding offer is sequentially rational for C if*

$$\mu(H) \geq \frac{c_C + c_T^L}{c_C + c_T^H}.$$

*If an alliance has been formed between A and P, then making the more demanding offer is sequentially rational for C if*

$$\mu(H) \geq \frac{c_C + c_P^L + k}{c_C + c_P^H + k}.$$

Although C does not know P's war costs, it does know that it can avoid war by minimizing its demand. Since C runs the most risk of being rejected and inducing war if P has low war costs, C could avoid war by offering  $x \leq 1 - p + c_T^L$  in the case of no alliance, or  $x \leq 1 - q + c_P^L + k$  if an alliance does exist. If C demands more than this, then the payoff is a lottery between receiving  $1 - p + c_T^H$  (no alliance) or  $1 - q + c_P^H + k$  (with alliance) if the cost is indeed high, or receiving the war payoff,  $1 - p - c_C$  (no alliance) or  $1 - q - c_C$  (alliance). We've determined that C's belief that P's costs are high is  $\mu(H)$ , so we can calculate this value to determine what value of  $\mu(H)$  would make the lottery worth the demand. Looking first at the case of no alliance, C's belief is determined by the following,

$$\mu(H)(1 - p + c_T^H) + (1 - \mu(H))(1 - p - c_C) \geq 1 - p + c_T^L,$$

which means that

$$\mu(H) \geq \frac{c_C + c_T^L}{c_C + c_T^H}.$$

In the case of an alliance, C's belief is determined by

$$\mu(H)(1 - q + c_P^H + k) + (1 - \mu(H))(1 - q - c_C) \geq 1 - q + c_P^L + k,$$

which means that

$$\mu(H) \geq \frac{c_C + c_P^L + k}{c_C + c_P^H + k}.$$

The existence of the alliance, and more specifically the deductible that A induces on P, impacts C's beliefs in a way that affects its behavior. This allows us to make the following statement:

**Proposition 2.** *The probability of war decreases with the deductible.*

$\mu(H)$  represents the belief that C must possess to make risking war with a higher demand sequentially rational. This belief, as seen in the values in Lemma 5, is represented by the ratio of the gains C makes from the sure thing (numerator) to the gains it could potentially make via the lottery (denominator). If the ratio is sufficiently low, meaning that C has enough confidence in the lottery relative to the sure thing, then it is rational for C to risk war. However, given that the numerator is smaller than the denominator, adding  $k$  to both increases the value of the sure thing at a relatively greater rate than it does the lottery. Therefore, as  $k$  increases, the required belief rises (i.e. the ratio of the value of the sure thing rises relative to the lottery), making C less inclined to risk war. Thus, relaxing the perfect information assumption does not alter the expectations of the model.

## 1.6 Conclusion

I began this paper by asking why the stronger power in these asymmetric alliances may intentionally leave its partner vulnerable to attack. This strategy appears counterintuitive; however, if one examines the problems that a stronger ally in such a relationship faces in forming, maintaining, and managing such a potentially advantageous alliance, the reason for such behavior becomes clearer. The security-autonomy tradeoff that underpins these

alliances creates inherent incentives for the protégé to act with risk to the stronger ally by pursuing interests counter to the ally's policy objectives, creating a moral hazard problem.

I have argued that the ally can manage the protégé's incentives to behave with risk by manipulating the amount of costs the protégé will have to endure on the condition that conflict is initiated (i.e. by inducing a deductible). The formal model presented in this paper indicates that the greater the difference between the policy interests of the states, the greater the incentive for the protégé to act with excessive risk, and therefore, the greater the deductible it takes to restrain such behavior. As the nations' policy preferences grow in alignment, a lower deductible is required. While the increase in the deductible restrains the protégé's aggressiveness, it follows that decreasing the deductible incentivizes boldness. The model suggests that the ally can set the deductible to an optimal level, preventing the protégé from being too bold, while simultaneously providing it with enough bargaining power that it does not concede to excessive demands. The challenger, in turn, is deterred from making such demands. In effect, the model suggests that the ally can pursue its own desired policy by controlling the costs the protégé is forced to endure, conditional on conflict initiation.

Using the case of the US-ROK alliance, I find the US's behavior to be consistent with the model's predictions. First, we see that when the alignment gap between the states was significant, the US took measures to ensure the ROK would pay a high deductible before the US provided greater protection. Second, the case demonstrates that when alignment between the states grew, the US lowered the deductible, providing greater protection against an initial attack. These findings indicate that the model's predictions have been demonstrated in at least one asymmetric alliance.

The findings also provide opportunities for future research. While I have demonstrated the existence of the theory in one specific case, further research using quantitative data is merited to examine the prevalence of the mechanism across asymmetric alliances. Such a study could utilize existing data on power indices and ideal policies as the independent

variables, as well as statistics on troop deployments, arms sales, and military aid as possible dependent variables; however, more work would be required to collect data on the specific policy alignments that matter to each alliance.

This novel theory provides a fresh perspective on alliance behaviors and dynamics that may impact future studies on arms sales and restrictions, the size and purpose of tripwire force contingents, and how extra-contractual agreements and structures may affect alliance behaviors. Future studies may also investigate the conditions under which a stronger ally may better achieve its policy objectives by using a strategy of ambiguity, a strategy of enforcing a deductible, or a combination of these strategies. For example, in an asymmetric alliance between states that are close in geographic proximity, the stronger ally may have a difficult time making a credible threat to delay troop deployments. In such a case, the stronger ally may find it more advantageous to be ambiguous about the level of or trigger for its military support. On the other hand, if it is clearly within the vital national interest of the ally to join the protégé in a fight if conflict is initiated or if maximum deterrence is required, the deductible may be the best strategy. These and other factors should be examined to determine a state's optimal mechanism for managing an asymmetric alliance.

Asymmetric alliances have proven to be the most durable (Morrow 1991) and among the most deterrent (Siverson & Tonnefoss 1984, Benson 2011) of all alliances, consequently providing stability to the international system. This study provides a new mechanism for managing and maintaining these valuable alliances, whereby contributing the the greater peace.

## 2 Conceptualizing and Measuring Polar Tension in the International System

### 2.1 Introduction

In attempting to describe the international order after World War II, political scientists began using the term “bipolar,” a word borrowed from other scientific fields, such as geography, geometry, and physics. The fact that two world superpowers came out of the war, the United States and the Soviet Union, made this an appropriate analogy. It invoked the idea of two opposing poles along an axis exerting pressure on one another. There was indeed a great degree of tension between the two nations, which were not only the most powerful states in the world but were also diametrically opposed to one another ideologically. The combination of power parity and international policy opposition created a dynamic in which each power perceived the other as a threat to its own interests. As scholars and political practitioners adopted the concept of polarity to help explain the international structure of former periods, they saw fit to alter the prefix, distinguishing a period of bipolarity from an earlier era of multipolarity and eventually to the later period of unipolarity, creating a typology based on the number of great powers or clusters of nations that held sway in the international order. This categorical, often binary, use of polarity caused scholars to concentrate on the number of significant actors, or poles, in the system to explain state behavior, rather than focusing on the threat that is realized from the dual aspects of power and policy.

Using polarity to categorize eras in the world’s systematic structure is useful. It has provided scholars with the opportunity to explore how these categorical structures affect issues such as global stability (Waltz 1964, Deutsch & Singer 1964), trade (Gowa & Mansfield 1993, Mansfield 1995), and alliance behavior (Christensen & Snyder 1990). However, in reducing polarity to a typology, political scientists have limited the theoretical mechanisms available to explain behaviors in the international system. For example, considering system structure in terms of the number of poles leads to mechanisms such as the number of opportunities for interactions and misunderstandings between major powers (Waltz 1964). This view of

polarity also influences the number of available protectorate states with whom weaker states may ally (Gowa and Mansfield 1993) and the clarity of the role of any particular major power in maintaining the international order (Snyder 1984). While these structural mechanisms are viable options for attempting to explain state behavior, they are not comprehensive.

I propose a new concept of *polar tension*, defined as the degree to which international policy space is contested. This term moves beyond the categorical function and considers power, or the ability to successfully pursue national interests, alongside a second dimension, the degree of polarization, or the level of conflicting interests between the most powerful and influential players. Because it considers both the *desire* and the *ability* of opposing forces to pursue their own interests, this conceptualization highlights the level of tension or general threat that nations perceive in the environment. This tension, or threat, produced by the two-dimensional view of system structure provides an alternative mechanism for explaining the impacts of system structure on state behavior.

Polar tension also returns to the scientific roots of the term “bipolar”, which was used analogously to convey the idea that there were two poles at which the maximum force was exerted on the space between them along a single axis (Keermaeker 2017). I conceptualize the poles not as nations or clusters of nations, but as the opposite extremes of the ideology or underlying political philosophy that motivates international policy preferences. Therefore, the space between these poles, analogous to the polar axis, comprises the continuous one-dimensional range of policy space. When the ideal policy positions of powerful states are opposed to one another, then these powerful states perceive the threats that the others could pose to their interests, creating tension in the environment. If only a single nation has the power to affect the status quo, then the policy space is uncontested. However, the greater parity there is among nations with the most power, the more potential tension that exists in the policy space (Organski 1968, Organski & Kugler 1980, Gilpin 1981). Likewise, if nations, despite possessing great power and having power parity with one another, are in alignment regarding policy preferences, then no nation in the set would have an incentive or

intent to oppose the other nation(s), and policy space would remain uncontested (Organski 1968). However, the greater the disparity between the ideal policy preferences of nations, the greater the incentive to oppose differing policies. Therefore, if nations possess a great deal of power parity and hold maximally opposing policy preferences, it induces a condition where policy space is greatly contested, or where there is a high degree of polar tension.

The questions that naturally arise are: 1) Why is polar tension an important measure in international relations? and 2) What can it tell us that the binary measure of polarity cannot? To answer the first question, the measure of polar tension can serve as an additional explanatory variable. For example, the measurement can help scholars determine if the polar tension in the system leads to more or less trade, tighter or looser alliance bonds, or more or less interstate or intrastate conflict. The answer to the second question follows in that the outcomes of studies relying on categorical poles depend on the specific mechanisms involved. Introducing a new mechanism may supply different theoretical outcomes. Furthermore, the variation in polar tension from year to year also allows a much more nuanced picture of the interaction of major powers within the categorical time periods and may help indicate or predict transitions from one categorical era to another.

This study contributes to the rich literature on polarity, defining it in terms of the tension it creates and providing a scoring system with which to measure the tension in the international system. Additionally, the topic has strategic relevance; understanding variation within periods and between them can aid practitioners in employing relevant policy tools in pursuit of national interests.

## **2.2 Polarity and Polarization as Structural Aspects of the International System**

### **System Structure**

Before examining the various concepts and measurements of polarity, it is essential to define what is meant by system structure. Although the terms system and structure are often used interchangeably, Ray (1990) provides a definition of each that helps explain the

analogy of polarity. He considers a system to be “an aggregation of social entities that... are sufficiently interdependent to have the actions of some consistently affect the behavior and fate of the rest”. The structure, on the other hand, is “the way in which relationships are arranged” (Ray 1990, 99). By these definitions, the system is affected by changes to the structure. The major players in the system have the greatest impact on others, and the dynamics between these players (e.g. their relative power and the degree to which they are hostile or friendly with one another) determine the severity and nature of that impact. Ray further asserts that structure is two-dimensional in that it is based on “comparisons”, which he refers to as the *vertical* domain, and “links or bonds”, which he considers to be the *horizontal* domain. Similarly, Clark identifies “power” and “conflict” as being important notions in the international order (Clark 1980, 13).

Wayman and Morgan (1990) describe the way that the concept of polarity has been divided between these dimensions, labeling them “polarity” and “polarization”. The vertical, or comparative dimension has become synonymous with the overall concept and has therefore been labeled “polarity”. They consider this aspect to be “the distribution of power in the international system.” (Wayman & Morgan 1990, 142). As we will see, this domain has become the sole factor that many scholars consider when calculating polarity.

The lateral dimension of international structure, while often ignored, is identified by Wayman and Morgan as “polarization”, or “the pattern of alliance bonds within the system” (Wayman & Morgan 1990, 142). The fewer alliance bonds between the major powers, or the most powerful and influential states in the international system, the greater the polarization. Rapkin et al. (1979) expand this concept beyond alliances, considering other bonds, such as “common memberships in organization” or “transactions in trade or aid” (Rapkin et al. 1979, 272). These interactions act to operationalize the degree of alignment between states or blocs of states. Wayman admits that the “ideal data would be scaling of hostility and friendship levels gleaned from content analysis of diplomatic documents”, but asserts that data on alliances is the best available data to capture the concept (Wayman 1984, 68).

Keersmaecker reinforces the point that the salient issue is about policy differences, noting that the lack of interaction between the East and West during the Cold War was the “essence of the political and ideological divide” and was “the result of an ideologically inspired policy” (Keersmaecker 2017, 61). The lateral dimension is therefore best articulated as the political spatial alignment (or lack of alignment) that affects the degree of hostility, or policy conflicts, in the international environment.

Therefore, structure is best understood as having two dimensions: the vertical dimension of relative power and the horizontal dimension of ideological policy alignment. These two aspects of international system structure determine both the *ability* and the *desire* of nations to contest international policy space. Examining the vertical aspect without the horizontal would reveal only the ability to contest international policy without regard for a desire to do so. Conversely, examining only the horizontal aspect would reveal the desire to contest policy without accounting for a nation’s ability. For policy to be contested, both domains must be considered.

### **Polarity as a Way to Describe System Structure**

Since the end of World War II, the analogy of “polarity” has been the primary device used to describe the structure outlined above. The first scholar to use the term “polarity” to describe the international system was Fox in 1944, who understood it to cover both dimensions described above. Even before the end of World War II, he envisioned the possibility that two nations—the US and USSR—would come out of the war with significantly more power than the others.<sup>5</sup> In this sense there would be two poles. However, it was only because of the potential that the two would be “polar *opposites* in the postwar world” that he coined the term “bipolar” to describe it (Fox 1944: 97). While he allowed for the possibility of a “tripolar” world if Germany remained strong and outside the “power-nuclei” of the East or

---

<sup>5</sup>Fox refers to the US, USSR, and Great Britain as potential superpowers; however, due to the ideological similarities between the US and Great Britain, he identifies only two poles, the US and Great Britain at one pole and the USSR at the other.

West, he indicated that polarization depended on the degree of opposition in competing ideologies (Fox 1944: 97-100). Therefore, just as the term “polar” indicates two opposite poles in the scientific sense, Fox’s use of polarity to describe the international system depended not only on the *vertical* dimension of power (i.e. the capability that nations have to affect the international order) but also the *horizontal* dimension (i.e. the opposing ideological nature of the order that these powerful nations would wish to construct).

Keersmaeker notes that as metaphors move from one scientific discipline to another, they sometimes change meaning (Keersmaeker 2017, 15). To a large degree, this has been the case with the term polarity. While first used appropriately to describe the post-World War II structure, polarity soon became a categorical term to describe the number of “poles” in the system. It was Kaplan that first used the term as a comparative, or categorical way of describing the international order, emphasizing only the vertical aspect of international structure (Keersmaeker 2017). In categorizing international systems, Kaplan (1957) uses the term bipolarity to describe two of the six possible international structures, both of which were dependent on there being only two great powers, or blocs of power. Although he does not use other prefixes, such as multi- or uni-, to describe the structural categories that contain more or less than two strong actors, his use of the term “polarity” as a typology soon became the norm among international security scholars. Waltz (1964, 1979) is arguably the most influential in further developing and reinforcing the categorical concept. He differentiates periods of time when there was a bipolar or a multipolar world in order to theorize and assess which is the most stable. While he and other authors, such as Deutsch and Singer (1964), disagree over how opportunities for interaction and the degree of certainty induced by the different polar structures affect stability, there became a near consensus among international relations scholars in their view of polarity as a categorical measure.

This understanding of polarity as a categorical measure affects the way system structure is viewed and measured. To fit the typological use, polarity came to be understood as a number rather than a degree. Most international relations scholars rely solely on the relative power of

states to determine international structure, bounding potential explanatory mechanisms to those accounted for by the number of poles in the system (Deutsch & Singer 1964, Waltz 1964, Rosecrance 1966, Wolfers 1962, Knorr 1966, Morgenthau & Thompson 1985). Essentially, these scholars consider only the vertical domain, without regard for the horizontal. In fact, Kersmaeker notes that by the end of the Cold War, the categorical understanding of polarity was so dominant that Wayman and Morgan’s horizontal concept of polarization had all but “disappeared from the research agenda” (Keersmaeker 2017, 19).

It could be argued that this uni-dimensional understanding of polarity assumes power is an indicator of a misalignment of interests. According to this line of thinking, power operationalizes both domains, being descriptive (vertical domain) and resulting from relationships (horizontal domain). For centuries before the idea of polarity was introduced to international relations, realists relied on the ‘balance of power’ concept to explain system structure (Buzan 2004, Wagner 1993). According to this theory, states gained power or formed alliances for the explicit purpose of contesting other states that they perceived as powerful. These alliance-induced shifts in power carried intent; they were either for the purpose of revising the status quo or for resisting those who wished to do so (Waltz 1979, Walt 1985, Morgenthau 1985). McGowan and Rood make this clear regarding alliances, stating that the balance of power system depends on alliances being “fluid and flexible,” to the point of ignoring “ideology” and even being “independent of past alliances and alignment patterns” (McGowan & Rood 1975, 69). Waltz even criticizes research that considers the separate issue of alignment (the horizontal aspect of Ray’s structural description) claiming they are nonstructural (Waltz 1979, 41-49).

However, this realpolitik perspective that nations ally solely for the purpose of balancing power faces some problems. First, it is not indicative of the Cold War or post-Cold War eras where alliances are based more on alignment than specific threats, making them more stable and durable (Kuo 2021, Wagner 1993). The fall of the Soviet Union is particularly problematic for the balance of power theory that supports the ‘power-only’ narrative of

polarity. Waltz and others who subscribed to this conception of polarity were initially convinced that the US becoming the sole superpower was a “unipolar moment” and that a power would soon rise to meet it. However, the theory has no explanation for America’s endurance as a sole superpower (Buzan 2004). This calls into question the idea that states gain power and form alliances only to oppose others. Second, even if we consider individual states instead of alliances, we see that a state’s increase in power is not exclusively militant. Other factors, such as their population, economy, and natural resources also increase a nation’s power (Fox 1944, 6-7; Organski 1968; Waltz 1979, 131, Singer & Small 1972, Beuno de Mesquita 1975). This implies that states may gain power out of a desire to prosper, without doing so to match a threat or with revisionist intentions. Therefore, the threat that great powers pose to one another cannot be based solely on the way in which power is distributed; it must also consider the degree to which they oppose one another’s interests.

