Dynamic Commitment Problems and Military Effectiveness: Resolve, Adaptation, and Flexibility in the Use of Force

by

Todd C. Lehmann

A dissertation submitted in partial fulfillment
of the requirements for the degree of
Doctor of Philosophy
(Political Science)
in the University of Michigan
2021

Doctoral Committee:

Professor James D. Morrow, Co-Chair
Associate Professor Yuri M. Zhukov, Co-Chair
Professor Emeritus Robert Axelrod
Associate Professor John D. Ciorciari

Todd C. Lehmann
tlehmann@umich.edu
ORCID iD: 0000-0003-2934-7545

To Brooke, Mila, Corinne, and Rose.

ACKNOWLEDGMENTS

Writing a dissertation requires tremendous commitment. Even more so is the level of commitment required to complete a dissertation during a pandemic with three young children at home. I could not have found the resolve to see this project through to completion without the enduring support from my wife, Brooke. She has withstood numerous evenings and weekends managing our three boisterous girls while I completed this dissertation. Without her sacrifice, this project would have been a Sisyphean task. I am also grateful for all the support provided by family, friends, and neighbors while Brooke and I have tried to keep so many plates spinning. In particular, I am especially thankful to Brooke's grandmother, Barbara Sloop, for her exceptional encouragement and support for both my graduate studies and for our family.

The members of my committee have provided outstanding training and mentorship throughout my time in graduate school. Bob Axelrod, Jim Morrow, and Yuri Zhukov have each included me as an active participant and colleague in separate research projects, and through these experiences they have taught me a wealth of substantive and methodological material. Additionally, John Ciorciari has been a reliable source of both expertise and encouragement. I am incredibly grateful for the opportunity to work with each of them and grow intellectually in the process.

My data collection would not have been possible without the financial support from Rackham Graduate School at the University of Michigan. Neither would the data have been complete without the invaluable efforts from the team of research assistants who were recruited for various parts of this project: Justin Ballard, Connor Bradbury, Tuhin Chakraborty, Mark Dovich, Victoria Guevara, Angelica Kalogeridis, Joesph Lacy, Manuel Melendez, Daniella Raz, Kate Ruehrdanz, and Julia Stavreva. These individuals were all instrumental in assisting with the tedious but critical tasks of data collection, cleaning, and coding, and devoted many hours to their efforts, for which I am extremely grateful.

TABLE OF CONTENTS

LIST OF FIGURES

FIGURE

LIST OF TABLES

TABLE

ABSTRACT

This dissertation explores the causes and consequences of military commitment problems and evaluates how they impact military effectiveness. Military organizations regularly encounter dynamic, heterogenous environments in which conditions can change both quickly and substantially over time, which can give rise to commitment problems. I investigate three factors necessary for military organizations to be effective in such situations: resolve—the willpower to continue with a course of action despite setbacks; adaptation—the ability to learn from and adjust to novel situations; and flexibility—the ability to respond quickly to different situations. Each of these factors is related to a different commitment problem that military forces often have to confront. First, high signals of resolve within an army can make the commitment to fight credible, such that commanders and troops believe fighting in combat is their best option, rather than fleeing or surrendering. Second, a high level of commitment to a conflict by political leaders can create better conditions for the military to adapt to novel situations and improve their doctrines. Finally, high flexibility through improved force projection capabilities can make security commitments to other states credible, as it allows military forces to respond to crises more quickly and efficiently. I demonstrate that organizational solutions to commitment problems are directly tied to military effectiveness, and along with other types of commitment solutions, provide a better framework for understanding military effectiveness than existing approaches.

CHAPTER 1

Introduction

> To try to measure the various dimensions of the effectiveness of armed forces involves, because of the self-generated momentum of modern war, a measurement of effectiveness in relation to a continual, kaleidoscopic shifting of purposes. Measuring effectiveness becomes almost impossible when the goals to be effected are incorrigibly protean.
>
> —Russell Weigley,
> *Military Effectiveness Vol. III*
> (Weigley, 1988)

Military organizations regularly encounter dynamic, heterogenous environments in which conditions can change both quickly and substantially over time. For example, in the quest for security, states develop military capabilities and improvements in technology, logistics, organizational management, force employment, tactics, and other related features of military power. Alternatively, states can also increase their security by establishing formal alliances (Morrow, 1993) or more limited forms of security cooperation, such as joint military exercises and intelligence sharing (Kinne, 2018). Together, these actions contribute to the security dilemma, whereby one state's pursuit of security causes another state to respond by making its own security improvements, thus creating a perpetual and dynamic process (Jervis, 1978). Likewise, in war, military forces face dynamic situations both during combat engagements and across the entirety of the conflict, as each side competes for an advantage through maneuver and engagement. Consequently, over the course of history, war has become increasingly complex and required more sophisticated force employment against an opponent in order to ensure military effectiveness (Biddle, 2004).

Rapid changes in power can create commitment problems (Fearon, 1995; Powell, 2004, 2006), in which a party to an agreement reneges on the terms of the deal, or chooses not to make a deal because it is unwilling or unable to commit. In such situations, enforcement mechanisms are either nonexistent or weakened due to the changes, which creates uncertainty about an actor's intentions in the future (Schelling, 1960; Axelrod, 1984). States and their military forces often confront a variety of commitment problems in dynamic situations when conditions shift and changes in power alter the terms of an agreement. For example, in an evolving security environment, shifts in capabilities can occur as states innovate and develop new modes of military power, which can challenge the credibility of an alliance commitment. Alternatively, in a high-intensity conventional conflict, changes in military advantage through maneuver or combat attrition can challenge soldiers' commitment to continue fighting cohesively. Similarly, in a protracted insurgency conflict, shifts in information necessary to defeat the insurgents—or the lack thereof—can challenge a state's commitment to continue fighting altogether.

Commitment problems are often understood within the context of bargaining in coercive diplomacy, in which military capabilities can sometimes be used as bargaining power (Schelling, 1966). Military commitment problems are a special class of this more general phenomenon. Military organizations operate in a hierarchical chain of command that regulates decision-making for national security under the threat of war (within coercive diplomacy) or during a war (when coercive diplomacy breaks down). Consequently, military commitment problems can arise at different levels in the chain of command, depending on the context. Scholars and practitioners generally identify four levels within the military command and control structure: political, strategic, operational, and tactical (Millett, Murray, and Watman, 1988). At the tactical level, commitment problems arise from individual soldiers' decisions about how to conduct themselves in battle. The credibility of soldiers' commitment to fighting, rather than surrendering or fleeing, directly impacts a military's ability to fight effectively in war. Separately, the credibility of states' commitment to fighting a war depends upon how well their militaries can adapt doctrine to local conditions, which is driven by feedback that runs throughout the chain of command, from the tactical level up to the political, and back down again. Finally, at the political level, commitment problems arise from political authorities' grand strategy decisions with regard to global force posture: the credibility of alliance commitments therefore contributes to how effectively states can pursue their security.

This dissertation explores the causes and consequences of military commitment problems in order to answer the following question: in dynamic environments where conditions can change quickly and substantially over time, why do military commitment problems arise, how do they impact states' ability to deter threats and fight wars, and how can states overcome them? To answer this question, I investigate three factors that are necessary for military organizations to be effective in dynamic situations: resolve—the willpower to continue with a course of action despite setbacks; adaptation—the ability to learn from and adjust to novel situations; and flexibility—the ability to respond quickly to different situations. Resolve, or the willpower to fight despite the temptation to back down (Kertzer, 2017), is a critical feature of military effectiveness. Without the conviction to continue fighting among both leaders and troops, a military unit will not operate cohesively, leaving it susceptible to coercion and eventually resulting in its defeat.[1] Adaptation is equally important because it ensures better fitness with the environment. Without successful adaptation, a military organization will apply the wrong choices against an opponent and have lower effectiveness as a result. Finally, flexibility is also essential for military effectiveness because it allows military forces to respond to dynamic situations both more quickly and more efficiently.

Each of these three factors is related to a different commitment problem that military forces often have to confront. First, high signals of resolve within an army can make the commitment to fight credible, such that soldiers believe fighting in combat is their best option, rather than fleeing or surrendering. Second, a high level of commitment to a conflict by political leaders can create better conditions for the military to adapt to novel situations and improve their doctrines. Finally, high flexibility through improved force projection capabilities can make security commitments to other states credible, as it allows military forces to respond to crises more quickly and efficiently.

Many existing studies employing game theory have focused on different tools for solving commitment problems. These tools include invoking the actor's reputation, creating costly signals, and establishing commitment devices. The U.S. war in Vietnam is an example where U.S. reputation was invoked as a way to commit itself to the conflict: Secretary of State Henry Kissinger explicitly stated that U.S. commitment in Vietnam was motivated by an effort to signal its credibility to other countries throughout the world (Kissinger, 1969, p. 218–219). Alternatively, costly signals can be used to communicate commitment by demonstrating that only a committed actor would pay the cost of the signal, whereas a less committed actor would not. For example, a state can pay costs by deploying troops to an ally, or tie its hands by invoking audience costs through a formal treaty, as a way to make an alliance commitment credible (Fearon, 1997). Relatedly, commitment devices are a way for an actor to commit itself to a course of action beforehand by making the costs of reneging too high. For example, if a military

[1] "Resolve" can also be considered "morale" when judged as the willingness to fight.

unit in battle is cornered into a position near a bridge, burning the bridge down so that there is no way to retreat is a way to establish a commitment device, which ensures for the soldiers in the cornered unit that fighting is more attractive than fleeing (Schelling, 1966). Some military organizations, such as the Soviet Red Army during World War II, have employed units as "blocking detachments" that effectively serve as bridge burners: these units are placed behind the front lines and ordered to detain—and in some cases shoot—any deserting or retreating soldiers (Lyall and Zhukov, 2020). Because actors might employ several different mechanisms to solve commitment problems, some of these other commitment "solutions" will appear throughout my analysis, though they are not my primary focus.

Instead, I explore three organizational factors related to dynamic commitment challenges—resolve, adaptation, and flexibility—as a way to explain military effectiveness. Accordingly, the organizational solutions to commitment problems I consider are directly tied to military effectiveness, and along with other types of commitment solutions, provide a better framework for understanding military effectiveness than existing approaches.

Previous work on military effectiveness has typically focused on topics related to organizational efficiency. For example, in the three-volume work entitled *Military Effectiveness*—a collection of historical accounts that evaluate conventional military effectiveness during the first half of the twentieth century—organizational efficiency is the central theme: effectiveness is evaluated in terms of how well a military was able to achieve its goals given its capabilities and constraints. Thus, measuring effectiveness, according to this work, is a matter of answering various questions about the extent or degree to which military organizations' actions are consistent with their goals: "To what extent do military organizations have access to manpower in the required quantity and quality," or "To what degree are strategic goals and courses of action consistent with force size and structure," or "To what extent are tactical concepts consistent with operational capabilities," and so on (Millett, Murray, and Watman, 1988).

While evaluating organizational efficiency is one aspect of military effectiveness, it is not the only dimension. Additionally, if we take such a context-specific approach to explaining military effectiveness, we do not allow for much, if any, generalization across cases. Furthermore, if we assume military organizations are comprised of rational actors, then any inefficiencies that exist are puzzles to be explained, rather than simply ways to describe ineffectiveness. Therefore, it is not particularly useful for explanatory purposes to treat efficiency *per se* as the only, or even most important, element of military effectiveness.

Another common approach attempts to quantify military effectiveness with the loss-exchange ratio (LER), which computes relative combat attrition by dividing the number of opposing troops killed by the number of friendly troops killed (Biddle, 2004; Cochran and Long, 2017). Thus, if a military is able to inflict large losses on its opponent relative to the number of losses it suffers, then the idea is that it must be more effective. However, the LER is a narrow measure, because it only considers one dimension—the number of troops killed. Furthermore, it is not a measure that all armies are inherently responsive to. For example, soldiers typically expect casualties in war, and therefore often display high resilience despite their units experiencing attrition in combat, making them less responsive to casualties than the LER would predict. Alternatively, soldiers may have little idea about the relative casualties of each side in a battle, and may be less responsive to the LER as a result. Therefore, in either case, the LER is not a very powerful measure of military effectiveness.

Other common explanations attempt to evaluate military effectiveness by quantifying battles as decisive, or coding battles according to which side was victorious (Biddle, 2004; Grauer and Horowitz, 2012; Stam, 1999). Yet such approaches lack explanatory power due to their subjective and post hoc nature, as others have also noted (Cochran and Long, 2017). In particular, scholars employing such explanations use information about battlefield outcomes to inform the values they choose for the independent variables—such as by coding a battle as victory or defeat, and then evaluating which type of strategy or operational forces were employed—which creates problems with both selection bias and overall measurement bias.

To assess military effectiveness, I instead investigate the causes and consequences of military commitment problems at different levels of analysis in the chain of command.

1.1 Plan of the dissertation

In each of the following chapters, I focus on a particular organizational factor within a given military context. Specifically, these chapters provide a demonstration of how resolve, adaptation, and flexibility contribute to military effectiveness in the context of conventional conflict, unconventional conflict, and international security, respectively. Below I briefly describe the approach for each chapter in more detail.

Chapter 2: "Until the Bitter End? The Diffusion of Surrender Across Battles"

Why do some armies fight until the bitter end, but others collapse and surrender? Existing research has highlighted the importance of battlefield resolve for the onset, conduct, and outcome of war, but has left these life-and-death decisions mostly unexplained. We know little about why battle-level surrender occurs, and why it stops.

In this chapter, I argue that surrender emerges from a collective-action problem: success in battle requires that soldiers choose to fight as a unit rather than flee, but individual decisions to fight depend on whether soldiers expect their comrades to do the same. As a result, surrender becomes contagious across battles because soldiers take cues from what other soldiers did when they were in a similar position. Conversely, where no recent precedent exists, mass surrender is unlikely.

I find empirical support for the diffusion of surrender using a new data set of conventional battles in all interstate wars from 1939 to 2011. The findings in this chapter advance our understanding of battlefield resolve, with broader implications for the design of political-military institutions and decisions to initiate, continue, and terminate war.

Chapter 3: "Adapting Counterinsurgency Doctrine in the Shadow of the Future"

Modern insurgencies create significant evolutionary pressures on conventional militaries to improve their doctrines and overall military effectiveness. Why are some militaries more effective at evolving their counterinsurgency doctrine than others?

This chapter develops a theory which argues that complex local conditions force militaries to optimize their organizations for the commitment horizon they possess in a conflict. This optimization drives militaries to take on particular organizational characteristics, such as the amount of delegation and the tolerance for experimenting with new tactics, which systematically affects their performance.

Using an agent-based model, I evaluate the theory by simulating different organizational characteristics and evaluating how doctrinal effectiveness changes based on different commitment horizons. Results from the simulations suggest a typology of counterinsurgent militaries according to their commitment horizon. Contrary to existing arguments that counterinsurgency effectiveness is determined by a particular set of strategies or by a fixed, preexisting military culture, the results also suggest that doctrinal effectiveness follows a U-shaped curve as the commitment horizon increases. This emphasis on a spectrum of commitment, rather than a simple high/low dichotomy, provides more nuance to political and military behavior and counterinsurgency conflicts. Biddle (2021) takes a similar approach by evaluating the continuum of nonstate actors' military behavior.

To illustrate the logic of the model, I describe the typology and its implications in several historical counterinsurgency cases: the U.S. in Iraq and Afghanistan; the British in Malaya and Iraq; and the French in Indochina and Algeria.

Chapter 4: "Projecting Credibility: Alliance Commitments and Foreign Policy Pursuits"

How does the ability to quickly project large-scale military force—such as with transport aircraft and naval aircraft carriers—affect states' alliance commitments and the likelihood of militarized disputes with other states? In this chapter, I argue that states with greater capacity to project force can deploy fewer troops to allies for deterrent purposes, instead combining their capacity to respond swiftly in a crisis with public words and actions—such as statements by political leaders—to demonstrate their commitment to allies. However, greater force projection capacity is a double-edged sword: although it helps deter threats against allies, it also increases a state's likelihood of initiating militarized disputes in other parts of the world. Using data on troop deployments and a new index for force projection capacity from 1985 to 2018, I show that as force projection capacity increases, states deploy fewer troops, but to more countries. Furthermore, I show that increased force projection capacity works to effectively deter militarized disputes against allies, but it also increases the likelihood of initiating militarized disputes with other states. Together, these results help explain states' behavior with respect to alliances and other foreign policy pursuits in the changing international security environment.

I conclude by summarizing the findings from each of these chapters, and reevaluate how the factors contributing to dynamic military commitments relate to overall military effectiveness. Finally, I consider some directions for future research and the policy implications from my research.

CHAPTER 2

Until the Bitter End? The Diffusion of Surrender Across Battles

> Once fear strikes, it spreads like an epidemic, faster than wildfire. Once the first man runs, others soon follow.
>
> —WWII American paratrooper, *Seven Roads to Hell* (Burgett, 1999)

Note: this chapter was co-authored with Yuri M. Zhukov and was published in *International Organization*, Volume 73, Issue 1, Winter 2019.

Across a sequence of battles, surrender and desertion can cascade through an army, undermining unit resolve and hastening a military's disintegration.[1] During the Battle of Sailor's Creek in the U.S. Civil War, eight Confederate generals and 7,700 troops surrendered to the Union army, following a string of similar events in the Appomattox Campaign. Analogous episodes occurred during the Italian campaign of World War II, Israel's conquest of the Sinai Peninsula in 1967, and, recently, the fall of Ramadi, Fallujah, and Mosul to the Islamic State in Iraq.

Decisions to raise the white flag of surrender have consequences far beyond the battlefield. Besides the obvious—loss of territory, shifts in the lo-

[1] We define *resolve* as a unit's ability to continue fighting as an organized, cohesive force.

cal balance of power—surrender reduces the costs of war for the opponent, making conquest easier and military action more attractive. It is difficult to signal resolve, deter aggression, or compel the opponent to stop fighting if one's own troops will not fight. Surrender is also individually costly—many political authorities consider it high treason, and establish political-military institutions to prevent it. Given the gravity of such decisions, the choice to lay down one's arms is not trivial. Why do soldiers surrender en masse in some battles, but not others?

We argue that battlefield surrender emerges from a collective-action problem within military organizations. Battlefield success requires that soldiers fight as a unit rather than flee, but individual decisions to fight depend on whether soldiers expect their comrades to do the same. When they receive information about recent acts of surrender—within the same army, or in other armies fighting the same opponent—soldiers expect their own unit's resolve to be low, and become less likely to fight. These dynamics are not unlike those driving the diffusion of labor strikes, protests, and insurgency: actors learn from the experience of others and update their beliefs about what others will do in similar situations. Where no recent precedent exists, surrender is unlikely to occur.

Using a new battle-level data set of all conventional wars from 1939 to 2011, we show that surrender is indeed contagious across battles. Soldiers are much more likely to surrender to the enemy if other soldiers have done so recently. This effect holds after we account for alternative explanations of surrender, like military effectiveness and expectations of high losses. We also consider the role of principal-agent dynamics in this process, and show that low expectations of punishment by commanders make soldiers' collective-action problem even worse.

This study advances our understanding of surrender in several ways. On a theoretical level, existing research has highlighted the importance of battlefield resolve for the onset, conduct, and outcome of war, but has left these life-and-death decisions mostly unexplained. International conflict literature has traditionally treated the military as a unitary actor, and a direct, cohesive extension of the state.[2] More direct examinations of battlefield surrender have studied this phenomenon largely in the context of war termination, investigating how surrender affects higher-order political decisions, but not why surrender occurs in the first place (Ramsay, 2008; Weisiger, 2016) (though see Grauer (2014)). Other works have attributed surrender to macro-level institutional features (Belkin, Clark, Gokcek et al., 2002; Castillo, 2014; Lyall, 2014; McLauchlin, 2010; Reiter and Stam, 1997), like regime type, state-society relations, and treaty membership—most of which are relatively static and cannot explain why units from the same military behave differently across battles.

Compounding these theoretical challenges is the reliance of most previous empirical research on highly aggregated, macro-level data, with entire conflicts—rather than individual battles—as units of analysis. This macro-level perspective has conflated the concept of battlefield surrender with war termination, limiting our understanding of how battle dynamics influence decisions to capitulate, and why battlefield surrender occurs in the first place. With a handful of exceptions (Grauer, 2014; Ramsay, 2008; Reiter and Stam, 1997), political scientists have mostly avoided looking below the aggregate level of war, in large part because of the selection problems and limited scope of existing battle-level data sets.[3] Despite the recent proliferation of "micro-comparative" studies of civil war, similarly disaggregated data have been mostly absent from research on conventional war. As a result, quantitative scholars continue to treat wars as unitary black-box events, and qualitative approaches continue to dominate research on surrender.[4]

[2]Research using Jones, Bremer, and Singer (1996)'s militarized interstate disputes (MID) data set, for example, generally implicitly assumes that state leaders purposefully initiate all disputes, while the military faithfully carries out its orders. Even civil-military relations literature, which explicitly questions assumptions of a unitary state, often treats the military itself as a unified entity. See, for example, Feaver (2003).

[3]The most common existing battle-level data set is the U.S. Army's CDB90, otherwise known as HERO. See Dupuy (1984); Helmbold and Kahn (1986). While CDB90 provides a useful baseline for disaggregating wars into battles, its selection of battles is an ad hoc convenience sample: primarily Western front battles in World War II, the Arab-Israeli wars, and the Vietnam War.

[4]In a survey of over 100 academic articles on the topic published in leading political

We build on this previous work by conceptualizing battlefield surrender as a collective-action problem, and test the validity of this perspective with new battle-level data. Our core finding—that surrender can have a cascade effect—challenges macro-level explanations by showing that information about previous battles, rather than the attributes of states fighting them, drive decisions to surrender (Castillo, 2014; McLauchlin, 2010; Reiter and Stam, 1997). To this end, our collective-action approach provides a more comprehensive understanding of battlefield dynamics than existing rational choice approaches, which assume soldiers in the same army act independently from each other (Grauer, 2014; Reiter and Stam, 1997). Finally, our findings open a new empirical frontier for research on intrawar bargaining (Powell, 2004; Slantchev, 2003), by treating resolve not as an exogenous cause of war termination, but as an outcome of primary theoretical interest.

2.1 Surrender as a collective-action problem

Existing research on surrender in battle assumes that soldiers want to survive, and make rational decisions according to the costs and benefits of fighting versus surrendering (Grauer, 2014; Reiter and Stam, 1997). Previous theoretical models, however, assume individualistic utility calculations, and overlook group dynamics that exist in war. From an individual standpoint, fighting is costly. These costs may be outweighed by the benefits of battlefield success, but success is impossible if many soldiers abandon the fight. Surrender, of course, is also not costless. Militaries harshly punish insubordination and desertion, and opponents often do not treat prisoners well. Yet in deciding to fight or flee, soldiers also consider what others are likely to do. If they expect others to flee, they will view success as less likely and opt to surrender rather than die fighting. As an American paratrooper in World War II recalled, "Once fear strikes, it spreads like an epidemic, faster than wildfire. Once the first man runs, others soon follow" (Burgett

science and policy journals in the last twenty-five years, we found that 55 percent used only qualitative methods like process-tracing, 41 percent used quantitative or mixed methods, and 4 percent used formal models.

(1999), quoted in Hamner (2011, p. 79)).

The choices soldiers make on the battlefield are part of a broader class of collective-action problems that drive participation in conflict, violence, and other contentious politics. In a typical threshold model of collective action, a group of individuals decides whether or not to participate in an activity (e.g., riot, strike, protest), depending on how many others are already participating (Granovetter, 1978; Kim and Bearman, 1997; Kuran, 1991; Macy, 1991). Most such models have explored the dynamics of initial mobilization, since groups involved in civil conflict and protest often lack extensive organizational structures initially. These "start-up" challenges are less of a concern for military units in battle, where the state has already overcome initial mobilization problems, and is instead seeking to maintain resolve in the face of outside pressure.

The pre-existence of an organizational structure settles the mobilization challenge, but also adds a layer of complexity highlighted by principal-agent models: the soldiers' collective-action problem unfolds in a hierarchical context, where principals seek to maintain control over their agents' behavior. Yet when agents are part of an organized group and rely on each other's coordinated actions to improve their chances of success and survival, the principal-agent dynamic alone may not fully explain the agents' choices (Holmstrom, 1982). In addition to the threat of punishment from above, soldiers face a more proximate and variable danger on the battlefield, the scope of which depends on whether they expect others to fight or flee. Each choice implies a safety in numbers. A standard principal-agent framework overlooks these collective-action dynamics.

Although many studies have considered how social movements expand and transform, questions of how and why groups decline have traditionally received less consideration in the literature (Koopmans, 2004).[5] Several recent efforts have used global games to model the cohesiveness of a group's

[5]However, there are exceptions. Several studies have explored whether repression (Davenport, 2015; Francisco, 2004; Siegel, 2011), leadership decapitation (Cronin, 2006; Johnston, 2012), and organizational features (Edwards and Marullo, 1995) contribute to groups' decline.

actions in the face of external coercion.[6] These models examine incentives to manipulate information to either prevent or enable an uprising (Bueno de Mesquita, 2010; Edmond, 2013), and the effect of information flows on coordination problems facing both dissidents and the regime (Casper and Tyson, 2014).

The central insight of the collective-action literature—that information about past collective action drives future collective action—implies a diffusion process, where the occurrence of a new event in one context alters the probability of a similar event happening elsewhere (Brooks, 2007; Simmons and Elkins, 2004). In the context of decisions by commanders and troops in war, such processes typically involve the transfer of information from one battle to another, and the updating of prior beliefs about the wisdom of a given action. As armed actors consider the choices before them—the most basic of which is to continue fighting or surrender—they draw lessons from this previous experience. Initially uncertain about the appropriateness of a given action to their situation (i.e., surrender), soldiers examine how previous battles developed, and the decisions people fighting in them made. If surrender has been widespread, this uncertainty declines because soldiers come to expect similar dynamics in the current case, and adjust their own behavior. These information cascades are missing from most existing individualistic rationalist accounts of surrender.

Despite an abundance of empirical research on conflict diffusion, no study has yet examined battlefield surrender as a dynamic, self-reinforcing process.[7] By analyzing these phenomena in a diffusion framework, we can

[6] Carlsson and Van Damme (1993) introduce the global game formulation in a 2x2 one-shot setting, where players face a coordination problem with incomplete information and must choose a strategy based on a noisy signal. Other research extends this approach to a dynamic setting, with a large number of players interacting over multiple rounds; see Angeletos, Hellwig, and Pavan (2007).

[7] The study of diffusion has a long tradition in conflict research. Since the pioneering work of Most and Starr (1980), theoretical and empirical models of diffusion have produced new insights about the onset of interstate and ethnic conflict (Hammarström, 1994; Weidmann, 2015), the spread of innovations and military technologies (Goldman and Andres, 1999; Horowitz, 2010a), the proliferation of tactics like suicide terrorist attacks (Horowitz, 2010b), and the effect of coercion on the spread of insurgent violence (Toft and Zhukov,

potentially account for the endogenous dynamics of learning and updating of beliefs based on prior experience in battles, explain how such processes emerge, and predict if a given case of surrender is likely to spark a general breakdown of war-fighting resolve.