While most have not considered alignment as a structural attribute, some scholars have (Singer & Small 1968, Haas 1970, Wallace 1973, Bueno de Mesquita 1975, Maoz 2006). However, most of these scholars have only considered alignment as an alternative to the power polarity described above. Rather than considering it as a horizontal domain to complement the vertical, they consider “cluster polarity” as a way to count poles and thus categorize the international structure. They examine the tightness of clusters (usually determined by alliance bonds) and the discreteness between clusters to delineate the poles. Wayman (1984) is among the few who consider both power and cluster polarity, but even he does so independently. He understands that the two measures could produce a different set of poles, where the clusters do not coalesce around the most powerful nations; however, he does not combine the measures in a way that demonstrates a single measure of tension that the poles produce.

In addition to categorization limiting the polar concept to a single domain, it also neglects one of the most descriptive aspects of the original analogy: the *degree* of polarity. When we only look at categories, such as bipolar and multipolar systems, we cannot answer the

question of ‘How polar is the system?’, or ‘How contested is the environment?’ Instead, we lose any variation within the time period. To answer these questions, we require a new concept and a new measure.

### 2.3 The Concept of Polar Tension

Therefore, I propose a novel concept, *polar tension*, which I define as *the degree to which international policy space is contested*. The term “polar tension” distinguishes it from the differentiated and familiar concepts of “polarity” and “polarization” discussed above while incorporating both ideas into a single construct. The concept captures the horizontal dimension of alignment in the relative positions of states’ interests in the international policy space while the vertical dimension of power is characterized by the ability of states to contest opposing policy preferences to achieve their own interests. Therefore, the high degree of polar tension in the system is created by two factors: (1) opposing interests (a large alignment gap), creating a *desire* to contest opposing policy interests, and (2) power parity (a small power gap), creating the *ability* to contest opposing interests. The tension that exists between such states is indicative of the degree to which each of their interests is threatened by these factors and could serve as a motive for state behavior.

The concept of threat perception is key to understanding states’ behaviors and to understanding why the analogy of polarity was originally used to describe the international system structure. Threats are not perceived merely from the power another nation possesses. Instead, they are perceived from a combination of that power and a sense of motivation or “intent” to use the power (Singer 1958, 93-94; Cohen 1979, 4-5). This intent does not need to be explicitly declared but may be implied by each nation as the other nation pursues interests counter to its own. Therefore, there is congruity between the concepts of system structure, consisting of the dimensions of *power* and *alignment*, and the concept of threat perception, consisting of *capability* and *intent*. As this perception or tension is generated from the ability and desire of major powers to oppose the interests of others, it creates a

situation in which international policy space is contested, or polar tension.

The dual aspects of this concept also find support in power transition theory, which holds that war is most likely to occur as a nation that is not satisfied with the international order seeks to surpass the dominant nation in power (Organski 1968, Organski & Kugler 1980, Gilpin 1981, Allison 2017). Kim (1991) develops a measure of satisfaction by examining the similarities between  $\tau b$  coefficients, which operationalize international policy preferences.<sup>6</sup> This implies that the “satisfaction” aspect of this theory is equivalent to the concept of policy alignment. The threat of conflict increases as a nation nearing parity and desiring opposing policy outcomes seeks to change the international order. Organski writes that nations that are “satisfied with the present international order and its working rules... feel that the present order offers them the best chance of obtaining the goals they have in mind” (Organski 1968, 366). These nations, even if they are rising in power, have no reason to threaten the interests of the dominant power. He illustrates this point, citing the peaceful power transition between Great Britain and the United States in the twentieth century when power parity was reached (and surpassed) without an alignment gap (Organski 1968, 362). The policy space, in this case, was never threatened. A dissatisfied nation, on the other hand, seeks to enforce opposing policies when it has the power to do so. The threat of conflict requires not only power parity (i.e. the capability to enforce policy) but also an alignment gap, implying the intent to challenge the existing international order.

In the context of international system structure, threat perception is not to be understood as concern over opposing specific or particular interests but over differences in policy preferences that are driven by deeper ideological distinctions (e.g. preferences relating to such interests as democracy, sovereignty, economic freedom, and human rights). Snyder draws a line between “particular” and “general” interests, describing the former as those that pertain to specific issues between states and the latter as more strategic in nature and not pertaining to any specific threat (Snyder 1984, 464). In this sense, system structure does not create

---

<sup>6</sup>Bueno de Mesquita (1975) also used the  $\tau b$  measure to calculate the degree of alignment within alliances.

the crises that bring nations into conflict with one another over particular issues. Rather, system structure may create tension between major powers as the perception of threats to their general interests leads them to prepare for such conflicts.

The international structure of the post-World War II era inherently induced a high degree of polar tension by this measure. There were two superpowers with diametrically opposed ideologies and interests that determined all international policy decisions. In fact, Noguee noted that a transition out of what was considered bipolarity was marked by the fact that global issues were no longer “resolved exclusively according to the priorities of the United States and the Soviet Union” (Noguee 1975, 1198-1199). What created tension in this relationship was not that the two nations existed as superpowers, but that both understood the threat the other posed in areas where their interests overlapped. Fox recognized this as the primary concern when introducing the concept. He doubted that either of the two powers would challenge the other directly, but that the combination of power and opposition may draw them into conflict if and when their interests did collide (Fox 1944, 100). The perceived threat drove their policies and behaviors in a way that affected the whole system. Any measure that is to capture the tension that exists due to polarity, or to answer ‘How polar is the environment?’, must feature both the capability (power) and intent (opposition) to contest international policy space.

Therefore, unlike the categorical measures of polarity in the context of the vertical domain mentioned above, polar tension exists on a continuous scale such that it describes the *degree* of tension or threat in the international environment. In other words, it is not simply which states have the most power that matters, but the difference in the power (power gap) and the difference in the alignment (alignment gap) that contribute to the degree of polar tension. This concept is displayed in a conceptual quad-chart in Figure 7.

In the following section, I discuss extant measures of power and alignment and then use aspects of those measures to operationalize the vertical and horizontal dimensions. I then combine them to create a measure of polar tension.

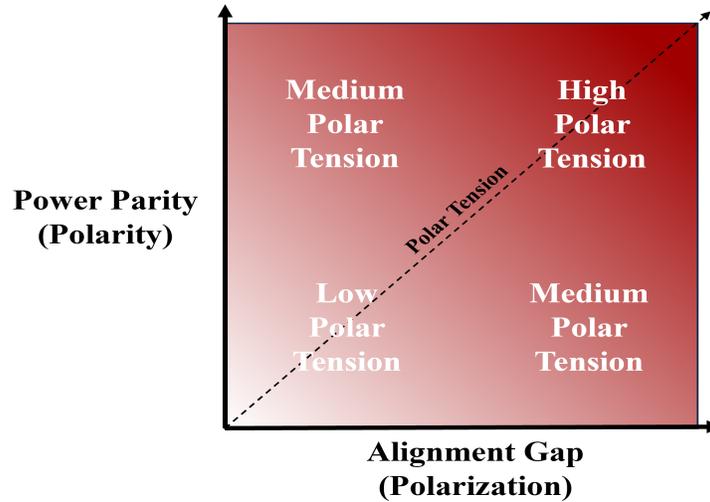


Figure 7. The Vertical and Horizontal Dimensions of Polar Tension

## 2.4 Operationalizing and Measuring Polar Tension

### 2.4.1 Defining the Actors

While the majority of scholars agree with Waltz and base polarity on the number of poles in the system, as mentioned above, there is variation in the way these poles are conceptualized and measured. Some scholars consider clusters, or blocs, of nations as the most significant actors in the system (Bueno de Mesquita 1975, Ronscrance 1963, Rapkin et al. 1979, Wallace 1973). This concept hinges on the idea that nations within an alliance give up their autonomy, so alliances act together and represent a single entity in the way they affect the system (Bueno de Mesquita 1975). Therefore, alliances are the most relevant actors composing the international structure (Deutsch & Singer 1964, Bueno de Mesquita 1975). While examining the tightness and discreteness of alliances may theoretically help us delineate one cluster from another, actual clustering by alliances has proven problematic when using it to determine the number of poles. Bueno de Mesquita (1975) uses alliances to cluster nations into poles and finds that there are significantly more poles in the twentieth century than in the nineteenth, a result that runs counter to intuition. Similarly, Maoz uses social network analysis to cluster alliances and measure polarity, and he concludes that the Cold War period “is remarkably

less polarized than the early and mid-19th century”, a result that also fails the test of face validity (Maoz 2006, 403). Both researchers found that the number of nations gaining independence and entering the system, along with several cross-cutting alliances frustrated this measurement and its utility. Furthermore, cluster power, when scholars are able to use it as a differentiating tool, only serves to count clusters and reveals nothing about the power gap between them or the extent to which they disagree on policy. To be used as a measure of polarization, one would have to make lofty assumptions that the clustered poles were equal in power and opposite in interests.

Rather than considering clusters as poles, most scholars recognize only great powers, often referred to as major powers, to be the poles (Deutsch & Singer 1964, Waltz 1964, Rosecrance 1966, Wolfers 1962, Knorr 1966, Morgenthau & Thompson 1985). They posit that the great powers, and not weaker states, are the determinants of change in the international system and are therefore the most significant. This assumption is illustrated by Snyder and Diesing, who considered non-major powers to “serve as objects of the Great Power competition,” (Snyder & Diesing 1977, 419-420). This is a reasonable assumption. Morrow finds that in asymmetric alliances, the weaker power exchanges its autonomy, or “the degree to which it pursues desired changes in the status quo”, for the security guarantee that the stronger nation provides (Morrow 1991, 908-909). Therefore, the weaker states make minimal impact on the system outside the interests of their stronger partners. For those considering states as the primary units in the system, using power to delineate the most significant and meaningful actors from the non-influential states has a strong theoretical consensus.

There is some variation in which major powers matter most to the system. Some envision all the major powers being significant, while others consider only the most powerful, or polar powers, to be the poles. This is why, even as Waltz recognizes multiple major powers during the Cold War era, he considers it bipolar, seeing only the US and USSR as polar powers (Wagner 1993). However, Mansfield (1993) points out that several studies have shown that non-polar major powers have had a significant impact on international policy outcomes.

Focusing only on the superpowers during the post-World War II era, and dismissing other major powers, such as Great Britain, France, and China, who have permanent seats at the UN Security Council, overlooks forces that could potentially increase or mediate the tension and hostility in the international system.

For this reason, I use all the major powers in the international system as the primary actors affecting polar tension. The first step in calculating polar tension is determining which states qualify as major powers. These should be the states whose actions have the greatest potential impact on the system and those who have the potential to contest policies they oppose. To determine which nations qualify as *major powers*, I use the Correlates of War (COW) Major Power data set (COW State System Membership 2017), which lists the years in which any nation qualifies as a major power in the international system. The measure acknowledges the general consensus among scholars that certain nations hold military, political, and financial power. Since the end of World War II, this list has consisted of the permanent members of the UN Security Council. These include the United States, United Kingdom, France, Russia (Soviet Union), and China, with China only being considered a major power after the communist party's victory in 1949, though it did not have voting rights in the UN until 1971 when the Peoples Republic of China replaced the Republic of China. Two more states, Germany and Japan, were recognized as major powers in 1991, after the end of the Cold War, due to their political and financial power.

Because these major powers were selected somewhat arbitrarily by experts, without any truly objective measurements, I compared this list to a list prepared using a power index I introduce later in this paper. From 1946-1991, the COW recognizes 5 nations as major powers. The 5 most powerful nations in the world were identified by the COW's index consistently from 1950 to 1960 when West Germany entered the top five. Japan surpassed one of the COW's listed major powers in 1970, and the two powers remained among the most powerful until they were recognized as great powers in 1991 when these nations were less constrained by Cold War politics. This expanded the COW major power membership from

5 to 7. The COW again accurately identified the most powerful 7 nations until 2013 when India became one of the strongest states. While there may be merit to including India, I elected to use those states that are generally recognized as the most influential by the COW, understanding that the subjective list has objective merit.

#### **2.4.2 Measuring Power and Parity**

After establishing the major powers, I turn to determining the vertical aspect of structure by estimating their relative power. Power within the international system is viewed systematically rather than dyadically and is consistent with the neorealist view, which considers the national resources of the state that could potentially enable a nation to successfully pursue its interests in the international system (Waltz 1979; 131, Bueno de Mesquita 1975, 202-203, Organski 1968). Comparing the power of states allows us to determine which ones have the greatest ability to successfully pursue their own international policy interests.

There are a couple of issues with the way scholars have viewed power in the past. The first pertains to power distribution. Whether considering clusters, all major powers, or only polar powers to be the most significant, most scholars conceive of the power being roughly distributed evenly between the poles (Mansfield 1993, Snyder & Diesing 1977). This means that these authors dismiss the dynamics that may occur due to major powers gaining or losing power relative to one another. It also assumes parity between the strongest (polar) powers. Despite the prevalence of theories that assume power is distributed evenly among the great powers, this is hardly the case; a significant power gap existed even between the US and USSR throughout the “bipolar” Cold War period (Mansfield 1993). To account for this power disparity, Mansfield (1993) borrows from economics literature on market forces to contrive a polarity measure that includes all major powers and their power inequalities. He views the important aspect of power to be its ‘concentration,’ which emphasizes the imbalance of power between major powers and allows for the influence of non-polar major powers. If these states are not considered, and even distribution of power is assumed, then

one would assume a maximum degree of tension in the system during a time of bipolarity, such as the Cold War. However, if we consider that one polar power may be stronger than the other and that other states possess a degree of influential power as well, then we see that power is not strictly concentrated at the extremes. Additionally, changes in power that may indicate an impending shift may be detected when considering changes that occur in levels of power. This is a much different, and clearer, picture of polarity, even though both perspectives would be considered the Cold War period to be “bipolar”, categorically.

The second issue pertains to the way in which power is measured. One of the most prolific systematic measures of power is the COW National Material Capabilities dataset, which defines power as “the ability of a nation to exercise and resist influence,” and includes three power indicators: demographic, industrial, and military (Singer 1987, 2). These indicators are captured into the Composite Index of National Capability (CINC) and scored by aggregating six measures of national material capabilities (two for each of the indicators listed above) into a single value (0 - 1) per year, according to the portion of the global capabilities each possesses. There are several issues that scholars have noted in using this index.<sup>7</sup> Beckley (2018) highlights one important point, in particular, emphasizing how the CINC score inflates the power of highly populated but inefficient countries by only taking gross figures for production rather than net. Using case studies of extended major power rivalries, large-n data predictions of war outcomes, and the retesting of previous studies, he demonstrates that using Gross Domestic Product (GDP) multiplied by the Gross Domestic Product Per Capita (GDPPC) as a measure of net power is superior to using CINC. Beckley’s net power indicator (GDP X GDPPC) is 8% better at predicting war outcomes (Beckley 2018, 38).

Beckley’s net power indicator accounts for some shortcomings in CINC’s demographic and industrial measures; however, it does not account for any military measures. Therefore, I develop a hybrid measure, which includes the two military CINC components (the

---

<sup>7</sup>Carrol and Kenkel (2019) present an alternative to the CINC model that improves its predictive power; however, it is designed for dyadic predictions. Their model could predict that A defeats B, B defeats C, and C defeats A. While this is useful, it does not provide a strictly spatial power ranking for nations relative to all others in the system.

number of military personnel (*milpers*) and military expenditures (*milexp*)), but replaces the demographic (total population and urban population) and industrial (iron and steel production) measures with Beckley’s net power indicator (GDP X GDPPC). This captures the advantages of using Beckley’s efficiency index and not over-inflating the gross measurement aspects that were previously noted while maintaining a very important component of power perception, the military. To calculate this hybrid measurement of power, I start with the military power from the CINC index. Just like the calculation in CINC, I take each military measure (*milexp* and *milper*) as a portion of those of all the major powers. I then average the measures by adding them and dividing by two. Therefore, the equation is

$$M_{it} = \frac{\frac{milexp_{it}}{\sum_i^n milexp_{it}} + \frac{milper_{it}}{\sum_i^n milper_{it}}}{2}, \quad (6)$$

where  $M$  represents military power,  $i$  represents each major power, and  $t$  represents a given year. The next calculation is Beckley’s net power. Using the Varieties of Democracy (V-Dem) dataset, I take the GDP of each major power in a given year multiplied by its GDPPC and then divide by the sum of this measure of all the major powers to arrive at the net power (Coppedge et al. 2022). The equation is

$$B_{it} = \frac{(GDP \times GDPPC)_{it}}{\sum_i^n GDP \times GDPPC_{it}}, \quad (7)$$

where  $B$  indicates Beckley’s net indicator. I then weight the military aspects equally with Beckley’s index to arrive at the hybrid power measurement,  $P$ , as seen below.<sup>8</sup>

$$P_{it} = \frac{M_{it} + B_{it}}{2} \quad (8)$$

Since this measure has not been used previously, I test its validity against both the

---

<sup>8</sup>The weighting is somewhat arbitrary; however, when the leaders of one nation estimate the power of another, it is reasonable to believe that their consideration of the other’s relative immediate military capabilities and its potential to sustain conflict may be weighted differently depending on the context and idiosyncratic details of a possible conflict. Without any further consideration to drive the weighting in either direction, I default to equal weighting.

CINC measurement and Beckley’s measurement using one of the same tests Beckley used to demonstrate the superiority of his measurement, a large-n data prediction of war outcomes. The results, shown in Table 1, demonstrate that the hybrid power measure out-performs the other measures by 4.6 – 5.6%, in predicting war outcomes since 1816 and is just as good as Beckley’s model at predicting militarized interstate disputes.

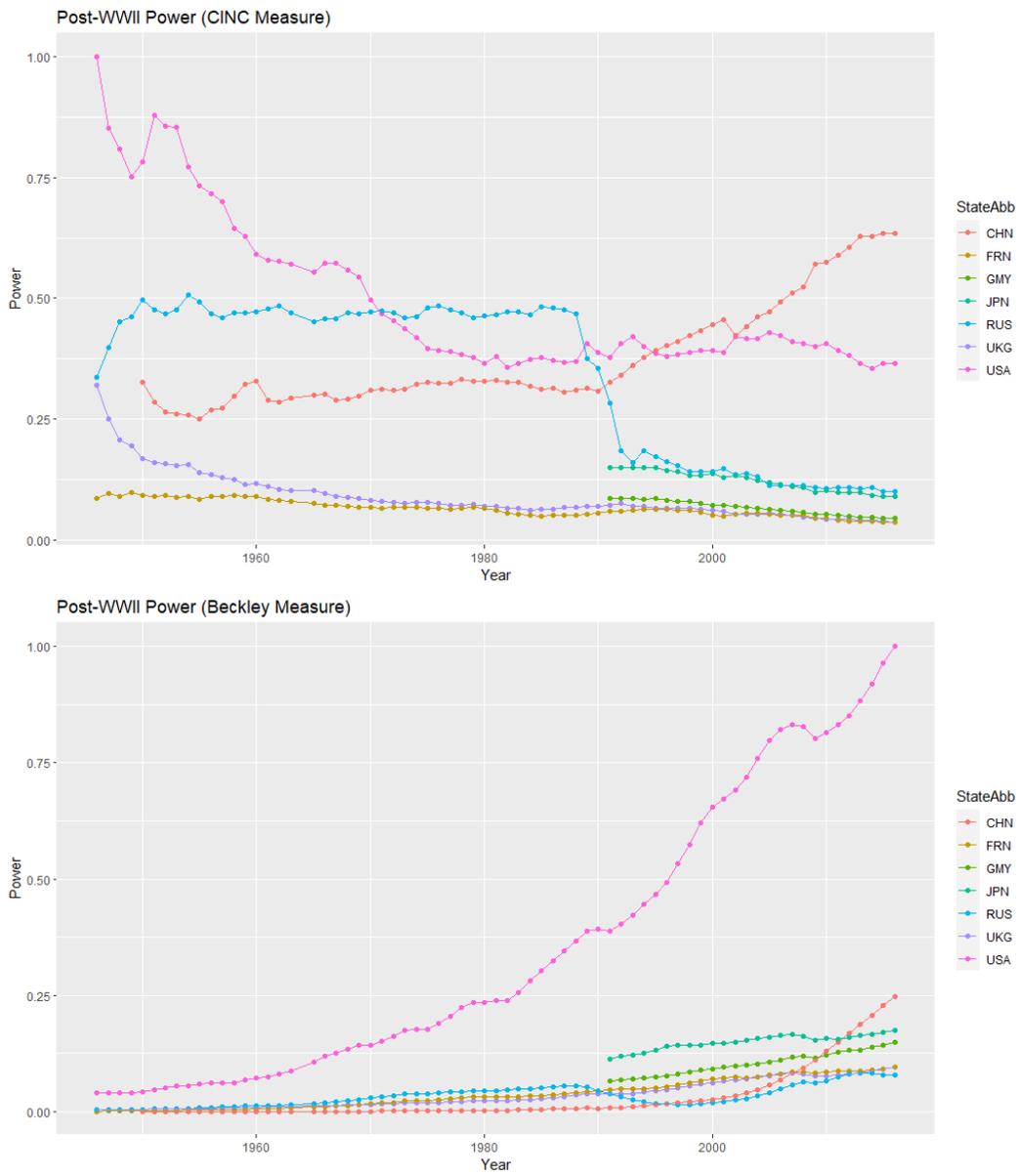
**Table 1.** Power Measures as Predictors of War and Dispute Outcomes, 1816-2010 (percentage of outcomes predicted correctly)

Measurement Method	Wars	MIDs
CINC	71.7%	63.4%
Observations	46	265
GDP x GDPPC	72.7%	66.3%
Observation	44	255
Hybrid	77.3%	66.3%
Observations	44	255

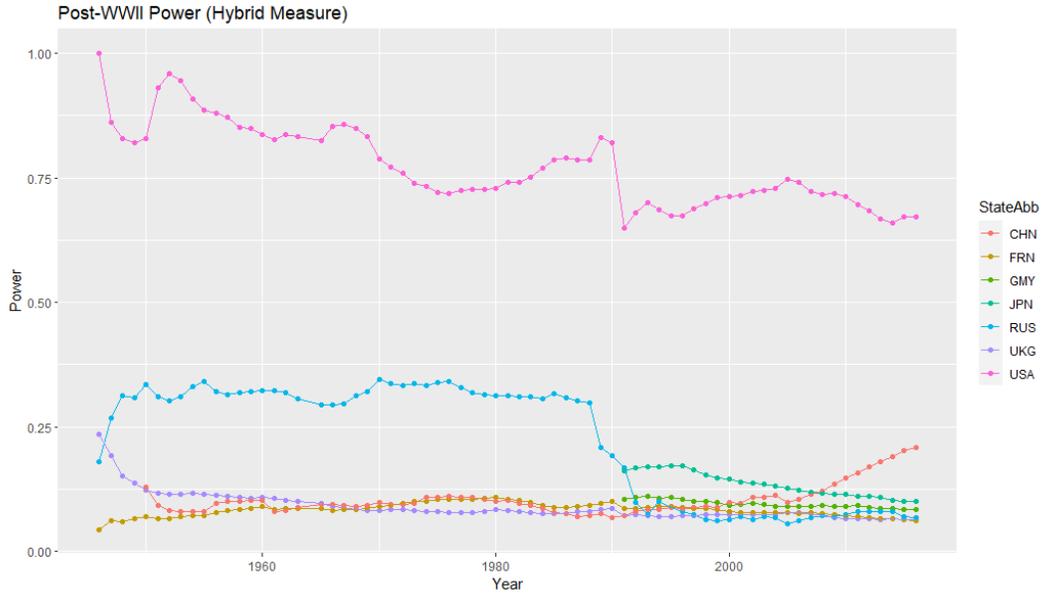
For a visual comparison, Figure 8 displays the power of each of the major powers from 1946 to 2016, as measured by the CINC index and Beckley’s net power index, respectively. While the CINC scores show the sharp drop in Soviet power at the end of the Cold War, as we would expect, it also indicates that the US’s relative power has been in near-constant decline since the early 1950s and that China’s power surpassed that of the US in the mid-1990s. The power inflation issue the Beckley pointed out with this measurement reveals itself in the lack of face validity demonstrated by this model.

However, Beckley’s net power index also fails face validity. Counter to the CINC model, it shows that US power has grown in a near-exponential manner since the end of World War II relative to the other great powers, while Soviet/Russian power was never comparable to it. In fact, there is very little variation in the Soviet/Russian power according to this index, even during the fall of the Soviet Union.

By contrast, the hybrid power index, displayed in Figure 9, clearly demonstrates face



**Figure 8.** CINC Index and Beckley Index as Measures of the Power of the Major Powers in the International System



**Figure 9.** Power using the Hybrid (milexp, milper, and GDP x GDPPC) measure

validity in that the US and the Soviet Union were the two leading powers during the Cold War and that Soviet/Russian power dropped significantly from 1988-1992, during the break up of the Soviet Union. The drop in US power in 1991 is due to the relative nature of introducing two new powers into the system (Germany and Japan). The graphic also displays the rise in China’s power over the last decade.

There are other aspects of military power that are not captured in this data set, such as military proficiency, leadership, and training; however, these factors are generally not known between states and therefore do not affect their perceptions of one another. The dataset does provide knowable information about not only the initial capabilities a nation can bring to bear through its current military personnel and equipment (approximated by expenditures), but also is capacity to grow and sustain forces. It also demonstrates a degree of economic power in the GDP aspect of the measure. Therefore, while not a perfect measure of power, this hybrid index provides an adequate estimate of how states may assess one another’s power.

To determine how power translates into a state’s ability to successfully pursue its own policy interests or to oppose another’s, we must calculate how each state’s power stacks

up against the strongest state's. If no state is comparable in power to the strongest state, then the structure, which would be considered hierarchical or unipolar, is such that policy is not contested, and therefore, there is low polar tension. If there is a state (or states) that are similar in power to the strongest, then there is potential contestation, depending on the alignment of the powers. Therefore, we must measure the power parity between the strongest state and the other major powers in the system.

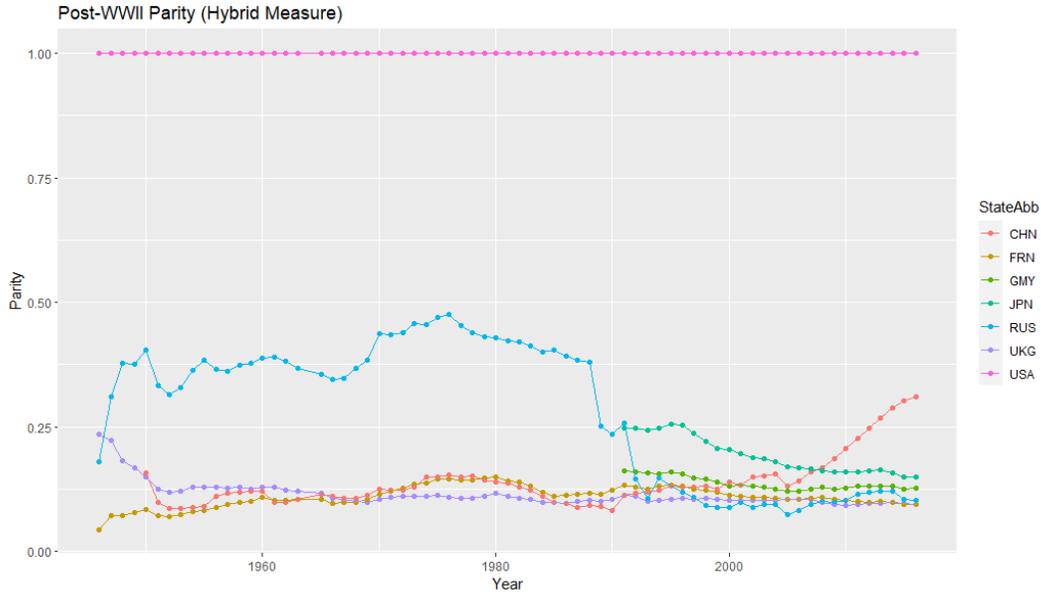
Having a power score for each nation, I can measure the parity between the nations by dividing each major power's power score,  $P_{it}$ , by the maximum power score in the year,  $P_{it\ max}$ , which is always the United States in the given timeframe. Therefore, the parity equation is

$$PAR_{it} = \frac{P_{it}}{P_{it\ max}} \quad (9)$$

The resulting parity is displayed in Figure 10. Not surprisingly, the greatest parity was with the Soviet Union (Russia) until its fall between 1988 and 1991. There is a rise in that parity in the 1970s as the USSR continued its trend of yearly increases in military spending while also increasing its troop levels. There has been a much lower degree of parity since then; however, it is worth noting that Japan entered the group of major powers during an industrial boom, giving it the greatest parity and potential to rival US policy. Figure 10 also reveals China's growth over the last decade, giving it the greatest potential to oppose US interests.

### 2.4.3 Measuring Alignment

As noted previously, power parity only raises the potential for contested policy space. Even if nations have the ability to contest the strongest nation's policies, their desire to do so depends on the extent to which their own policies are threatened. Therefore, after establishing the power parity of the major powers, I assess the degree to which each major power aligns with the strongest power. While the majority of researchers have focused



**Figure 10.** Parity between the Major Powers (Greatest Power [US] = 1.00)

primarily on the vertical structural aspect of power to define polarity, several have recognized the need for a measure of opposition, or alignment, and even a need for a measure of the degree of polarization (as opposed to the typology of polarity) (Wayman 1984, Bueno de Mesquita 1975, Rapkin et al. 1979, Noguee 1974, Buzan 2004). Buzan even notes that “Polarity logic plays different in a system dominated by friends than in one dominated by enemies” (Buzan 2004, 78). There have been several attempts to measure the concept of alignment, such as measuring joint participation in Inter-Governmental Organizations (Wallace 1975), measuring alignment by trade flows (Alker and Puchala 1968), and using an indicator based on issue distance between foreign policy actions flows (Altfeld 1984). Some of the most significant studies that have attempted to capture alignment have considered greater polarization to occur when there is greater tightness within alliances and greater discreteness between them (Bueno de Misquita 1975, Maoz 2006). However, even with the recognition that alignment is an important factor, these measurements have been primarily for the purpose of identifying groups, or clusters of nations in order to again identify the number of poles.

One of the first and most influential measures of alignment was introduced by Bueno de

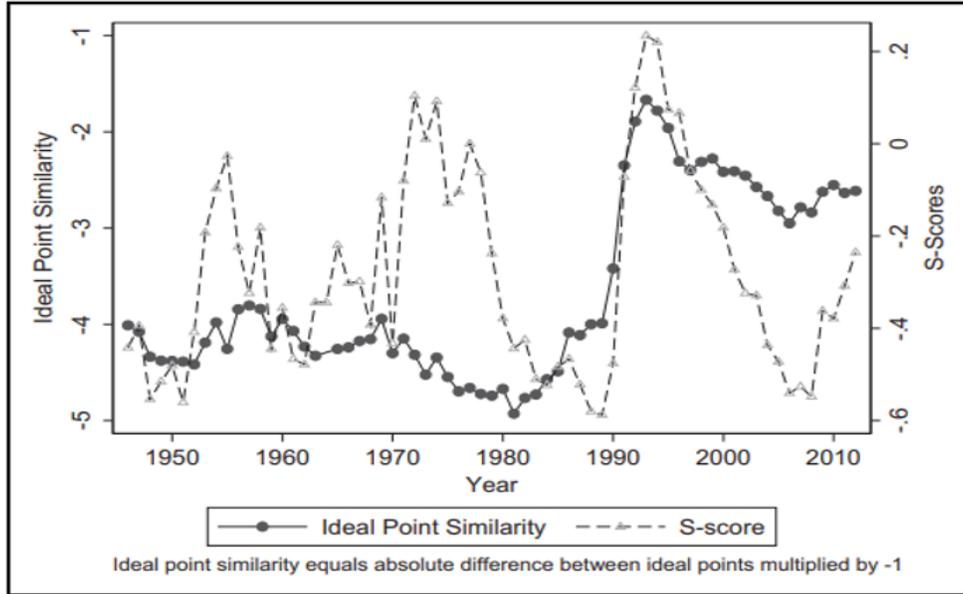
Mesquita (1975). As a method for identifying alliance clusters, he proposes an indicator of alignment on policy issues based on the dyadic similarities between the alliance portfolios of two countries. He creates a matrix for each dyad of nations in the system, comparing each of their alliance types (defensive, neutrality, entente, or no alliance). He then uses Kendall's rank correlation measure (Kendall 1938),  $\tau b$ , to rank them (on a scale from 1.0 (all aligned) to -1.0 (none aligned)). He then makes another matrix of the  $\tau b$  scores and uses typal analysis to group them into clusters. The assumption is that the tighter the cluster or the more common allies a pair of nations has, the more they align on international relations issues. One of the weaknesses of this measure is that it fails to account for *why* nations are not allied; these reasons could range from nations being hostile to one another to nations being so close that no alliance is needed (Signorino & Ritter 1999). Another weakness is that it assumes alignment within alliances without addressing any particular issues.

A second prolific measure of alignment, the S-score, was created with the intent of shoring up these weaknesses in measuring alliances. Rather than basing alignment solely on an alliance, the S-score is a measure intended to accommodate other available and relevant data. More specifically, Signorino and Ritter base the measure on the dyadic similarity or difference in vote choices in the United Nations (UN). Rather than examining an alliance portfolio and assuming it represents policy alignment, as Bueno de Mesquita did, by examining UN General Assembly (UNGA) votes, these authors consider the "*policy portfolio*", which represents "a point in (foreign) policy space" (Signorino & Ritter 1999, 126). They consider how close the states are to one another in actual policy positions to determine their similarity (S), or alignment, regardless of their alliance status. While this system is generally seen as an improvement over  $\tau b$ , and has been used extensively, it also has some important limitations. For example, it does not weight any issues over others, it doesn't consider the number of redundant votes on similar issues, and it does not account for agenda changes in the UN.

Recently, Bailey, Strezhnev, and Voeten (2017) produced another measure that can be used to estimate alignment. Much like the S-score, they calculate ideal points by examin-

ing each UN member's desired international policy based on UN General Assembly voting. Rather than using a dyadic comparison between states to determine their alignment, this dataset offers a single score per nation per voting session, which occurs once per year, by determining where their votes fell along a single dimension of policy space based on the US-led liberal international order. They accomplish this by categorizing the issues within each measure, accounting for UN agenda changes and polarity shifts in the wording, and by tracking the consistency through identical measures proposed multiple times across different sessions. These follow spatial theory and Item Response Theory models that help establish cut points and use votes to determine latent preferences (Poole 2005). This overcomes weaknesses in the S-score model and gives an accurate account of how each vote reflects a nation's policy preferences.

The liberal order used in this scale is established along three lines: democratization, financial liberalization, and government ideology. The ideological policy space related to the liberal order has a high correlation to the most significant and relevant international policy issues, such as democracy, human rights, colonialism, and trade and economics (Bailey et al. 2017, 444). Furthermore, the authors demonstrate the ideal point's measurement superiority to the S-score by examining which one reflects a more accurate representation of the relationship between the United States and the Soviet Union during the Cold War. They included a graphic, shown in Figure 11, that compares the two measurements. The S-score shows a great deal of variation in the alignment between the two nations during the Cold War period, even indicating times of significant alignment in the mid-1950s and throughout the 1970s. Bailey et al.'s ideal point similarity model, on the other hand, by taking the absolute difference between the two nations' ideal points, indicates a consistently broad gap between the two prior to the fall of the Soviet Union. The face validity is further demonstrated by examining the changes in ideal points of each of the major powers. Figure 12 depicts these ideal points. The US has consistently been at or near the top, leading the liberal order while Great Britain and France have also repeatedly been near the top. The Soviet Union owned

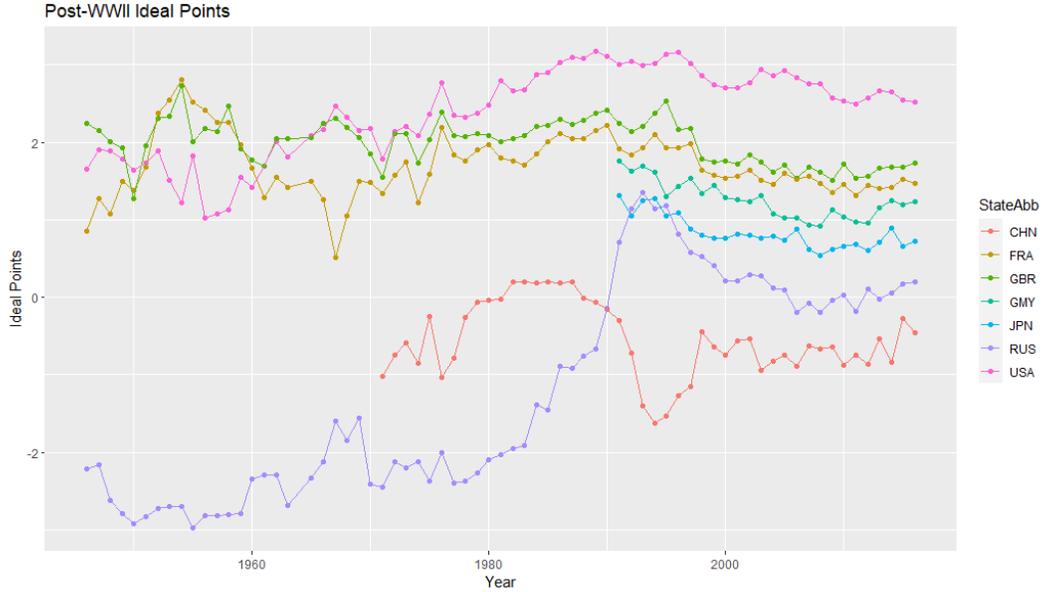


**Figure 11.** UN voting similarity between the USSR/Russia and the US (Bailey et al. 2017, 439)

the lowest ranking prior to its fall, and then China became the least liberal-leaning nation in its international policy preferences. The coverage of salient international policy issues and the measurement improvements over previous methods make this measurement an effective proxy to estimate the dyadic alignment of nations.