2.1.1 Theoretical expectations

We assume that a military unit's resolve in battle depends on its ability to fight effectively as a team toward some predefined objective.[8] Soldiers within the unit can choose either to *fight* (i.e., contribute an individual effort to the battle and support other soldiers), or *abandon* (i.e., surrender, desert, or defect).[9] Each battle can result in one of two states: *success*, in which a critical mass of soldiers fights and the military maintains its organizational resolve and effectiveness, or *failure*, where organizational resolve breaks down and a critical mass of soldiers choose to abandon.[10] In this sense, "success" and "failure" are conceptually distinct from military "victory" and "defeat," but are not completely orthogonal.[11] Crucially, if enough battles end in "failure" because a critical mass of soldiers abandoned, then political leaders may need to negotiate an end to hostilities on unfavorable terms.

The payoffs to fighting and abandoning are different under the two states. If a soldier chooses to fight when a sufficiently large proportion of others also fight ("success" state), the soldier pays some personal cost

2012).

[8]For a formal derivation of our theory using the framework of global games (Angeletos, Hellwig, and Pavan, 2007; Carlsson and Van Damme, 1993), see the appendix.

[9]Although the relative payoffs between surrendering, deserting, or defecting are likely different in many battles, we argue that the basic process is not. To reduce the argument to its core components, we focus on a binary choice set rather than a multinomial one.

[10]Formally, *resolve* is the maximum level of abandonment a unit can withstand while still being able to effectively fight.

[11]For "success," the actual battle outcome could be victory, stalemate, or defeat. Consequently, a regime may be forced to negotiate an end to hostilities given a series of "success" outcomes if these battles did not fully achieve their strategic objectives, such as gaining territory. However, because of sustained organizational resolve, the terms of ceasefire should be more favorable to the regime in this case. In the "failure" case, however, the actual battle outcome does result in a defeat.

for fighting, but also receives a positive benefit for maintaining resolve and contributing to the effort. If instead he chooses to fight and most others abandon ("failure" state), he receives no positive benefit, but still pays the cost of fighting. If the soldier abandons, in either state, he receives no positive benefit, and pays a different kind of cost, which may include punishment by commanders, harsh treatment as a prisoner of war, or both.

In deciding to fight or abandon, soldiers make inferences about the battle's likely state, using surrender rates in past battles as a noisy signal about their own unit's resolve. If soldiers see that many of their comrades surrendered in recent battles, they will reason that a "failure" state is likely in the current battle, and that payoffs from fighting under these circumstances are likely to be worse than if one abandoned. If past surrender rates were low, soldiers will instead expect a "success" state, where payoffs to fighting are considerably higher. As more battles occur, soldiers receive more information, update their priors, and converge in their beliefs. Thus, we can establish the following testable hypothesis:

> *H1: Battlefield surrender is increasing in the amount of information soldiers receive about high rates of surrender in previous battles.*

2.1.2 Alternative explanations

While past surrender may influence battlefield decisions, soldiers may also look to other types of information to assess whether their comrades will fight or flee. We now survey ten explanations advanced by past research on combat motivation, and consider their implications for our theoretical model and empirical analysis. These explanations range from small-group dynamics within individual units, to macro-level, national institutions. As independent causes of surrender, many of these explanations compete with each other. As we argue, however, nearly all of these explanations are consistent with the collective-action framework, either in influencing expectations of resolve or in shaping individual incentives in battle.

Alternative explanation 1: Mutual surveillance. Expectations of battlefield resolve depend on the observability of battlefield behavior—the ability of commanders to monitor and direct their troops, and of soldiers to monitor each other (Hamner, 2011). For this purpose, in part, soldiers have historically fought in tightly grouped, closed tactical formations (Keegan, 1976). Besides an increased volume of fire, tight formations make abandonment more costly and more visible, compared to dispersed formations, where soldiers are more isolated and unable to observe each others' actions (Hamner, 2011). Although combat tactics have evolved away from tight formations, the mechanism at play—mutual surveillance between soldiers—has imperfect counterparts on the modern, dispersed battlefield.[12] The development of two-way radios and modern communications equipment in the twentieth century, for instance, has allowed isolated groups on the battlefield to coordinate and share information, while giving commanders greater visibility over their actions. Depending on the direction and pace of this technological diffusion—and its consequences for communications capabilities in battle—we may expect different baseline rates of surrender for different combatants, in different wars.

In the context of the collective-action model, mutual surveillance affects soldiers' coordination problems. Increased surveillance lowers information uncertainty and improves coordination, but the effect of this coordination on surrender could conceivably be in either direction. For example, while tight formations can provide visual assurances that others will fight, direct observation of troops abandoning the battlefield could swiftly lead to organizational breakdown. Similarly, increased communication among dispersed soldiers could make it easier to coordinate both fighting and surrendering as a group.

[12]The benefits of tight formations typically exceeded their cost in pre-modern warfare, but changes in the accuracy and destructiveness of weaponry have since turned tightly grouped troops into clear targets for enemy fire. To increase soldiers' survival chances, modern tactics have evolved toward increased dispersion.

Alternative explanation 2: Training and discipline. Some military scholars attribute surrender to problems of military discipline and training. Here, expectations of battlefield resolve stem not from operational experience, but organization-wide standards and procedures. Where military training and discipline are rigorous, "prowess and personal courage all but disappear beneath an armor-plated routine" (McNeill, 1982, p. 130). Where these qualities are lacking, surrender becomes pervasive.

Historically, the emphasis on training and discipline emerged out of efforts to improve battle outcomes. Driven by the increased rate of fire from matchlock firearms, and drawing inspiration from Roman tactics, Maurice of Nassau introduced a series of reforms to Dutch military training in the late sixteenth century, emphasizing smaller units, constant drills, and a clear operational chain of command. These reforms enhanced control over soldiers' actions in battle, decreasing uncertainty over decisions to fight or flee. Whereas medieval "crowd" armies relied on mass and individual talents to win battles (Keegan, 1976), Maurice showed that a smaller, more professional army could consistently defeat a much larger force. Other armies soon took notice and adopted similar tactics and procedures, which they passed on to their institutional successors.[13]

One limitation of military discipline as a cause of battlefield surrender is that discipline tends to vary mostly at the national or organizational level, and—given the time needed to implement new training standards—it changes relatively slowly. While discipline may explain variation in surrender across combatants and across wars, it may be too static to explain variation across individual battles.

Alternative explanation 3: Social cohesion. Following World War II, leading social science explanations of combat motivation attributed battlefield resolve to the strength of within-unit social bonds (Marshall, 1947; Shils and Janowitz, 1948). According to this view, soldiers are less likely to flee

[13]Sustainable improvement in discipline is a perennial challenge for military organizations because of personnel turnover and the individual incentives that compete with organizational purposes.

if strong bonds of mutual trust and loyalty exist between them and their comrades, and more likely to flee if they are socially isolated.[14] As in the collective-action model, mutual expectations about what others will do in battle are central to the social cohesion story. Where the approaches diverge is on the origins of these expectations: within-group social structures drive expectations in the social cohesion school, not information on recent behavior by other units and groups.

While this literature speaks mainly to the internal dynamics of small units, one empirical implication is that recruitment methods matter: surrender should be less likely where units consist of volunteers rather than conscripts (McLauchlin, 2015). In volunteer armies, interpersonal relationships are generally less conflictive, and within-unit social integration is greater (MacCoun, Kier, and Belkin, 2006; Siebold, 2007). In armies staffed by long-service professionals, soldiers may therefore expect a higher baseline of resolve.

Alternative explanation 4: Ideological cohesion. One challenge to the social cohesion perspective is that unit composition can change dynamically through combat and attrition, yet soldiers often continue to fight—even after initial unit social structures collapse. Drawing on the experience of World War II, Bartov advances an alternative explanation for combat motivation, attributing surrender not to mutual expectations of battlefield behavior, but to the ideology instilled within soldiers by political authorities (Bartov, 1992). Where this indoctrination is more extreme and uncompromising (e.g., German troops opposing the Soviet army on the Eastern Front of World War II, or Japanese troops fighting the Allies in the Pacific theater of World War II), soldiers should expect higher resolve in their army, and will therefore be more reluctant to surrender.

[14] More recent research on unit cohesion has shifted away from social structures, and toward units' commitment to specific combat missions and tasks; see, for example, MacCoun (1993). This view holds that trust among soldiers stems not from social bonds, but from soldiers' performing their jobs to accomplish their common mission. In many ways, task cohesion in an army is closely related to military training and discipline, outlined in alternative explanation 2.

The empirical implication of the ideological cohesion school is straightforward: surrender rates should increase as ideological cohesion breaks down. The problem is that, in war, state ideology tends to change slowly, if at all. While ideological differences may help explain variation between countries and between wars, changes in political ideology may not occur frequently enough to explain organizational breakdown within war. For example, there was no coinciding shift in the Nazi regime's ideology when Wehrmacht troops began surrendering en masse in 1945.

Alternative explanation 5: Aggregate military power. While we can group several of the previous explanations under the general rubric of "troop quality," rates of surrender may also depend on aggregate preponderance in capabilities, and more general perceptions of the balance of power. At the macro level, bargaining models of war—both the "costly lottery" (Fearon, 1995; Powell, 1996) and "costly process" (Slantchev, 2003; Smith and Stam, 2004) variants—assume that the probability of military victory follows the dyadic balance of power. A lopsided balance should therefore increase resolve expectations in the more powerful army. Conversely, soldiers in weaker armies should anticipate that more of their comrades will lay down their arms, rather than fight a hopeless battle.

Alternative explanation 6: Offensive advantages. One criticism of macro-level preponderance is that a smaller force can concentrate its strength at the weak point of the adversary, creating local superiority despite aggregate disadvantage. As a result, macro-level perceptions of the balance of power may have little bearing on battle-level outcomes. This insight lies at the core of literature on offensive advantages (Van Evera, 1998) and the "3:1 rule" (Mearsheimer, 1989). Within wars, attackers generally begin battles with numerical superiority, to offset the challenges of fighting defenders in prepared positions. As a result of local preponderance—and other first-mover advantages like speed, initiative, and surprise—expectations of resolve may be higher among attacking troops.

Alternative explanation 7: Principal-agent problems. From the standpoint of military leadership, battlefield surrender represents a principal-agent problem (Gates and Nordås, 2016): commanders delegate direct orders to lower-ranking personnel, but cannot perfectly observe that personnel's performance. To prevent surrender, commanders often institute harsh punishment for insubordination and desertion. During World War II, for instance, Stalin's order number 270 stated that Soviet personnel "who surrender to the enemy shall be considered malicious deserters, whose families are liable to be arrested" (Zolotarev, 1997, p. 58–60). Soviet General Georgy K. Zhukov had a reputation for publicly executing deserting troops to deter others from fleeing. The implication for the collective-action framework is that where monitoring is less effective and punishment is less severe, soldiers' incentives to surrender may be higher.

Commanders' ability to monitor and punish subordinates depends on many things, including institutional factors like discipline and training, and tactical considerations like mutual surveillance. Empirically, one circumstance where principal-agent problems are arguably most acute is that of a breakdown in leadership: a commander who surrenders in battle is one who cannot effectively monitor or punish surrendering troops. To the extent that such principal-agent dynamics may help solve soldiers' collective-action problem, we should expect higher troop-surrender rates in armies where senior officers had surrendered in recent battles.

Alternative explanation 8: International law. The expected costs of surrender depend not only on internal dynamics within one's own military, but also on the opponent's likely treatment of detainees. All else equal, soldiers are more likely to surrender if they believe the opponent will treat prisoners well (Grauer, 2014; Reiter and Stam, 2002). One potentially informative signal of humane treatment is the ratification of treaties stipulating basic rights for wartime prisoners, like the Geneva Conventions (Morrow, 2014). Where the opponent has made such commitments under international law, soldiers may expect the cost of surrendering to be lower than that of fighting, and anticipate that fewer of their comrades will stay. Like other macro-level factors, however, treaty ratification is a relatively static variable, better suited for explaining cross-national variation than battle-level outcomes.

Alternative explanation 9: Political regime type. Another national-level signal of humane treatment is political regime type (McLauchlin, 2010; Reiter and Stam, 1997). Because of autocracies' comparative lack of transparency, repressive institutions, and weaker records on human rights, soldiers may doubt these regimes' commitments to protecting prisoners from abuse. For this reason, soldiers fighting against more democratic armies may expect more of their comrades to surrender, while those fighting against more repressive regimes may expect surrender to be rare. Yet regime type is another macro-level variable that typically remains constant over the course of a war. While it may explain why troops in some armies have different incentives for fighting than troops in other armies, it is less informative of why troops from the same country, fighting the same opponent, are more likely to surrender in some battles than others.

Alternative explanation 10: Military effectiveness. Finally, because troops often surrender, at least in part, because they are either losing or expect to lose, information about past surrender may simply be a proxy for broader conceptions of military effectiveness: whether one's army "won" or "lost" a battle, and how well others had fared against the same opponent. Because definitions of "winning" and "losing" tend to be subjective and battle specific (Biddle and Long, 2004), quantitative measures of military effectiveness have tended to focus on relative casualties inflicted by each side. One example is the loss-exchange ratio (LER), or the number of enemy troops killed divided by the number of friendly troops killed (Biddle, 2004; Cochran and Long, 2017). In this case, if troops enter a battle knowing that others in their position have suffered significant casualties while inflicting little damage on the opponent, they may see success as unlikely.[15] This expectation alone may be enough to make them lay down their arms.

These explanations highlight the extensive scholarly debate on the determinants of battlefield resolve. At their root is an inherent tension between a soldier's individual motivation to survive, and the physical danger of taking or defending some political objective through force. The question is why soldiers are sometimes able to overcome their survival instincts and other times not. Most of these approaches agree that the answer depends on what soldiers expect their comrades to do in battle: fight if they expect others to fight, flee if they expect others to flee.

The first nine explanations are not inconsistent with a collective action framework. Mutual surveillance helps alleviate coordination problems in battle. Discipline, social cohesion, ideology, aggregate power, and attacker advantages all affect expectations of resolve. Principal-agent problems, international law, and political regime type shape individual incentives in various ways.

Despite this overlap, the collective-action perspective diverges from existing accounts in two important respects. First, unlike ideology, military discipline, regime type, and international law—which assume a relatively static set of expectations in an army over the course of a war—our model allows these expectations to either remain firm or change as soldiers receive more information about what others have done. Second, unlike unit cohesion and mutual surveillance—where surrender is primarily an intra-unit and intra-battle phenomenon—we allow these dynamics to extend across units and battles, with past surrender in one unit affecting future surrender in other units and even other countries' armies.

In sum, the collective-action framework conceptualizes surrender as a process that unfolds endogenously across battles, depending on the dynamic flow of information. Competing with this approach, however, is the

[15] To evaluate this possibility, we will explore several different ranges over which a soldier makes this judgment.

view that surrender is a by-product of military effectiveness: since soldiers observe information not just about surrender, but also about their military's general performance in battle, expectations of relative casualties could be driving decisions to surrender.

2.2 Data

To enable the empirical study of surrender, we developed a new battle-level data set of conventional wars composed of every major battle in interstate conflicts from 1939 to 2011. To overcome the selection problems present in CDB90/HERO and other previous battle data sets, we opted to collect data for the full population of interstate conflicts since and including World War II, using Correlates of War (Singer, 1979; Sarkees and Wayman, 2010) to enumerate the population of cases for which battle data were to be collected. For each interstate conflict, we assembled a chronological list of battles from historical encyclopedias (Clodfelter, 2008; Showalter, 2013).

Since wars are hierarchical enterprises conducted by hierarchical organizations, their disaggregation requires some non-trivial decisions about what constitutes an individual battle. For our purposes, we define a battle as a major engagement involving at least two opponents fighting over some clearly defined overarching military objective. This definition does not require disaggregation down to every skirmish between small units, since such actions are typically part of larger efforts. Rather, we collected data for discrete campaigns, disaggregating them further if they entailed multiple distinct operational objectives and are detailed as such in historical records. For example, we coded the Normandy D-Day landings by Allied forces on 6 June 1944 as a single battle, rather than dividing it into sub-objectives like the Gold, Juno, Sword, Omaha, and Utah Beaches. However, we include separate battles for Caen and St. Lô, since these D-Day objectives saw subsequent fighting distinct from the Normandy landings.[16]

[16]In this sense, some battles enter the data set as distinct events because of how military efforts unfolded over time, rather than because they were independently planned objec-

In all, our data include 597 battles from eighty-two conflicts, covering 83 percent of interstate conflicts in Correlates of War between 1939 and 2011.[17] We collected data for each battle participant, including each coalition member fighting on the attacking and defending sides, yielding a total sample size of 1,720 battle-dyads.[18]

We collected location data for each battle from historical maps and military atlases, recording the geographic coordinates for the towns or geographic features where fighting took place. We used the distribution of these locations to construct convex hull polygons encompassing the largest extent of area over which forces were engaged.[19] Figure 2.1 illustrates the spatio-temporal distribution of battles in our data set. We used these data to calculate deployment distances to each battle, as well as the geographic size of the front, and the temporal sequence of events.

We used Clodfelter as the primary source for data on battle participants, troop numbers and casualty statistics, including killed (KIA), wounded (WIA), missing in action (MIA), prisoners of war (POWs), defections, and desertions.[20] To capture military commanders' influence on subsequent events, we coded separate binary variables as 1 if a flag officer (i.e., those ranked in the general or admiral grade or equivalent) surrendered, defected, or was captured or killed in battle.[21] In addition to raw counts of ca-

tives. Since we are focused on battle outcomes, rather than causes, this inclusion criterion is appropriate for our needs.

[17]Interstate wars for which we do not currently have data include: Franco-Thai War of 1940–1941, Offshore Islands War of 1954, Ifni War of 1957–1959, Taiwan Straits War of 1958, War of Attrition of 1969–1970, Sino-Vietnamese Border War of 1987, Kargil War of 1999.

[18]The dyads here are directed. For example, the USSR-Germany dyad for Stalingrad enters the data more than once—first with Germany as the focal combatant (attacker vs. defender), once with the USSR (defender vs. attacker), and with additional observations for Italy and other Axis members fighting the USSR in the battle.

[19]For naval battles, we used approximate geographic coordinates to encompass areas of water in which ships were attacked or sunk.

[20]Clodfelter (2008) provides a relatively comprehensive account of force strength and losses, but organizes this information primarily in narrative form, rather than as tables of statistics.

[21]We rely on Clodfelter's narratives to indicate whether a flag officer surrendered, defected, or was captured or killed in battle. Since these events are high profile and are typically prominently highlighted in historical records, we assume that Clodfelter captured

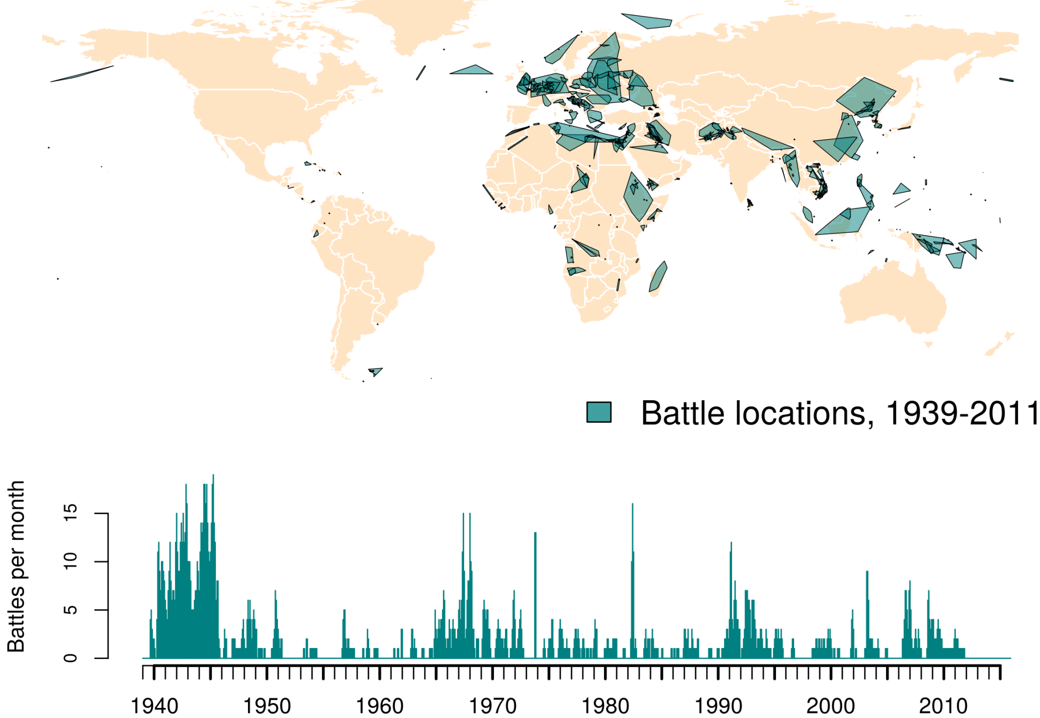


Figure 2.1: Spatio-temporal distribution of battles in data

sualties and prisoners, we calculated the loss-exchange ratio (LER) for each battle participant (i.e., enemy casualties divided by friendly casualties)—a standard measure of relative attrition. To account for relative differences in personnel surrendering from smaller and larger formations, we created an ordinal measure of battle size.[22]

most of these events in his narratives.

[22]The levels, based on the size of the largest force participating in the battle, are (1) 0–5,000 troops, (2) 5,001–20,000, (3) 20,001–100,000, (4) 100,001–400,000, (5) 400,001–1,000,000, (6) 1,000,001–10,000,000. These correspond, roughly, to (1) brigade and below, (2) division, (3) corps, (4) army, (5) army group, (6) theater. We do not use per capita surrender rates for two reasons. First, the number of personnel directly engaged in combat can vary greatly over the course of a battle. Second, per-capita surrender rates can be inherently misleading: the forces engaged in large battles are often only a small proportion of the total force size present, whereas those engaged in smaller battles are more likely to include a larger proportion of the total force present.

To test alternative hypotheses proposed in past literature and control for other potential confounders, we supplemented this battle-level information with country-year-level variables from other sources. To account for political regime type, we used a modified version of the Polity index (Marshall, Gurr, and Jaggers, 2014).[23] To account for perceptions of the overall balance of power, we measured relative military capacity, using the Composite Index of National Capabilities (CINC) (Singer, Bremer, and Stuckey, 1972).[24] Because national-level measures of relative power potentially mask significant local imbalances, we included an indicator of which side is on the offensive, as well as a measure of the local force ratio at the start of the battle. We also created a dummy variable for a state's primary means of recruitment (from Toronto (2005))—coded 1 if a state relied only on volunteers and 0 if it also relied on conscripts. Finally, we considered whether each side had ratified the Geneva Conventions (International Committee of the Red Cross, 2016).[25]

Table 2.1 lists the descriptive statistics for battle-level and country-level variables.

2.3 Data analysis

We model the determinants of battlefield surrender as follows:

$$y_{ijk} = \rho W(y) + \gamma Z_k + \beta X_{ij} + \alpha_i + \zeta_m + \theta_{\tau(k)} + u_{ijk} \tag{2.1}$$

Our unit of analysis is the battle-dyad, where m indexes the war (e.g., World War II), k indexes the battle (e.g., Stalingrad), i indexes the focal combatant (e.g., USSR) and j indexes the opponent (e.g., Germany). The dependent

[23] Because Polity often assigns scores of -66 (foreign interruption), -77 (interregnum), or -88 (transition) for country-years at war, we converted these missing values to the regime's most recent Polity score prior to its -66/-77/-88 value.

[24] CINC captures states' combined population, military personnel and expenditures, iron and steel production, and energy consumption as a proportion of the world total.

[25] Prior to 1950, when the first state ratified the 1949 Geneva Conventions, we coded the variable based on the 1929 version of the treaty.

Variable	Mean	SD	Min	Max	# Missing
START YEAR	1966.27	23.63	1939	2011	0
POW (PERSONNEL)	13,829.11	80,023.16	0	1,200,000	876
MIA (PERSONNEL)	360.63	2916.20	0	60,000	1005
KIA (PERSONNEL)	6557.61	31,845.47	0	458,080	932
WIA (PERSONNEL)	10,900.44	83,661.08	0	1855,605	1077
BATTLE SIZE	2.42	1.27	1	6	387
FORCE RATIO	6.03	127.18	0	3909.09	775
POLITY 2	2.79	7.80	-10	10	392
MORE DEMOCRATIC	0.60	0.49	0	1	834
CINC	0.08	0.10	0	0.38	385
MORE POWERFUL	0	0.12	-0.33	0.33	735
LER SIDE A	1747.91	6463.96	0	26,000	1069
LER SIDE B	16,202.99	51,305.89	0	179,000	911
COMMANDER SURRENDERED	0.03	0.16	0	1	2
PROFESSIONAL ARMY	0.17	0.38	0	1	452
GENEVA (OPPONENT)	0.65	0.46	0	1	2
DEPLOYMENT DISTANCE	5959.38	7595.41	0	35,279.91	329
INITIATOR	0.56	0.50	0	1	2

Note: $N = 1720$.

Table 2.1: Descriptive statistics: battle-combatant variables

variable y_{ijk} is the logged number of soldiers from combatant i who surrendered to opponent j in battle k.

The parameter of primary theoretical interest is ρ, which captures the influence of past surrender rates on surrender in the current battle. We specify the set of combatants and past battles that influence i's decision to surrender with an information flow network, $W(y)$. We consider two types of information: instances of past surrender by combatant i to all opponents in war m (SAME COMBATANT), and past surrender by all other combatants to opponent j during war m (SAME OPPONENT). Following Zhukov and Stewart (2013), we estimate the ρ coefficient in separate models for each diffusion measure.[26]

We assume that soldiers place greater weight on more recent and geographically proximate cases of surrender. We specify the *temporally weighted* diffusion term as

$$W(y)^{[\text{same combatant}]} = \sum_{t}^{\tau(k)} \frac{y_{i.t}}{(1+r)^{\tau(k)-t}} \tag{2.2}$$

$$W(y)^{[\text{same opponent}]} = \sum_{t}^{\tau(k)} \frac{y_{.jt}}{(1+r)^{\tau(k)-t}} \tag{2.3}$$

where $\tau(k)$ is the start date of battle k, and t indexes the start dates of previous battles in war m, involving either the same combatant (Eq. 2.2) or other combatants fighting opponent j (Eq. 2.3).[27] The temporal discount rate is $r \in (0,1)$, with higher values placing a greater weight on more recent battles, and $r = 0$ placing equal weight on all past battles in war m. Because we do not have a strong prior on r, our empirical models automatically select values that minimize the Akaike Information Criterion (AIC).[28] We also

[26]Zhukov and Stewart (2013) show that including multiple, partially overlapping networks in a single model can yield biased estimates of autoregressive parameters.