Bueno de Mesquita rejected the use of UN votes as a measure of alignment because of the equal weighting of all issues, regardless of importance, and because of the short period of time they covered (Bueno de Mesquita 1975, 193). However, the spatial model introduced by the ideal point measurement resolves his first critique. Additionally, nearly fifty years have been added to the data since he made his assessment, giving the measure greater analytical potential. Therefore, I use the ideal point (IP) measurement to calculate the horizontal structural aspect of the alignment gap between the great powers by measuring the difference in their ideal points.

I start by determining which nation is the most powerful in the system, following Organski's notion that the existing international order is largely set by the dominant nation in the system (Organski 1968, 366). This makes the most powerful nation the most relevant

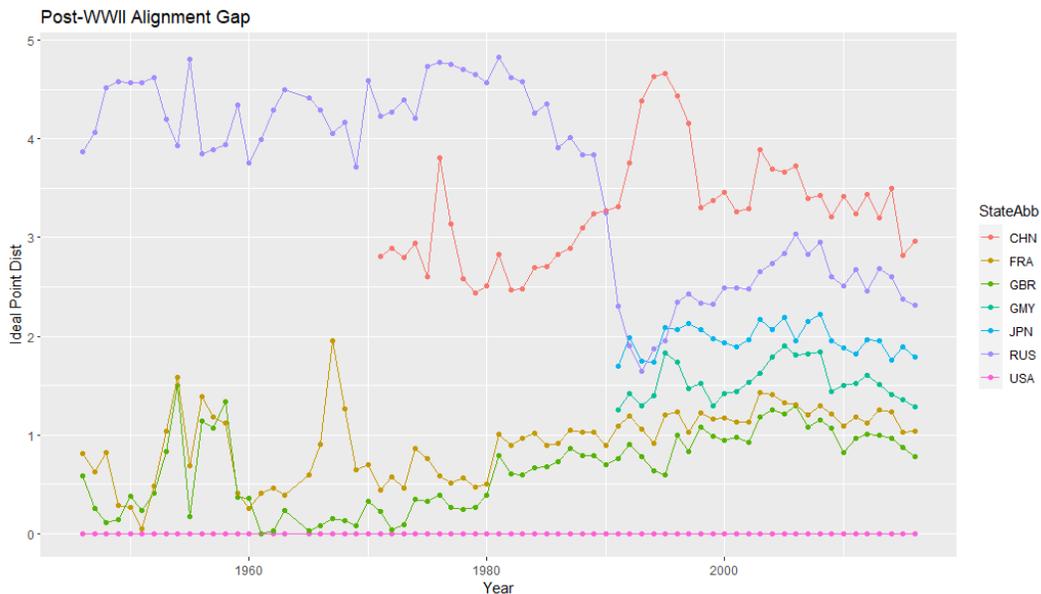


**Figure 12.** Ideal Points for each of the Major Powers since World War II

to the polar tension. Therefore, I assess the policy preferences of the other major powers relative to those of the dominant nation. The absolute value of the difference between the ideal points of the most powerful major power and each of the other major powers gives us their positional distance from one another.

Since the concept of polar tension does not consider the nations as poles, or extremes, it is important that we do not measure them as such. The values exist in policy space, between the two poles, or extremes of the policy space. Therefore, instead of normalizing the positional difference between the strongest major powers and the other major powers by dividing them by the maximum difference, it makes theoretical sense to divide the value by the distance between the poles. However, Bailey et al.’s scale does not include a minimum or maximum degree of inclination toward a liberal world order. To overcome this, I use the minimum and maximum positions that any nation has taken over the entire time period to define the poles, which I label as  $D_{ij\ max}$ .<sup>9</sup> Therefore, the distance,  $D$ , between the strongest state,  $i$ , and each of the other major powers  $j$ , in a given year,  $t$ , is written

<sup>9</sup>The ideal point dataset uses comparable measures, year-over-year. This allows me to use the minimum and maximum ideal points over the breadth of the dataset to determine the lower and upper boundaries, or “poles.”



**Figure 13.** Ideal Point Distance between Major Powers (Greatest Power [US] = 0)

$$D_{ijt} = \frac{|IP_{it} - IP_{jt}|}{D_{ij\ max}}, \quad (10)$$

where  $IP$  is the ideal point. These distances are displayed in Figure 13.

Having already established that the United States is the most powerful country throughout the timeframe, it is therefore the baseline from which the distance is measured. Again, we see that the Soviet Union presented the greatest opposing interests to the liberal order until its fall. Immediately after the Soviet collapse, China moved significantly farther from the liberal order to become the most opposed major power. It is also worth noting that since the end of the Cold War, there has been less spatial distance between the positions of the major powers overall. We see that the Western powers of Great Britain and France had begun moving away, albeit not significantly, in the 1980s and have continued to support less liberal interests than the US since.

#### 2.4.4 Measuring Polar Tension

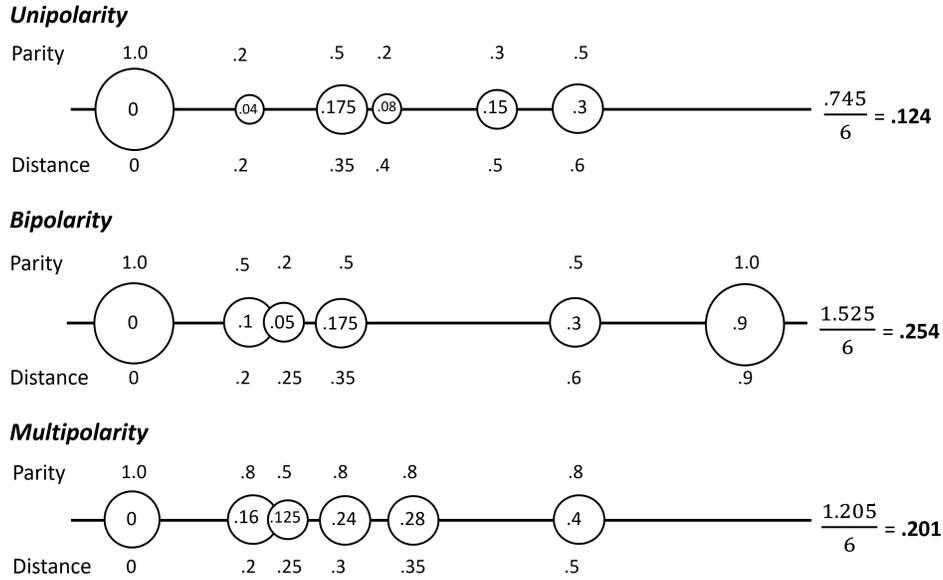
I combine the values for power parity and spatial distance (or the alignment gap) to obtain a measure of polar tension, which assesses how power is distributed over the continuum of desired international policy outcomes.

To arrive at a single value for polar tension for a given year, I first take the sum of the parity scores from Equation 9 multiplied by the alignment difference score from Equation 10. I then divide these scores by the number of major powers,  $N$ , to estimate polar tension,  $POL_t$ .

$$POL_t = \frac{\sum_i^n (PAR_{it} \times D_{it})}{N_t} \quad (11)$$

This method allows us to estimate a nuanced degree of tension regardless of whether the environment is considered to be unipolar, bipolar, or multipolar. Figure 14 illustrates this point by depicting a representation of each of the categories as most scholars tend to envision them and estimate their discrete polar tension scores. Along the top of each system are parity scores. Along the bottom of each depicted system is the spatial distance. The first system, labeled “Unipolarity”, has one polar power with the other weaker major powers at various distances from it. The system labeled “Bipolarity” has two evenly matched polar powers with near maximum spatial distance between them. Finally, the system labeled “Multipolarity” has several near-evenly matched polar powers with various levels of alignment between them. To the right of each system are the results given by calculating Equation 11, which reveal the polar tension score for each system. This measure of polar tension is consistent with the amount of contested policy space we may expect, given the type of system. Here, the bipolar system represents the one in which policy space is most contested and the unipolar system represents the order with the least contestation.

To demonstrate how variation within a single categorical period might exist, I examine three hypothetical systems, displayed in Figure 15, that those who count polar powers would



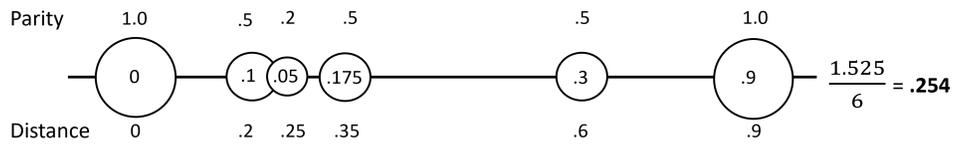
**Figure 14.** Hypothetical Polar Tension scores comparing typical Polarity categories

consider to be bipolar. This figure displays the impact that the ideological alignment gap has on polar tension. It shows the same system in three different ways, only changing the spatial distance between the polar powers. We see that there is a great deal more polar tension in System A when the polar powers have greater differences in their policy preferences, than in either of the two other systems displayed. System C demonstrates that there can be very low polar tension in a bipolar system if the two polar powers are closely aligned. This demonstrates the importance of including both the vertical and horizontal dimensions of the international system structure when assessing the tension that can rise, even in the same categorical era.

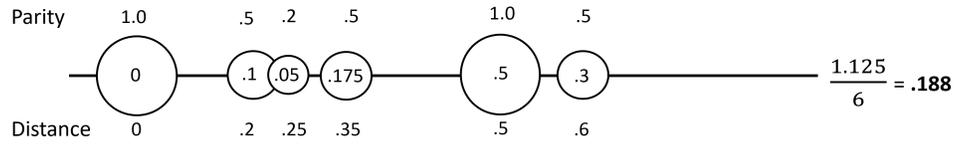
For further illustration, I map the actual calculated average parity (vertical domain) and average alignment gap (horizontal domain) for each year onto a Cartesian plane (displayed in Figure 16). This allows us to see the distribution of the component parts of polarity—power parity and the alignment gap—over time. This figure demonstrates that there is variation in both dimensions, and it shows that at times when both aspects were high, the resulting polarity was at its highest.

To demonstrate what this looks like in a given year, I map actual parity and distance from

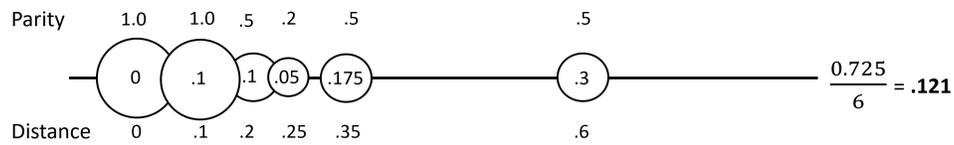
**System A**



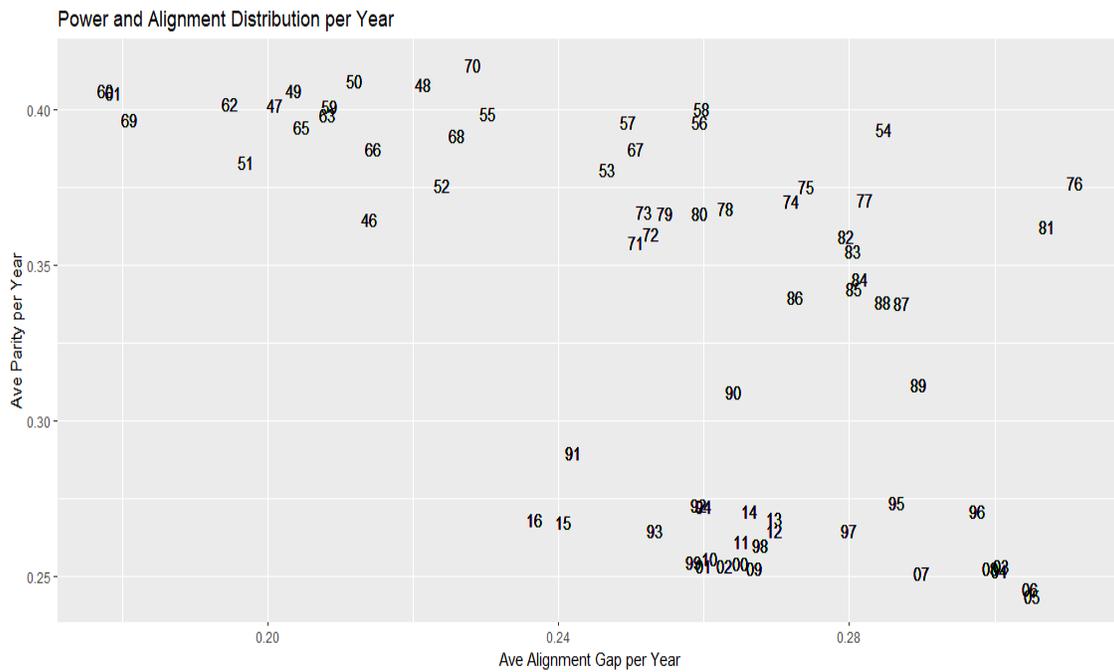
**System B**



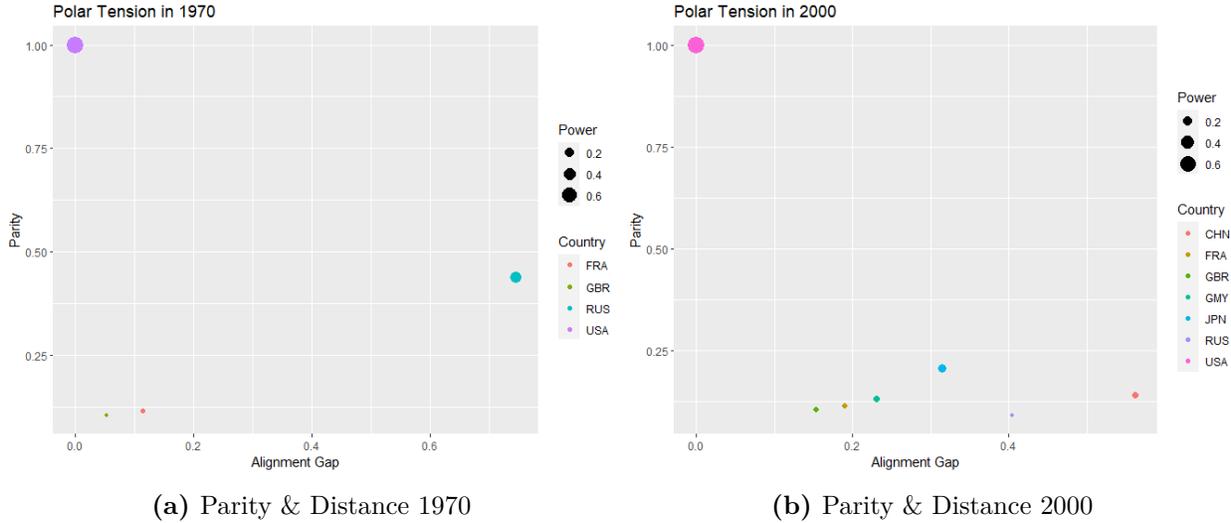
**System C**



**Figure 15.** Polar Tension Variation in Bipolar System based on Spatial Distance



**Figure 16.** Polarity Components (Parity and Alignment Gap) per year

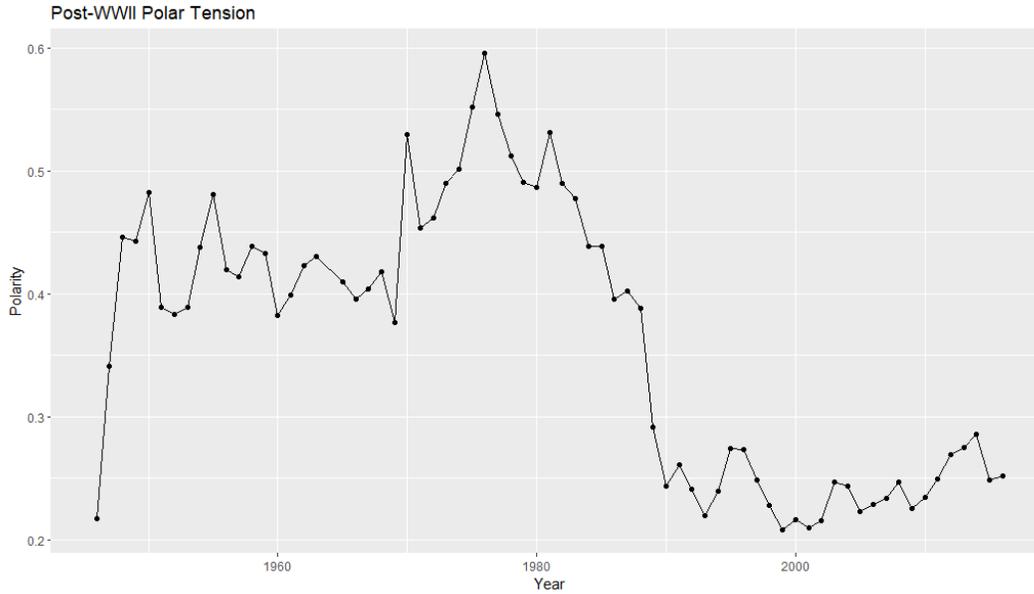


**Figure 17.** Comparing Polar Tension in ‘Bipolarity’ (1970) and ‘Unipolarity’ (2000)

a time of bipolarity (1970) and a time of unipolarity (2000) onto a Cartesian plane as well. Figure 17 illustrates how both the parity and alignment gap play into polar tension. The 1970 plot shows the United States positioned at the top left as it is the baseline for parity and alignment. The Soviet Union (labeled RUS) represents the greatest parity, but significantly, it also has an alignment gap above 0.6. The other major powers have very little parity and are closely aligned with the US. The USSR is therefore in a position to contest US policy. In 2000, on the other hand, we see that China has a significant alignment gap, positioned well above .5; however, it has little parity with which to challenge any policy discrepancies with the US. Japan shows the greatest parity in 2000 but is significantly more aligned with the US than China, giving it less desire to challenge policy. This comparison demonstrates the way in which both parity and alignment combine to determine polar tension.

## 2.5 Assessing Polar Tension

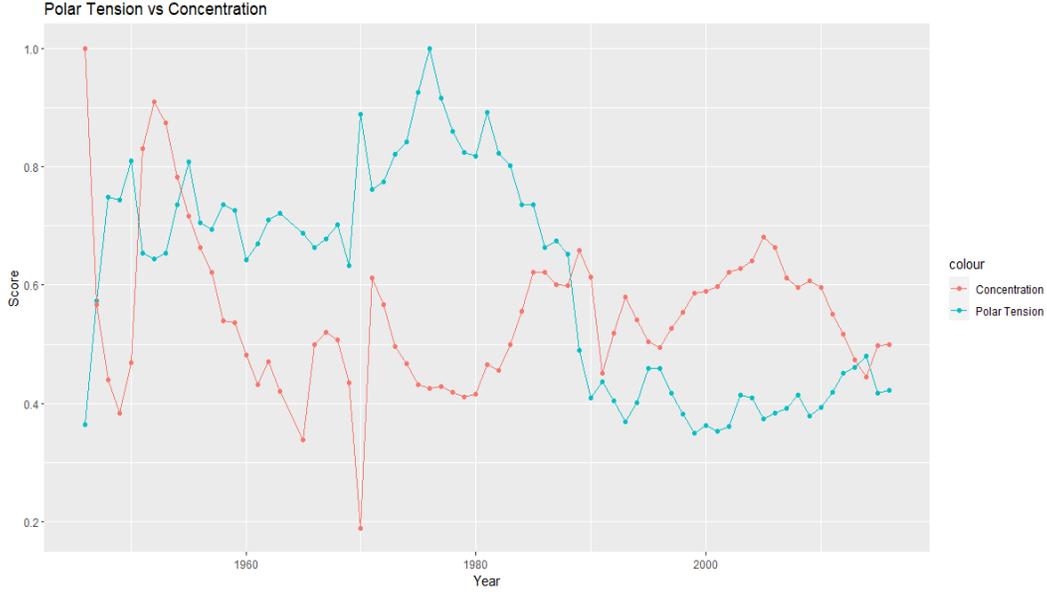
The polar tension scores for the 1946 to 2016 timeframe are displayed in Figure 18. The figure reveals face validity in that it is consistent with what we would expect the tension to be over the seventy-one-year period. We see an extreme increase in polar tension at the beginning of the Cold War and an extreme decline in polar tension between 1988 and 1991,



**Figure 18.** Polar Tension in the International System

which corresponds to the fall of the Soviet Union. The increased tension in the 1970s, on the other hand, is somewhat counter-intuitive, considering this was a time of détente between the United States and the Soviet Union. However, we could also consider that détente is most necessary and likely in cases where tensions are highest. In fact, the US-Soviet détente was a response to the theory of “mutually assured destruction” (MAD), which held that the tension between the nations combined with their nuclear arsenals threatened the existence of both nations. Furthermore, President Reagan stated that “détente’s been a one-way street that the Soviet Union has used to pursue its own aims” (Reagan 1981). The increased parity during this time, revealed in Figure 10 and caused by the Soviet’s continued increase in military expenditures and personnel, gives legitimacy to the President’s claims. Other events during this time, such as the Communist Vietcong capturing South Vietnam and the US normalizing relations with China, may also have contributed to greater misalignment between the superpowers.

To further validate this measure, I compare it to one of the only continuous measures of polarity, the power concentration method used by Wayman (1984) and Mansfield (1993). They used the measure to evaluate the difference in stability (likelihood of war) during



**Figure 19.** Comparing Concentration Scores to Polar Tension

times of bipolarity and multipolarity. Unfortunately, this was a timeframe that the measure of polar tension does not cover. However, I use the concentration equation,

$$CON_t = \sqrt{\frac{\sum_i^{N_t} (H_{it})^2 - 1/N_t}{1 - 1/N_t}}, \quad (12)$$

to measure the polarity within the same timeframe covered by the polar tension measure and then compare its face validity. Figure 19 reveals that the power concentration measure,  $CON_t$ , is somewhat sporadic and is inconsistent with common intuition about the tension involved in the international environment over this time. The polar tension measure is overlaid on this figure for the purpose of comparison. The concentration measure lacks the face validity possessed by the measure of polar tension.

## 2.6 Conclusion

This novel measure of polar tension focuses on an aspect of the international system structure that has been overlooked or assumed by scholars who perceive of a categorical concept of polarity. Combining the vertical and horizontal aspects of system structure,

which correspond to the components of threat perception, reveals the amount of hostility in the global environment. Conceptualizing this hostility as the contestation of policy space allows us to identify and measure the motivation for the animosity. The power that others have relied on to describe polarity acts as an enabler for those nations that have opposing policy preferences to promote their own policies while opposing the others'. The combination of power and alignment paint a clear and nuanced picture of tension and help answer the question, 'How threatening is the environment?'.

Additionally, the measure provides an annual assessment of the polar tension, so that assessments of variation in polar tension can be made within the classical categorical eras. The measurement shows a great deal of variation in the degree to which international policy space is contested within the Cold War and post-Cold War period, not just between them. This intra-era variation can be used to assess the effects of rising polar tension on various outcomes.

The measurement also has some limitations. While the timeframe is greater than it was when Bueno de Mesquita dismissed UN votes as a method of assessing alignment, it remains limited to 1946 and beyond. This means that the measure cannot be used to compare the multipolar period to the bipolar, as most studies in the past have attempted to do. However, there is validity to measuring polar tension and conducting studies that include the timeframe from 1946 on, as the post-World War II system is truly global in nature and has introduced a new era of extremely devastating military capabilities and global superpowers.

Since this is a measure that combines the effects of power parity and alignment gaps on the global system, studies that pertain to the general threat perception that arises from a combination of these factors may consider using it. However, studies that deal with polarity and do not include both of these elements should not. That is to say that if the study is primarily based on the power differences, the number of poles, or the degree of disparity between policies only and is not related to the degree to which these factors create contested policy space, then the polar tension measure is of little to no use.

However, the measure may offer an alternative mechanism for studies whose hypotheses are based on other polar factors. For example, Gowa and Mansfield (1993), in considering how polarity affects the way major powers view and cope with the negative externalities resulting from free trade in alliances, theorize that a larger number of poles creates more opportunities for allied nations to leave one alliance for another. Using polar tension, on the other hand, would introduce the mechanism of threat perception as a factor that determines trade with allies rather than the opportunities presented by the number of poles, which may produce a different conclusion altogether. Waltz (1964) also theorizes that the increased number of poles provides more opportunities for interaction and miscommunication that could lead to conflict. Again, using polar tension as an alternative mechanism for potential conflict could produce different results.

The primary purpose of this paper is to introduce a new concept of polar tension as a measure of the degree to which international policy space is contested, offering a different theoretical mechanism through which to assess state behavior. Future research should seek ways to improve upon the way in which this concept is measured. However, any valid measure of tension should consider both power and alignment. Dismissing either of these measurements risks failing to capture the tension and hostility in the international system that the initial analogous meaning of term “polarity” was intended to reveal.

### 3 The Effects of Polar Tension on Policy Alignment in Asymmetric Alliances

#### 3.1 Introduction

When nations join one another in a military alliance, they exchange a degree of autonomy for the benefits that the relationship affords them, which usually constitutes an increase in security. However, in asymmetric alliances, the security-autonomy tradeoff is especially pronounced and uneven; the stronger nation provides a security guarantee to the weaker state, and in return, the weaker state relinquishes a potentially great deal of autonomy to its protectorate. The alliance is beneficial to both parties, but in different ways; the weaker state gains protection for its otherwise vulnerable interests while the stronger state may further its interests by gaining diplomatic power, increased status, or increased military reach through basing rights. The different benefits also point to potentially differing interests. While states come together based on common interests, they may hold other interests that are disparate, or that even stand in opposition to one another. In the case of asymmetric alliances, the benefits of the alliance may come through the protection of interests that are different in scope and magnitude. The stronger ally, or major power, is likely to have a much broader and strategic perspective, being more in tune with the shifts in the international system structure (i.e. to changes in the way the power and policy relationships are arranged among the major powers). The weaker state, or protégé, on the other hand, is likely to have particular interests that are localized and proximate in nature and to be less affected by and concerned about changes in the international system. If changes in the system structure increase the degree to which international policy space is contested, a concept I outline in the previous chapter called *polar tension*, then the major powers may perceive that their interests are being threatened and desire to increase their relative strength. Smaller states, perceiving less threat from international structure changes, are less likely to react to polar tension in this way. So, the fact that these states are allied, despite their different perspectives and security concerns raises a question about their behavior when the general interests of the

major powers are threatened due to changes in the international system structure.

This study, therefore, asks the question ‘How does change in the international system structure affect the alignment of international policy preferences in asymmetric allies?’ More specifically, as polar tension increases in the international environment, raising the perceived threat to the general interests of the major powers, how does it affect the policy preference alignment of small-state asymmetric allies?<sup>10</sup>

I theorize that (1) as polar tension rises, weaker states increase their alignment with their major power allies, and (2) the degree to which weaker states increase alignment in times of increased polar tension depends on the prioritization that the weaker ally places on the common general interests that it shares with its protectorate state. I further argue that not only is there variation between protégé states in their prioritization of common general interests but that this variation exists chiefly along the line between multilateral and bilateral alliances, making the differentiation binary in nature. Protégés that place a relatively high priority on the general interests of the major power protectorate are selected into multilateral alliances, which are constructed to endure and adapt to general threats. Conversely, protégés that place a relatively low priority on the general interests of their major power ally are selected into bilateral alliances, which are designed primarily to address a narrow range of particular interests.

I use selection on observables to evaluate the effects of polar tension on the alignment of asymmetric alliances as well as the effect of having a multilateral asymmetric alliance on that relationship. I find that, as polar tension increases, protégés increase their alignment overall, while, more specifically, multilateral alliances increase alignment as polar tension increases and bilateral alliances do not.

---

<sup>10</sup>I treat these alliances as exogenous, meaning that I’m not considering how polarization affects alliance formation. Regarding theory, I’m considering alliances that are based on written agreements, indicating some degree of intended duration (Morrow 2000), especially in the post-WWII period considered in this study, where alliances tend to endure (Kuo 2021). Furthermore, regarding precedence, when assessing the impact of international structure on alliances, Bueno de Mesquita (1975) and Maoz (2006) use a measure of increasing alignment between major powers and their allies, not the number of allies gained or lost, to determine polarization, indicating that they also treat alliances as exogenous.

This study contributes to the literature on international system structure by examining not only its effects on asymmetric alliances but also the unique effects that may be attributed to the differences between multilateral and bilateral relationships. It may also have policy implications, as it addresses the value of asymmetric alliances during times of increased tension in the international system.

### **3.2 Literature Review and Theoretical Underpinnings**

To understand how international system structure affects asymmetric alliances, we must first clarify what is meant by system structure and how changes in the structure may represent a threat to the general interests of the major players in the system. Ray defines structure as “the way in which relationships are arranged” between states in the international system (Ray 1990, 99). Changes in the structure affect the entire system, which he defines as interdependent entities whose actions “affect the behavior and fate of the rest” (Ray 1990, 99). Ray asserts that there are two dimensions to this structure (i.e. two ways in which the relationships between nations are arranged) that have the greatest impact on the system: “comparisons” and “links and bonds” (Ray 1990, 99). Clark makes a similar claim, calling the two dimensions “power” and “conflict” (Clark 1980, 13). The first of these aspects, power, determines which of the nations have the greatest power and influence (i.e. which are the major powers), and their relative power to one another (i.e. their degree of power parity). This measure determines which states may challenge the policy positions of another. The second aspect, conflict, addresses the degree to which states are aligned or misaligned in their policy preferences. When we combine these structural dimensions, we may observe their impact on the system. As the policy preferences of major powers grow in opposition to those of the strongest power, they have a desire to challenge those positions. If, at the same time, their power parity is such that they are capable of challenging the strongest state’s policy positions, then the tension between the states increases (Organski 1968). This dynamic is what I refer to as *polar tension*, or the degree to which international

policy space is contested.

This tension, or hostility in the global environment, drives states' behavior because it creates a perception of threat among the major powers. The dual aspects of *power parity* and *policy opposition* that define high levels of polar tension have the same impact as the two underlying factors that compose a threat: *capability* and *intent* (Singer 1958, 93-94; Cohen 1979, 4-5). If one powerful nation views another rising in power and pursuing policy outcomes contrary to its own, it may perceive such a shift as a general implicit threat to its own interests (Cohen 1979, Organski 1968). In the context of system structure, power parity and opposing policy preferences are perceived as threats to a major power's general interests.

It is important to note that polar tension produces a threat to general interests, not to any specific interest that might trigger an immediate crisis. Snyder provides a helpful discussion on this matter by categorizing states' interests as being either "general" or "particular" (Snyder 1984, 464). He considered particular interests to be specific localized issues with the potential to bring conflict between specific states. General interests, on the other hand, are those that are strategic in nature and do not pertain to any specific threat. This general, or ideological view of interests is what Fox (1944) envisioned as he introduced the concept of bipolarity before the Cold War was underway. He predicted that the United States and the Soviet Union would come out of World War II as the most powerful nations in the world and that if the two had opposing ideologies that drove their policy preferences, it would create tension, much like the tension between two poles. He further claimed that the general opposition would not be enough to generate conflict, but that it would set the conditions such that if the two powers' conflicting interests overlapped in a particular area, it may draw them into conflict (Fox 1944, 100). Therefore, polar tension is seen as a general threat to states whose interests are broad and strategic.

The threat induced by polar tension has the greatest impact on the major powers. Their broad general interests make them sensitive to changes in the status quo, which may alter

their ability to pursue their own interests. However, the question remains as to how this affects the relationship between the major powers and their weaker allies. To understand how system structure, or more specifically, polar tension, affects asymmetric alliance dynamics, it is important to understand the nature of asymmetric alliances and the differences in global perspectives and interests of these disproportionate bedfellows. While major powers are strong enough to impact the international system, the weaker powers in these asymmetric relationships are not. Defining a “small power”, Rothstein writes that such a state “recognizes that it can not obtain security primarily by the use of its own capabilities and that it must rely fundamentally on the aid of other states, institutions, processes, or developments to do so” (Rothstein 1968, 29). Morrow (1991) describes how this security differential plays into the unique dynamics of asymmetric alliances. Unlike major power-to-major power or small power-to-small power alliances, where both partners tend to surrender a similar amount of autonomy for the additional security their ally supplies, the security-autonomy tradeoff in asymmetric alliances is lopsided. The weaker protégé gains a security guarantee from its stronger, more capable major power ally. This is a tremendous boon for the weaker state; however, as Rothstein notes, “borrowing someone else’s strength, like borrowing his money, has advantages as well as disadvantages” (Rothstein 1968, 45). The protégé, in return for the boost in security, agrees to surrender a potentially great deal of its autonomy, which Morrow defines as “the degree to which [a nation] pursues desired changes in the status quo” (Morrow 1991, 908-909). This forfeited autonomy allows the stronger state to control some or all of its protégé’s international policy positions. The smaller state may also provide military basing or even accept its protectorate’s heavy hand in its domestic political affairs. So, while the protégé benefits from the security provided by its protectorate major power, it also pays a price.

The unique tradeoffs in these asymmetric relationships help determine the way in which polar tension affects the behavior of each ally. Liska writes that “Alliance ideology merges with the rationalization of the struggle which has brought it about” (Liska 1962, 61). In other

words, alignment occurs when states are made to focus on the common interests on which they formed their alliance. When polar tension is low and there is little perceived threat to the status quo of the international environment, the protégé states are more likely to focus exclusively on their own particular and localized interests, potentially drifting farther away from aligning with their major power allies on the broader, more general issues. However, as polar tension rises, the alliance's shared general interests gain relevance to the protégé, causing greater alignment. This leads to our first hypothesis:

**H1:** *Protégés are likely to increase their alignment with their major power protectorates during times of increased polar tension.*

There is reason to believe, however, that there is variation in this tension-induced alignment. The degree to which rising polar tension leads to increased alignment depends on the *prioritization* that the protégé places on the shared interests it has with its major power ally. In other words, the degree to which the shared general interests of the major power “gain relevance,” as stated above, depends on whether the protégé places enough priority on the general interests to increase its support of those interests when they are threatened. The asymmetric nature of the alliance creates conditions on which these factors vary. In addition to the differences in the tradeoffs that stronger and weaker nations make in their unbalanced relationship, their status as stronger and weaker states creates variation in the nature of their interests (Rothstein 1968, Morrow 1991). Rothstein, in describing the differences between major powers and smaller states, writes that they are “different in kind, not merely in degree” (Rothstein 1968, 1). Whereas major powers' interests tend to be general, global, and distant, the interests of small states are more likely to be particular, localized, and pressing. Rothstein stated that “Great powers tend to ally in terms of a threat to the balance of the whole system, while the small powers in terms of a threat to its local balance” (Rothstein 1968, 62). Walt shares this perspective, holding that major powers balance against other major powers while smaller powers tend to balance against threats (Walt 1987, 148). These differences make major powers more sensitive to the shifts in the international structure than

smaller states; however, they do not exclude weaker states from sharing the general concerns of their major power allies.