[27]In wars with only two combatants, the "same combatant" and "same opponent" measures should converge to the same value.

[28]Using the Bayesian Information Criterion (BIC) will give the same result as using the AIC, because we are only modifying one parameter value of the model, and the model's complexity remains constant.

provide sensitivity analyses for all $r \in (0,1)$.

We specify the *geographically weighted* diffusion term as

$$W(y)^{[\text{same combatant}]} = \sum_{t}^{\tau(k)} y_{i.t} \frac{d_k^{-r}}{\sum_t^{\tau(k)} d_k^{-r}} \tag{2.4}$$

$$W(y)^{[\text{same opponent}]} = \sum_{t}^{\tau(k)} y_{.jt} \frac{d_k^{-r}}{\sum_t^{\tau(k)} d_k^{-r}} \tag{2.5}$$

where d_k is the geographic distance, in kilometers, between battle k and all previous battles in war m, involving either combatant i (Eq. 2.4) or other combatants fighting opponent j (Eq. 2.5). The spatial discount rate, $r \in (0,1)$, is selected by AIC, with higher values assigning greater influence to past battles closer to k.

In addition to the information flow network, our model includes a set of battle-level (Z_k) and dyad-level covariates (X_{ij}). These include essential control variables like battle size, and variables needed to account for additional explanations of surrender, like recruitment (i.e., whether i has a professional army), regime type (i.e., whether i has a higher Polity2 score than j), and treatment of prisoners (i.e., whether j has ratified the Geneva Conventions), as well as controls for relative power (i.e., difference in CINC scores between i and j; local force ratio between i and j), offensive and defensive battles, logistics (i.e., i's deployment distance), and time (i.e., year in which the battle began). We also include fixed effects for each combatant (α_i), war (ζ_m), and season of the year ($\theta_{\tau(k)}$), and an i.i.d. error term (u_{kij}). These fixed effects help us account for relatively static, macro-level drivers of surrender, like ideology, discipline, and technological change from war to war.

2.3.1 Results

Tables 2.2 and 2.3 report the empirical determinants of surrender, with temporally and geographically weighted diffusion terms, respectively. The first two models in each table estimate the effect of information on past surrender by the SAME COMBATANT (Model 1) and other combatants fighting the SAME OPPONENT (Model 2). The remaining models incorporate battle-level and combatant-level covariates. Because parameter estimates are sensitive to scales of measurement, we report standardized coefficients—representing estimated standard deviation (SD) changes in the outcome following a standard deviation increase in the explanatory variable.

Surrender is contagious across battles. The analysis reveals strong evidence for our hypothesis: surrender is more intense following other recent cases of surrender. According to Model 1 in Table 2.2, a standard deviation increase in recently surrendered troops from the same army increases the logged number of troops surrendering in the current battle by .27 SD (95% confidence interval: .21, .32). This figure is slightly smaller, .24 SD (95% CI: .18, .31), for surrender from other armies fighting the same opponent (Model 2).

Parameter estimates are of similar relative magnitude for the geographically weighted diffusion measures in Table 2.3, which represent the influence of past surrender in nearby battles. A standard deviation increase in surrendering troops in spatially proximate battles produces a .23 SD rise (95% CI: .16, .30) in logged POWs if the surrendering troops were from the same army (Model 1), and a .19 SD increase (95% CI: .12, .26) if they were from armies fighting the same opponent.

Figure 2.2 shows a graphical representation of this relationship, through simulations based on Models 1 and 2 in Table 2.2 (temporal weights), with fixed effects for Russia/USSR fighting an average summer battle in World War II. Following a hypothetical increase from 0 to 300,000 recent prisoners of war from the same combatant—roughly equivalent to Soviet POW rates during the 1941 Battle of Smolensk—the expected number of surrendering troops rises by 330 percent per battle, on average (95% CI: 245, 560), from 68,880 to 296,556.[29] The rise is a smaller, but still formidable 139 percent (95% CI: 119, 182), from 92,768 to 222,161 per battle, following an identical

[29] Predictions based on Model 1 in Table 2.2.

	Dependent variable:					
	log(POWs)					
	(1)	(2)	(3)	(4)	(5)	(6)
SURRENDER IN PREVIOUS BATTLES $W(y)$						
W(SAME COMBATANT)	0.27***		0.21***		0.21***	
	(0.21, 0.32)		(0.13, 0.29)		(0.12, 0.29)	
W(SAME OPPONENT)		0.24***		0.09*		0.10*
		(0.18, 0.31)		(0.004, 0.18)		(0.01, 0.19)
BATTLE LEVEL COVARIATES Z_k						
BATTLE SIZE			0.27***	0.28***	0.27***	0.28***
			(0.17, 0.37)	(0.18, 0.39)	(0.17, 0.37)	(0.17, 0.39)
LOG(FORCE RATIO)					−0.03	−0.01
					(−0.08, 0.03)	(−0.07, 0.04)
DEPLOYMENT DISTANCE			−0.01	−0.05	0.11	0.07
			(−0.20, 0.19)	(−0.25, 0.15)	(−0.10, 0.31)	(−0.14, 0.28)
INITIATOR			−0.63***	−0.69***	−0.61***	−0.67***
			(−0.81, −0.45)	(−0.87, −0.50)	(−0.80, −0.43)	(−0.86, −0.48)
START YEAR			−0.02	0.07*	−0.01	0.08*
			(−0.09, 0.05)	(0.01, 0.13)	(−0.08, 0.06)	(0.01, 0.14)
DYAD AND NATIONAL-LEVEL COVARIATES X_{ij}						
MORE POWERFUL			−0.06	−0.13	−0.06	−0.12
			(−0.20, 0.09)	(−0.28, 0.02)	(−0.22, 0.09)	(−0.28, 0.04)
PROFESSIONAL ARMY			−0.02	−0.24	−0.31	−0.56
			(−0.78, 0.73)	(−1.01, 0.53)	(−1.22, 0.61)	(−1.50, 0.38)
MORE DEMOCRATIC			−0.31	−0.14	−0.10	0.002
			(−1.33, 0.70)	(−1.19, 0.92)	(−1.18, 0.98)	(−1.12, 1.12)
GENEVA (OPPONENT)			0.78	0.51	0.74	0.46
			(−0.09, 1.64)	(−0.38, 1.39)	(−0.13, 1.60)	(−0.42, 1.35)
Seasonal fixed effects	✓	✓	✓	✓	✓	✓
War fixed effects	✓	✓	✓	✓	✓	✓
Combatant fixed effects	✓	✓	✓	✓	✓	✓
Observations	844	844	426	426	411	411
Adjusted R^2	0.34	0.30	0.49	0.46	0.50	0.47
Log Likelihood	−1,092.13	−1,119.79	−518.64	−530.11	−498.13	−508.75
UBRE	0.81	0.87	0.69	0.73	0.68	0.72
RMSE	0.73	0.75	0.68	0.69	0.67	0.69
AIC	2184.25	2239.59	1037.27	1060.22	996.26	1017.51

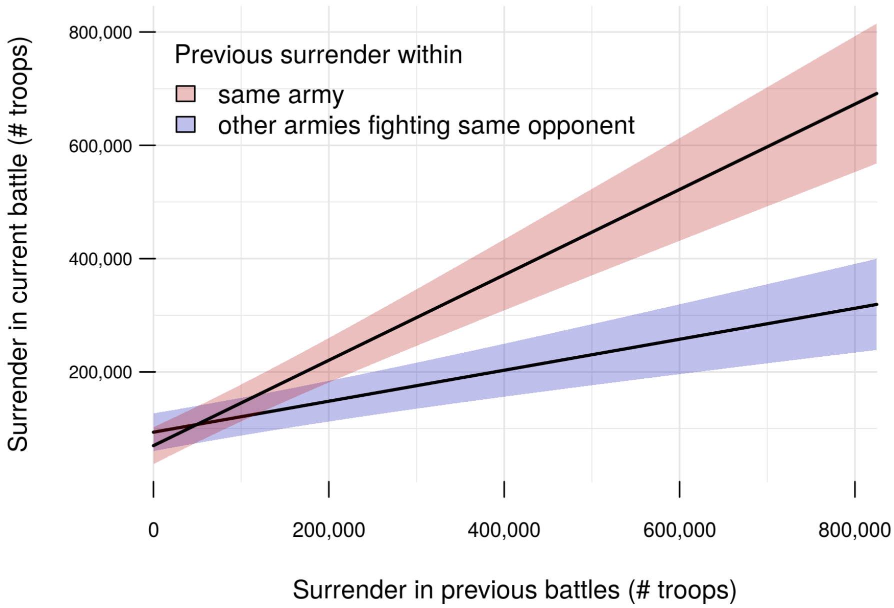
Notes: Standardized coefficients. 95% confidence intervals in parentheses.
$^{*}p<0.05$; $^{**}p<0.01$; $^{***}p<0.001$

Table 2.2: Determinants of battlefield surrender (temporal weights)

	Dependent variable:					
	log(POWs)					
	(1)	(2)	(3)	(4)	(5)	(6)
SURRENDER IN PREVIOUS BATTLES $W(y)$						
W(SAME COMBATANT)	0.23***		0.13**		0.09	
	(0.16, 0.30)		(0.04, 0.22)		(0.00, 0.18)	
W(SAME OPPONENT)		0.18***		0.10*		0.09*
		(0.11, 0.25)		(0.03, 0.18)		(0.005, 0.17)
BATTLE LEVEL COVARIATES Z_k						
BATTLE SIZE			0.29***	0.29***	0.29***	0.29***
			(0.19, 0.39)	(0.19, 0.39)	(0.18, 0.39)	(0.18, 0.39)
LOG(FORCE RATIO)					−0.02	−0.02
					(−0.08, 0.03)	(−0.08, 0.03)
DEPLOYMENT DISTANCE			−0.04	−0.02	0.06	0.08
			(−0.24, 0.16)	(−0.22, 0.18)	(−0.15, 0.27)	(−0.13, 0.30)
INITIATOR			−0.72***	−0.72***	−0.71***	−0.71***
			(−0.90, −0.54)	(−0.90, −0.54)	(−0.90, −0.53)	(−0.89, −0.53)
START YEAR			0.06	0.07*	0.07*	0.08**
			(−0.01, 0.12)	(0.01, 0.14)	(0.01, 0.13)	(0.02, 0.15)
DYAD AND NATIONAL-LEVEL COVARIATES X_{ij}						
MORE POWERFUL			−0.15*	−0.16*	−0.17*	−0.17*
			(−0.30, −0.01)	(−0.30, −0.01)	(−0.31, −0.02)	(−0.32, −0.02)
PROFESSIONAL ARMY			−0.30	−0.30	−0.52	−0.55
			(−1.07, 0.46)	(−1.07, 0.47)	(−1.46, 0.41)	(−1.49, 0.38)
MORE DEMOCRATIC			−0.39	−0.36	−0.17	−0.15
			(−1.43, 0.65)	(−1.40, 0.68)	(−1.29, 0.94)	(−1.26, 0.97)
GENEVA (OPPONENT)			0.54	0.46	0.50	0.43
			(−0.34, 1.42)	(−0.43, 1.34)	(−0.39, 1.38)	(−0.45, 1.32)
Seasonal fixed effects	✓	✓	✓	✓	✓	✓
War fixed effects	✓	✓	✓	✓	✓	✓
Combatant fixed effects	✓	✓	✓	✓	✓	✓
Observations	844	844	426	426	411	411
Adjusted R^2	0.29	0.28	0.47	0.47	0.47	0.47
Log Likelihood	−1,126.69	−1,132.76	−528.07	−528.64	−509.22	−508.94
UBRE	0.88	0.89	0.72	0.73	0.72	0.72
RMSE	0.76	0.76	0.69	0.69	0.69	0.69
AIC	2253.38	2265.52	1056.13	1057.27	1018.43	1017.88

Notes: Standardized coefficients. 95% confidence intervals in parentheses.
$^{*}p<0.05$; $^{**}p<0.01$; $^{***}p<0.001$

Table 2.3: Determinants of battlefield surrender (geographic weights)



Note: Simulations based on Models 3 and 4 in Table 2.2. Fixed effects: ζ_m = World War II, α_i = Russia/USSR, $\theta_{\tau(k)}$ = summer.

Figure 2.2: Impact of past surrender on surrender in the current battle

increase in POWs among other armies fighting the same opponent.[30]

A greater sensitivity of troops toward past surrender rates in their own army is not surprising. From a theoretical standpoint, signals soldiers receive through the SAME COMBATANT network should be less noisy than those from the SAME OPPONENT network. With pre-existing social networks, communication channels, and rumor mills, troops are likely to be better informed about the conduct of units within their own military than other countries' armed forces—even if the latter are part of the same coalition. Troops may also see previous surrender within their army as a more indicative signal of how their own comrades will behave.

Contagion effect is stronger if soldiers think opponent treats prisoners well. Are soldiers more likely to surrender to opponents who have signed

[30] Predictions based on Model 2 in Table 2.2.

treaties on the humane treatment of prisoners of war? The evidence here is more mixed. Tables 2.2 and 2.3 suggest that—if such an effect does exist—it cannot explain battle-level variation on its own. The coefficient for GENEVA (OPPONENT) is statistically insignificant in Models 3 and 4.

There is, however, tentative evidence for an interactive relationship between Geneva ratification and past surrender, supporting existing arguments that expectations of humane treatment increase surrender levels.[31] As Table 2.4 shows, the contagion effect is stronger if the opponent has ratified the Geneva Conventions. Here, a standard deviation increase in past surrender within the same army yields an increase of between .22 (Model 3) and .31 SD (Model 1). Where the opponent had not ratified (about 30 percent of cases), past surrender has no effect.

This apparent heterogeneity is not surprising, given the logic of collective action. Where the opponent has ratified the conventions, soldiers can reasonably expect the costs of surrender—narrowly defined as the probability of harm or death in captivity—to be relatively low. Where opponents have not ratified, soldiers are more uncertain about these costs, and are more hesitant to pay them. As a result, ratification of the Geneva Conventions can amplify the contagion effect.

Troops are more likely to surrender if senior officers recently surrendered. Can some actions by military leaders potentially accelerate the tide of surrender? The role of such principal-agent problems is difficult to empirically establish without battle-level data on monitoring and punishment of deserting troops. To the extent that more autocratic governments can institute more draconian forms of punishment than democracies (Castillo, 2014), we could assume that the costs of surrendering are higher in the armies of more repressive regimes. Yet the negative and insignificant coefficients on MORE DEMOCRATIC in Tables 2.2 and 2.3 are not what we would expect to find if such regimes succeeded in deterring troops from surrendering. More-

[31] Here, our evidence more strongly supports Morrow (2014)'s argument that treaty ratification shapes battlefield behavior, than Grauer (2014)'s and Reiter and Stam (2002)'s rejections of this claim.

	Dependent variable:			
	log(POWs)			
	(1)	(2)	(3)	(4)
INTERACTION $W(y) \times X_{ij}$				
INTERACTION (SAME COMBATANT)	0.58***		0.45*	
	(0.30, 0.86)		(0.02, 0.87)	
INTERACTION (SAME OPPONENT)		0.26		0.58
		(−0.29, 0.80)		(−0.08, 1.25)
SURRENDER IN PREVIOUS BATTLES $W(y)$				
W(SAME COMBATANT)	−0.27*		−0.23	
	(−0.54, −0.01)		(−0.64, 0.17)	
W(SAME OPPONENT)		−0.01		−0.51
		(−0.54, 0.52)		(−1.16, 0.15)
GENEVA CONVENTIONS X_{ij}				
GENEVA (OPPONENT)	0.02	0.11	0.36	0.61
	(−0.35, 0.39)	(−0.29, 0.52)	(−0.61, 1.32)	(−0.28, 1.50)
Battle-level covariates			✓	✓
Dyad-level covariates			✓	✓
Seasonal fixed effects	✓	✓	✓	✓
War fixed effects	✓	✓	✓	✓
Combatant fixed effects	✓	✓	✓	✓
Observations	844	844	426	426
Adjusted R^2	0.36	0.30	0.49	0.46
Log Likelihood	−1,083.61	−1,121.22	−518.16	−530.26
UBRE	0.79	0.87	0.69	0.73
RMSE	0.72	0.75	0.67	0.69
AIC	2167.22	2242.44	1036.31	1060.52

Notes: Standardized coefficients reported, 95% CI in parentheses.
$^{*}p<0.05$; $^{**}p<0.01$; $^{***}p<0.001$

Table 2.4: Interaction between past surrender and opponent's ratification of Geneva Conventions

over, while regime type changes relatively slowly, commanders' treatment of subordinates can vary greatly over the course of a war.

By way of an indirect test, we examined the impact of past *surrender by commanders* on surrender by rank-and-file troops in subsequent battles. Our reasoning here is that, when a commander has previously abandoned the battlefield, subordinates are likely to significantly discount the leadership's monitoring and punishment capacity. As a result, future commanders' threats to punish insubordination, surrender, and desertion lose credibility. If surrender is indeed less likely where monitoring and punishment capacity is high, then we should expect it to be more likely where commanders have themselves recently surrendered.

Table 2.5 reports the results of these additional analyses, with Models 1 to 4 estimating the impact of past surrender by commanders on surrender in the current battle by rank-and-file troops. These results confirm that soldiers are significantly more likely to surrender if commanders have recently done the same. A standard deviation increase in surrender by commanders within the same army yields an increase in the logged number of surrendering troops of between .13 (95% CI: 0.04, 0.2) and .25 SD (95% CI: 0.2, 0.3). Unsurprisingly, the actions of commanders in the soldiers' own army have a more substantial impact than commanders surrendering from other armies fighting the same opponent.

	Dependent variable:							
	log(POWs) *GLM link: identity*				Commander surrenders *GLM link: logistic*			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
SURRENDER BY COMMANDERS IN PREVIOUS BATTLES $W(x)$								
W(SAME COMBATANT)	0.24*** (0.18, 0.30)		0.16*** (0.08, 0.24)		2.26*** (1.35, 3.17)		1.20** (0.39, 2.01)	
W(SAME OPPONENT)		0.16*** (0.10, 0.23)		0.05 (-0.02, 0.13)		2.58*** (1.58, 3.58)		1.73*** (0.70, 2.76)
Battle-level covariates			✓	✓			✓	✓
Dyad-level covariates			✓	✓			✓	✓
Seasonal fixed effects	✓	✓	✓	✓	✓	✓	✓	✓
War fixed effects	✓	✓	✓	✓	✓	✓	✓	✓
Combatant fixed effects	✓	✓	✓	✓	✓	✓	✓	✓
Observations	844	844	426	426	1,718	1,718	813	813
Adjusted R^2	0.31	0.27	0.48	0.46	0.15	0.13	0.33	0.35
Log Likelihood	−1,111	−1,135	−524	−532	−372	−371	−183	−182
UBRE	0.85	0.90	0.71	0.74	−0.57	−0.57	−0.55	−0.55
RMSE	0.75	0.77	0.69	0.70	0.14	0.14	0.17	0.16
AIC	2223	2271	1049	1063	744	743	366	364

Notes: Standardized coefficients reported, 95% CI in parentheses.
$^{*}p<0.05$; $^{**}p<0.01$; $^{***}p<0.001$

Table 2.5: The impact of past surrender by commanders

If surrender by commanders helps drive surrender by their troops, a natural question arises: why do commanders surrender? Our data suggest that the collective-action problems facing soldiers may be part of a broader problem in military organizations that reaches across ranks. As Models 5 to 8 in Table 2.5 show, commanders are more likely to surrender if other commanders have recently surrendered. In making this choice, furthermore, commanders take cues not only from their own colleagues but also from other armies fighting the same opponent.

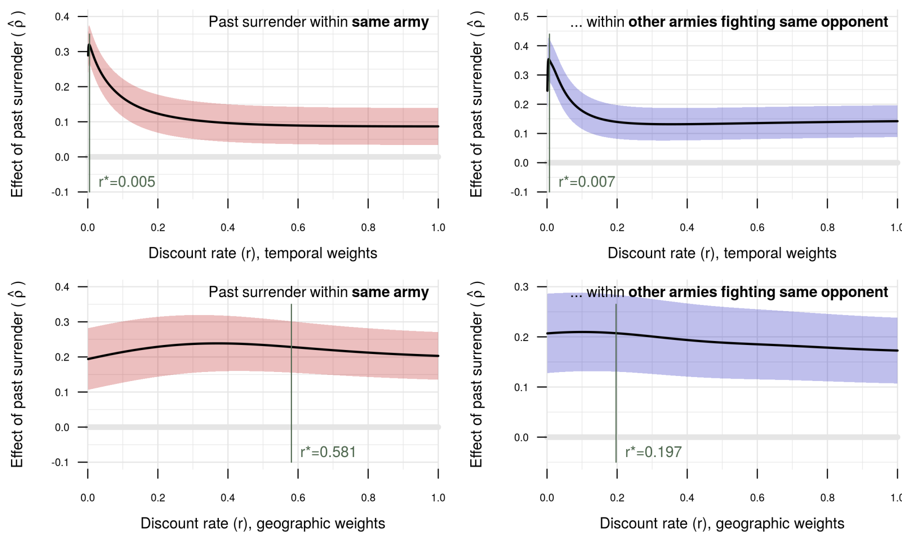
Macro-level state characteristics are poor predictors of surrender. While these results provide tentative evidence that surrender is contagious across battles, past surrender rates are, of course, not the only potential drivers of soldiers' decisions. Unsurprisingly, surrender rates are higher in larger battles, where more troops are potentially at risk.[32] Surrender rates are also lower for attacking troops—potentially as a result of offensive advantages in numbers, speed, and surprise.

Consistent with other recent research, we find most other macro-level variables to be poor predictors of surrender (Grauer, 2014). Aggregate national power (MORE POWERFUL), regime type (MORE DEMOCRATIC), and conscription (PROFESSIONAL ARMY) explain virtually none of the battle-level variation in surrender. The direction of these estimated effects is consistent with what we might expect from past literature. Fewer troops surrender from more materially capable armies. Armies staffed by long-service professionals are less likely to see higher rates of surrender than conscript armies. Troops are less likely to surrender if the opponent is less democratic than their home state. Once we account for battle-level factors like past surrender and battle size, however, these effects disappear.

2.3.2 Sensitivity analysis

The evidence so far has been supportive of the collective-action model: troops are more likely to surrender if, based on recent battlefield experience, they expect others to do the same.

[32] We are grateful to an anonymous reviewer for raising this point.



Note: Values shown are estimates of $\hat{\rho}$ (effect of surrender in past battles) at different levels of r (discount rate). Top row replicates Models 1 and 2 from Table 2.2, bottom row replicates Models 1 and 2 from Table 2.3. r^* are values of r used in original models.

Figure 2.3: How discount rate affects the contagion of surrender

How sensitive are our results to soldiers' discount rates (r in Eq. 2.2 to 2.5)? In the preceding analyses, we used values of r that optimized AIC. As we show in Figure 2.3, these values were relatively low for temporal discount rates ($r^* = .005, .007$ in Models 1 and 2 of Table 2.2) and intermediate for geographic discount rates ($r^* = .581, .197$ for Models 1and 2 of Table 2.3). These choices have implications for the scope of our findings: Table 2.2 assumes that soldiers weighed recent and past battles about equally, while Table 2.3 assumes they focused on battles that occurred nearby.

To ensure that our findings hold under a broader set of time and geographic horizons, we replicated Models 1 and 2 in Tables 2.2 and 2.3 with values of r between 0 and 1. Figure 2.3 shows how estimates for $\hat{\rho}$ gradually decrease and level off as r increases in the temporal weights models, while remaining steady in the geographic ones. Overall, however, the value of r does not fundamentally change our results. The impact of past surrender remains positive and significant in all four sets of models, regardless of how heavily one discounts long-ago events, or far-away battles.

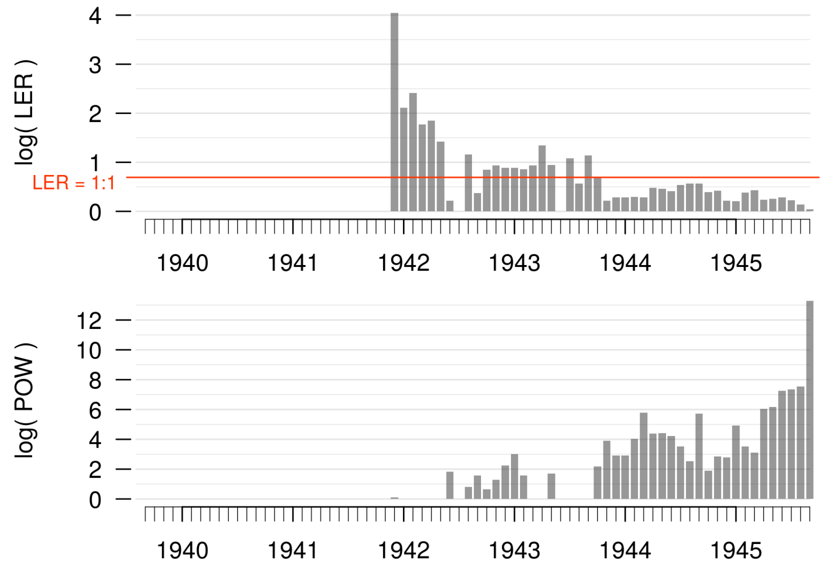
Another potential objection to our analysis is that dynamics of surrender are different for ground battles than air and sea battles, but our data set pools these events together. Because modern sailors and airmen typically surrender after the destruction of their ship or aircraft, past surrender is less salient to their decisions.

To address this concern, we reran the models in Tables 2.2 and 2.3 with a restricted data sample that includes only land warfare. The results are consistent with those we reported earlier. In the geographically weighted network, the contagion effect even increases, to .24 SD (95% CI: .17, 32) for the same combatant, and to .23 SD (95% CI: .16, .30) for other combatants fighting the same opponent. This increase makes intuitive sense: since it is more difficult for airmen and sailors to surrender mid-battle, keeping these battles in the sample should attenuate the estimated effect of past surrender.

2.4 Previous surrender or military effectiveness?

Could more general expectations of military success be driving the contagion of surrender? So far, we have seen little evidence that troops surrender at lower rates to militarily weaker opponents. As Models 3 and 4 show, combatants with higher CINC scores than their opponents (MORE POWERFUL) have few discernible advantages in this area. Yet because aggregate national capabilities do not vary across individual battles in a given year, they are a poor proxy for military effectiveness. Local numerical preponderance as measured by LOG(FORCE RATIO), meanwhile, has no apparent effect (Models 5 and 6).