States must have some shared interests on which to form an alliance. In fact, Liska claims that “All association depends on the existence of identical interests” or that the associates are at least “capable of merging in a higher, more inclusive common interest” (Liska 1962, 27-28). He explains that, in addition to these identical interests, allied nations typically have disparate and even conflicting interests as well. The asymmetric nature of the relationship, and the resulting difference in perspectives, creates conditions where even if interests are shared, there is friction between the shared general interests and shared particular interests. The question of how much a weaker state will align with a major power ally is determined by the degree to which the protégé prioritizes the common general interests over common particular interests and interests that diverge or conflict. Liska claims that the “chances of convergence are best when there are complexes of interests” and when the identical interests motivate states to adjust their opposing, particular interests (Liska 1962, 28). He illustrates this point by referencing the example of the United State alliance with Pakistan during the Cold War. The two had virtually identical interests concerning a potential Sino-Soviet alliance, which affected the international system structure. Meanwhile, Pakistan’s concerns over the Kashmir issue with India, which were more localized and narrow in scope, were somewhat conflicting with US interests. Liska claims that the alliance endured due to Pakistan prioritizing the shared interest with the US (Liska 1962, 29).

The shared complex of interests between the two allies also increases the value of the alliance for each partner, in turn increasing the prioritization of common interests. Fang, Johnson, and Leeds (2014) demonstrate that the reliability of an ally in a time of crisis increases with the degree to which the ally values the alliance. Certainly, the more a protégé benefits from the alliance, the more it will prioritize its partner’s policy desires. As the interests become entangled with those of the major power, a risk to the major power’s interests, even if they are somewhat disparate, may jeopardize their own. Considering the

different perspectives of the states in asymmetric alliances, this implies that weaker states that prioritize the shared general interests with the major power ally over their particular interests are more likely to feel threatened by the increase in polar tension and therefore align their interests with the stronger partner.

However, this is not the case for those allies who prioritize their particular interests. In fact, the protégé may worry that if its local interests affect the larger balance of power between the major powers, the major powers may be prone to sacrifice the weaker ally's particular interest. Even if the local issue does not affect the larger system structure, it may get ignored altogether (Rothstein 1968, 62). This means that those states that prioritize their own localized interests over the shared general interests of the major powers have reason to be concerned about growing tension in the global environment. As long as things are stable, the major power may be willing to expend resources to support the protégé's shared local interests; however, under conditions of increased polar tension, when the major power is forced to consider its broader stake in the world, it may indeed de-prioritize its protégé's particular interests. This means that those who do not prioritize the general interests are not incentivized to increase their alignment in times of increased polar tension when the general interests are threatened. This leads us to a second hypothesis:

**H2:** *Protégés that highly prioritize their common general interests with major powers are likely to increase their alignment as polar tension increases, and protégés who place a relatively low priority on common general interests are not.*

### 3.3 Formal Model

A simple formal model helps clarify how the prioritization of general interests contributes to the variation in alignment between asymmetric allies.

We start by examining the protégé's,  $i$ , utility,  $U$ , for the alliance:

$$U_i(\text{alliance}) = \sigma B_i(\rho) + (1 - \sigma)b_i - C_i(\rho)$$

In this utility equation,  $\sigma \in [0, 1]$  represents the protégé's prioritization of general interests and  $(1 - \sigma)$  represents its prioritization of particular interests.  $B_i(\cdot)$  is a differentiable and increasing function representing the security benefits guaranteeing the protection of *general* interests by the major power protectorate, and  $\rho$  represents polar tension. In this case, the security benefit that the protégé receives is a function of the degree of polar tension. Because  $B_i(\cdot)$  is increasing, when polar tension rises, the value of the general interests security benefit rises as well. The security benefits guaranteeing *particular* interests are represented by  $b_i$ , which is not a function of polar tension.  $C_i(\cdot)$  is a differentiable and increasing function representing the cost that the protégé must pay to receive the benefit. With increasing  $\rho$ , the cost associated with the alliance increases.

Before, examining the impact of prioritization, we can also gain some insight into the relationship proposed in H1, where an increase in polar tension increases the utility or value of the alliance, which we assume increases alignment. To do this, I take the partial derivative of the utility equation with regard to polar tension,  $\rho$ :

$$\frac{\partial U}{\partial \rho} = \sigma B'_i(\rho) - C'_i(\rho)$$

Since  $B_i(\cdot)$  and  $C_i(\cdot)$  are increasing, their derivatives are positive, which means that both the benefits and the costs are positive and increase with polar tension. The utility, then, is determined by the marginal effect of polar tension on the benefit compared to its marginal effect on the cost. I assume that for a member to desire to be in an alliance, its utility from membership must be greater than its utility from not being in the alliance, which I normalize to 0. So, we can create the inequality  $C'_i(\rho) \leq \sigma B'_i(\rho)$ . Assessing this inequality, we see that the hypothesis is dependent on two factors. First, it is dependent on the value of  $\sigma$ , which leads us to the second hypothesis that we will examine below. Second, if we assume for a moment that the prioritization is not a factor (i.e. we place the value of  $\sigma$  to 0), then we can also assume that for an alliance to exist, the marginal effects of  $\rho$  on the benefit must be greater than or equal to the marginal effects of  $\rho$  on the costs. Therefore, we find support

for H1, given that the assumption that  $B'_i(\rho) \geq C'_i(\rho)$  holds.

To examine the relationship proposed in H2, where the impact of polar tension is moderated by the prioritization of general interests, I take the cross partial derivative of the utility equation with regard to prioritization,  $\sigma$ , and polar tension,  $\rho$ .

$$\frac{\partial^2 U}{\partial \sigma \partial \rho} = \frac{\partial}{\partial \sigma} \left( \frac{\partial U}{\partial \rho} \right) = \frac{\partial U}{\partial \sigma} (\sigma B'(\rho) - C'(\rho)) = B'(\rho)$$

Once again, since  $B_i(\cdot)$  is increasing,  $B'_i(\cdot)$  is positive. Therefore, this comparative static reveals that the combined effect of prioritization and polar tension on utility is positive, supporting the concept in H2 that higher prioritization increases the positive impact of polar tension, creating higher utility and therefore greater alignment.

### 3.4 Prioritization in Multilateral and Bilateral Alliances

The degree to which weaker allied partners prioritize the general interests of their major power allies exists along a continuum and is difficult to measure; however, these allies are also selected into a binary categorization of multilateral or bilateral alliances according to their prioritization of interests. Selection into these different alliance types creates a discontinuity in the would-be continuous prioritization variable. Protégés that place a high priority on general interests, after being selected into multilateral alliances, become further invested in the interests of the major power, while those selected into bilateral are less entangled. In fact, due to the difference in prioritization, those in bilateral alliances may fear that the major power would sacrifice or ignore their particular interests. Therefore, the variable becomes binary with weaker nations selected into multilateral alliances finding greater benefits and being incentivized to further prioritize the major power's general interests.

The logic of why these alliances become categorized begins with Riker's (1962) size principle, which states that political institutions grow until they become powerful enough to form a winning coalition and then grow no further. Fordham and Poast (2014) test the principle on alliance formation and demonstrate that multilateral alliances exhibit this be-

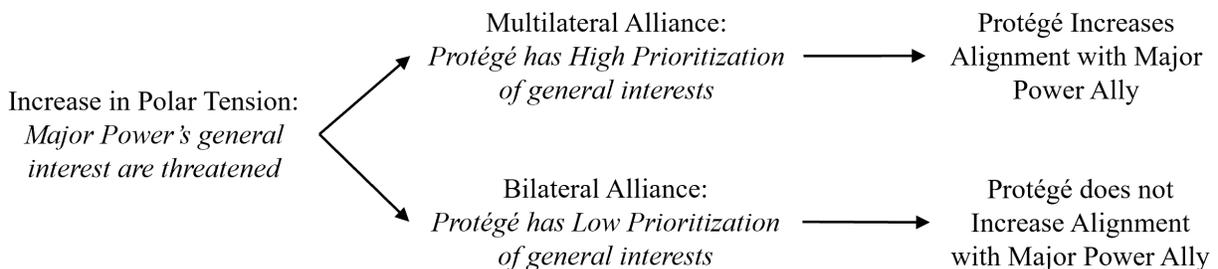
havior. While their study focuses on the power and capabilities of nations and considers all alliances, not just those that are asymmetric, it also shows that common foreign policy interests are positively and significantly correlated with multilateral alliance formation. One reason the authors give for nations that place high priority on particular interests being selected into bilateral rather than multilateral alliances is that these states' security concerns may "entrap other alliance members in wars where they have no stake" (Fordham & Poast 2014, 846). This insinuates that (1) there is an abundance of potential allies that could be included in a multilateral alliance, allowing the major power to be selective and choose the nations it deems most likely to prioritize its general interests, and (2) multilateral alliances share common interests that other nations with particular interests may not share. Major powers that build their relationship on similar particular interests with states left out of the multilateral arrangement tend to form bilateral alliances, as their lack of prioritization of the general interests prevents these weaker states from being selected into the multilateral alliance.

The notion that the two alliance types represent different prioritization of common interests is supported by the types of provisions and commitments found in each. Multilateral and bilateral alliances serve somewhat different purposes and are thus designed in accordance with their differing ends. Rothstein notes that some alliances are designed to "cope with distant or potential military threats," while others are formed to deal "with an immediate military threat" (Rothstein 1968, 51). The former of these alliances are intended to endure and adapt to the more general or ideological threats in the international environment, such as those that arise from shifts in the system structure. Kuo (2021) describes this as an "integrative approach" to alliance formation in which allies have extensive coordination measures and institutions that help them reconcile differences and adapt to evolving threats. The latter alliance type, which Kuo refers to as the "realpolitik approach", occurs when particular interests are threatened, and they are marked by limited institutionalization, conditional commitments, and "occasionally 'predatory' features" (Kuo 2021, 7).

Therefore, we can associate multilateral alliances with nations that prioritize the general interests of the major power and other members over their own individual, localized, and particular interests. These alliances, since they are designed to deal with potential conflicts that may grow from the general threat that brings them together, are designed with features that make them more durable and responsive to changes in the international system. Bilateral alliances, on the other hand, are more predisposed to prioritize particular interests and are less responsive to structural changes in the international system. In a later section, entitled “Testing the Mechanism”, I demonstrate the strength of this claim, showing that there are significant and suggestive differences in the commitments and provisions of each alliance type, therefore validating the use of the binary alliance type variable to operationalize the prioritization of the general interests. The division of the protégés’ prioritization level into distinct alliance categories allows us to restate the second hypothesis:

**H2’:** *Protégés in multilateral alliances with major powers are likely to increase their alignment as polar tension increases, and protégés in bilateral alliances are not.*

The mechanism behind this hypothesis is displayed in Figure 20, proposing that the theorized effect of polar tension on the alignment of an asymmetric alliance is moderated by the prioritization each protégé places on the general interests of its major power ally. It also illustrates how this prioritization is operationalized and categorized into the moderating effects of multilateral and bilateral alliances.



**Figure 20.** Mechanisms for Polar Tension’s Impact on Alignment in Asymmetric Alliances

### 3.5 Research Design

To test these hypotheses, I analyze the behavior of asymmetric alliance partners over time, and I start by establishing the unit of analysis as a dyad-year, drawing from a population of asymmetric dyads in an alliance in a given year. I use the Alliance Treaty Obligations and Provisions (ATOP) dataset (Leeds 2022) to determine which nations have alliances in each year examined. In this dataset, Leeds defines an alliance as “a formal agreement among independent states to cooperate militarily in the face of potential or realized military conflict” (Leeds 2022, 6). The key elements are that it includes a formal written agreement and that the agreements include military cooperation. Although the definition leaves out some relationships due to a lack of a formal agreement, such relationships merely imply commitment, whereas written agreements make commitment explicit and clear. Additionally, formal alliances define provisions that help outline the security-autonomy relationship inherent in asymmetric alliances (Morrow 2000). Therefore, Leeds’s exclusion of these informal partnerships comports with the understanding of alliances in this study. Additionally, I omit alliances whose only commitment to one another is a non-aggression pact, or the promise not to engage in military conflict with one another, as this does not require active military cooperation and could potentially be an agreement between foes rather than allies.

I refine the ATOP dataset by filtering for only asymmetric alliances. I do this by first integrating the Correlates of War (COW) Major Power dataset (COW State Membership 2017) to establish which of the nations are classified as major powers (i.e. those whose behavior has the greatest impact on the international system structure). Based on a general understanding that some nations have greater military, political, and financial power than others, this dataset lists the nations that qualify as major powers in the international system each year. In the post-World War II era, this included the permanent members of the United Nations Security Council (the United States, the United Kingdom, France, Russia (the Soviet Union), and China). China was only added to the list after the communist party victory in 1949. After the Cold War, in 1991, Germany and Japan were added to the list,

increasing the total number of major powers to seven. Because the criterion for becoming a major power is ambiguous, I tested the list by comparing it to the most powerful nations in each year according to a power index I describe later in this section. The COW list accurately identified the strongest nations each year with very few exceptions, confirming the legitimacy of its classification of major powers. After establishing which nations qualify as the major powers, I then filtered the dataset to include only asymmetric alliances, which, for the purposes of this study, I define as those alliances formed between these major powers and non-major powers.

**Dependent Variable: ALIGNMENT.**

The dependent variable that we are seeking to test is the alignment of the foreign policy preferences of the alliance dyads. Several attempts have been made to operationalize and measure this alignment. These include joint participation in Inter-Governmental Organizations (Wallace 1975), dyadic trade flows (Alker & Puchala 1968), foreign policy action flows (Altfeld 1984), alliance portfolio similarity (Bueno de Mesquita 1975), and similarity of United Nations General Assembly (UNGA) votes (Signorino & Ritter 1999). Each of these measures had important limitations. However, significant improvements were made when Bailey, Strezhnev, and Voeten (2017) produced a spatial measure using Item Response Theory (IRT). While this measure, like that of Signorino and Ritter, considers UNGA votes, Bailey et al. categorize the issues that were included in each initiative and accounted for the number of repeated votes, UN agenda changes, and shifts in wording to determine how each nation's preference scores on a single dimension each year. The scale indicates the degree to which states align with US-led liberal international order and their unique placement, which the authors refer to as each state's "ideal point," indicates their policy preferences. The policy space is highly correlated with significant international policy issues, such as democracy, human rights, colonialism, and trade and economics, making it a good indicator of general international policy preferences (Bailey et al. 2017, 444).

To arrive at the spatial distance between the preferred policies of the major power and

its protégé and create the ALIGNMENT variable, I use a similar method to the one I used to establish the alignment of the major powers to one another in estimating polar tension in the previous chapter.<sup>11</sup> I take the absolute value of the difference between the ideal points,  $IP$ , of each state. I then divide this by the maximum difference. Bailey et al.'s scale does not include a minimum or maximum degree of inclination toward a liberal world order, so I use the difference between the minimum and maximum positions that any nation in the dataset has taken over the entire time period as the maximum spatial distance, which I label as  $D_{ij\ max}$ .<sup>12</sup> Therefore, the estimated general policy distance,  $D_{ijt}$ , between the strong state,  $i$ , and its protégé,  $j$ , in a given year,  $t$ , is

$$D_{ijt} = \frac{|IP_{it} - IP_{jt}|}{D_{ij\ max}}.$$

Since Bailey et al.'s dataset relies on UNGA votes, its earliest time period corresponds with the first yearly session of the UN, 1946, which limits the overall timeframe covered by the dataset used in this study. Additionally, the 19th session of the UNGA, which corresponds to 1964 is missing from the Bailey et al. (2017) dataset with no explanation. Therefore, I omit that year from this study as well.

### **Key Independent Variables: POLAR TENSION and MULTILAT.**

The previous chapter explains the theoretical basis for the concept and measure of Polar Tension. It measures the degree to which international policy space is contested, a concept that creates the perception of a threat to the general interests of the major powers in the international system. The concept and measure consist of two factors. First is the alignment between the major powers, based on Bailey et al.'s (2017) Ideal Point dataset in the manner described above, only between major powers.

---

<sup>11</sup>The measure of polar tension includes a similar alignment variable drawn from Bailey et al.'s (2017) ideal point measure; however, the measure of polar tension only measures the spatial distance between the major powers. This ALIGNMENT variable takes the spatial distance between major powers and their non-major power allies.

<sup>12</sup>I am able to do this because the methodology used by Bailey et al. makes the ideal positions of each country comparable over time (Bailey et al. 2017, 435).

The second factor is power. The measure most often used to examine power is the Composite Index of National Capability (CINC) index provided by the COW's National Material Capabilities dataset (Singer 1987); however, as noted by Beckley (2018), this measure, while widely used, is often called into question. He points out that the measure tends to over-inflate highly populated but inefficient nations because it accounts for gross industrial production rather than net. Beckley proposes another measure that captures power in terms of how efficient a nation is with its resources. He demonstrates, using several different methods, that a net industrial efficiency measure, calculated by taking the product Gross Domestic Product and Gross Domestic Product Per Capita ( $GDP \times GDPPC$ ), is a better predictor of war outcomes than CINC (Beckley 2018, 38). However, Beckley's measure does not account for any source of military power, which is widely regarded as an important factor when considering how power is perceived as a threat (Waltz 1979, Morgenthau 1978, Singer 1987, Beuen de Mesquita 1975, Organski 1968). For this reason, I pull two of the six factors that compose the CINC measurement (the number of military personnel (milpers) and military expenditures (milexp) and combine these with Beckley's measure, pulling GDP and GDPPC from the Varieties of Democracy (V-Dem) dataset (Coppedge et al. 2022), to arrive at a measure that better captures the military and industrial power of nations.

In the previous chapter, I demonstrate how I arrive at this hybrid measure that includes both Beckley's measure of industrial efficiency and a ready military power measure, equally weighted. I further demonstrate that this measure is 4.6 to 5.6% better than either the CINC or Beckley method at predicting war outcomes since 1816 and is 3% better than CINC and just as good as Beckley's model at predicting militarized interstate disputes. Therefore, I rely on this measure to calculate the power of each nation in each year in the dataset.

The measure of polar tension captures both the measure of alignment and power. The power aspect is captured in a measure of power parity, which is obtained by dividing each major power's individual power score by that of the strongest nation.

$$PAR_{it} = \frac{P_{it}}{P_{it\ max}}$$

The polar tension score is then calculated by taking the sum of the products of power parity and policy distance and dividing them by the number of major powers. The following equation displays this calculation:

$$POL_t = \frac{\sum_i^n (PAR_{it} \times D_{it})}{N_t},$$

The greater the parity and the gap in alignment, the more policy space is contested, therefore, the higher the polar tension. The creation of this variable produces one measure of polar tension per year, beginning in 1946 and excluding 1964 due to its reliance on Bailey et al.'s (2017) index.

The other primary independent variable is MULTILAT. To develop this dummy variable, I divide the distinct asymmetric alliance dyads into those that have a multilateral relationship listed in the ATOP dataset and those that only have a bilateral alliance listed. Therefore, if a dyad has both a bilateral and a multilateral alliance in any given year, it is counted as multilateral. Only in the absence of a multilateral agreement is an alliance considered bilateral. The variable is coded 1 if the asymmetric dyad has at least one multilateral alliance in a given year and 0 if it has a bilateral alliance but no multilateral alliance.

Table 2 displays a breakdown of the asymmetric alliances between the years 1946 and 2015. While only 11.6% of all alliances and 6.7% of all alliance year observations are bilateral, 35.3% of asymmetric alliances and 23.4% of asymmetric alliance year observations are bilateral. This table also shows the average overall alignment of the types of alliances, with multilateral alliances showing about 10% greater alignment overall.

**Table 2.** Asymmetric Alliance Data Summary: 1946-2015

	Total		Asymmetric		Ave. Alignment b/w Protégé & Maj Pwr
	Num.	Obs.	Num.	Obs.	
Alliances	2,098	62,819	283	8,545	72.8%
Bilateral	244	4,225	100	2,003	67.7%
Multilateral	1,854	58,594	183	6,542	74.4%

**Control Variables.**

I include a list of control variables that are potentially confounding, with the potential to affect the alignment between nations and their alliance status.

**CONTINUITY.** While some neighboring nations may find conflicting interests over particular concerns such as border issues, they may have reasons to share regional and larger geopolitical interests. This variable, which I extract from the COW Direct Contiguity dataset (Stinnett et al. 2002), is coded 0 if the states are not contiguous, 1 if they are separated by 400 miles of water or less (but more than 24), and 2 if the states are separated by a land or river border or by 24 miles of water or less.

**In DISTANCE.** Geographic distance has a similar effect to that of contiguity. Nations that are close to one another geographically tend to interact often and are also apt to share more general geopolitical interests; both factors that may affect alignment. For this variable, I use the natural log of the geographic distance between the capital cities of each state, taken from the CEPII database (Mayer & Zignago 2011).

**POWER DIFF.** Fordham and Poast (2014) contend that power is a significant factor in multilateral alliances. Additionally, in an asymmetric relationship, the difference in power between the two states may be indicative of the amount of autonomy that the weaker ally has had to surrender. The loss of autonomy, in turn, may affect the protégé's desire and

willingness to maintain alignment in times of increased pressure, revealing its prioritization of general versus particular interests. I use the same measure of power for this variable as I did to calculate polar tension, combining elements of the CINC index and Beckley's efficiency model.

There are three more controls explained below that are likely to be correlated with alignment and alliances, but whose causal direction may be questionable.

**WAR.** If two nations have been at war with one another in the previous year, it highlights the fact that there are some salient conflicting interests between the two, and it may significantly affect their agreement on international policy issues. I use the COW Militarized Interstate Dispute (MID) (Palmer et al. 2020) data to identify if any asymmetric allies in the base dataset have been at war with one another in the previous year. To create this dummy variable, I assign a 1 to any dyad if a conflict between the nations composing it rose to the point of "use of force" or was categorized by the MID dataset as "war". If conflict did not rise to this level between the two states, the variable is coded 0.

**DISPUTE.** Similar to the WAR variable, disputes that do not lead to open conflict between nations may nonetheless be indicative of a fissure in the relationship that could affect their alignment with one another and their alliance status. To create this variable, I rely on the MID dataset once again, but this binary variable is coded as 1 if the MID data considered the highest hostility level as either "threat to use force" or "display of force", and is coded 0 if it did not rise to that level.

**TRADE.** Trade is also a factor that may be correlated with the alignment of states. The co-dependence of trade may increase the degree to which states align on general issues such as democracy and free trade. To create this variable, I use the COW Trade dataset (Barbieri & Keshk 2017), which examines the total trade flow between the allies each year.

### **Analysis.**

Given the combination of the data summarized above, I use an ordinary least squares (OLS) regression to examine the effects of polar tension on the alignment of asymmetric

allies from 1947 to 2015. The time period is limited on the front end by Bailey et al.'s (2017) ideal point dataset and on the back end by the MID dataset. I cluster the standard errors at the dyad level to account for potential autocorrelation within the dyads.<sup>13</sup> I also lag the independent and control variables by one year to reduce the risk of reverse causality; however, the last three control variables mentioned above (WAR, DISPUTE, and TRADE) have long-term trend implications which may open them up to be affected by alignment rather than only affecting it. I do not include the last three controls in the basic regressions; however, due to the possibility and extent to which these factors do impact alignment, I include them in a final regression.

The basic regression takes the form of

$$\begin{aligned}
 ALIGNMENT_{ijt} = & \alpha + \beta_1 * POLAR\ TENSION_{t-1} + \beta_2 * CONTIGUITY_{ijt-1} + \\
 & \beta_3 * \ln\ DISTANCE_{ijt-1} + \beta_4 * POWER\ DIFF_{ijt-1} + \epsilon.
 \end{aligned}
 \tag{13}$$

This regression examines the effect of polar tension on alignment, testing the validity of H1. I build on this basic regression to examine the validity of H2. To do this, I run the regression again, adding the MULTILAT variable and interacting it with the POLAR TENSION variable to examine the impact of having a bipolar alliance relationship on the relationship between POLAR TENSION and ALIGNMENT. In the final regression, I also add the variables WAR, DISPUTE, and TRADE.

### 3.6 Results

Table 3 displays the regression results. Column 1 presents the results from the regression Equation 13 above. We see here that POLAR TENSION in the previous year has a positive

---

<sup>13</sup>Clustering standard errors at the dyad level does not account for the autocorrelation that may occur in multilateral alliances. However, clustering on alliances is problematic, as some dyads belong to as many as seven different alliances in a single year. Therefore, I include a regression table in Appendix A that includes standard errors clustered at the level of the largest alliance each member is a part of in each year, accounting for the largest degree of autocorrelation. Clustering at this level is aggressive, drastically increasing and likely overestimating the the standard errors (Abadie et al. 2017).

**Table 3.** The Effects of Polar Tension on Policy Alignment of Asymmetric Alliances

	<i>Dependent variable:</i>		
	ALIGNMENT <sub>ijt</sub>		
	(1)	(2)	(3)
POLAR TENSION <sub>t-1</sub>	0.263*** (0.059)	-0.083 (0.152)	-0.102 (0.152)
MULTILAT <sub>ijt-1</sub>		-0.049 (0.051)	-0.050 (0.052)
CONTIGUITY <sub>ijt-1</sub>	0.025 (0.019)	0.034* (0.019)	0.042** (0.017)
ln DISTANCE <sub>ijt-1</sub>	-0.045*** (0.013)	-0.027** (0.013)	-0.028** (0.013)
POWER DIFF <sub>ijt-1</sub>	-0.478*** (0.078)	-0.572*** (0.076)	-0.556*** (0.077)
WAR <sub>ijt-1</sub>			-0.110*** (0.037)
DISPUTE <sub>ijt-1</sub>			-0.094** (0.047)
TRADE <sub>ijt-1</sub>			-2.3 x 10 <sup>-7</sup> (1.5 x 10 <sup>-7</sup> )
POLAR TENS <sub>t-1</sub> * MULTILAT <sub>ijt-1</sub>		0.412** (0.161)	0.409** (0.164)
Constant	1.061*** (0.101)	0.975*** (0.113)	0.987*** (0.113)
Observations	6,661	6,661	6,661
R <sup>2</sup>	0.246	0.281	0.289

Note: SE clustered at dyad level.

\*p<0.1; \*\*p<0.05; \*\*\*p<0.01

and significant correlation with ALIGNMENT, even with strong and significant effects from the control variables. This supports H1, demonstrating that as polar tension rises, states in an asymmetric alliance increase their policy preference alignment.

Column 2 presents the results when adding the MULTILAT variable and interacting it with POLAR TENSION. The result shows that the interaction term has a positive and significant coefficient, meaning that the impact of polar tension on alignment is dependent on the alliance being multilateral. A Wald test reveals that when adding the POLAR TENSION variable and the interaction variable, we still get a highly significant result, demonstrating that the marginal effect of polar tension for multilateral alliances is strong and significant. The results displayed in Column 3 are not substantially different than those displayed in Column 2. Columns 2 and 3 support H2, corroborating the theory that as polar tension increases, protégés in multilateral alliances are likely to align with their major power protectorate, and protégés in bilateral alliances are not.

### **3.6.1 Testing the Mechanism**

The results shown in Table 3 accord with our expectations based on the theoretical mechanisms proposed. However, if the mechanism is correct, we should also observe fundamental differences in bilateral and multilateral alliances that drive the difference in alignment. Alliance design should match the purpose of the partnership. That means that if the weaker state in the alliance prioritizes the general interests of the major power, it is less concerned about a present particular threat than a general threat that may develop into a particular threat in the future. These common general interests lead to alliances that are designed to endure and adapt to the potential conflicts that may materialize from the threat to general interests (Rothstein 1968, Kuo 2021). Since some states are selected into multilateral alliances based on their shared interests while other states are selected into bilateral alliances for reasons of particular interests, the multilateral alliances likely have more measures that allow them to endure and adapt to threats to general interests. Kuo (2017) suggests some

measures that alliances may include if their purpose is to endure and adapt, such as the existence of a coordination body to help adjust to evolving threats. Similarly, these alliances are likely to have requirements for military coordination and provisions for non-military cooperation. Along the same lines, they are likely to include conflict mediation structures to help the alliance members work through their conflicting interests. They are also likely to have fewer specific threats addressed in their agreements. Finally, the reduced realpolitik nature of these alliances also makes them less predatory, having greater equity within and fewer asymmetric obligations.

To discover if there are indeed notable differences in the way multilateral and bilateral alliances are constructed regarding the commitments and provisions listed above, I use ATOP variables to examine each of these alliance agreement characteristics. The results displayed in Table 4 show the percentage of bilateral and multilateral alliance observations that contain each of the six substantial commitments and provisions. By comparing the percentages of alliance year observations that include these treaty measures, we can see that: (1) there are substantial differences in the alliance commitments and provisions between the types of alliances, and (2) multilateral alliances are much more likely than bilateral alliances to possess the qualities for an integrative alliance that we would expect to see among nations that prioritize shared general interests.

It could be that there are some unique factors pertaining to specific dyads that cause this variation in alliance features. To minimize dyad variation, I examine only those alliance dyads that transition from a bilateral alliance to a multilateral one or vice versa. In selecting this data, I filter out those dyads that had a break of one year or more between the transition from one to the other, as this may be indicative of differing international circumstances or regime changes. I also eliminate those that were only in a bilateral or multilateral alliance for two years or less. There are a total of 867 observations. The results, displayed in Table 5, are not substantially different than those in Table 4, further supporting the validity of the mechanism for H2.

**Table 4.** Bilateral versus Multilateral Asymmetric Alliance Agreements: 1946-2016

	Bilateral Alliance		Multilateral Alliance	
	Obs	Percent	Obs	Percent
Coordination Body	808	40.3%	6,665	97.6%
Military Coordination	541	27%	4,082	60%
Non-military Cooperation	1,369	68.3%	6,468	94.7%
Mediation Mechanism	158	7.9%	2,959	43.3%
No Specific Threat	132	6.6%	1,098	16.1%
Asymmetric Obligations	358	17.9%	0	0%
Total Observations	2,003		6,829	

**Table 5.** Transitioned Asymmetric Alliance Agreements: 1946-2016

	Bilateral Alliance		Multilateral Alliance	
	Obs	Percent	Obs	Percent
Coordination Body	226	55.5%	406	88.1%
Military Coordination	154	37.8%	303	65.7%
Non-military Cooperation	312	76.7%	456	99%
Mediation Mechanism	39	9.6%	69	15%
No Specific Threat	18	4.4%	108	23.4%
Asymmetric Obligations	66	16.2%	0	0%
Total Observations	407		461	

### 3.7 Conclusion

Asymmetric alliances provide unique advantages to both the major power and its weaker protégé partner. These advantages are particularly salient when the threat for which the alliance was formed increases. As the advantages increase, so does the value of the alliance. In an alliance where a protégé depends on its ally for its security, as polar tension increases and the threat to general interests grows, the relevance of the alliance increases for the protégé, causing it to increase its alignment with its major power ally. Testing this hypothesis suggests that it has validity. As polar tension increases, alignment grows between asymmetric allies.

This conclusion, however, is determined largely by the degree to which the protégé prioritizes its common general interests with its ally. The more the protégé prioritizes its common general interests with its ally, the more relevant the alliance becomes when those interests are threatened. Shared general interests also act as the centerpiece of multilateral alliances, which are designed to oppose challengers to these general interests and are concerned with shifts in the international system structure. Those states that prioritize general interests over their particular issues are apt to be selected into these multilateral alliances, while those that prioritize their common particular interests with the major powers are more likely to land in bilateral alliances. Herein, we find variation in the degree to which nations find advantages in their alliance. The protégés connected to major powers via multilateral alliances are more likely to prioritize their shared general interests, thus gaining greater advantage from the alliance as polar tension increases, and are incentivized to increase their alignment. Testing this hypothesis revealed that, indeed, as polar tension rises, protégés in multilateral alliances increase their alignment while those in bilateral alliances do not.

The measures of alignment and polar tension that allow us to test these hypotheses are limited by the available data on which the measures are built; therefore, we are unable to explore the validity of this theory in any period prior to the Cold War. Given a more comprehensive measure of polar tension in the future, this study should be updated to confirm its general applicability. However, in the post-World War II era, marked by growing

global interaction and increasingly harmful military capabilities, this study has current policy applicability; as polar tension shifts, states should more carefully consider the benefits they receive from their alliances and shore up those that share common general interests and provide the most advantages. With the anticipated rise of China (US DOD 2022, Allison 2017), the international structure may be increasing in polarity in the near future. In a recent meeting of the Group of Seven (G-7), the US, UK, France, Germany, Italy, Canada, and Japan acknowledged the growing military threat from China. The *Wall Street Journal* reported that Secretary Yellen called for the US to “friendshore,” suggesting that it should shore up its relationship with its allies (Duehren & Ip 2023). This concern demonstrates a heightened threat perception and increased relevance of alliances. The suggested response is consistent with the theory and results presented above and demonstrates their relevance to current global trends. As polar tension in the international system shows signs of shifting, the value of multilateral asymmetric allies may increase.

## 4 References

Abadie, Alberto, Susan Athey, Guido W. Imbens, and Jeffrey Wooldridge. 2017. “When Should You Adjust Standard Errors for clustering?”. *National Bureau of Economic Research, Working Paper 24003*.

Alker, H. and D. Puchala. 1968. “Trends in Partnership in the North Atlantic Area: 1928-1963.” In J.D. Singer, ed. *Quantitative International Politics*. New York: Free Press.

Allison, Graham T. 2017. *Destined for War: Can America and China Escape Thucydides’s Trap?* Boston: Houghton Mifflin Harcourt.

Altfeld, Michael F. 1984. The Decision to Ally: A Theory and Test. *The Western Political Quarterly*. 37(4), pp.523-544.

Altfeld, Michael F. 1984b. “Measuring Issue-Distance and Polarity in the International System: A Preliminary Comparison of an Alliance and an Action Flow Indicator.” *Political Methodology*, 10(1): 29-66.

Arrow, K. J. 1971. Political and Economic Evaluation of Social Effects and Externalities. *Frontiers of Quantitative Economics: Papers Invited for Presentation at the Econometric Society Winter Meetings*. New York, 1969.

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

Barbieri, Katherine, Omar M. G. Keshk, and Brian Pollins. 2009. “TRADING DATA: Evaluating our Assumptions and Coding Rules.” *Conflict Management and Peace Science*, 26(5): 471–491.

Barbieri, Katherine and Omar M. G. Keshk. 2017. *Correlates of War Project Trade Data Set Codebook, Version 4.0*. Online: <http://correlatesofwar.org>.

Barnett, D. Sean, Yvonne K. Crane, Gian Gentile, Timothy M. Bonds, Dan Madden, and Katherine Pfrommer. 2020. *North Korean Conventional Artillery: A Means to Re-*

*taliate, Coerce, Deter, or Terrorize Populations*. Santa Monica, CA: RAND Corporation.  
<https://www.rand.org/pubs/researchreports/RRA6191.html>

Beckley, Michael. 2018. "The Power of Nations: Measuring What Matters." *International Security*, 43(2): 7-44.

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

Benson, Brett V. Adam Meirowitz, and Kristopher W. Ramsay. 2014. Inducing Deterrence through Moral Hazard in Alliance Contracts. *Journal of Conflict Resolution*. 58(2), pp.307-335.

Benson, Brett V. and Bradley C. Smith. 2022. Commitment Problems in Alliance Formation. *American Journal of Political Science*, pp.1-14.

Bowers, Ian and Henrik Stalhane Hiim. 2020. Conventional Counterforce Dilemmas: South Korea's Deterrence Strategy and Stability on the Korean Peninsula. *International Security*, 45(3), pp.7-39.

Bueno de Mesquita, Bruce. 1975. "Measuring Systemic Polarity." *Journal of Conflict Resolution*, 1((2): 187-216.

Burns, Arthur Lee. 1957. "Balance to Deterrence: A Theoretical Analysis." *World Politics*, 9(4): 494-529.

Buzan, Barry. 2004. *The United States and the Great Powers: World Politics in the Twenty-First Century*. Malden, MA: Polity Press.

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

Carroll, Robert J. and Brenton Kenkel 2019. "Prediction, Proxies, and Power". *American Journal of Political Science*, 63(3): 577-593.

Cavendish, Richard. 2004. Troop Withdrawals from Korea. *History Today*, 54(8).

Christensen, Thomas J. and Jack Snyder. 1990. "Chain Gangs and Passing Bucks: Predicting Alliance Patterns in Multipolarity". *International Organization* 44(2): 137-68.

Clark, Ian. 1980. *Reform and Resistance in International Order*. Cambridge: Cambridge University Press.

Cohen, Rayomong. 1979. *Threat Perception in International Crisis*. Madison, WI: The University of Wisconsin Press.

Coppedge, Michael, John Gerring, Carl Henrik Knutsen, Staffan I. Lindberg, Jan Teorell, David Altman, Michael Bernhard, Agnes Cornell, M. Steven Fish, Lisa Gastaldi, Haakon Gjerløw, Adam Glynn, Sandra Grahn, Allen Hicken, Katrin Kinzelbach, Kyle L. Marquardt, Kelly McMann, Valeriya Mechkova, Pamela Paxton, Daniel Pemstein, Johannes von Römer, Brigitte Seim, Rachel Sigman, Svend-Erik Skaaning, Jeffrey Staton, Eitan Tzelgov, Luca Uberti, Yi-ting Wang, Tore Wig, and Daniel Ziblatt. 2022. "V-Dem Codebook v12" *Varieties of Democracy (V-Dem) Project*.