To more directly account for perceptions of battlefield success and failure, we reran our models with several "placebo" diffusion terms, capturing information about total dead and wounded in previous battles, and previous loss exchange ratios (i.e., enemy dead and wounded divided by friendly dead and wounded). Higher loss exchange ratios (LER) indicate superior military effectiveness in the narrow sense of being able to inflict



Note: Values shown are monthly averages of the logged loss-exchange ratio (top pane) and average logged surrender rates (bottom pane).

Figure 2.4: Japan's military effectiveness and surrender rates against U.S.

heavy losses on the opponent with minimal casualties of one's own. If coefficient estimates on these placebo terms are positive, then the tendency to surrender may simply reflect expectations of higher losses, rather than any precedent set by previous surrendering troops.

To illustrate this possibility, Figure 2.4 shows Japan's average monthly LER in World War II (logged), along with Japan's monthly surrender rates (logged). The plots suggest an inverse relationship. Early in the war, Japan's military effectiveness was high and surrender rates were low. Beginning in late 1943, LER dropped below parity (red line), and surrender rates grew. From this picture, one may conclude that Japanese troops became more likely to surrender not as a result of cases of past surrender, but because of an increasingly untenable military situation.

Table 2.6 reports the results of our placebo tests. In each specification, the confidence interval on the diffusion coefficient covers 0. The high uncertainty around these placebo effects provides further evidence in favor of

	Dependent variable:			
	log(POWs)			
	(1)	(2)	(3)	(4)
CASUALTIES IN PREVIOUS BATTLES $W(KIA + WIA)$				
W(SAME COMBATANT)	0.05 (−0.02, 0.12)			
W(SAME OPPONENT)		0.04 (−0.05, 0.13)		
LOSS-EXCHANGE RATIOS IN PREVIOUS BATTLES $W(LER)$				
W(SAME COMBATANT)			0.10 (−0.02, 0.22)	
W(SAME OPPONENT)				0.05 (−0.25, 0.35)
Battle-level covariates	✓	✓	✓	✓
Dyad-level covariates	✓	✓	✓	✓
Seasonal fixed effects	✓	✓	✓	✓
War fixed effects	✓	✓	✓	✓
Combatant fixed effects	✓	✓	✓	✓
Observations	426	426	426	426
Adjusted R^2	0.46	0.46	0.46	0.46
Log Likelihood	−531.50	−532.29	−531.11	−532.62
UBRE	0.74	0.74	0.73	0.74
RMSE	0.70	0.70	0.70	0.70
AIC	1063.01	1064.59	1062.21	1065.25
Notes:	Standardized coefficients reported, 95% CI in parentheses. *p<0.05; **p<0.01; ***p<0.001			

Table 2.6: Placebo tests: determinants of surrender

our preferred interpretation of the "past surrender" result. Surrender is neither more nor less likely in battles that, based on past experience, soldiers should expect to lose. Nor is the expectation of death by itself predictive of surrender. If political authorities wish to maintain the resolve of their armies in battle, these results indicate that they should worry less about how dangerous a combat environment is likely to be, and more about recent precedents for mass surrender.

2.5 Conclusion

Our results offer several contributions to research on interstate conflict. We demonstrate that battlefield surrender can be contagious because of a collective-action problem within military organizations. Success in battle requires that soldiers fight as a cohesive unit, but individual decisions to fight depend on whether soldiers expect their comrades to do the same. As troops learn of past decisions to surrender within their own army, they lose confidence in their unit's resolve, and decide to flee rather than fight. This pattern is particularly acute if the expected costs of surrender are also low—either because troops believe the opponent will treat prisoners well, or because senior officers have recently surrendered, shaking the credibility of threats to punish desertion and surrender by the rank and file.

In addition to diffusion, we examined several alternative explanations of surrender. We found tentative, if mixed, support for a few factors that might affect the parameters of the collective action model—like international law, principal-agent problems, and offensive advantages. However, we found no evidence that surrender depends on political regime type, recruitment methods, or relative national power. Although data limitations prevent us from directly testing several other explanations—mutual surveillance, discipline, and ideology—we sought to at least account for them econometrically, through combatant and war fixed effects. We also demonstrated that it is information specifically on past surrender, rather than military effectiveness generally, that drives soldiers' decisions.

The determinants of surrender are theoretically important because these life-and-death choices tend to resonate well beyond individual battles. Although previous research suggests that combatants acquire information about war-fighting resolve through battle outcomes, scholars often treat resolve as an exogenous cause of war termination. In our approach, by contrast, battlefield resolve is of primary theoretical interest. If wars are a continuation of political bargaining, reconciling informational asymmetries through the use of force, then understanding the mechanisms and processes influencing battlefield resolve is crucial for explaining and predicting bargaining outcomes. Our results illustrate that wartime resolve does not depend solely on political leaders' assessment of probabilistic battlefield outcomes. Instead, military officers and their troops are the primary actors mutually influencing each other's behavior. Because soldiers' choices in future battles depend on precedents set by others in the past, it is these cascading battlefield decisions that ultimately shape and constrain leaders' choices.

Our study opens several future avenues of research. For example, although we have demonstrated that surrender can have a contagion effect across battles, we do not analyze how this process begins within battles, and what critical events must occur to jump-start surrender and its subsequent diffusion. While our focus has been on inter-battle dynamics, a more explicit focus on intra-battle behavior is needed to understand the conditions leading to initial organizational breakdown.

Further research is also needed to understand how different political-military institutions affect whether the diffusion process occurs, or whether it can be reversed. We know little about why some military organizations can absorb losses and adapt to changing circumstances, while others are unable to recover from battles in which soldiers surrendered en masse. By disaggregating wars into battles and stepping away from the classical approach of treating the military as a unitary actor, we can better understand how collective action dynamics affect battlefield outcomes and, ultimately, decisions to initiate, continue or terminate war.

2.6 Appendix

This section formalizes the collective-action model of surrender, which we described qualitatively earlier. We develop a basic theoretical framework for a global game using Angeletos, Hellwig, and Pavan (2007)'s game structure and apply it to the domain of battlefield surrender. In the model, survival-oriented soldiers choose either to fight, thus increasing their unit's chances of success, or abandon. Soldiers' decisions depend on what they expect others to do, based on private information and observation of previous battles. We begin by specifying a baseline static model, and then discuss the dynamic version separately.[33]

The game unfolds as a series of battles in discrete time, indexed by $t \in \{1, 2, ..., T\}$.[34] At t, each soldier $i \in \{1, 2, ..., N\}$ in an army simultaneously chooses to *Fight* ($a_{it} = 0$) or *Abandon* ($a_{it} = 1$). We denote the proportion of soldiers abandoning at time t as $A_t \in [0, 1]$.[35] The payoffs associated with fighting and abandoning depend on the resultant battle's state: *success* or *failure*. The state depends on whether the proportion of soldiers abandoning exceeds the army's level of organizational resolve, $\theta \in \mathbb{R}$. We can interpret θ as the maximum level of abandonment an army can withstand while still being able to fight as a cohesive force.

If the abandonment rate is low ($A_t < \theta$), the battle will end in a success state, and each soldier who fights will receive payoff B. If, instead, aban-

[33]The static, one-shot version of the game is partly analogous to an inverted version of the classical threshold model of collection action from Granovetter (1978). We can think of each soldier's individual threshold as determining the level at which they will surrender, such that a cascade of surrender can trigger within a single battle. This is useful for thinking about how many soldiers may surrender based upon local signals. However, there are several limitations of the classical threshold model for our purposes: first, it is not a model of a coordination game with incomplete information; second, it does not address noisy signals; and finally, it does not describe the diffusion of information across a sequence of locally cascading events. The dynamic global game framework incorporates all these elements.

[34]The game ends at an undetermined time T, which we assume is determined by political leaders. We do not directly analyze the decision to terminate war here.

[35]We assume that N is relatively large, such that an individual soldier's contribution is negligible as a proportion of the entire effort.

donment rates are high ($A_t > \theta$), a failure state will occur, and each soldier who fights will receive a lower payoff η. A soldier who abandons will receive payoff z in the success state and v in the failure state. Payoffs z and v depend on both the level of punishment abandoning troops receive from their own army—which is particularly salient in the success state payoff, z—and the opponent's treatment of prisoners of war.

In a success state, soldiers prefer fighting to abandoning ($z < B$). In a failure state, they prefer abandoning to fighting ($\eta < v$). Soldiers also prefer successful fighting over abandoning in failure ($v < B$) and, by transitivity, prefer fighting in a success state to fighting in a failure state ($\eta < B$). The value of z relative to v can vary, based on expected punishment with one's own army and expected treatment by the opponent.

The relative cost of fighting for each soldier is $c = \frac{v-\eta}{B-z+v-\eta} \in (0,1)$.[36] This cost is increasing in the payoffs to surrendering, v and z. In line with previous research, we should expect v (and c) to be higher when opponents have ratified treaties on the humane treatment of prisoners. Armies who increase their opponents' v therefore increase the relative cost of fighting against them, which makes abandoning more attractive. Similarly, we should expect z (and c) to be lower when an army can effectively punish its own surrendering troops. Consequently, armies who decrease z reduce the cost of fighting (since soldiers then avoid punishment), making fighting more attractive. Table 2.7 summarizes the payoff structure, with soldiers' choices in the rows and the battle's state in the columns.[37]

2.6.1 Static equilibrium analysis

Following Angeletos, Hellwig, and Pavan (2007), when θ is perfectly known by all soldiers, there are two pure strategy equilibria for $\theta \in (0,1]$: all sol-

[36] Because a soldier finds it optimal to fight if and only if s/he expects success, we can interpret c as the probability of a failure state, in which the proportion of soldiers abandoning is above the threshold value of organizational resolve ($A_t \geq \theta$).

[37] To simplify notation, we express the payoffs by their differences and normalize them between zero and one. We thank Scott Tyson for suggesting this simplification.

	$A < \theta$ Success	$A \geq \theta$ Failure
Fight ($a_{it} = 0$)	$1 - c$	$-c$
Abandon ($a_{it} = 1$)	0	0

Table 2.7: Payoff structure

diers fight ($A_t = 0 < \theta$) or all soldiers abandon ($A_t = 1 \geq \theta$).[38] When θ is imperfectly known and there exists heterogeneous information about organizational resolve, the decision to fight or abandon depends on signals that each soldier receives. In this case, Nature draws an initial common prior, $\theta \sim N(\omega, \frac{1}{\alpha})$, where α indicates the common prior's precision. Each soldier receives a private signal:

$$x_i = \theta + \epsilon_i$$

where $\epsilon_i \sim N(0, \frac{1}{\beta})$ indicates noise, i.i.d. across soldiers and independent of θ, and β describes the signal's precision.

Let $\hat{x} \in \mathbb{R}$ be a threshold, such that a soldier abandons when $x_i \leq \hat{x}$. Given this threshold, the proportion of soldiers who abandon is decreasing in θ:

$$A(\theta) = Pr(x \leq \hat{x}) = \Phi(\sqrt{\beta}(\hat{x} - \theta))$$

where Φ is the CDF of the standard Normal distribution. This observation dovetails with previous research: the proportion of soldiers abandoning is decreasing in the level of organizational resolve, which is related to factors such as attacker advantages.

Organizational failure occurs when $\theta \leq \hat{\theta}$, where $\hat{\theta}$ solves $\theta = A(\hat{\theta})$: $\hat{\theta} = \Phi(\sqrt{\beta}(\hat{x} - \hat{\theta}))$. The posterior of θ given x is distributed

[38] If $\theta \leq 0$ or $\theta > 1$, there is one pure strategy equilibrium: all abandon or all fight, respectively.

$N\left(\frac{\beta}{\beta+\alpha}x+\frac{\alpha}{\beta+\alpha}\omega, \frac{1}{\beta+\alpha}\right)$. The resulting probability of failure is

$$Pr(\theta \leq \hat{\theta}|x) = 1 - \Phi(\sqrt{\beta+\alpha}(\frac{\beta}{\beta+\alpha}x + \frac{\alpha}{\beta+\alpha}\omega - \hat{\theta}))$$

This probability is decreasing in x. Consequently, a soldier will find it optimal to abandon when $x \leq \hat{x}$, where $\hat{x}$ solves $Pr(\theta \leq \hat{\theta}|\hat{x}) = c$

$$1 - \Phi(\sqrt{\beta+\alpha}(\frac{\beta}{\beta+\alpha}\hat{x} + \frac{\alpha}{\beta+\alpha}\omega - \hat{\theta})) = c$$

A monotone equilibrium $(\hat{x}, \hat{\theta})$ exists for all ω iff $\beta \geq \frac{\alpha^2}{2\pi}$.

Given a particular α, when β is smaller (i.e., private information is less precise), there are dominance regions where subsets of soldiers prefer one action over the other, depending on their individual value for x. However, as $\beta \to \infty$, the threshold $\hat{\theta}$ converges to $\theta_\infty \equiv 1 - c$. When this occurs, the proportion of soldiers abandoning converges to 1 for all $\theta < \theta_\infty$ and to 0 for all $\theta > \theta_\infty$.

While this analysis describes only the static, one-shot version of the game, the same basic mechanisms operate in the dynamic model.[39]

2.6.2 Dynamic model

In each period $t \geq 1$, each soldier receives a private and public signal about organizational resolve, θ. Furthermore, in $t \geq 2$, each soldier also receives a public and private signal about soldiers abandoning in previous battles, A_{t-1}. In the dynamic game, private information evolves over time as soldiers receive more information and update their beliefs. We specify these

[39] For a full proof of the dynamic game, see Angeletos, Hellwig, and Pavan (2007, p. 731–733).

signals as follows:

$$\text{Private, }\theta\text{: } x_{it} = \theta + \epsilon_{it}$$
$$\text{Public, }\theta\text{: } z_t = \theta + \xi_t$$
$$\text{Private, }A_{t-1}\text{: } X_{it} = S(A_{t-1}, v_{it})$$
$$\text{Public, }A_{t-1}\text{: } Z_t = S(A_{t-1}, \zeta_t)$$

where ϵ_{it} and v_{it} are idiosyncratic noise terms, ξ_t and ζ_t are common noise terms, and $S : [0,1] \times \mathbb{R} \to \mathbb{R}$. Each of the noise terms is distributed normally with mean zero and variance specified as follows, independent of θ, serially uncorrelated, and i.i.d. across all i for private noise terms (Ibid.):

$$\epsilon_{it} \sim N(0, 1/\eta_t^x)$$
$$\xi_t \sim N(0, 1/\eta_t^z)$$
$$v_{it} \sim N(0, 1/\gamma_t^x)$$
$$\zeta_t \sim N(0, 1/\gamma_t^z)$$

The past period's signals condition posterior beliefs similarly to the static game.

Two cases illustrate the novelty of modeling the information structure in this way. First, consider a case where soldiers observe public and private signals only about the value for θ, but not the precise size of past levels of abandonment, A_t. If a soldier observes that abandonment has occurred (without knowing the size), and sees that it did not lead to organizational failure, she will update her beliefs *upward* about the value for θ. In other words, by recognizing that the organization was able to sustain some unknown level of abandonment without total failure, the soldier comes to see that the army may be *more* resolved than previously believed, making future abandonment less likely. Consequently, when expected resolve is high, an army becomes more resilient against individual bouts of surrender.

In a second case, where soldiers also observe past levels of abandonment, A_{t-1}, the dynamic changes. Here, separate signals about the proportion of soldiers abandoning in previous battles may counter the effect of knowing that the organization did not fail, based on the public signal about θ. When this happens, the likelihood that soldiers will abandon in the next battle can rise, potentially leading to a cascade effect across battles. These dynamics establish the microfoundations of the diffusion process posited by our main hypothesis: the flow of information from previous outcomes affects soldiers' decisions in battle, and future surrender increases with information about past surrender.

CHAPTER 3

Adapting Counterinsurgency Doctrine in the Shadow of the Future

> The Afghan War has lacked high-level American commitment for years now. If there is any surprise, it is that for eight years of Barack Obama and four years of Donald Trump, the United States persisted in a conflict that most senior officials in those administrations regarded with pessimism and distaste.
>
> —Eliot A. Cohen, *The Atlantic* (Cohen, 2021)

Since 1945, conventional militaries have a poor record of waging counterinsurgencies. Only about one third of post-World War II insurgency conflicts have been won by military forces, while the remaining conflicts have ended in either outright victory for the insurgents (33 percent) or negotiated concessions to their demands (34 percent). For conflicts in which the state is an external occupier, military victory drops to 24 percent, and insurgent victory climbs to 62 percent (Lyall and Wilson III, 2009). Despite having significant advantages in resources, technology, and training, militaries struggle to win modern asymmetric conflicts. What makes effective counterinsurgency doctrine such a vexing puzzle for militaries, and why are some militaries better able to adapt their doctrine than others?[1]

Several possible explanations have been offered for why states are successful in some counterinsurgencies but not others. Macro-level explanations emphasizing regime type focus on the role that domestic pressures place on a state's ability to fight insurgent conflicts (Byman, 2016; Getmansky, 2013; Lyall, 2010; Mack, 1975; Merom, 2003). However, several of these studies find no direct relationship between regime type and insurgency outcome, making regime-focused explanations relatively weak. Another common explanation is that military culture leads to better or worse counterinsurgency by creating organizations that are more or less open to learning (Long, 2016; Nagl, 2002). Yet existing cultural explanations, which tend to rely on organizational stickiness over time, are unable to explain variations in counterinsurgency effectiveness within the same military, such as the British army's variable effectiveness between Malaya and Iraq. Other explanations consider characteristics internal to the conflict to better understand their outcomes, using them to evaluate the effectiveness of tactics (Johnston, 2012; Pampinella, 2015) and military strategies (Arreguín-Toft, 2001; Enterline, Stull, and Magagnoli, 2013; Lyall and Wilson III, 2009; Paul, Clarke, and Grill, 2010). However, these existing studies overlook the paradox in supposing a single explanation according to a particular strategy or tactic: opponents can imitate or adjust their strategies or tactics in response, distorting their effectiveness across other cases. Relocating parts of the population, for example, was a successful strategy for the British in Malaya, but did not succeed for either the French in Algeria or the U.S. in Vietnam. Explanations that emphasize a particular approach to counterinsurgency fail to convincingly account for such variation.

In any conflict, military organizations are Darwinian actors facing com-

[1] I define doctrine as the comprehensive set of political and military decisions about how to fight a war. A military's doctrine is the result of learned experiences and observations, all formed under particular organizational and policy constraints.

petition for the survival of the fittest. Adaptation, therefore, is a matter of life or death and can determine the difference between victory and defeat. As the British military theorist Liddell Hart declares, "Adaptability is the law which governments survival in war as in life—war being but a concentrated form of the human struggle against the environment" (Hart, 1991, p. 330). Therefore, to explain counterinsurgency effectiveness, I model a military's ability to adapt its doctrine to local conditions.

When fighting an insurgency, a military faces high uncertainty about how best to defeat its opponent because complex local conditions create enough variation in which strategies and tactics are most effective that they cannot be identified *ex ante*—they can only be learned through experience. Other scholars have made similar arguments about counterinsurgencies. Kilcullen (2010), for example, argues that the counterinsurgent's imperative is "to understand each environment, in real time, in detail, in its own terms, in what that would be understood by the locals—and not by analogy with some other conflict, some earlier war, or some universal template or standardized rule-set" (p. 2). Thus, where adaptability is better, a military can learn which strategies and tactics are more successful than others, and implement them across the organization in a coordinated and cohesive way. Conversely, where adaptability is worse, a military will apply an ineffective doctrine, or will identify and implement improvements too slowly, and will be less likely to defeat an insurgency as a result. Therefore, by identifying organizational features that promote or hinder adaptability, we can better explain and predict insurgency conflict outcomes.

Contrary to fixed military culture explanations, I use the model to show that doctrinal adaptability arises from a state's commitment horizon in a conflict. The commitment horizon is the extent to which the future matters for the present—i.e., how much the "shadow of the future" (Axelrod, 1984) weighs on a military's decision-making—and is related to the concept of resolve, i.e., the "firmness or steadfastness of purpose" (Kertzer, 2017).[2] In the model, the commitment horizon can formally be understood as the dis-

[2] I operationalize commitment horizon qualitatively in Section 3.3 below.

count factor. I argue that the basis for a state's commitment horizon varies by case and is thus not a systematic feature of an organization. When a state has a higher commitment horizon, the military will persevere despite setbacks, search for better solutions, and make more effective decisions that account for complexities in local conditions. When a state has a lower commitment horizon, on the other hand, the military may search for and adopt quick improvements, but these may be less effective overall.

However, this is not the whole story. The commitment horizon also directly affects the optimal choice of organizational characteristics—namely, the level of military autonomy, and the tolerance for political and military experimentation—and it is these characteristics, which are manifestations of military culture, that drive the tendency to adapt and thereby improve effectiveness. Moreover, the model's results indicate that particular organizational characteristics adopted are nonmonotonic in the commitment horizon level. Consequently, I show that doctrinal effectiveness follows a U-shaped curve, corresponding to the organizational characteristics adopted as the commitment horizon increases.

Using the model's results, I develop a typology of counterinsurgent militaries according to similar commitment horizon levels and corresponding organizational characteristics, which can be used to explain a military's counterinsurgency effectiveness. I then illustrate the logic of the model and the implications of this typology in three sets of historical counterinsurgency cases: the U.S. in Iraq and Afghanistan; the British in Malaya and Iraq; and the French in Indochina and Algeria. I show that the variation in counterinsurgency effectiveness across these cases can best be explained by the variation in commitment horizons and corresponding organizational features, rather than alternative explanations such as fixed military culture or specific counterinsurgency strategies.

This paper contributes to our understanding of the determinants of military doctrine and battlefield effectiveness in counterinsurgency operations by providing a novel approach to explaining military adaptation during wartime. Furthermore, I build on previous work by connecting the findings in the literature on military adaptation into a more unified understanding of how doctrinal evolution varies across military organizations. Finally, the explanation I provide for counterinsurgency effectiveness has important policy implications for both conflict dynamics and outcomes.

3.1 Military adaptation and complex conflicts

Existing studies on military adaptation and innovation can be grouped into those that focus on top-down processes of adaptation, and those that focus on bottom-up processes. Top-down explanations focus on identifying the processes through which political or military leaders identify needs for change and implement organizational change. These explanations have focused on issues such as the role of intervention by civilian policymakers (Posen, 1984), intra-service competition for promotion (Rosen, 1991), institutional structures of delegation (Avant, 1993), changes in the external security environment induced by an adversary (Zisk, 1993), military leader replacement (Sorley, 1999), and bureaucratic flexibility (Horowitz, 2010a).

Conversely, other scholars have examined the importance of bottom-up innovation and adaptation for explaining battlefield effectiveness, which can involve experimentation with new tactics (Farrell, 2010; Gudmundsson, 1995; Kollars, 2015) or the horizontal sharing of information (Foley, 2012). Bottom-up adaptation may fail to occur, however, when ideas are lost or forgotten before they are more fully implemented (Catignani, 2012), or if there is a culture that discourages experimentation and improvisation (Milner, 1984), such as when subordinates are discouraged from questioning policies or ideas are rejected by higher commanders (Nagl, 2002).

More recently, research on the process of military change has focused on synthesizing the top-down and bottom-up processes into more structured frameworks (Horowitz and Pindyck, 2020). Some studies, for example, describe bottom-up learning as adaptation, whereas top-down learning is described as organization-wide innovation, e.g., by institutionalizing practices into doctrine (Murray, 2011; Ucko, 2009). Russell (2011) and Adamsky and

Bjergam (2012) make similar arguments, describing the development of military thought as a dialectical process between soldiers adapting in wartime and military thinkers innovating doctrine primarily during peacetime. Similarly, the evolution of doctrine—when it occurs—has been described as an iterative process, up and down the chain of command, where changes are disseminated quickly throughout the organization (Barno and Bensahel, 2020).

More broadly, the role of military culture has been used to explain military preferences and doctrinal development over time. Military culture, in short, is "the ideas about how a military should fight" (Morrow, 2014, p.270), and therefore provides a basis for explaining the origins of a military organization's preferences (Legro, 1995). Military culture has been invoked to explain issues such as the preference for offensive military operations leading up to the outbreak of World War I (Snyder, 1984; Van Evera, 1984), differences in preference between offensive and defensive doctrines (Kier, 1997), the extent to which military organizations comply with international laws of war (Legro, 1995), variation in the levels of individual initiative among commanders during wartime (Muth, 2011), and differences in approaches to counterinsurgency operations (Long, 2016). These existing explanations point to important socialization processes that cause beliefs to become ingrained within the military organization over time, but they are unable to explain doctrinal change without some exogenous shock to the organization. Consequently, the existing studies on military culture tend to push the question of doctrinal evolution one step beyond the organization, rather than trying to explain how cultures and corresponding doctrines evolve over time through processes within the organization itself.

Some of the literature on civil-military relations has addressed endogenous organizational issues, especially with regard to the power distribution and divisions of control between political leaders and the military, though scholars studying these topics are typically less concerned with explaining doctrinal evolution. Since Huntington (1957), much of the focus on the civil-military power distribution has been on debating the proper (i.e., normative) balance between objective and subjective civilian control. Yet regardless of the normative implications or prescriptions, understanding the causes and consequences of civilian oversight is ultimately one that can be studied objectively and systematically, e.g., by framing the issue as a principal-agent problem (Avant, 1994; Feaver, 2003). Moreover, evaluating different levels of military autonomy can be useful for explaining not just the managerial dilemmas that political leaders face, but also how different levels of decision-making control affect the development of doctrine and the military's ability to learn, adapt, and change the doctrine over time.

Underlying the studies of military adaptation, innovation, and organizational processes is a recognition that modern conflict is highly complex. Consequently, these studies generally imply that modern battlefield effectiveness requires a complex pattern of force employment, which Biddle (2004) terms the "modern system." In the range of possible types of conflict, many scholars argue that counterinsurgency is particularly complex (see, for example, Galula (1964); Jones (2017); Kilcullen (2010); Trinquier (1961)). In his classic treatise on insurgency, Mao describes the importance of adaptation as such: "Guerrilla commanders adjust their operations to the enemy situation, to the terrain, and to prevailing local conditions. Leaders must be alert to sense changes in these factors and make necessary modifications in troop dispositions to accord with them" (Mao, 1961, p. 101). Similarly, the U.S. Army's current field manual on counterinsurgency operations, FM 3-24, refers to a "complex" environment or the "complexity" of counterinsurgency operations 27 times throughout the document (United States Army, 2014). Counterinsurgency conflicts therefore provide a useful substantive context to evaluate military adaptability.

Most existing work on military adaptation and learning has attempted to describe adaptive processes qualitatively. To my knowledge, no existing study has attempted to formally model the structure of organizational interactions to explain doctrinal evolution. Therefore, in what follows, I present a stylized model of doctrinal adaptation in order to provide a more systematic analysis of how variations in organizational features affect evolutionary doctrinal outcomes in highly complex environments.