COW State System Membership. 2017. "Frequently Asked Question, Version 2002.1." (PDF). <https://correlatesofwar.org/wp-content/uploads/State-FAQ.pdf>.

Deutsch, Karl, and J. David Singer. 1964. "Multipolar Power Systems and International Stability." *World Politics*, 16(3): 390-406.

Duehren, Andrew, and Greg Ip. 2023. "U.S., Allies Weigh How to Reduce Economic Ties with China." *The Wall Street Journal*. April 17. <https://www.wsj.com/articles/u-s-allies-weigh-how-to-reduce-economic-ties-with-china-8f5321ab>.

English Standard Version Bible. 2001. Wheaton, IL: Crossway.

Fang, Songying, Jesse C. Johnson, and Brett Ashley Leeds. 2014. To concede or to Resist? The Restraining Effect of Military Alliances. *International Organization*. 68, pp.775-809.

Fearon, James D. 1995. Rationalist Explanations for War. *International Organizations* 49(3), pp.379-414.

Fearon, James D. 1997. Signaling Foreign Policy Interests: Tying Hands versus Sinking Costs. *The Journal of Conflict Resolution*, 41(1), pp.68-90.

Fordham, Benjamin and Paul Post. 2016. "All Alliances are Multilateral: Rethinking

Alliance Formation.” *Journal of Conflict Resolution*, 60(5): 840-865.

Fox, William T. R. 1944. *The Super-Powers: The United States, Britain, and the Soviet Union—Their Responsibility for Peace*. New York: Harcourt, Brace and Company.

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

Goemans, H. and Spaniel, W., 2016. Multimethod research: A case for formal theory. *Security Studies*, 25(1), pp.25-33.

Goertz, Gary. 2017. *Multimethod Research, Causal Mechanisms, and Case Studies: An Integrated Approach*. Princeton, NJ: Princeton University Press.

Gowa, Joanne and Edward D. Mansfield. 1993. Power Politics and International Trade. *The American Political Science Review*, 87(2): 408-420.

Grudo, Gideon. 2016. Iron Dome. *Air & Space Force Magazine*. March 28. <https://www.airandspaceforces.com/article/iron-dome/>

Haas, Mark L. *Frenemies: When Ideological Enemies Ally*. Ithica, NY: Cornell University Press.

Heo, Uk and Terence Roehrig. 2019. *The Evolution of the South Korea-United States Alliance*. New York: Cambridge University Press.

Horton, Jake. 2021. Israel-Gaza: How Much Money Does Israel Get from the US? *BBC News*. May 24. <https://www.bbc.com/news/57170576>.

Kane, Tim, 2006. Global U.S. Troop Deployment, 1950-2005. *The Heritage Foundation*. <https://www.heritage.org/defense/report/global-us-troop-deployment-1950-2005>.

Kaplan, Morton. 1957. *System and Processes in International Politics*. New York: John Wiley & Sons.

Keersmaeker, Goedele De. 2017. *Polarity, Balance of Power and International Relations Theory: Post-Cold war and the 19th Century Compared*. Cham, Switzerland: Palgrave Macmillan.

Kendall, Maurice G. 1938. “A New Measure of Rank Correlation”. *Biometrika*, 30(1/2):

81-93.

Kim, Kyong-ok. 2002. The Necessity of USFK and the Strengthening of US-ROK Relations. *Security Window*.

Kim, Tae-Hyung. 2008. North Korea's Missile Development and its Impact on South Korea's Missile Development and the ROK-US Alliance. *Korea Observer*, 39(4), pp.571-602.

Kim, Brian. 2021. US Lifts Missile Restrictions on South Korea, Ending Range and Warhead Limits. *Defense News*. May 25. <https://www.defensenews.com/global/asia-pacific/2021/05/25/us-lifts-missile-restrictions-on-south-korea-ending-range-and-warhead-limits/>

Knorr, Klaus. 1966. *On the Uses of Military Power in the Nuclear Age*. Princeton, NJ: Princeton University Press.

Krasner, Stephen D. 1976. "State power and the structure of international trade". *World Politics*, 28(3): 317-347.

Krauthammer, Charles. 1990. "The Unipolar Moment". *Foreign Affairs* 70(1): 23-33.

Kuo, Raymond C. 2021. *Following the Leader: International Order, Alliance Strategies, and Emulation*. Stanford, CA: Stanford University Press.

Lee, Hangsuck, Minha Lee, and Jimin Hong. 2021. Optimal Insurance Under Moral Hazard in Loss Reduction. *North American Journal of Economics and Finance*. 60, pp.1-16.

Lee, Jeong-Ho. 2022. South Korea says THAAD Missile Shield 'Not Negotiable' with China. *The Japan Times*. August 11. <https://www.japantimes.co.jp/news/2022/08/11/asia-pacific/south-korea-china-missiles/>

Leeds, Brett Ashley. 2022. "Alliance Treaty Obligations and Provisions (ATOP) Codebook." Online, <http://www.atopdata.org>.

Liska, George. 1962. *Nations in Alliance: The Limits of Interdependence*. Baltimore: Johns Hopkins Press.

Lorentzen, P., Fravel, M.T. and Paine, J., 2017. Qualitative investigation of theoretical models: the value of process tracing. *Journal of Theoretical Politics*, 29(3), pp.467-491.

Mansfield, Edward D. 1993. "Concentration, Polarity, and the Distribution of Power". *International Studies Quarterly*, 37(1): 105-128.

Mansfield, Edward D. 1995. *Power, Trade, and War*. Princeton, NJ: Princeton University Press.

Maoz, Zeev. 2006. "Network Polarization, Network Interdependence, and International Conflict, 1816-2002". *Journal of Peace Research*, 43(4): 391-411.

Mattes, Michaela. 2012. Reputation, Symmetry, and Alliance Design. *International Organization* 66, pp.679-707.

Mayer, T. and S. Zignago. 2011. *Notes on CEPII's distances measures: the GeoDist Database, CEPII Working Paper 2011-25*.

McGowan, Patrick J. and Robert M. Rood. 1975. "Alliance Behavior in Balance of Power Systems: Applying A Poisson Model to Nineteenth-Century Europe." *American Political Science Review*, 69: 850-70.

Memorandum. 1957. DECLASSIFIED Memorandum of Discussion at the 320th Meeting of the National Security Council, Washington, April 17, 1957. Eisenhower Library, Whitman File, NSC Records. Top Secret; Eyes Only. *Foreign Relations of the United States*.

Memorandum. 1959. DECLASSIFIED Memorandum of Discussion at the 411th Meeting of the National Security Council Washington, June 25, 1959. *Foreign Relations of the United States*.

Memorandum. 1962. DECLASSIFIED Memorandum From Michael V. Forrestal of the National Security Council Staff to President Kennedy. Kennedy Library, National Security Files, Countries Series, Korea, General, 8/62-3/63. Top Secret. Washington, September 24, 1962. *Foreign Relations of the United States*.

Minutes. 1970. DECLASSIFIED Minutes of a National Security Council Meeting. Washington, March 4, 1970. *Foreign Relations of the United States*.

Morgenthau, Hans J. 1946. *Scientific Man vs. Power Politics*. Chicago: University of Chicago Press.

Morgenthau, Hans J. 1978. *Politics Among Nations: The Struggle for Power and Peace, 5th Ed.* New York: Knopf.

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

Morrow, James D. 1991. "Alliances and Asymmetry: An Alternative to the Capability Aggregation Model of Alliances." *American Journal of Political Science*, 35(4): 904-933.

Morrow, James D. 1994. Alliances, Credibility, and Peacetime Costs. *Journal of Conflict Resolution* 38 (2), pp.270-297.

Morrow James D. 2000. Alliances: Why Write Them Down? *Annual Review of Political Science*, 3, pp.63-83.

Mutual Defense Treaty. 1953. Mutual Defense Treaty. *American Foreign Policy 1950-1955. Basic Documents Volumes I and II. Department of State Publication 6446. General Foreign Policy Series 117.* Washington, DC: U.S. Governemnt Printing Office, 1957.

Nogee, Joseph L. 1974. "Polarity: An ambiguous concept". *Orbis*, 18: 1193-1224.

Oberdorfer, Don. 2001. *The Two Koreas: A Contemporary History.* New York: Basic Books.

Oneal, John R., and Bruce M. Russett. 1997. "The Classical Liberals Were Right: Democracy, Interdependence, and Conflict, 1950-1985." *International Studies Quarterly*, 41(2): 267-94.

Organski, A. F. K. 1968. *World Politics, 2nd Ed.* New York: Knopf.

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

Palmer, Glenn, Roseanne W. McManus, Vito D'Orazio, Michael R. Kenwick, Mikaela Karstens, Chase Bloch, Nick Dietrich, Kayla Kahn, Kellan Ritter, Michael J. Soules. 2020. "The MID5 Dataset, 2011-2014: Procedures, Coding Rules, and Description." *Conflict Management and Peace Science*, 39(4): 470-482. DOI: <https://journals.sagepub.com/doi/full/10.1177/0738894221995743>.

Pauly, Mark V. 1968. The Economics of Moral Hazard: Comment. *The American Economics Review*. 58(3-1), pp.531-537.

Pauly, Mark V. 1974. Overinsurance and Public Provision of Insurance: The Roles of Moral Hazard and Adverse Selection. *The Quarterly Journal of Economics* 88(1), pp.44-62.

Pinkston, Daniel. 2012. The New South Korean Missile Guidelines and Future Prospects for Regional Stability. *International Crisis Group*. October 25. <https://www.crisisgroup.org/asia/north-east-asia/koreanpeninsula/new-south-korean-missile-guidelines-and-future-prospects-regional-stability>

Poole, Keith T. 2005. *Spatial Models of Parliamentary Voting*. Cambridge, UK: Cambridge University Press.

Powell, Robert. 2006. War as a Commitment Problem. *International Organization* 60, pp.169-203.

Rapkin, David P., William R. Thompson, and Jon A. Christopherson. 1979. "Bipolarity and Bipolarization in the Cold War Era: Conceptualization, Measurement, and Validation". *Journal of Conflict Resolution* 22(2): 261-295.

Ray, James Lee. 1990. "The Measurement of System Structure". in *Measuring the Correlates of War*. Ed. J. David Singer and Paul F. Diehl. Ann Arbor, MI: The University of Michigan Press.

Reagan, Ronald. 1981. "The President's News Conference, January 29". *The American Presidency Project*. <https://www.presidency.ucsb.edu/documents/the-presidents-news-conference-992>.

Rees, Ray and Achim Wambach. 2008. The Microeconomics of Insurance. *Foundations and Trends in Microeconomics*. 4(1-2), pp.1-163.

Riker, William H. 1962. *The Theory of Political Coalitions*. New Haven, CT: Yale University Press.

Rosecrance, Richard N. 1963. *Action and Reaction in World Politics: International Systems in Perspective*. Boston/Toronto: Little, Brown & Company.

Rosecrance, Richard. N. 1966. "Bipolarity, Multipolarity, and the Future". *Journal of Conflict Resolution* 10:314-327.

Rothstein, Robert L. 1968. *Alliances and Small Powers*. New York: Columbia University Press.

Rowell, David and Luke B. Connelly. 2012. A History of the Term "Moral Hazard." *The Journal of Risk and Insurance*. 79(4), pp.1051-1075.

Sartori, Anne E. 2005. *Deterrence by Diplomacy*. Princeton, NJ: Princeton University Press.

Shavell, S. 1979. Risk sharing and incentives in the principal and agent relationship. *Bell Journal of Economics*, 10(1), pp.55-73.

Signorino, Curtis S. and Jeffrey M. Ritter. 1999. "Tau-b or Not Tau-b: Measuring the Similarity of Foreign Policy Positions." *International Studies Quarterly*, 43: 115-144.

Singer, J. David. 1958. "Threat-Perception and the Armament Tension Dilemma". *Journal of Conflict Resolution*, 2(1): 61-62.

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

Singer, J. David, and Melvin Small. 1968. "Alliance Aggregation and the Onset of War." In J. David Singer, ed., *Quantitative international politics*. New York: Free Press: 147-186.

Smith, Alastair. 1995. Alliance Formation and War. *International Studies Quarterly* 39 (4), pp.405-25.

Smith, Bradley C. 2021. Military Coalitions and the Politics of Information. *Journal of Politics*. 83(4), pp.1369-1382.

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

Snyder, Glen. H. and P. Diesing. 1977. *Conflict Among Nations: Bargaining, Decision Making, and System Structure in International Crises*. Princeton, NJ: Princeton University Press.

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

Strassler, Robert B. 1996. *The Landmark Thucydides: A Comprehensive Guide to The Peloponnesian War*. New York: The Free Press.

Taylor Jr, William J., Jennifer A. Smith, and Michael J. Mazarr. 1990. US Troop Reductions from Korea, 1970-1990. *The Journal of East Asian Affairs*, 4(2), pp.256-286.

US Department of Defense. 2022. *Military and Security Developments Involving the People's Republic of China*.

Vaughan, Emmett.J. 1997. *Risk Management* New York: Wiley.

Wagner, R. Harrison. 1993. "What was Bipolarity?" *International Organization*. 47(1): 77-106.

Walt, Stephen M. 1985. "Alliance Formation and the Balance of World Power". *International Security*. 9(4): 3-43.

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

Wallace, Michael D. 1973. "Alliance Polarization, Cross-Cutting, and International War, 1815-1964: A Measurement Procedure and Some Preliminary Evidence". *Journal of Conflict Resolution*, 17(4): 575-604.

Wallace, Michael D. 1975. "Clusters of nations in the Global System; 1865-1964." *International Studies Quarterly*, 19: 67-110.

Waltz, Kenneth N. 1964. "The Stability of a Bipolar World." *Daedalus*, 93(3): 881-909.

Waltz, Kenneth N. 1970. The Myth of National Interdependence. In *The International Corporation*, edited by Charles P. Kindleberger, 205-23. Cambridge, Mass.: MIT Press.

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

Wayman, Frank Whelon. 1984. "Bipolarity and War: The Role of Capability Concentration and Alliance Patterns among Major Powers, 1816-1965". *Journal of Peace Research*.

21(1): 61-78.

Wayman, Frank Whelon and T. Clifton Morgan. 1990. "Measuring Polarity in the International System". in *Measuring the Correlates of War*. Ed. J. David Singer and Paul F. Diehl. Ann Arbor, MI: The University of Michigan Press.

Wohlforth, William C. 1999. "The Stability of a Unipolar World." *International Security*, 24(1): 5-41. Singer, J. D. 1965. "Threat-Perception and the Armament Tension Dilemma." *Journal of Conflict Resolution* 2(1): 61-62.

Wolfers a. 1962. *Discord and Collaboration*. Baltimore, MD: Johns Hopkins University Press.

Wolford, Scott. 2014. Showing Restraint, Signaling Resolve: Coalitions, Cooperation, and Crisis Bargaining. *American Journal of Political Science*. 58(1), pp.144–56.

Work, Clint. 2020. How to Constructively and Safely Reduce and Realign US Forces on the Korean Peninsula. *The Henry L. Stimson Center: 38 North*. <https://www.38north.org/2020/08/cwork082520/>

Wright, Timothy. 2021. US and South Korea Scrap Ballistic-Missile Range Limits. *IJSS*, June 2. <https://www.iiss.org/online-analysis/online-analysis//2021/06/us-south-korea-ballistic-missile-range-limit>

Zeckhauser, Richard. 1970. Medical Insurance: A Case Study of the Tradeoff Between Risk Spreading and Appropriate Incentives. *Journal of Economic Theory*, 2(1), pp.10-26.

## 5 Appendix A

Clustering at the dyad level, as reported in Table 3, does not account for the autocorrelation due to random shocks within alliances with more than just the two members (i.e. multilateral alliances). Therefore, clustering at the dyad level, while incorporating bilateral alliances and correcting for a degree of autocorrelation, likely underestimates the true standard errors. However, clustering at the alliance level is problematic since many of the allied dyads belong to more than one alliance of varying size in a given year. To enable clustering at the alliance level, I took the most conservative path and aggregated data at the highest level, assigning each dyad to the largest alliance to which it was a party in each year (Cameron & Miller 2015). Large groups that endure for a long period of time create the potential for a great deal of in-group autocorrelation, which makes this a very aggressive method that likely overestimates the standard errors (Abadie et al. 2017).

The results contained in Table 6 contain higher standard errors, reducing the significance of the results. The p-value for POLAR TENSION in model 1 goes from 0.0008 in the previous model to 0.29, while the p-values for the interaction term move from 0.04 to 0.14 and from 0.04 to 0.15 in model 2 and model 3, respectively. Given the more liberal standards applied by clustering at the dyad level and the overly stringent standards associated with applying clustering to the largest possible alliances, the true standard errors likely exist somewhere between the results given in Table 3 and Table 6.

**Table 6.** The Effects of Polar Tension on Policy Alignment of Asymmetric Alliances

	<i>Dependent variable:</i>		
	ALIGNMENT <sub>ijt</sub>		
	(1)	(2)	(3)
POLAR TENSION <sub>t-1</sub>	0.198 (0.146)	-0.086 (0.159)	-0.101 (0.158)
MULTILAT <sub>ijt-1</sub>		-0.012 (0.099)	-0.014 (0.099)
CONTIGUITY <sub>ijt-1</sub>	0.020 (0.019)	0.030* (0.015)	0.035* (0.017)
ln DISTANCE <sub>ijt-1</sub>	-0.053*** (0.014)	-0.034** (0.011)	-0.034** (0.011)
POWER DIFF <sub>ijt-1</sub>	-0.507** (0.183)	-0.610** (0.167)	-0.599** (0.168)
WAR <sub>ijt-1</sub>			-0.116*** (0.030)
DISPUTE <sub>ijt-1</sub>			-0.103* (0.050)
TRADE <sub>ijt-1</sub>			-2.3 x 10 <sup>-7</sup> (1.5 x 10 <sup>-7</sup> )
POLAR TENS <sub>t-1</sub> * MULTILAT <sub>ijt-1</sub>		0.330 (0.221)	0.331 (0.222)
Constant	1.165*** (0.146)	1.039*** (0.104)	1.051*** (0.103)
Observations	7,265	7,265	7,265
R <sup>2</sup>	0.284	0.320	0.326

Note: SE clustered at largest alliance level.

\*p<0.1; \*\*p<0.05; \*\*\*p<0.01