3.2 The model

Military organizations are large-scale, distributed systems comprised of heterogeneous units with competing preferences that interact within a hierarchical structure. When faced with a complex decision space, the interactions between political and military levels, and between military units in the hierarchy, can lead to adaptive and coevolutionary outcomes for the entire organization iteratively over time, which lead to an evolution in military doctrine. Agent-based models (ABMs) are a useful method for investigating the ways in which such micro-level interaction rules affect emergent macro-level outcomes (Schelling, 2006), especially when the interactions are computationally complex, making a reduced-form mathematical solution extremely difficult (Axelrod, 1997). Consequently, ABMs are a useful approach for modeling the dynamics of organizational decision-making and learning that leads to the emergence of a group-wide doctrine. I model the dynamics of doctrinal evolution with an ABM and gather results through simulation to evaluate the effects that military autonomy, tolerance for experimentation, and commitment horizon have on doctrinal effectiveness over time.

The model of organizational decision-making that I develop draws inspiration from the literature on organizational learning (Levitt and March, 1988). A common approach to modeling this type of learning uses "hill-climbing" as an optimization technique to model the search for improvements, where a search is conducted over a fitness landscape, representing possible solutions to a problem, in order to find the highest peak on the landscape, which represents the optimal solution. For example, March (1991) uses a hill-climbing algorithm to evaluate the tradeoffs between exploration and exploitation in organizational learning.[3] One particular hill-

[3] March's concept of an "organizational code" has comparable qualities to that of a doctrine.

climbing model that has gained prominence in the study of complex adaptive systems is the *NK*-model, which describes an adaptive walk process on a rugged landscape (Kauffman and Levin, 1987). Originally designed to study biological processes of adaptation and selection in epistatic interactions, variants of *NK*-models have flourished in numerous organizational and management strategy studies (see Baumann, Schmidt, and Stieglitz (2019) for an overview). Properties that make *NK*-models fruitful for studying adaptive decision-making processes include the ability to incorporate bounded rationality in the search process, and the ability to vary landscape ruggedness to evaluate different levels of complexity in local conditions.

I model a set of political and military agents using the *NK*-model framework, where the agents are organized into a hierarchical network, such that only agents that are directly connected interact. The agents begin with an existing shared organizational doctrine, represented as a binary string of decision variables. Furthermore, the agents share an interest in optimizing their doctrine to a new environment of unknown complexity. The level of complexity in the environment affects the interdependence between agents' doctrinal decisions, meaning that a change in one decision can influence another decision's value, where increasing interdependence implies an increasingly complex environment.

The agents attempt to optimize their doctrine cooperatively by delegating mutually exclusive subsets of the doctrine across the organization and exploring a decision landscape for their assigned subsets. Subsets are divided between agents having superior authority over the hierarchy, representing political decisions, and those within the remainder of the hierarchy, representing military decisions. The agents then search for changes in decisions that improve the fitness of their subsets. Political decisions are implemented with certainty into the doctrine, whereas bottom-level military agents share their information up the chain of command. As information is filtered up the hierarchy, higher-ranking agents aggregate their subordinates' information with wider evaluation scopes so that the combined information leads to improved fitness. The resulting decision at the senior military level incorporates all the military decision subsets, and combined with the political decisions, becomes an updated doctrine for the entire organization. The process is then repeated iteratively until an equilibrium is reached, in which no agent prefers to make any changes to its assigned subset. The equilibrium decision string then represents the final evolved doctrine.

An ABM specifies properties of an environment, a set of agents, the interaction topology, and the action sequence. I explain each of the model's components in detail below. In the appendix, I provide an example of a single iteration of the model.

3.2.1 Environment

Doctrine D is characterized as a string of length N comprised of attributes $D_i \in \{0, 1\}$ for $i \in [1, N]$. Each D_i represents a decision and is randomly assigned an initial value. The initial string can therefore be interpreted as the status quo doctrine existing at the beginning of the conflict. For each D_i decision, there are $K \in [1, N-1]$ dependencies on other decisions within the doctrine. The K dependencies are also randomly assigned and remain fixed for a particular simulation. Each D_i decision can therefore take on 2^{K+1} possible fitness levels, which depend on the binary value for D_i and the binary values of its K other dependent decision values. For the total doctrine, the $2^{K+1}N \in [0, 1]$ fitness levels are randomly assigned and normalized by N.

The environment defines the ruggedness of the fitness landscape. For any given landscape, there is always one global optimum that represents the ideal set of decisions for the environment. Furthermore, for a fixed N, complexity increases with K because decisions become more interdependent, implying that the fitness value for a particular decision can become "frustrated" by the values of the decisions upon which it depends. Thus as K increases, the decision landscape becomes more rugged, increasing the number of local optima across the entire landscape and making the environment more complex.

3.2.2 Agents and interaction topology

The agents in the model represent individuals or close-knit communities of individuals that share the same behavior and can be represented as a single agent, such as combat units or a command staff. To focus the analysis, I treat agents as units, but refer to them as individuals or agents for ease of discussion. Agents are organized and interact within a fixed hierarchical network, which represents the chain-of-command structure that commonly exists in armed groups to help control the information flow and decision-making between upper and lower echelons (Millett, Murray, and Watman, 1988).[4]

The interaction topology defines the role that each agent has in the doctrinal decision-making process. Each role, furthermore, is a specialized focus on a limited subset of the organization's war-fighting mission. To develop this structure, I define the interaction topology of an ideal-type hierarchy as a balanced tree, with an additional node at the top of the tree to represent the commander-in-chief, referred to hereafter as the political leader.[5] The hierarchy in the model is designed to correspond to the different levels of a conflict: political, strategic, operational, and tactical. This hierarchy is illustrated in Figure 3.1. The aggregation of all the decisions made across these levels corresponds to the comprehensive doctrine.

In the model, I represent the military structure as a two-level tree with two subordinates per node. Above the tree, the top-level agent in the interaction structure represents the political leader. The political leader is thus directly connected only to the top agent in the tree, which represents the senior-level military commander and whose role is to determine the strategic decisions for the military. The senior commander has two subordinates that serve as mid-level commanders and are in charge of determining operational decisions. Each of these operational-level commanders also has

[4]Additionally, because military hierarchies are fractal—i.e., the structure between individuals within a single unit is the same as the structure between units across the organization—the model can apply across different levels of analysis.

[5]In a balanced tree, each node has a fixed number of subordinate nodes, and there are a fixed number of levels in the tree.

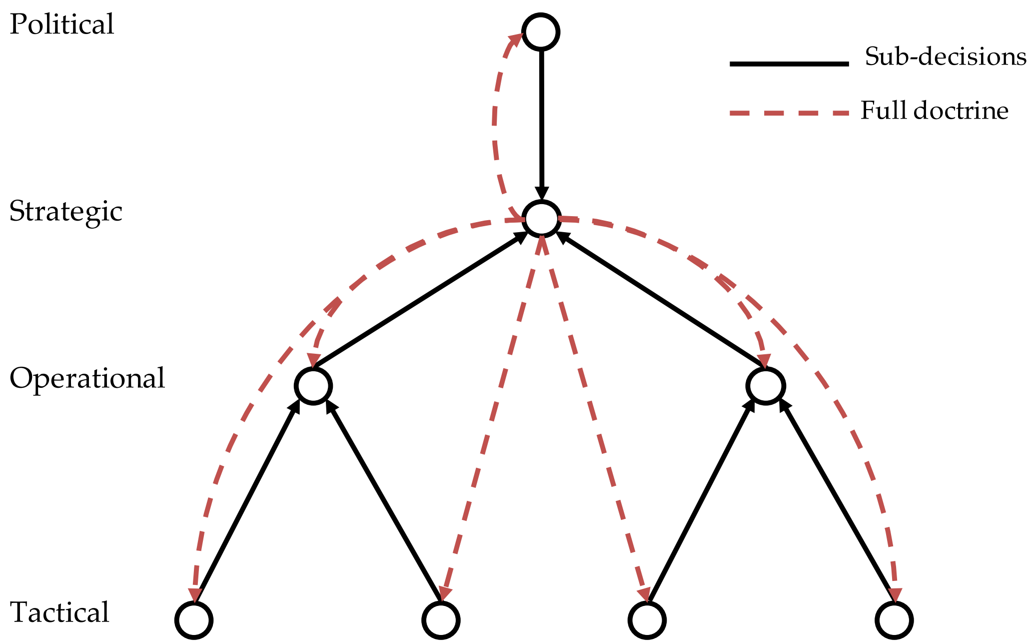


Figure 3.1: Hierarchy levels and information flows in the model

two subordinates, which represent its bottom-level units and are in charge of tactical decisions.

The flow of the model's decision between agents is illustrated in Figure 3.1. Decisions that comprise the doctrine are divided between political and military levels. The political leader in the structure has ultimate decision-making authority over her own decisions, the *political requirements*. The remaining decisions, the *military requirements*, are then divided equally and assigned as sets of sub-decisions down to the tactical-level units. Thus the political leader and the tactical-level agents directly explore the landscape by "experimenting" with how fitness levels change with different values for each of their assigned decision sets. However, only the political leaders' decisions are chosen with certainty for the doctrine. Each tactical-level agent, instead, makes a recommendation for his decision(s) to his direct superior, an operational-level commander. Each operational-level commander is responsible for evaluating aggregations of his subordinates' delegated sub-decisions, which represent the operational decisions, and selecting the higher fitness combination of operational decisions to recommend to the senior commander. Finally, the senior commander is responsible for evaluating aggregations of the operational recommendations to select the total military requirements with the highest overall fitness level, which represent the strategic decisions. The decisions at the political and strategic levels aggregated together comprise the full doctrine, which is then distributed to all agents.

3.2.3 Action sequence

Agents act according to the following sequence:

1. Political requirements are represented by P, the subset of political decisions. The political leader compares the political requirements' status quo fitness level $f(P_{SQ})$ to the fitness level resulting from a random change, or mutation, in a single political requirement value $f(P_M)$. If $f(P_M) > f(P_{SQ})$, then the political leader updates P to incorporate the requirements with P_M, otherwise she retains P_{SQ}.

2. Decisions delegated to the tactical level are represented by T. Each tactical-level agent compares his assigned decisions' status quo fitness level $f(T_{SQ})$ to the fitness level resulting from a random mutation in one of his delegated decision values, $f(T_M)$. If $f(T_M) > f(T_{SQ})$, then the tactical-level agent recommends T_M to his operational-level commander. Otherwise, he recommends T_{SQ}. The recommendation is represented by T_R.

3. Each operational-level commander $j \in \{1,2\}$ compares the fitness level for his subordinates' (T_1 and T_2) combined status quos with the fitness levels of all possible composites of T_1^j and T_2^j's recommendations. The operational-level commander then recommends the composite with the highest overall fitness for this decision subset, repre-

sented by O_R. The comparisons consist of the following composites:[6]

- $f(T^j_{1,SQ}, T^j_{2,SQ})$
- $f(T^j_{1,R}, T^j_{2,SQ})$
- $f(T^j_{1,SQ}, T^j_{2,R})$
- $f(T^j_{1,R}, T^j_{2,R})$

4. The senior commander compares the fitness level for the operational-level commanders' (O_1 and O_2) combined status quos with the fitness levels of all possible composites of O_1 and O_2's recommendations in the same fashion as in Step 3. The senior commander then selects the military requirements S with the highest overall fitness within the decision subset, which represent the strategic decisions.

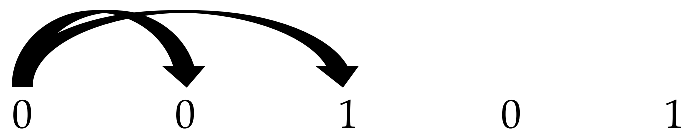
5. Aggregated political requirements and military requirements (P, S) comprise the updated doctrine D for the entire organization, which are distributed to each of the agents. Repeat from step 1 until the last iteration is completed.

Figure 3.2 illustrates an example of an agent's decision process. In this illustration, doctrine D has five components, and complexity K is two—i.e., each decision has two other decisions upon which it depends. An agent responsible for evaluating D_1 will compare the status quo fitness level to the mutated fitness level, taking into account how the fitness level changes *given the other decision values upon which it depends*, which is determined by the landscape. Thus, in this example, the agent determines that the mutated value for D_1 has a higher fitness level than the status quo value.

3.2.4 Observations about the model

Bounded rationality, complexity, and delegation. Agents try to maximize the fitness level of the decisions for which they are responsible according to

[6]Note that if both T^j_1 and T^j_2 recommend the status quo, all of the comparisons will be the same and no change will occur. Likewise, if one agent recommends the status quo but not the other, there will only be two comparisons. This holds for Step 4 as well.



$$f_{SQ}(001) = 0.02 \qquad < \qquad f_M(101) = 0.03$$

Figure 3.2: Example of experimental search in the *NK*-model

their role in the structure, which corresponds to an area of expertise. The political leader and the tactical-level agents are each responsible for a mutually exclusive subset of the entire doctrine and are thus limited in their rationality relative to the interests of the entire organization, because each is concerned only with maximizing its own subset. The operational-level commanders each consolidate their subordinates' recommendations and try to maximize these composites, but they are also limited in their rationality because they are only concerned with maximizing their tactical aggregated subsets. The senior commander, likewise, tries to maximize the operational-level commanders' composites. The resulting doctrine therefore entails the political leader's subset and the senior commander's subset, which is an aggregate of all the delegated tactical-level subsets after they have been filtered through the operational level.

The limits of expertise create additional challenges for agents' searches for improved fitness levels at higher levels of interdependence (i.e., complexity). In particular, at high complexity, the decision(s) in areas outside of agents' expertise have a greater effect on decisions within their expertise, and vice versa. One way to manage this problem is by varying how much decision-making autonomy the military has compared to the political leader. By lowering autonomy, more decisions are retained at the political level, whereas increasing autonomy increases the number of decisions that are delegated and thus evaluated at the tactical, operational, and strategic levels. To address this possibility I parameterize military autonomy as a variable in the simulations, which I discuss in more detail below.

Adaptation and doctrinal experimentation. At each iteration, the political-level agent and the tactical-level military agents "experiment" by randomly changing a sub-decision to evaluate whether fitness improves for their respective subsets. These experiments correspond to efforts by units within the organization to develop better solutions to problems and adapt to the environment. When the political leader successfully finds a higher fitness level, the political requirements are mandated for the next doctrinal update. For tactical-level agents, however, their successful experiments do not always get incorporated into the doctrine: a successful tactical-level experiment must pass the check by both operational- and senior-level commanders to ensure that the aggregated fitness levels improve at the operational and strategic levels, respectively. Therefore in the model, experimentation is a necessary but not sufficient condition for doctrinal adaptation. Agents must be able to experiment to search for better solutions, but the military components of the doctrine will only adapt if higher-order agents believe the improvements are valuable for a wider scope of the organization.

Although experimentation is necessary for change, it may not always be the case that experimentation is preferred within an organization. March (1991) shows that organizations face tradeoffs between exploration and exploitation. Thus in some cases, experimentation will increase organizational volatility, whereas leaders might prefer stability in order to exploit existing capabilities. Therefore, in the simulations, I parameterize the level of experimentation at both the political and military levels in order to evaluate how preferences for experimentation affect doctrinal evolution.

Competing preferences. Although the agents act cooperatively to search for doctrinal improvements, agents' individual preferences differ, which creates competing interests both horizontally and vertically in the hierarchy. Horizontally, agents have competing interests when their preferred subset of decisions cannot be aggregated with either the status quo or a recommended subset that is submitted for consideration up the chain of command. Vertically, agents have competing interests in two regards. First, preferences for political requirements and strategic military requirements may differ, creating a tension between political and military decisions. Second, within the set of military agents, competing preferences exist up and down the chain of command, because each level has a different scope of consideration for making decisions.

The military autonomy parameter models the effect of different weights on political leaders' decisions relative to the military overall. To narrow the focus of my analysis, I do not directly model the effects of differing preferences and competing interests between military components, and therefore weight each recommendation equally. However, an extension in the model could allow for the effects arising from certain phenomena observed empirically, such as interservice rivalries and competition between units, and disagreements or distrust between upper and lower military echelons, which would skew the weights favoring one agent's decisions over another. Because these additional features would lead to straightforward results without providing additional theoretical value (e.g., heavily weighting one agent, level, or branch's decisions will favor their preferences over another agent, level or branch), I do not include them in the model.

Similarly, to maintain the focus on my analysis, I do not evaluate variations in hierarchical structures. Thus I do not model an unbalanced tree—i.e., where there is an uneven number of agents between vertical branches or horizontal levels—nor do I model a flatter, more networked organization. However, future research could extend the model to explicitly evaluate how changes in organizational design and structure affect doctrinal consensus.

Inefficiency of the equilibrium doctrine. Every landscape has one global optimum, which represents the most effective doctrine for the landscape. However, high complexity means that the number of local optima is also high. Because of the organizational dynamics associated with bounded rationality, adaptation, and competing preferences, it is possible that the doctrine settles on an equilibrium that is only a local optimum. Therefore, the equilibrium may be inefficient, leading to a suboptimal outcome for the organization and lower overall fitness.

3.2.5 Complexity and organizational feature parameters

I evaluated the model through a series of simulations, in which I ran parameter sweeps across a range of possible values for a set of variables of interest. The range of parameter values I evaluated are:

$$\begin{aligned} N &= 10 \\ \textsc{Complexity (K)} &\in \{k \in \mathbb{N} : 1 \leq k \leq 9\} \\ \textsc{Military Autonomy} &\in \{a \in \mathbb{N} : 1 \leq a \leq 9\} \\ \textsc{Pr(Political Experiment)} &\in [0,1] \\ \textsc{Pr(Military Experiment)} &\in [0,1] \end{aligned}$$

I keep fixed the total number of decisions in the doctrine, $N = 10$, because the results are the same for different levels when normalized by N. The main structural model variable therefore is the number of interdependent decisions, K, from 1 to $N - 1$, representing the level of complexity. Additionally, organizations have varying characteristics that affect the ways in which doctrinal decisions are made. The first characteristic I model is MILITARY AUTONOMY, the amount of relative decision-making latitude granted to the military. I model military autonomy as the extent to which political leaders delegate decision-making power from the political to the military level. This dimension captures the concept of military influence, which corresponds to the amount of decision-making power that the military has over political leaders without a direct forceful takeover, such as by coup (Feaver, 1999). When military autonomy is high, the military has greater control over the set of decisions, and and thus the military has substantial influence over policy-making. Conversely, when military autonomy is low, military decisions have far less influence on political requirements, and most decisions are made at the political level and passed down the chain of command from the political leader.

As part of the military autonomy process in the model, any decisions delegated to the military are divided equally down the chain of command, which corresponds to a division-of-labor, or specialization, effort. I assume that the political leader always retains a non-zero number of decisions for herself, and always delegates a non-zero number of decisions to the military. Therefore, the minimum number of decisions that can be delegated in the model is one, and the maximum number of decisions is nine. When the number of delegated decisions is not divisible by four, they are distributed between the lower levels as evenly as possible. Thus, for example, if MILITARY AUTONOMY $= 3$, the decisions will be delegated as a 2-1 split between the two operational commanders; the first two decisions will then be delegated between each of the two tactical agents for the former operational agent, and the single remaining decision will be assigned to one tactical agent for the latter operational agent.

Two additional variables affect the search for doctrinal changes: the probabilities of political and military experimentation, which I treat as independent values. These probabilities correspond to the tolerance for experimentation by agents at the political and military levels, respectively. The probability of an experiment increases in the tolerance for experimentation at the corresponding level. It is important to note that tolerance for experimentation affects the potential for change in the doctrinal decisions that a particular agent has control over, but not necessarily the impact of such changes. Thus, for example, if military autonomy is low and the tolerance for military experimentation is high, the military will be likely to make changes, but because it has little control over doctrinal decisions, it can only have a small impact on any potential changes that might occur. Likewise, if military autonomy is high but the tolerance for military experimentation is low, then few changes will likely occur because the military controls most of the doctrine but has little tolerance for change. The same dynamic applies to the political level.

Variable	*Mean*	*SD*	*Min*	*Max*
COMPLEXITY (K)	0.50	0.26	0.10	0.90
MILITARY AUTONOMY	0.50	0.26	0.10	0.90
PR(POLITICAL EXPERIMENT)	0.50	0.29	0	1
PR(MILITARY EXPERIMENT)	0.50	0.29	0	1
FINAL DOCTRINE FITNESS	0.86	0.09	0.28	1
FINAL DOCTRINE ITERATION	21.88	33.57	1	200

Number of observations: 810,000

Table 3.1: Descriptive statistics: simulation variables

3.3 Model analysis

To analyze the model, I conducted a parameter sweep across the range of values for each of the variables. For each pair of discrete K-MILITARY AUTONOMY values, I ran 1,000 simulations of 200 iterations (time periods) each.[7] In each simulation, the probability of political and military experiments were drawn independently from a random uniform distribution. This resulted in 810,000 simulations covering the entire range of parameter values. Descriptive statistics for the simulation data are provided in Table 3.1. Fitness measures are normalized by the maximum possible fitness per simulation (which depends on the simulation's particular landscape), and complexity (K) and MILITARY AUTONOMY are normalized by the number of doctrinal decisions, $N = 10$.

To evaluate the effect of the commitment horizon level (i.e., the discount factor), I computed a discounted final fitness level, $\delta^t F$, for each simulation. The commitment horizon, $\delta \in (0, 1)$, affects the time value of the final doctrinal fitness F, where t is the number of iterations needed to reach an equilibrium. Thus, the organization's total payoff depends on how much it discounts the future fitness value of the final doctrine: when δ is low, an organization will prefer a doctrine that may provide a lower fitness level sooner

[7]Within the full simulation results, the likelihood of failing to reach model convergence by the 200th iteration is less than 0.001.

rather than wait for a better doctrine later, whereas when δ increases, the organization is more willing to wait for a better doctrine that takes longer to achieve.

I then used unbiased recursive partitioning, a supervised learning algorithm, to find the parameter values that maximize the final doctrinal fitness value at each 0.01 increment for the commitment horizon level. Unbiased recursive partitioning is a tree-based classification algorithm that uses recursive partitioning of the parameters to create a conditional inference tree (Hothorn, Hornik, and Zeileis, 2006). This approach is useful for generalizing and predicting output variables in the data (Patel, Abbasi, Saeed et al., 2018).The results of the algorithm provide a partition (i.e., a range of values) of each model parameter that maximizes the correlation with the observed final fitness value.

Next, I selected the set of partitions that maximize the final doctrinal fitness value for each discount level increment, based on the expectation that a military will optimize their organizations for the commitment horizon they have in a conflict. I limited the focus of the optimization to high complexity landscapes—corresponding to the substantive focus of counterinsurgency conflicts—by selecting the partitions that had no upper limit on K.[8]

Once I determined the optimal parameter values for each commitment horizon level at high complexity, I then evaluated how the results grouped together into similar characteristics using k-means clustering. This approach is a form of unsupervised learning, which iteratively minimizes the within-class sum of squares for a set of clusters (Hartigan and Wong, 1979; MacQueen, 1967). Because the computed clusters may vary based on how the algorithm initializes, a wide variety of indices have been developed to evaluate clustering results. Therefore, I used a method which computes 30 such indices for determining the relevant number of clusters and proposes the clustering scheme that is chosen by the plurality of the indices (Charrad, Gahhazli, Boiteau et al., 2014).

[8]The 95% confidence interval for the lower limit of complexity for this restriction is $K_{lower} = (0.70, 0.74)$.

3.3.1 Moving from exogenous parameters to endogenous equilibria

Ever since John Nash's pioneering concept of strategy equilibrium (Nash, 1950), game theorists have understood an equilibrium as a set of decisions from which players will not deviate. Because each player's optimal decision depends on every other player's decision, an equilibrium is inherently endogenous. However, in the agent-based *NK*-model, the parameters (such as *K*, autonomy, and experimentation) are exogenous, and therefore one may reasonably ask whether it makes sense for organizations to lack a conscious choice about the parameters they adopt. Yet as I explore the parameter space in the model, I evaluate all possible solutions in order to find those that are optimal, based on the assumption that organizations will choose only those sets of parameters that provide an optimal solution. This is somewhat analogous to defining the full game tree in which all possible decisions are mapped out, even though multiple decisions may ultimately be off the equilibrium path. Therefore, I make an explicit assumption that the exogenous features I have modeled are endogenously determined in reality.

3.4 Results

Figure 3.3 visualizes the results of the model analysis. Each data point in the figure plots the average (undiscounted) optimal final fitness value for a specific commitment horizon level, with 95% confidence intervals for each fitness value. I plot the undiscounted fitness values because doing so provides a standardized way to compare outcomes across commitment horizon levels. Moreover, the undiscounted fitness values represent the realized effectiveness that the organization achieves, regardless of how long it takes to reach the doctrine. Consequently, comparing undiscounted fitness values allows us to evaluate what the expected final observed effectiveness is for different commitment horizons.

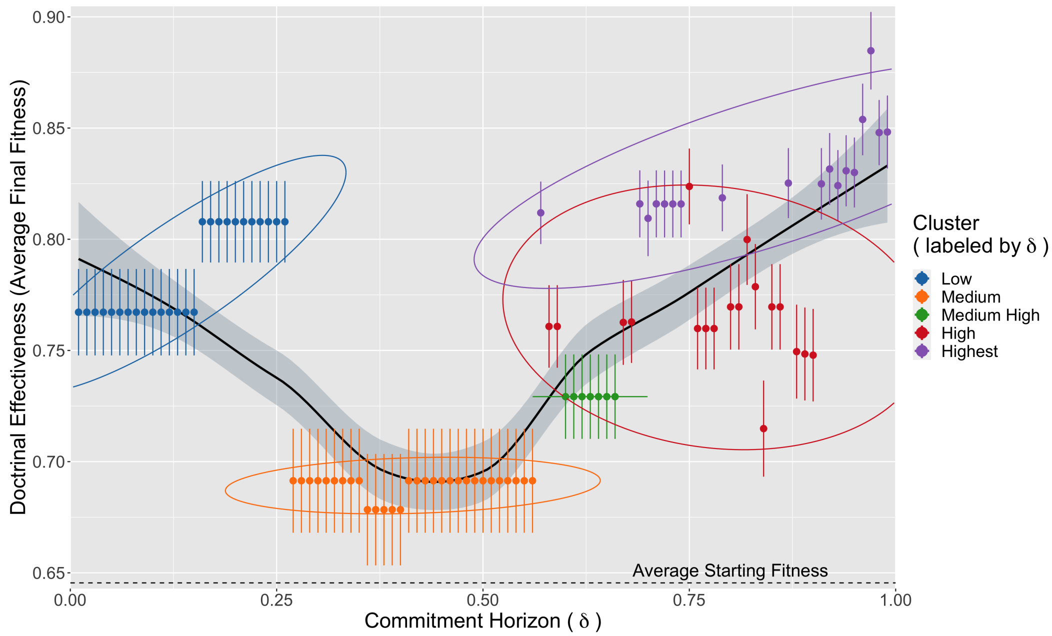
The results from the *k*-means clustering algorithm indicate that there are five distinct clusters across the range of commitment horizon values. These clusters, depicted in Figure 3.3, correspond to distinct sets of organizational characteristics from the model, which are listed in Table 3.2. Because the clusters are broadly distinguishable by the commitment horizon level, it makes sense to discuss them in reference to these levels. Therefore, the commitment horizon defines the categorization, and the levels of military autonomy and tolerances for political and military experimentation are predicted by this categorization. Thus, for example, at a low commitment horizon, a cluster exists which corresponds to medium military autonomy, low tolerance for political experimentation, and high tolerance for military experimentation. These clusters suggest a typology of counterinsurgent militaries according to their commitment horizon, which I discuss below.

Commitment Horizon (δ)		*Military Autonomy*		*Political Experim. Tolerance*		*Military Experim. Tolerance*	
0.01 – 0.26	Low	0.43 – 0.53	Medium	0.00 – 0.01	Low	0.65 – 1	High
0.27 – 0.56	Medium	0.02 – 0.13	Low	0.00 – 0.005	Low	0.00 – 1	Any
0.60 – 0.66	Medium High	0.30 – 1	Med/High	0.71 – 1	High	0.00 – 0.003	Low
0.58 – 0.90	High	0.37 – 0.49	Medium	0.00 – 0.01	Low	0.44 – 0.91	Med/High
0.57 – 0.99	Highest	0.72 – 0.96	High	0.13 – 0.50	Medium	0.82 – 1	Very High

Note: Row-wise cluster characteristics, computed by k-means clustering, for each type illustrated in Figure 3.3. Numeric ranges are the minimum and maximum values for δ, and the average minimum and maximum partition values for all other variables.

Table 3.2: Cluster characteristics

Another important feature to observe in Figure 3.3 is the trend in doctrinal fitness across commitment horizon levels—and thus, across clusters. The results in Figure 3.3 suggest that doctrinal effectiveness follows a non-monotone, U-shaped curve as the commitment horizon increases. The model generates this surprising result because of the ways that the optimal organizational characteristics vary according to the commitment level. In particular, the least effective organizations are *not* those with the lowest commitment, because these organizations will urgently try to "shore up" a quick solution. Rather, the least effective are those with a lukewarm commitment: they have neither the urgency to find quick fixes, nor the endurance to search for better solutions over a longer duration. This result has important implications for predicting how each of the organizational types



Note: Effect of commitment horizon (δ) on doctrinal effectiveness (F) in high complexity environments ($\lim \frac{K}{N} \to 1$), clustered by type (see Table 3.2 for details), with 95% confidence ellipses for each type. F is the average of the simulated final fitness for the optimal organizational characteristics at each δ (formally, $\arg\max F \; \forall \; \delta$, as determined by unbiased recursive partitioning), with 95% confidence interval lines plotted for each fitness level. Locally estimated scatterplot smoothing (LOESS) curve with 95% confidence interval band is also plotted to illustrate general trend pattern across δ levels.

Figure 3.3: Overall fitness and commitment horizon level (i.e., discount factor), clustered by type

will perform in counterinsurgency conflicts, as I will describe in the next section.

3.4.1 A typology of counterinsurgency militaries

The results from the analysis suggest a typology of counterinsurgency militaries according to their commitment horizon. In what follows, I describe each type, using mythological archetypes as a way to help identify each type.

Low commitment horizon: the Proteus. Proteus was a god in Greek mythology. If someone wished to consult him, they had to surprise and capture him. However, he would attempt to escape capture by shapeshifting.[9] Therefore, in personifying adaptation and quick escape, Proteus represents an archetype of a military that possesses some adaptability, but will try to "cut and run" rather than commit to a long struggle.

Proteus resembles a military that has a low commitment horizon. From Table 3.2, we see that a low commitment horizon corresponds to medium military autonomy, low tolerance for political experimentation, and high tolerance for military experimentation. These characteristics translate to a military with moderate decision-making autonomy that has a strong willingness to search for doctrinal changes that will improve its effectiveness. Concurrently, while the political leader retains some decision-making control, she will rarely make any changes to her decisions, preferring to maintain the status quo. Consequently, any adaptations that occur will likely come predominantly from the military in response to tactical-level inputs based on encounters with the local environment. Moreover, due to a low commitment horizon, these changes will tend to occur quickly.

A low commitment horizon and its corresponding military characteristics will result in moderate counterinsurgency effectiveness, relative to other types. A military facing a low commitment horizon will essentially jury rig its doctrine: it tries to improvise "on the fly," and does so quickly when facing unanticipated situations. These efforts are done in an attempt to snatch victory from the jaws of defeat, but the lack of an enduring commitment means that any results achieved will be moderate at best.

Medium commitment horizon: the Goliath. Goliath was a quintessential ancient warrior, described in the Bible as a fearsome, heavily armed and armored giant with enormous strength. However, Goliath was one-dimensional, relying on brute force for victory in combat. Thus, he believed he could win a one-on-one battle against David, a young shepherd, through

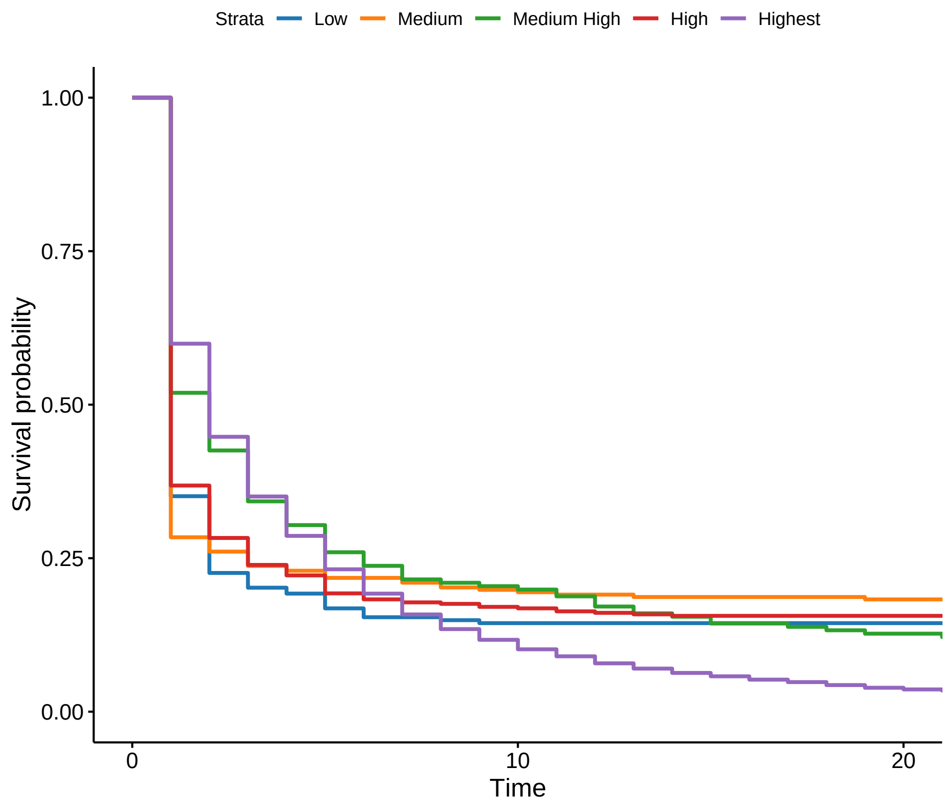
[9] *Britannica*. "Proteus." https://www.britannica.com/topic/Proteus-Greek-mythology.

sheer power. Yet David recognized that Goliath's size and armor reduced his mobility. Therefore David, wearing no armor and lacking advanced weaponry, avoided close combat and slayed Goliath where he was most vulnerable, by slinging a rock at his forehead.[10] Goliath therefore represents an archetype of a traditional large-unit military: it is large and powerful, but nevertheless vulnerable because of an incorrect belief that fighting a pitched battle—which may have led to success in the past—is the surest path to victory regardless of context. Such a military is unprepared for unconventional combat, because it lacks the flexibility to adapt to local conditions.

Goliath corresponds to a military that has a medium commitment horizon in counterinsurgency conflict. With a medium commitment horizon, the optimal organizational characteristics feature low military autonomy, low tolerance for political experimentation, and *any* level of tolerance for military experimentation. Consequently, the political leader retains control over most decisions and is unlikely to make many changes to the status quo doctrine. Moreover, even if the military has a high tolerance for experimentation, any changes it makes from tactical-level inputs are unlikely to have much impact because the military has so little control over the overall doctrine. Additionally, any changes that do occur are more likely to occur at the political level, because that is where most decision-making power is retained. Therefore, the optimal characteristics for a Goliath type imply that few doctrinal changes will occur with a medium commitment horizon.

Furthermore, a medium commitment horizon will result in poor counterinsurgency effectiveness, as illustrated in Figure 3.3. Counterintuitively, moderately increasing the commitment horizon above the Protean type leads to both stifled adaptation and lower overall effectiveness. This result emerges from ambiguity and uncertainty about the level and duration of commitment. The Goliath type lacks the same urgency that drives the Protean type to find the quickest solution, but also lacks a commitment to a longer time horizon and the coinciding patience to search for doctrinal

[10] *Britannica*. "Goliath." https://www.britannica.com/biography/Goliath-biblical-figure.



Note: Survival probability is the likelihood that the doctrinal equilibrium has *not* been reached.

Figure 3.4: Short-term survival curves plotted by type

improvements over an extended period. Consequently, the payoffs from solutions at or near the initial status quo are greater than payoffs from solutions that are either too quick and short-term focused or too slow and long-term focused.

The dynamic that drives the distinction between the Goliath type and those with other commitment horizons can be seen in Figure 3.4, which plots each type's survival curves for the first 20 time steps, i.e., the short term. Both the Protean type (corresponding to the blue, low commitment horizon curve), and its sister type, the Military Sisyphus (depicted as the red, high commitment horizon curve), initially have a higher doctrinal survival probability—i.e., higher likelihood of *not* having reached a doctrinal equilibrium—than the Goliath type, depicted by the orange (medium commitment horizon) curve. However, a few time steps later, the blue curve drops below the orange, but the red remains above and then even with the orange curve. Thus, initially, the Goliath type is most likely to have reached equilibrium; shortly thereafter, the Protean type is most likely, whereas the Military Sisyphus type is less likely. It is this dynamic between the survival curves—in conjunction with the relative expected payoff for each solution—that help determine the types' optimal characteristics. Consequently, the lack of urgency to find a quick solution, combined with a shorter commitment horizon than the high type, leads to less incentive for the Goliath type to adapt. As a result, the Goliath type ultimately optimizes by adopting characteristics that encourage stagnation, leading to poor overall effectiveness. Notably, the Goliath type also faces the longest total survival curve, implying that Goliath types may be more likely to experience high-duration conflicts than other types, assuming duration is correlated with the doctrinal survival curves.

Medium-high commitment horizon: the Political Sisyphus. In Greek mythology, Sisyphus was condemned to an eternal punishment in Hades where he was forced to roll a monstrous boulder up a hill. As he rolled up the boulder, it would appear to almost fall over to the other side, but each time he tried, the boulder would tumble back down again under its own weight.[11] Thus, Sisyphus "would heave, would struggle to thrust it up, sweat drenching his body, dust swirling above his head" (Homer, 2002, p. 269). Sisyphus exemplifies a military archetype that exhibits high endurance and persistence, yet lacks sufficient ability to achieve its goal, which may be in sight but is ultimately beyond its ability to reach.

The Political Sisyphus type is one of two variations that are observed in the model at medium-high and high commitment horizons. The other variation, the Military Sisyphus, will be described next. As Figure 3.3 illustrates,

[11] *World History Encyclopedia*. "Sisyphus." https://www.ancient.eu/sisyphus/.

the two Sisyphus types have completely overlapping clusters, suggesting that they are organizational variants of a single type.

The characteristics of a Political Sisyphus type—a military with a medium-high commitment horizon in a counterinsurgency—are medium or greater military autonomy, high tolerance for political experimentation, and low tolerance for military experimentation. With these features, an organization optimized for medium-high commitment in a counterinsurgency will delegate control of some doctrinal decisions to the military, but with the understanding that most of these decisions will not be subject to change. Instead, most of the changes will occur at the political level; in fact, the Political Sisyphus is the only type that exhibits high tolerance for political experimentation. As such, we would expect to see an organization with a medium-high commitment horizon regularly changing its political approaches in an attempt to find a political solution to the conflict, rather than focusing on adapting militarily as the primary means for improving. These changes will occur over time, as the organization possesses a moderately high level of commitment in the conflict. However, the total effectiveness for a Political Sisyphus is only moderate at best. While it is better off than the Goliath type, its fitness is below other types that could be achieved, both at low and higher commitment levels.

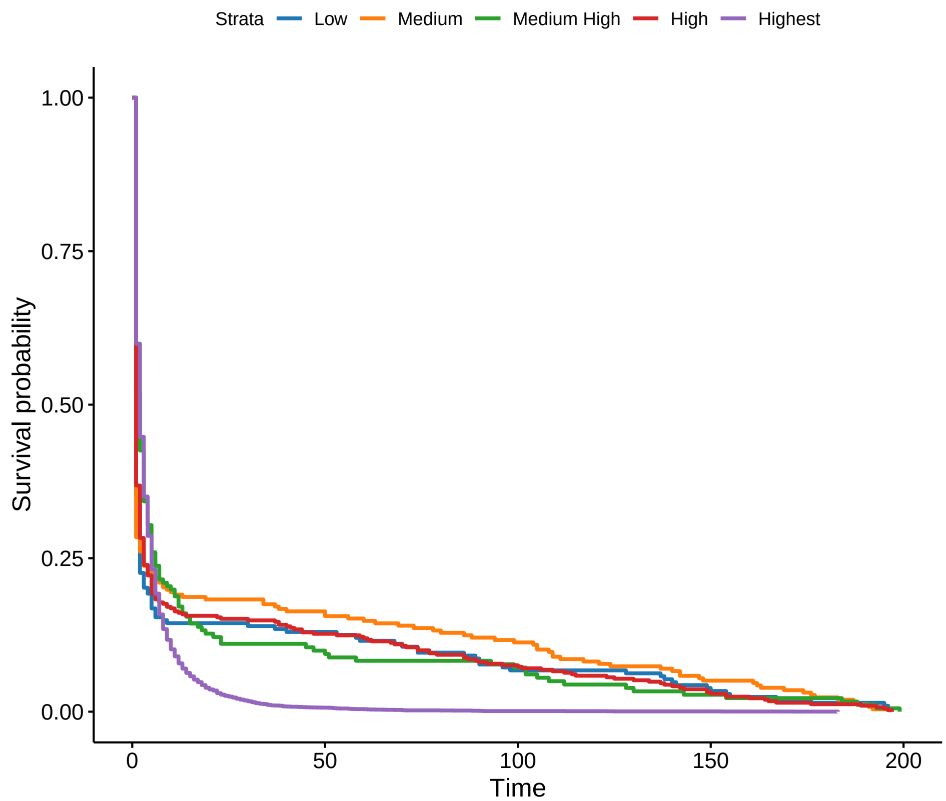
High commitment horizon: the Military Sisyphus. The Military Sisyphus type is like its political counterpart in terms of similar commitment level and overall effectiveness. Instead, the primary distinction between the two types is in where changes tend to occur: the Military Sisyphus type features low tolerance for political experimentation and medium or greater tolerance for military experimentation. Consequently, most adaptations and improvements to the doctrine will occur over an extended period of time at the military level, rather than the political level. Yet the total effectiveness for a Military Sisyphus is approximately the same, if only slightly better, than the Political Sisyphus. Nevertheless, compared to the Protean type, the Military Sisyphus' expected total fitness is similar.

Note, also, that the optimal organizational characteristics for the Protean and the Military Sisyphus types are nearly identical. This suggests that a military which organizes by these characteristics will match both a low and a high commitment horizon for counterinsurgency. Yet even if the same military is involved in a low commitment conflict and transforms to one with a high commitment, either in the same conflict or elsewhere, its total effectiveness will be approximately the same. Therefore, while a naïve actor might think that a shift from a low to a high commitment would imply greater effectiveness, in fact the U-shaped aspect of the effectiveness curve suggests that the overall doctrinal changes will not result in demonstrable improvements. Thus, the Military Sisyphus type toils with greater endurance, but fails to achieve its goals to any greater effect.

Highest commitment horizon: the Atlas. Atlas was portrayed in Greek mythology as having eternal endurance, responsible for holding the weight of the heavens or the earth on his shoulders, or holding up pillars in the Atlantic Ocean that separated the heavens from the earth.[12] Atlas therefore embodies an archetype of a military with the greatest possible endurance that is fully committed to reaching its goal until it is achieved, regardless of the challenges it encounters.

The Atlas type coincides with a military possessing the highest commitment horizon in a counterinsurgency. Such an organization will be marked by high military autonomy, medium tolerance for political experimentation, and very high tolerance for military experimentation. As such, an Atlas type military will have high decision-making autonomy, with a very strong willingness to to search for doctrinal changes at the military level, as well as a strong willingness to search for changes at the political level. Given the very high commitment horizon, one or more of these changes may occur slowly over time. Nevertheless, the political and military levels act in concert, adapting to each other's changes and working together toward comprehensive improvements over the long run.

[12] *World History Encyclopedia*. "Atlas." https://www.ancient.eu/Atlas/.



Note: Survival probability is the likelihood that the doctrinal equilibrium has *not* been reached.

Figure 3.5: Long-term survival curves plotted by type

A very high commitment horizon with Atlas-type characteristics will exhibit the highest overall effectiveness of all the counterinsurgency types. Intuitively, this makes sense: when an organization is willing to search for improvements and is committed to a long time horizon, it is more likely to achieve a strong solution. This result also corresponds with March's model, which demonstrates that the best way for an organization to optimize over the long run is to engage in high exploration before exploiting the discoveries (March, 1991).[13] The tradeoff, of course, in my model is that it may

[13]See also the theory of simulated annealing, which is a useful optimization technique for finding a global optimum when time considerations are less important.

take longer to reach a better outcome: the organization must endure the potentially difficult process of adapting to local conditions, searching for solutions, and changing—possibly many times—as new and better approaches are discovered. As depicted in Figure 3.5, the Atlas type (highest commitment horizon, purple line) has the highest survival curve for much of the early search process. Notably, however, as time continues the Atlas type's survival curve drops below the other types. This result indicates that the duration of a conflict involving an Atlas type may sometimes be shorter than a conflict involving the lower commitment types. Thus, contrary to what one might expect, a military with a very high commitment horizon may actually experience a shorter counterinsurgency duration than a military with a lower commitment horizon. Therefore, commitment horizon level and conflict duration are not necessarily equivalent.

3.5 Illustrative cases of counterinsurgency types

To illustrate the logic of the model and evaluate the typology of counterinsurgent militaries I have developed, I next consider the historical record from several modern counterinsurgencies. For each of these cases, I briefly discuss which counterinsurgency type the military resembles and how well the model's predictions fit the case. While I will describe the relevant components of each case, much has already been written about these cases, and therefore I refer the reader to outside historical literature for additional details.

The typology of counterinsurgent militaries describes a set of organizational equilibria that share similar features across commitment horizon levels. Yet if states are rational actors, why would they want to adopt a less effective strategy profile if they recognize that there are better solutions possible? For example, why remain as a Goliath type profile, given that it is strictly dominated by all others? First, it is important to emphasize that my model implies that political-military systems optimize their organizational features according to the commitment horizon, which then explains their overall effectiveness. Therefore, commitment horizons are the primary driving force of all the other features. However, I argue that commitment horizons vary by case, and therefore are not systematic features of an organization.[14] The organizational features adopted follow from the commitment horizon, and these features (i.e., the strategy profile) will tend to remain stable throughout the conflict—so long as it is not too costly for the organization to change. Therefore, it is possible to change characteristics and induce change if the benefits sufficiently outweigh the costs, though only within the bounds of what the nearby commitment horizon will allow.[15] Nevertheless, in most of the cases I discuss below, the strategy profiles tend to remain stable, and thus the costs of reaching a new equilibrium by adopting new organizational features may often be too high.

The cases I will evaluate are: the U.S. in Iraq and Afghanistan; the British in Malaya and Iraq; and the French in Indochina and Algeria. I selected each of these cases according to the method of concomitant variation, which is based on the principle that the cause of changes in a variable across a set of cases can be inferred by identifying other variables that also change in a systematic way across the cases (George and Bennett, 2005). Thus, I have identified a set of cases where there is variation in commitment horizons (and corresponding organizational characteristics) that can be used to explain variation in counterinsurgency effectiveness. Within these cases, several variables are held constant to isolate them as possible causes: namely, the cases all involve Western democratic militaries, after World War II, fighting counterinsurgencies outside their own borders.[16]

[14]To evaluate the qualitative features of commitment horizons empirically, I operationalize the parameter below by considering political and military leaders' beliefs and expectations about the conflict, as well as exogenous features such as geographic distance.

[15]Thus, as I will explain below, the Political and Military Sisyphus types can transform into each other, because they share overlapping commitment levels, and the only major organizational change is the relative emphasis on political versus military experimentation—an arguably low-cost organizational change.

[16]The one minor exception is the French military in Algeria, because at the time the French considered Algeria to be part of France—*Algérie française*—rather than a separate colony. Nevertheless, Algeria was distinct from *la Métropole* in virtually every way except for its sovereignty.

Furthermore, the selected cases include the same state fighting separate counterinsurgency conflicts against different opponents, and therefore I can compare variation not only between armies, but also within armies across different conflicts. This is important because it helps account for alternative explanations of counterinsurgency effectiveness, including strategy selection and fixed military culture. As I will show, because of the variation in outcomes within the same army across conflicts, fixed military culture is not an effective explanation for counterinsurgency effectiveness. Additionally, because the same strategy applied in separate conflicts has led to different outcomes, strategy selection is also not an effective explanation.

Operationalizing commitment horizon requires carefully evaluating political and military leaders' beliefs and expectations about the conflict in its earliest stages, without consideration for the conflict's total duration as understood *ex post*. To do so, I considered the following questions for each case, summing up the responses (lower $= -1$, higher $= +1$) to determine the overall commitment horizon. Table 3.3 provides a summary of each case's commitment horizon according to these scores:

1. Did the conflict begin because the counterinsurgents' military or civilian population was attacked?
   Yes = higher commitment. No = lower commitment.

2. Do political and military leaders expect a short-duration conflict?
   Yes = lower commitment. No = higher commitment.

3. Are there strong domestic constraints (selectorate opposition, domestic crises, etc.) to the counterinsurgency campaign?
   Yes = lower commitment. No = higher commitment.

4. Is there an emphasis on having an "exit strategy"?
   Yes = lower commitment. No = higher commitment.

5. Is the conflict geographically distant from the counterinsurgents' own country?
   Yes = lower commitment. No = higher commitment.

Case	*Attacked?*	*Expect short war?*	*Domestic constraints?*	*Exit strategy?*	*Distant war?*	*Score*	*Commitment Horizon*
U.S. in Iraq	No (−)	Yes (−)	Yes (−)	Yes (−)	Yes (−)	-5	Low
British in Iraq	No (−)	Yes (−)	Yes (−)	Yes (−)	Yes (−)	-5	Low
French in Indochina	Yes (+)	Yes (−)	Yes (−)	No (+)	Yes (−)	-1	Medium
U.S. in Afghanistan	Yes (+)	No (+)	No (+)	Yes (−)	Yes (−)	1	Med. High/High
French in Algeria	Yes (+)	Yes (−)	Yes (−)	No (+)	No (+)	1	Med. High/High
British in Malaya	Yes (+)	No (+)	No (+)	No (+)	Yes (−)	3	Highest

Table 3.3: Operationalizing commitment horizon for each case

3.6 The Proteus (low commitment horizon)

3.6.1 U.S. in Iraq, 2003–2011

The U.S. military had a low commitment horizon from the very first stages of the Iraq war: in November 2001, U.S. military officers who began planning for an invasion of Iraq assumed that the war would end after Saddam Hussein was deposed (Barno and Bensahel, 2020). Moreover, domestic opposition to the war was quite high even before it began, with public support for the war hovering between 52 and 59 percent (Smith and Lindsay, 2003). When the war began in March 2003, political and military leaders alike thus expected that U.S. forces would swiftly overwhelm the Iraqi military through "shock and awe," the Iraqi people would install a friendly democratic regime, and U.S. forces would exit the country soon thereafter. Initially, events reinforced these expectations. In just three weeks, the U.S. military managed to defeat Iraqi forces, conquer Baghdad, and overthrow the Saddam Hussein regime. Soon after, however, an insurgency began to emerge and quickly escalate, marked by improvised explosive devices, car bombs, mortars, and suicide attacks against occupying forces and civilians alike throughout the country. In 2003, Iraqi civilians suffered an estimated 7,300 fatalities; by 2006, that number would rise to 34,500 (O'Hanlon and Livingston, 2011).

Thus, U.S. forces quickly found themselves in a highly complex environment requiring adaption to local conditions. Interdependencies between doctrinal decisions that arose due to competing sectarian groups, Muslim cultural beliefs that were foreign to many U.S. troops, political definitions (and the lack thereof) on lawful combatants, new insurgent tactics, and new technologies of war all contributed to complexities that far exceeded what the U.S. military had previously trained for. The insurgency that emerged involved dozens of groups, including Ba'ath Party loyalists of the Saddam regime, Sunni militias such as Al Qaeda in Iraq, and Shi'a militias such as the Mahdi Army. The complexity of Shi'a-Sunni relations also helped create a sectarian civil war that added to the conflict's challenges.

A Protean type military will optimize for its low commitment horizon in a counterinsurgency by adopting medium military autonomy, low tolerance for political experimentation, and high tolerance for military experimentation. As a result, it will quickly adapt its doctrine, but will achieve moderate total effectiveness. How well does the evidence from the U.S. efforts in Iraq support these predictions?

The U.S. military has historically possessed high autonomy over its doctrine, given the expertise it brings to bear on the problem of fighting conventional combat. Whether as AirLand Battle in the 1980s and 1990s or its successor, network-centric warfare, the conventional doctrine emphasizes large-scale, highly mobile operations that are designed to overwhelm an enemy's command and control structure and thereby defeat opposing forces. However, due to the inherently political nature of the conflict in Iraq, U.S. political leaders retained greater control over some of the doctrinal decisions. Consequently, after an initial period of high military autonomy, political leaders withdrew some decision-making control from the military, thereby creating a set of political requirements that remained for the duration of the conflict. As a result, in accordance with the Protean type, we can characterize the U.S. military as having medium autonomy over its doctrinal decisions in the Iraq war.

As conditions in Iraq deteriorated following the early conventional victory, therefore, the political interest in taking control of some doctrinal decisions increased. For example, senior military and political leaders came to recognize that the military's existing approach, emphasizing stability operations in Iraq, was proving insufficient. Within the first year of the war, there was greater discussion among the leadership at the Pentagon that waging a counterinsurgency would require changing policies and approaches (Ucko, 2009). Furthermore, a particularly high-profile decision was controlled at the political level that would create lasting political damage to the George W. Bush administration: namely, the authorization by the highest levels of government that military units could engage in "enhanced interrogation techniques"—euphemism for torture—which led to the inhumane treatment of detainees and events such as the torturing of Iraqi prisoners at Abu Ghraib prison (United States Senate, 2008).

Throughout the duration of the Iraq war, political experimentation was fairly sparse, supporting the model's expectation. Although Bush signed the Detainee Treatment Act in 2005, which codified the humane treatment of detainees and banned torture and waterboarding for the military, it was not until Leon Panetta became Director of the Central Intelligence Agency in 2009 that a similar ban was applied to CIA interrogators. Thus, some torture was tolerated for several years, even after the fall-out from media exposure about the treatment of Iraqi detainees. Separately, the Bush administration chose to dissolve the Coalition Provisional Authority in 2004 in order to hastily jump start Iraqi self-rule and end the U.S. occupation, though it ultimately did not have the desired effect. Similarly, both in 2005 and in 2010, the White House switched from favoring an Iraqi government chosen by elections to a power-sharing deal. Yet the single most apparent political change during the war was the "Surge," President Bush's decision to send an additional 30,000 soldiers to reinforce the U.S. presence in Iraq. Beyond this decision, high-level U.S. political leaders maintained a stable set of nation-building requirements for the military in Iraq, which focused on developing a self-determined democratic government and training security forces that could eventually maintain stability in the country without U.S. support. Ultimately, despite the lack of major progress toward its policy goals, U.S. political leaders remained steady in their requirements until U.S. troops were withdrawn from Iraq at the end of 2011.

Tolerance for military experimentation, on the other hand, was far higher. In the early stages of the insurgency, with troops isolated from the Iraqi population on large bases, the U.S. military's mission was to transition security to Iraqi control as soon as possible. U.S. forces would conduct counterinsurgency missions by patrolling outside of the base, driving around in heavily armored vehicles. Moreover, the most common approach taken by U.S. troops was purely focused on heavy-handed "kinetic" operations—missions involving lethal force. Yet as conditions deteriorated, some began to recognize that new approaches were needed. Evidence from numerous studies of the U.S. military in Iraq demonstrates that it experienced substantial bottom-up adaptation throughout the organization (Barno and Bensahel, 2020; Burton and Nagl, 2008; Kollars, 2015; Russell, 2010). As a result, for example, as early as 2003, then-Major General David Petraeus led his unit in an unprecedented effort to provide public goods to northern Iraqi cities, including restarting telephone service, paving roads, and trading Iraqi oil with Turkey and Syria for electricity in the region (Boot, 2013). By 2005, some other commanders, such as then-Colonel H.R. McMaster, led their units to enact tactics that further shifted the emphasis away from kinetic operations toward providing support to the local population, treating detainees with respect, and establishing effective intelligence about insurgents (Boot, 2013; Burton and Nagl, 2008). In December 2006, FM 3-24 (the Army's counterinsurgency field manual, authored by then-Lieutenant General Petraeus), was published, establishing the official doctrine of population-centric counterinsurgency (COIN) and reinforcing many of the solutions McMaster's and Petraeus' units had identified and implemented.

Beginning in 2007 (and coinciding with Petraeus' command of all U.S. forces in Iraq), the new doctrine was implemented in Iraq. U.S. troops were no longer kept isolated in their bases; instead, they were placed into smaller outposts throughout population centers. From there, they would patrol the local area, living among Iraqi civilians and gaining their trust. Rather than armored convoys, foot patrols became more common. The results from these changes were significant, coinciding with a sharp drop in overall violence within a few months (Boot, 2013).

Despite invading Iraq with no clear expectation of—or plan for—waging a counterinsurgency, the U.S. military thus managed to adapt its doctrine, transforming from a conventional to a counterinsurgent force, thereby improving its overall effectiveness (Burton and Nagl, 2008). As a Protean type facing a low commitment horizon, the U.S. military navigated the rugged landscape of social, political, and technological complexity in Iraq primarily through bottom-up military experimentation that discovered better solutions to the challenges driven by local conditions. The Iraqi "Anbar Awakening," the 2006 publication of FM 3-24, the political experimentation with the surge in U.S. forces, and the implementation of population-centric COIN, all converged around the peak of the insurgency and together led to the subsequent decline in insurgent violence (Biddle, Friedman, and Shapiro, 2012). By 2009, there were an estimated 3,000 civilian fatalities, further declining to 1,600 fatalities in 2011 (O'Hanlon and Livingston, 2011).

Nevertheless, despite its improvements as an organization, U.S. military effectiveness in Iraq was mostly short-lived. After U.S. forces were withdrawn at the end of 2011, sectarian violence continued. A year later, tens of thousands of Sunnis took to the streets to protest the predominantly Shi'a government. Soon after, ISIS emerged as a formidable extremist group, significantly threatening Iraqi security and stability. ISIS eventually waged an organized offensive to capture Anbar province, and U.S.-trained Iraqi forces collapsed in battle, leading to the fall of Ramadi, Fallujah, and Mosul in 2014 (Lehmann and Zhukov, 2019).

In short, the characteristics of the U.S. military in Iraq from 2003 to 2011 fit the predictions of a Protean-type military in a counterinsurgency. Its belief in a quick victory led it to adopt a low commitment horizon, and its organizational features corresponded to its type, as it adopted medium military autonomy, low tolerance for political experimentation and high tolerance for military experimentation. Furthermore, the U.S. military adapted quickly to local conditions, as it sought to quickly find a solution to an un-expected insurgency. Moreover, although it managed to improve upon its initial effectiveness, its overall success was temporary, and thus it achieved only moderate effectiveness in the long run. Therefore, the U.S. military's quick adaptability and overall effectiveness also matches a Protean type, further bolstering support for the model's predictions.

3.6.2 British in Iraq, 2003–2009

The British military in Iraq also fits the Protean type. Their commitment to Iraq was closely tied to the U.S. military's, and consequently their organizational characteristics also closely mirrored those of the U.S. Therefore, because of the close similarities between the two militaries, I will only briefly discuss the British case. The reason I include it, however, is because of a key distinction in the British military's beliefs: namely, that because of their past success in other counterinsurgencies—including in Malaya (discussed below), Kenya, and Northern Ireland—the British (and others) believed that they had discovered the keys to maximize counterinsurgency effectiveness. In particular, many have attributed the British military's success to a combination of a long-term, fixed military culture and the correct application of a "hearts and minds" formula (Porch, 2013). Yet rather than serving as a model for the U.S. and others to follow, the British would also demonstrate only moderate effectiveness in Iraq.

As with the U.S. military, the British invaded Iraq in 2003 with similar levels of public support and expecting a short conflict, and therefore also had a low commitment horizon. For example, the British intended to reduce the number of troops used for an invasion by half within the first six months of the war (Akam, 2018). Yet the situation began to quickly deteriorate in Basra, where the British military was responsible for operating. Like the U.S., the British military had organized for a conventional conflict, and was unprepared for what followed the initial invasion success. Without a plan for conducting a counterinsurgency, they failed to implement effective local solutions, leading to escalating insurgent violence and civil unrest. By 2006, at least one British general was openly remarking to journalists that the British should withdraw from Iraq (Dixon, 2009).

British military autonomy was somewhat less than that of the U.S. military. British political control of its doctrinal decisions was driven by risk aversion: political leaders required military approaches that would minimize casualties, given the unpopularity of the war domestically. For example, British political leaders mandated that British forces would not embed training teams with Iraqis, despite the fact that British policy "had pegged withdrawal to building up Iraqi security forces" (Akam, 2018, p. 227). Moreover, the decisions that British political leaders controlled did not change during the war, and thus there was minimal tolerance for political experimentation.

As with the U.S., however, the British military experienced high tolerance for military experimentation, especially as conditions in Iraq worsened. For example, the training group that prepared soldiers for deployments shifted from teaching conventional tactics in 2003 to an "emphasis on creativity, acceptance of ideas from the bottom up, and rapid response to reactions in the field" by 2008 (Akam, 2018, p. 327). Yet the trajectory of doctrinal experimentation and military effectiveness that the British experienced was different than the U.S. military. The situation in Basra remained relatively stable until 2005, when violence quickly escalated as the Mahdi Army militia gained influence in the city. For the next two years, the situation showed no signs of improvement. Eventually, the British military would experiment by negotiating with the Mahdi Army to help reduce violent attacks in the city. As part of this negotiation, the British military agreed to withdraw troops from the city and provide "overwatch" from the Basra airport (Mumford, 2011). The withdrawal from the city increased the militia presence in Basra, which then led to worsening conditions in the city. Soon after, operations by Iraqi security forces, supported by the U.S. military, would help clear the militia presence from parts of Basra (Ucko and Egnell, 2013). The British, for their part, would instead focus primarily on training Iraqi security forces until the political decision was made to withdraw most British troops from Iraq by 2009.

The British military achieved moderate counterinsurgency effectiveness in Iraq—and arguably even poorer than the U.S. military. Together, these two cases demonstrate that there is some variation in effectiveness even within military types. Nevertheless, compared to the Goliath type, it is clear that what the U.S. and British each achieved in Iraq can broadly be categorized as moderate effectiveness, largely due to their short-term adaptability.

3.7 The Goliath (medium commitment horizon)

3.7.1 French in Indochina, 1946–1954

After World War II ended, the French military deployed 50,000 troops to reacquire French Indochina, the French colonies comprising present-day Vietnam, Cambodia, and Laos. Soon after, an insurgency against French forces began, spearheaded by the Viet Minh, the Vietnamese national independence coalition led by Ho Chi Minh. Initially, French political and military leaders characterized the conflict as a "mop-up of outlaw bandits" (Fall, 1961). The combination of colonial claims and an incorrect belief in the Viet Minh threat therefore meant that the French commitment horizon in the conflict lacked any short-term urgency, as U.S. and British forces possessed in Iraq in 2003. However, neither did the French have a long-term commitment that signaled a willingness to invest the effort to finding a long-term solution to the conflict. As historian Martin Windrow describes,

> ...for much of the war the French were hampered by disunity between the military and civilian authorities, and a lack of any real continuity in strategic planning. Their home governments came and went, none of them able to pursue a coherent policy either for winning the war or for abandoning it, and in the face of soaring costs each one was less committed than the last. Every couple of years a new commander-in-chief arrived and appointed new regional commanders, each inheriting a disappointing situation from his predecessor, and each denied the scale of rein-

> forcements needed to achieve any real transformations in the strategic balance. (Windrow, 2004, p. 128)

French military doctrine at the time asserted the superiority of the defense, the origins of which can be traced back to the interwar period (Kier, 1997). Consequently, by late 1947, when Viet Minh insurgents had secured control over large areas of the countryside, the French military began a quest for a large set-piece battle (Fall, 1961). The French believed that if they could fight the Viet Minh in open battle, they would overwhelm the Viet Minh through firepower from defensive fortified positions. Throughout the remainder of the war, this doctrine would not change. The French thus built the De Lattre Line, a defensive line of over 900 concrete blockhouses, around the Red River Delta in northern Vietnam, where much of the insurgent activity occurred.

As a Goliath type, the French military was characterized by low military autonomy. Political leaders maintained that the French military mission was to "protect grateful colonial subjects from looting, murder, and . . . the threat of Communist slavery" (Windrow, 2004, p. 184). Yet despite this insistence, French political leaders failed to provide any guidance to the military on how to conduct the war (Ambler, 1966). However, the lack of political guidance does not necessarily imply high military autonomy. Rather, French political leaders simply expected the French military to carry out its mission without consideration for the unique aspects of unconventional conflict, including the importance of winning over the civilian population and understanding the interdependencies between political, strategic, and tactical issues. Therefore, the military maintained its scattered positions along the De Lattre Line, assigning over 80,000 troops to statically man the De Lattre Line and wait from fortified positions for the Viet Minh to attack. As a consequence, the French had few remaining troops to carry out offensive operations against insurgent strongholds or develop an effective intelligence network (Boot, 2013).

One reason that French political leaders remained so indecisive about the conflict is that French domestic politics at the time were markedly un-stable. The French Fourth Republic, formed after World War II ended, was faced with the difficult task of rebuilding the country socially and politically. With such pressing domestic concerns to address, the war in Indochina was viewed by French politicians as a distant conflict, and as such was never carefully managed. Consequently, the tolerance for political experimentation in the conflict was essentially non-existent. Thus, both political and military decisions remained mutually stagnant, reinforcing the existing doctrine and preventing any major adaptations in the way the war was conducted.

During the conflict, the French military experienced a mixture of stagnant tactics and sporadic, inchoate doctrinal experimentation that was never implemented across the organization. Overall, tolerance for military experimentation was low. For example, aerial forces would engage in indiscriminate bombing of villages with the intent to kill civilians, naval forces would ineffectively spray enemy gun positions with automatic weapon fire, and ground forces—lacking tactical surprise—would conduct limited "mop-up operations" in villages that often would either be ambushed along the route, or would not encounter any insurgents at all (Fall, 1961). Because so many static troops were dedicated to the De Lattre Line, the French failed to ever find an effective solution to the dilemma of choosing "between a deep penetration for a short time or a shallow penetration for a long time" (*ibid.*, p. 73).

Often, military experimentation that did occur would happen at the company or battalion level, as junior commanders came to recognize that existing conventional tactics were ineffective against a flexible insurgency (Kelly, 1965). Nevertheless, such experimentation remained isolated at these lower levels, and was never aggregated into a broader strategy across the military. However, two experimental programs were attempted with higher-level approval, but they also ultimately failed to impact the existing doctrine. The first program was the creation of 52 "commando" battalions of the Vietnamese National Army, the *Tieu-Doan Kinh-Quan* (TDKQ). The purpose of the TDKQs was to use Vietnamese units who already knew the terrain, culture, and people, so that they could win their hearts and minds and seek out and defeat the insurgents militarily. While the idea seemed promising, the Viet Minh recognized the potential threat that the TDKQs posed and crushed them in their first military encounters. These initial defeats ruined the TDKQ reputation among the local population, thereby preventing the program from building on any success, and eventually leading to its demise (Fall, 1961).

The second experimental program was a French commando unit, the *Groupement de Commandos Mixtes Aéroportés* (GCMA), later retitled as the *Groupment Mixte d'Intervention* (GMI). This unit was created based on the experiences of French partisans fighting against the Germans during World War II (Pottier, 2007). The early GCMA recruits in 1951 were noted for being trouble-makers and were "considered too individualistic by their unit commanders who were glad to get rid of them that way" (Fall, 1961, p. 242). Thus, in the midst of an organization that stifled most creativity and innovation, a unit was create that was composed of like-minded individuals who were far more innovative than their previous commanders could tolerate. The GCMA/GMI would succeed where the rest of the French army failed: members would infiltrate Viet Minh-controlled areas and live among the locals, recruiting partisans and leading groups up to a thousand strong. These groups would then sabotage Viet Minh supply and communication lines and create areas that were free of Viet Minh influence. By 1953, the program had achieved enough success that recruits were among the elite members of the French military, growing in size from 5,000 to 14,000 men. Nevertheless, the program remained unpopular among conventionally minded commanders who often lost airlift capability to the unit, as well as some of their best soldiers (Windrow, 2004). Ultimately, French investment in the program would prove to be too little and too late.

Although the GCMA/GMI achieved much success in a limited time, they were the exception. The conventional French units were far less adept at operating among the local Vietnamese population, lacking the requisite cultural and linguistic skills to establish relationships with civilians. Seemingly minor cultural interactions from a Western perspective could have major consequences: "a European's most natural response to the delightful village children was to hand out sweets, ruffle their hair, and perhaps compliment their parents. Nobody had told the soldiers that to touch an infant's head was to damn it to lifelong bad luck, or that to praise its beauty attracted the vengeance of envious spirits" (Windrow, 2004, p. 186). Moreover, because the French were mostly isolated from civilians except for brief encounters in villages, trust between French and Vietnamese locals rarely developed. Consequently, the Viet Minh steadily grew in strength and influence, so that by 1953 the French military estimated that they controlled only 25 percent of Vietnam (Boot, 2013).

Thus the French situation deteriorated to the point that a decisive military victory over the Viet Minh no longer appeared possible. Instead, when the French prime minister appointed General Henri Navarre to the Indochina command in 1953, Navarre's charge was to break the stalemate and find an "honorable solution" so that the French could negotiate a favorable peace agreement. Navarre therefore pivoted from a strategy that focused on the overstretched De Lattre Line to consolidating forces into a small number of "air-ground bases" placed in strategic locations. Notably, however, the French defensive doctrine did not change. As Navarre later testified, "We were absolutely convinced of our superiority in defensive fortified positions; this was considered in Indochina as dogma, and we were absolutely persuaded that a fortified position could hold out easily at odds of three to four against" (Windrow, 2004, p. 216).

In March 1954, the French would finally get their set-piece battle at a heavily fortified air-ground base in the Battle of Dien Bien Phu. However, the result was disastrous for the French: by that time, the Viet Minh had built up strong conventional capabilities, and they besieged the French fortress. After a two-month battle, the French positions were overrun, and nearly 12,000 soldiers surrendered. It was a decisive defeat for the French military, and the war ended soon after. The 1954 Geneva Accords, far from being an honorable solution for the French, led to a complete withdrawal of French forces and partitioned Vietnam along the 17th parallel, setting the stage for the subsequent U.S. war in Vietnam.

As this case illustrates, the French military in Indochina fits the predictions of a Goliath-type military in a counterinsurgency. The French lacked the urgency to find a quick solution to their doctrine's ineptitude, but they also lacked a long-term commitment, instead falling into the zone of a medium commitment horizon where the model expects doctrinal stagnation and poor effectiveness to occur. Indeed, this is what we observe from the historical record. With low military autonomy and lacking any political experimentation, the few higher-level military experiments that did occur (such as the TDKQ and the GCMA/GMI) never amounted to any substantial improvements in doctrine that could be adopted across the organization. Overall, the French military never adapted to local conditions, and consequently it achieved poor overall counterinsurgency effectiveness. Instead, the one-dimensional French military doctrine stagnated in Indochina, drawing out a conflict that eventually led to a decisive French defeat.

3.8 The Political Sisyphus (medium-high commitment horizon) and the Military Sisyphus (high commitment horizon)

As previously described, the Political Sisyphus and the Military Sisyphus types can be considered variants of each other. These two types have overlapping commitment horizons, broadly characterized as medium-high to high. Both types also tend to have medium military autonomy. Thus, the main distinction between the two types is in their tolerance for experimentation. The Political Sisyphus has a high tolerance for political experimentation and a low tolerance for military experimentation, whereas the Military Sisyphus is the opposite. The next two cases illustrate how one Sisyphean type can transform into another in the same conflict, leading to a change in the relative emphasis that is placed on political versus military experimentation.

3.8.1 French in Algeria, 1954–1962

Just four months after the ceasefire in Indochina, another war broke out in Algeria that would embroil France in eight more years of counterinsurgency operations. This time, however, the conflict was much closer to home. Additionally, the French governed Algeria as an extension of France itself, and therefore Algeria was more closely associated with France than its other colonies. Moreover, about ten percent of the Algerian population consisted of French and other European migrants and their descendants, and these *pieds noirs* strongly supported French rule. Consequently, from the very beginning of the conflict, France had a much higher commitment horizon in Algeria than it did in Indochina. However, as in Indochina, there was no sense of urgency at the beginning of the conflict (Kelly, 1965). Political and military leaders alike initially failed to recognize the growing insurgency as a powerful nationalist movement that could upend the status quo. Therefore, there was little incentive to find quick solutions to the insurgency.

Phase I: The Military Sisyphus, 1954–1958 In the first half of the war, the French military adopted characteristics that matched a Military Sisyphus type. In Algeria, the higher commitment horizon led to a noticeable expansion in French military autonomy compared to Indochina. Political decisions were limited to managing yearly budget appropriations and legislation addressing conscription, which restricted the number of troops deployed to Algeria in the conflict's early period (Kelly, 1965). Other issues controlled by political leaders involved colonial administration, but in this area especially, there was little tolerance for experimenting with new policies. For example, after the insurgency emerged in 1954, the French premier declared that Algeria would remain part of France, and that the government would implement the 1947 Algerian Statute. This statute was originally designed to create local reforms and increase Muslim Algerian participation in the Algerian government, but French and Algerians alike had opposed it in 1947; thus in 1954, Algerian nationalists considered it a particularly weak concession (Heggoy, 1972). Tepid French politicians, tied up by their own domestic challenges, therefore granted moderately high military autonomy but had low tolerance for political experimentation in the Algerian conflict.

Early in the conflict, the *Front de libération nationale* (FLN), the nationalist Algerian organization conducting the insurgency, was unconsolidated and lacked the level of organization the Viet Minh possessed in Indochina. Nevertheless, the French military was initially unprepared for an insurgency, as the first groups of soldiers deployed to Algeria were still trained in conventional approaches. Moreover, military commanders, like their civilian leaders, believed that a police action would suffice to restore order. Early operations focused on establishing fortified locations and conducting mop-up operations, to little effect (Heggoy, 1972). However, by 1955, as more soldiers returned from Indochina, their experiences informed new perspectives about how to fight a counterinsurgency conflict (Paret, 1964). Thus, after an initial period of stagnation, the French military then began adapting its practices toward a more comprehensive counterinsurgency focus. However, often the French thought the lessons they learned were perfectly transferable to the Algerian theater, and so, for example, would refer to the FLN as "les Viets," and believed that the FLN and Viet Minh efforts were both part of a wider Communist effort to establish global domination (Horne, 1977). Therefore, because the early practices applied were not implemented with concern for the local Algerian context, many of the initial efforts did not lead to successful adaptations right away.

Gradual changes made by the French military, however, led to further improvements. These adaptations developed over time, as soldiers discovered new lessons and began to piece together the complex interdependencies that existed in the conflict. A French military school established in Arzew, which was created to provide a 12-day class to deploying officers, illustrates how the military learned to understand the landscape of the Algerian context. Initially, the school emphasized operational practices against insurgents, but the curriculum evolved over the next several years to incorporate Mao's texts on revolutionary warfare, and then eventually to broader non-military issues such as cultural dimensions of Algerian society and Islam (Heggoy, 1972).

Other efforts led to additional doctrinal innovations as well. For example, in 1955 the French created the *Section Administrative Spèciale* (SAS), a civil-military program conducted by specially trained soldiers to provide hearts-and-minds style support to Algerian villagers (Kelly, 1965). In rural Algeria, small groups of insurgents would periodically ambush French patrols and terrorize local villages indiscriminately into providing support. Over 400 SAS detachments were initially created to counter these efforts, and as time went on they achieved some success at winning the support of the rural population (Horne, 1977). Another approach the French military undertook involved relocating thousands of Algerians from rural villages to regroupment camps—barracks-style camps built to better control and protect the population from the insurgents. From 1957 to 1961, over two million civilians would be relocated to these camps (Kelly, 1965).

The most visible event during the first half of the war occurred in Algiers in 1957. Insurgent efforts increased in Algiers through 1956, as the FLN and French engaged in tit-for-tat actions—the French arresting and executing FLN members, the FLN conducting terrorist attacks against local civilians, usually those of European origin—that further enraged each side. By January 1957 the French determined that they needed to reestablish control in Algiers, especially over the Casbah quarter of Algiers, where the FLN were holed up. General Jacques Massu's 10th Parachute Division was granted full police powers and responsibility for Algiers, and began systematically imposing control over the city in an effort to dismantle the FLN organization within it. These efforts included developing a system of *quadrillage*, which "divided the city up into sectors, sub-sectors, blocks and buildings, each bearing a number or letter" (Horne, 1977, p. 198). Within each block, "loyal" Muslim citizens were given responsibility for reporting suspicious activity to the French military. Algerians were required to show identification papers at checkpoints, and soldiers regularly patrolled the city day and night (Heggoy, 1972). Furthermore, the French recruited double agents to infiltrate the FLN, which sowed doubt within the FLN and led to selfdestructive purges within the organization (Boot, 2013).

Most infamously, however, Massu's soldiers engaged in systematic torture of captured Algerians in an expedient effort to root out further terrorist activity. These efforts were sanctioned by Massu and other French leaders based on the belief that they were necessary to prevent further violence (Horne, 1977). Indeed, one of the masterminds of French counterinsurgency doctrine at the time, Lieutenant Colonel Roger Trinquier, explicitly advocated the use of torture for counterinsurgency. Trinquier, who served as a member of the GCMA/GMI in Indochina and also in the Battle of Algiers, would publish *Modern Warfare* in 1961, which helped inform French thinking about counterinsurgency in the Algerian War and influenced subsequent counterinsurgency theories. In his book, Trinquier emphasizes important points about counterinsurgency, such as the need for a military to adapt to new conditions, but he also devotes an entire section to explaining and justifying the use of torture against prisoners during interrogation (Trinquier, 1961).

Ultimately, Massu and his men would manage to eradicate most FLN activity in Algiers. By October 1957, the Battle of Algiers had ended and the FLN was defeated in the city, an "undisputed success" for the French military by most accounts (Kelly, 1965). Outside of Algiers, the French military also achieved greater effectiveness. Oran Province, for example, became "almost impermeable to rebel terrorism" (Kelly, 1965, p. 177). Other parts of the country, however, including the more mountainous region of Kabylia, were still controlled by the FLN, and the insurgents also found safe havens in bordering Tunisia. Equally important, moreover, was the fact that the "pacification" efforts were imposed by the French military without the consent of the Algerian people. Thus in Algiers, for example, a sense of growing solidarity among Algerians arose as they experienced the months-long repressive operations of Massu's men (Heggoy, 1972). The French military's success would soon prove to be limited to the short term.

Phase II: The Political Sisyphus, 1958–1962 Coinciding with the first years of the Algerian War was a slowly unfolding French political crisis that had continued from the Indochina years, as the French Fourth Republic's domestic legitimacy faltered. As a result, in the first four years of the conflict, a civil-military relations imbalance existed that heavily favored the military. The military's failure in Indochina, moreover, had created resentment among soldiers toward their political leaders and the French civilian population (Windrow, 2004). Consequently, some military leaders, along with sympathetic colonial ministers, began to perceive a lack of political support for their efforts, and feared that the Indochina failure would be repeated in Algeria. Ultimately, political crisis would strike France in May 1958, when a group of military leaders drafted plans for a coup d'état in order to install Charles de Gaulle in power, whom they believed would support the military in Algeria. Under the threat of a coup, the French Parliament approved de Gaulle as its new leader, and he commissioned a new constitution for France to replace its weak political institutions, thereby establishing the Fifth Republic.

After returning to power, however, de Gaulle quickly determined that "permanently pacifying Algeria against the wishes of most of its populace would be too costly to contemplate" (Boot, 2013, p. 374). After the military's use of torture in Algiers became public, moreover, French public opinion soured toward the military's efforts in Algeria. Furthermore, de Gaulle recognized the military's excessive political power, and commenced a major purge of military leaders from leadership roles in Algeria in order to remove the activist influences from the military. Within ten months, approximately 1,500 officers in Algeria were transferred or had retired (Horne, 1977). Thus, 1958 marked a distinct phase transition for the military with a change from a Military to a Political Sisyphus type. By that time, moreover, the military had succeeding in reducing insurgent activity throughout Algeria, and de Gaulle began searching for a political solution to the war. Military experimentation declined as the existing efforts continued, and political experimentation increased.

De Gaulle's words and actions indicate a high level of experimenting with different approaches to the conflict. One French politician even referred to de Gaulle in 1960 as "this prince of ambiguity" (Horne, 1977, p. 373). For example, in October 1958, de Gaulle offered the Algerians a *paix de braves*—a peace deal designed to save face for the insurgents (Bernstein, 1993). After that failed to end the conflict, he broadcasted a radio address the following year in which he offered the Algerian people three options for self-determination: remain integrated as part of France, fully secede, or establish an independent association with France (Horne, 1977). Overall between 1957 and March 1960, he publicly changed his position on Algeria multiple times, as historian Alistair Horne documents:

> 1957: "Of course independence will come, but they are too stupid there to know it."
> June 1958: "*Je vous ai compris!*"
> June 1958: "*Vive l'Algérie française!*"
> June 1958: "*L'Afrique est foutue, et l'Algérie avec!*"
> October 1958: Algerian independence? "In 25 years, Delouvrier."
> March 1959: "The French army will never quit this country; and I, General de Gaulle, will never deal with those people from Cairo and Tunis."
> January 1959: "Algeria has chosen peace."
> April 1959: "I am the only person capable of bringing a solution to Algeria."
> May 1959: "a new Algeria bound forever to France . . . "
> August 1959: "Peace is a necessity. This absurd war."
> September 1959: "I deem it necessary that recourse to self-determination be here and now proclaimed."
> January 1960: "*une solution qui soit française*"
> January 1960: "How can you listen to the liars and the conspirators who tell you that in granting free choice to the Algerians, France and de Gaulle want to abandon you, to pull out of Alge-

> ria and hand it over to the rebellion?"
> March 1960: "Independence ... a folly, a monstrosity ... France must not leave. She has the right to be in Algeria. She will remain there ..." (Horne, 1977, p. 378)

As these quotes illustrates, de Gaulle switched between Algerian independence, French control of Algeria, and a variety of options blending these positions at various times during the latter part of the conflict. While these examples are public statements, they may not necessarily be honest representations of de Gaulle's position, but instead may have been tactical statements made to advance his position at different times. However, if de Gaulle did have some stable position throughout the conflict, he never made it clear enough to anyone that a doctrinal consensus could develop around it. Accounts from various public officials are consistent in portraying de Gaulle as either ambiguous in his position, or else as someone who advanced different positions to different people, possibly as a trial-and-error strategy. For example, General Massu believed that de Gaulle had always intended to withdraw; Pierre Mendès-France, one of the primary politicians opposed to de Gaulle, believed that de Gaulle had always preferred *Algérie française*; Army General Andrés Beaufre believed that de Gaulle had created a "formula" for a resolution to the conflict but discovered that it did not work; Minister for Algerian Affairs Louis Joxe believed de Gaulle had no clear position; and Bernard Tricot, one of de Gaulle's advisers, thought that de Gaulle never preferred integration, but also thought that he spent time searching for an optimal position (Johnson, 2000; Horne, 1977, p. 377). Thus, the totality of the historical record suggests that de Gaulle's behavior matches the political decision-maker in the agent-based model: he experimented with a new position, evaluated its payoff, and changed when he found what appeared to be a better position, even if it meant switching back to a previous option. As frustratingly ambiguous as it was politically for many at the time, de Gaulle's process can be explained as a search over a rugged landscape.

The French military, on the other hand, proceeded without any significant changes to its operations before de Gaulle's rise to power. It continued relocating rural villagers to the regroupment camps and built a long line of fortifications along the Tunisian border to limit insurgent cross-border movements. Additionally, the SAS program became more centralized under the French army as it institutionalized what it had learned, further reflecting how the French military had little tolerance for experimenting with new approaches after 1958, instead choosing to consolidate its existing approaches (Heggoy, 1972).

Overall, the French military achieved high counterinsurgency effectiveness in Algeria, but only in the limited military sense that the FLN struggled to operate within the country by 1958. Nevertheless, this success was limited to the short term, as effective solutions to the political and social elements of the conflict were never found. Moreover, the overall effectiveness of many French decisions are particularly questionable in hindsight, as they often encouraged an Algerian identity separate from France, rather than keeping it integrated. The regroupment camps tended to be overcrowded, which spread diseases and led to malnutrition among its inhabitants—Algerians were observed eating grass in the field, and some infants died from the cold. The regroupment policy also had the unintentional effect of disrupting the Algerian tribal structures and encouraging a shared sense of solidarity and hence national identity among the inhabitants (Horne, 1977). SAS officers tended to focus on turning Algerians into French collaborators, showing less concern for Algerian values and instead emphasizing French principles (Paret, 1964). The systematic use of torture in the Battle of Algiers may have helped the French military root out FLN insurgents in the city, but it created strong resentment among Algerians toward the French, increasing the nationalists' resolve. Furthermore, as Horne points out, if the French had lost the Battle of Algiers, they may have agreed to a peace settlement with the FLN and foregone the additional costs of four more years of conflict (Horne, 1977). Thus, the use of torture backfired against the French in multiple ways.

In the end, France lacked the commitment to Algeria that would have been needed to achieve greater effectiveness in the long run. Consequently, the French military held a tight grip on the Algerian population while waiting for an elusive political solution to appear—an approach that the French people (of the metropole) determined to be unsustainable indefinitely. The FLN found sanctuaries in Morocco and Tunisia and maintained a respectable and well-equipped army, while its leaders internationalized the conflict politically, helping to build international opinion against the French (Heggoy, 1972). De Gaulle eventually shifted toward disengagement from the conflict, but not before another civil-military relations crisis occurred: a cabal of French generals organized for a putsch in 1961, and after that failed, several broke away from the army and organized the *Organisation Armée Secrète*, which committed its own terrorist campaign in an attempt to keep the French in Algeria. Eventually, de Gaulle agreed to a cease-fire with the FLN, which was put into effect in March 1962. The final agreement was a total French capitulation, as it abandoned all of its prior bargaining positions, including French control over parts of Algeria, dual nationality for the *pied noirs*, and Algerian association with France. Algeria won its independence from France, the FLN took revenge on Algerians who had supported the French, and over one million European civilians fled to France within one year.

The case of the French military in Algeria illustrates how a Sisyphean type can achieve doctrinal adaptations that increase effectiveness, but overall its achievements tend to fall short of its goals. The French were highly committed to Algeria, but due to various constraints, they lacked the long-term commitment to maintaining *Algérie française* indefinitely. Consequently, the French adapted to the Algerian conflict—first militarily, and then politically—and found some success, but only for a while. Ultimately, without finding better solutions to the complex Algerian situation, the French chosen to abandon their efforts rather than endure without better effectiveness.

3.8.2 U.S. in Afghanistan, 2001–2021

The U.S. war in Afghanistan has followed a similar trajectory as the French counterinsurgency in Algeria, and thus also matches a Sisyphean type that transitioned from a Military to a Political Sisyphus over time. Given the circumstances of the U.S. invasion of Afghanistan, the U.S. has possessed a much higher commitment to the conflict than it ever had in Iraq: Osama bin Laden and his al Qaeda operatives had found political refuge with the Taliban regime in Afghanistan, which allowed them to plan and conduct the September 11, 2001 terrorist attacks. Therefore, U.S. political and military leaders were highly committed to rooting out al Qaeda, replacing the Taliban regime, and ensuring that another major terrorist attack could never be planned inside Afghanistan again. Nevertheless, the U.S. approach always included a focus on an "exit strategy," i.e., how to make the country safe so that the U.S. could withdraw its forces as soon as possible. Yet finding a solution to this problem was never fully resolved: namely, how to ensure Afghanistan did not become a terrorist safe haven again without stationing U.S. combat forces in the country indefinitely.

Because the U.S. military's organizational characteristics in Afghanistan match those discussed in the Iraq case (i.e., as the Protean type), I will only briefly discuss the U.S. in Afghanistan. Nevertheless, the dual cases of the U.S. in Iraq and Afghanistan fit precisely with the typology's expectations: i.e., that a Protean and a Military Sisyphus type both have matching levels of military autonomy and tolerances for political and military experimentation, and only differ in their commitment levels. Moreover, their overall doctrinal effectiveness levels will be approximately equal, as they both tend to achieve moderate effectiveness.

In Afghanistan, the U.S. remained a Military Sisyphus type from the 2001 invasion until approximately 2011. The transition to a Political Sisyphus is harder to pinpoint than the French case in Algeria, as the U.S. transition did not coincide with a dramatic change in political institutions. Nevertheless, a defining event that approximates the point at which a transition occurred is the U.S. military's mission to kill Osama bin Laden in May 2011. Before the bin Laden mission, military decision-making dictated most of the doctrinal approaches to counterinsurgency in Afghanistan. The Afghan insurgency was relatively quiet until 2006, when violence began to spike. U.S. efforts to counter the insurgency then focused primarily on military operations to disrupt and dismantle insurgent groups, which continued through 2009. That year, President Barack Obama decided on an Afghanistan "surge," sending an additional 70,000 U.S. forces to augment the 30,000 already deployed to the country (Peters and Plagakis, 2021). Yet even during this period, there was discussion of a withdrawal timetable. Thus, as with the U.S. in Iraq, even the major political change represented by the Afghan surge occurred against the backdrop of consistent motivations to bring troops home. In 2010, for example, a NATO summit in Lisbon established a plan to hand over full responsibility for Afghan security starting in July 2011, when NATO troops would begin withdrawing.

From 2009 to 2011, U.S. military operations continued at an even higher level than before the "surge," and the military, along with its coalition partners, achieved considerable success at reducing insurgent attacks. Throughout the first ten years of the war, the U.S. military demonstrated considerable experimentation with new approaches as it adapted to the Afghanistan conflict, attempting various efforts focusing on large-scale operations, smaller-scale provincial teams, "clear, hold, and build" approaches, and stationing troops at small outposts throughout the country (see, for example, Barno and Bensahel (2020), Farrell, Osinga, and Russell (2013), and Long (2016) for detailed descriptions). With the eventual success of the U.S. military's counterinsurgency efforts in Afghanistan, capped by the bin Laden mission—which provided a decisive military accomplishment against the world's top terrorist leader—there emerged a natural transition point to focus more on political rather than military solutions. Thus, one month after bin Laden was killed, Obama announced plans to reduce U.S. troop levels to pre-surge numbers by the summer of 2012.

By 2011, political leaders therefore took over most of the U.S. experimentation efforts, attempting different approaches to find improvements for how to manage and support Afghanistan's fledgling government and security forces. There were, however, some political experiments before 2011. For example, in March 2009, Obama declared a new strategy for the conflict by connecting stability in neighboring Pakistan to success in Afghanistan. Then in November 2009, Secretary of State Hillary Clinton tied all subsequent civilian aid to Afghanistan to improvements in Afghan President Hamid Karzai's efforts to fight corruption in his administration (Council on Foreign Relations, 2021). Yet after 2011, political experiments predominated, as what some characterized as "mission creep" expanded. Yet after the first decade of war, "misguided assumptions" and a lack of understanding about the country, its culture, and its political complexity meant that the search for further improvements would take years (Whitlock, 2019b). For example, after a period in the first decade of the war in which U.S. efforts aimed at centralizing Afghan authority, a shift was made to pursue decentralization to the provinces. Additionally, in 2014, the U.S. supported an Afghan power-sharing agreement between fierce political rivals Abdullah Abudullah and Ashraf Ghani after disputed election results. Overall, the U.S. has spent more money in Afghanistan on various nation-building projects than it spent on the Marshall Plan, adjusted for inflation. With these projects, political leaders have periodically shifted their emphasis between goals for Afghanistan between Western democratic governance, training competent security forces, developing free-market liberalism, and support for women's rights (Whitlock, 2019a). Negotiations with the Taliban since 2018 to end the insurgency led to minor gains over multiple rounds of meetings, but mostly stalled. In February 2020, the U.S. signed a precarious deal with the Taliban to end the war within 14 months. Yet even following the deal, uncertainty remained as to whether the agreement would survive.

In April 2021, President Joseph Biden announced that the U.S. would withdraw all remaining forces from Afghanistan by September 11, 2021, even without certainty about Afghanistan's long-term stability and security. At that point, the Taliban controlled at least 15 percent of the country's territory and still conducted attacks against Afghan targets while periodically negotiating with U.S. diplomats (Roggio and Gutowski, 2021). At least as of 2018, two thirds of Afghanistan's budget was still funded by international aid (Ruttig and Bjelica, 2018). Chronic corruption was rampant within the Afghan government. Overall, the U.S. achieved moderate counterinsurgency effectiveness in Afghanistan, first militarily and then politically. As the French decided in Algeria, the U.S. determined that enduring in Afghanistan without seeing further improvements was not worth the costs. It remains to be seen how Afghanistan manages its affairs after U.S. troops finally withdraw, yet the conflict between the Taliban and the Afghan government appears likely to continue (Cohen, 2021).

3.9 The Atlas (highest commitment horizon)

3.9.1 British in Malaya, 1948–1960

British Malaya was a loose affiliation of sultanates on the Malay Peninsula that had been under British control since the late 18th century. In September 1945, after the defeat of Japanese forces in World War II, British troops arrived to reoccupy the region. To simplify administration of the territory, the British decided to form the Malayan Union in 1946. However, the union was heavily opposed by ethnic Malay citizens, who resisted the British plan to grant suffrage and equal rights to Chinese and Indian minorities in the region, as well as by the traditional rulers, whose authority was subordinated by the British governor. Seizing on the simmering unrest, the Malayan Communist Party, over 90 percent of which was comprised of Chinese minorities, began an insurgency on the peninsula in 1948 (Hack, 2009).

Of all the counterinsurgencies evaluated thus far, the British in Malaya can be characterized as having possessed the highest commitment horizon. Beginning in 1948, the British colonial administration reorganized the states as the Federation of Malaya, which reinstated the rulers' previous positions as heads of state. From at least 1948, the British had intended that the

Malayan federation would be part of the British Empire's goal of decolonization by establishing the Commonwealth of Nations, the modern form of which was formally declared at a meeting in London in 1949. Therefore, the British had a very high commitment horizon for the emerging counterinsurgency campaign, as they expected that the region's political and economic association with the United Kingdom would continue indefinitely into the future. Thus, even though Britain was committed to Malaya's independence, it did so because British political leaders recognized that a decolonization strategy was in Britain's long-term interest, and therefore they sought to cultivate a long-term political and economic relationship with an independent Malaya that would be incorporated into the British Commonwealth.

Furthermore, the challenge for the British in Malaya was of a somewhat different order than the other cases previously considered. In Malaya, because the insurgents were primarily ethnic minorities, they lacked significant support from the civilian population throughout the conflict. Moreover, insurgent forces numbered around 8,000 at their peak—far fewer than the other cases. There is no reason to believe these factors alone determined British success from the start. Nevertheless, they likely contributed to British beliefs that they could more easily succeed in a counterinsurgency campaign, and therefore may have helped contribute to a very high commitment horizon.

Despite having a very high commitment level, for the first three years of the conflict, the British military lacked significant political investment and struggled to adapt. Initially, the British troops were sent to protect local tin mines and rubber plantations, which were used by the British colonials for resource extraction and served as targets for the insurgents. As with other militaries of different types, the British experienced the same challenges of transforming from a conventionally-minded military engaged in large-scale operations to one that was more flexible, responsive to local conditions, and that accounted for the various complex interdependencies between political, social, and operational demands. There were some lower-level tactical innovations that occurred, such as the creation of small-scale "Ferret Force" groups that emphasized "native trackers, small patrols, and the inclusion of interpreters and local natives in army operations" (Nagl, 2002, p. 69). For the most part, there was little progress to show for these efforts. Nevertheless, both at the military's strategic level, such as with the establishment of the 1950 Briggs Plan—which institutionalized resettlement villages—and at the political level, British leaders recognized that "we cannot expect any sudden or overnight transformation in the situation. The only possible thing to do is to take one state at a time and get it and keep it permanently dominated" (Nagl, 2002, p. 72). Thus there was no urgency to find a quick solution, but there was emphasis on making incremental improvements and achieving progress over time.

Notably, British leaders advocated no increase in British troop levels even when they faced a stalemate situation by 1951, believing that the solution lay in better training local police officers. Moreover, they recognized the need for effective government institutions early on, and by early 1952, the British government had declared that "the policy of Her Majesty's Government in Great Britain is that Malaya should in due course become a fully self-governing nation" (Nagl, 2002, p. 88). Coinciding with this declaration, Prime Minister Winston Churchill appointed General Sir Gerald Templer as British High Commissioner for Malaya, granting him extensive civil and military autonomy to defeat the insurgency. Therefore, as of 1952, the British military possessed high autonomy in waging the counterinsurgency, matching an Atlas type.

Furthermore, tolerance for political experimentation was moderate. As previously noted, British political leaders sought arrangements with the local Malayan leaders that would satisfy all parties—hence the shift from the unpopular Malayan Union to the more widely accepted Federation of Malaya. These political efforts were then integrated into the military's efforts as a primary goal. For example, a March 1953 Colonial Office Memorandum notes, "We all aim at helping the colonial territories to attain self-government within the British commonwealth. To that end we are seeking as rapidly as possible to build up in each territory the institutions which its circumstances require" (Nagl, 2002, p. 101). This policy matched Malayan demands for self-government and eventual independence (Clutterbuck, 1966). Once Templer was installed as High Commissioner for Malaya, he emphasized both a willingness and an urgency to make changes politically across the different levels of government, starting with the rural and municipal levels. Consequently, municipal elections helped create alliances and coalitions among different political parties, which had the dual effect of meeting British requirements for the region and helping to create a sense of national identity among the multiethnic Malayan population (Nagl, 2002).

Tolerance for military experimentation in the British military was very high, especially after Templer arrived. For example, in 1952, Templer devised the creation of a doctrinal document titled *The Conduct of Anti-Terrorist Operations in Malaya*, which he conceived of as a living document that would be open to "criticisms and improvements" for revised versions. Subsequent editions with updated material were published in 1954 and again in 1958 (Nagl, 2002). As part of these efforts, the British tried different approaches to isolating the insurgents in addition to the resettlement camps, including food denial programs that attempted to starve out the insurgency. One innovation that found particular success was a program that allowed insurgents to earn a bounty for surrendering. Those who surrendered would often later help lead patrols on raids against insurgent base camps, earning additional rewards and assisting the British in further disrupting the insurgents in the process (Nagl, 2002). In order to persuade insurgents to surrender, the Psychological Warfare ("Psywar") section of the British Information Services had to experiment with different approaches. One British officer remarked that "Psywar policy evolved largely as the result of trial and error," and various programs and platforms were used, including leaflets, films, press, and loudspeakers on jeeps and aircraft (Nagl, 2002, p. 93). In summary, as John Nagl notes: "The British army demonstrated a remarkable openness to learning during the years from 1952 to 1957. Bottom-up input was welcomed, from tactical innovations . . . through operational ones . . . District advisers, British army privates, and Surrendered Enemy Personnel were asked for ideas on better ways to accomplish the objectives of the organization" (Nagl, 2002, p. 105).

By 1955, insurgent leaders began making overtures for peace talks. Insurgency participation had dwindled from 8,000 people at its 1951 peak to 3,000 by the end of 1955; attacks had declined from a peak of 500 per month to 65 a month; and civilian casualties had dropped from 80 to 12 per month (Short, 1975). Furthermore, by 1955, a Malayan political coalition had consolidated gains in the legislature, and Chief Minister Tunku Abdul Rahman took over direction of the war. By 1956, most of eastern Malaya was in government control, and a date for Malayan independence and a draft constitution was established (Clutterbuck, 1966). The British government then granted independence to Malaya on August 31, 1957, thus removing the insurgents' primary motivating purpose of national self-determination, and the insurgency collapsed for all but the most committed members. It then continued in small concentrated areas until July 1960, when it was declared officially over.[17] In the end, not only did Malaya achieve independence, but it also produced political reconciliation between the Malay and Chinese ethnic communities (Ucko, 2010).

As an Atlas type, the British in Malaya exhibited a very high commitment horizon, and they adopted characteristics that allowed them to adapt and improve their effectiveness: high military autonomy, medium political experimentation, and very high military experimentation. They focused on implementing changes and making improvements, but they were not urging quick solutions in order to shore up effectiveness. Instead, they recognized the importance of patience, integrating political and military efforts, and persistent pressure on the insurgents. Consequently, despite having a very high commitment horizon, the British managed to achieve higher effectiveness in a lower-duration conflict than less committed types, which matches the model's predictions.

[17]However, a small-scale insurgency would continue along the Malaysian-Thai border until 1989 (Hack, 2009).

3.10 Commitment Horizon Optimization or Inefficient Optimization?

The implication of my model is that the commitment horizon drives military performance. I assume that all political-military systems optimize their organizational features according to the commitment horizon, which explains their overall effectiveness. An alternative explanation for (in)effectiveness is that the organization did not optimize its features given its commitment horizon. To consider whether militaries may sometimes simply be inefficient, thus leading to lower military effectiveness, I evaluated the organizational opportunity cost at each commitment horizon—i.e., which features the organization forgoes in an optimal arrangement—and consider what these inefficient features suggest about overall effectiveness.

Commitment Horizon (δ)		*Mil. Autonomy*		*Pol. Experim. Tolerance*		*Mil. Experim. Tolerance*		*Avg. Fitness*
0.01 – 0.26	Low	0.26 – 0.45	Med	0.0007 – 0.01	Low	0.31 – 1	Med/High	0.73
0.27 – 0.56	Medium	0.30 – 0.57	Med	0.00 – 0.18	Low/Med	0.42 – 1	Med/High	0.75
0.60 – 0.66	Medium High	0.39 – 0.49	Med	0.00 – 0.01	Low	0.66 – 0.97	High	0.79
0.58 – 0.90	High	0.59 – 0.89	High	0.04 – 0.66	Low	0.70 – 0.95	High	0.80
0.57 – 0.99	Highest	0.55 – 0.80	High	0.18 – 0.62	Medium	0.67 – 0.99	High	0.80

Table 3.4: Organizational opportunity costs (next-best characteristics for less-than-optimal organizations)

Table 3.4 lists the organizational characteristics for each commitment horizon type if it chose the next-best arrangement to an optimal one. Two items are noteworthy in this table. First, there is very little variation in each of the organizational features. This implies that if organizations are not actually optimizing by their commitment horizon, then we should see very little variation in their organizational features regardless of their commitment horizon—at most, according to Table 3.4, a moderate increase in military autonomy and military experimentation only at high commitment levels, coinciding with potentially greater political experimentation. However, as the illustrative cases demonstrate, the historical record provides evidence of substantial variation in military autonomy and the tolerance for political and military experimentation. This variation cannot be explained by organizational inefficiency.

Second, the average total fitness is increasing in commitment horizon for suboptimal organizations, rather than U-shaped. This implies that we should see a linear increase in overall effectiveness as commitment horizon increases. Again, however, the illustrative cases demonstrate that this is not what we observe empirically. For example, we would expect that the French had greater effectiveness in Indochina than the U.S. did in Iraq (or likewise that the U.S. had greater effectiveness in Vietnam than it did in Iraq). Clearly, however, this is not correct.

Therefore, given the variation in organizational characteristics that we observe across cases, as well as the U-shaped pattern in effectiveness across commitment horizon levels, we can conclude that organizations do indeed optimize for their commitment horizons, and their effectiveness cannot simply be explained by inefficiencies.

3.11 Conclusion

This paper contributes to our understanding of the determinants of military doctrine and battlefield effectiveness in counterinsurgency operations by providing a novel approach to explaining military adaptation during wartime. I have shown that militaries waging counterinsurgencies can be categorized according to a typology defined by their commitment horizons. Depending on how large the shadow of the future looms, militaries will adopt certain organizational features to optimize their performance in a complex environment. These features—the level of military autonomy and the tolerance for political and military experimentation—directly affect the extent to which learning and doctrinal adaptation will occur, and thus overall performance. As the results of my model indicate, overall counterinsurgency effectiveness follows a U-shaped curve as the commitment horizon increases. Those militaries with a moderate commitment horizon tend to perform most poorly, as they lack both the urgency to adapt and the drive to persist and search for better solutions over time. On the other hand, the militaries that experience the greatest overall effectiveness are those that have the highest commitment level, because the organizational characteristics adopted allow the political and military levels to work in conjunction to find improvements over time.

Two alternative explanations for counterinsurgency effectiveness do not hold up against the results of my model: fixed military culture and strategy replication. First, fixed military culture, although a widespread explanation in the literature on military effectiveness, lacks explanatory strength because it treats militaries as static, and therefore expects that certain militaries will always perform better than others in counterinsurgencies. Yet, as others have also pointed out, the British "school" of counterinsurgency, widely acclaimed by many because of British military's success in Malaya and other earlier insurgencies, was exposed as a myth based on its performance in Iraq, as well as Afghanistan. Similarly, the "American way of war" is not a good explanation for the U.S. military's failure in Vietnam or its early struggles in Iraq. The U.S. military did, in fact, adapt and improve in Iraq and Afghanistan, and it avoided the same outcome that it experienced in Vietnam. Likewise, the French military experienced two consecutive insurgencies but performed quite differently between the two of them. In short, despite its persistence in the literature, fixed military culture is a weak explanation for counterinsurgency effectiveness. Instead, my results demonstrate that military culture—as manifested in the level of military autonomy, and the tolerance for political and military experimentation—can develop through processes within the organization itself in response to a state's commitment horizon.

Strategy replication—what Taber (2002) refers to as the "methods fallacy"—is also a poor explanation for counterinsurgency effectiveness. What works in one conflict will not necessarily work in another. For example, the British in Malaya are are often viewed as the model for counterinsurgency success that should be emulated by other militaries. Indeed, the British did achieve high counterinsurgency effectiveness. However, too often the wrong lessons from the conflict have been drawn. It is not the particular solutions the British found that should have been emulated, but their very high level of commitment and the willingness to search for a solution based upon local conditions that mattered most.

In other words, the complexity and uniqueness of local conditions requires militaries to learn through experience what strategies will work best, and adapt as they identify solutions. This critique has been made before in more qualitative terms. For example, in 1966, around the time 400,000 U.S. troops were deployed to Vietnam, Tilman (1966) wrote: "the Malayan Emergency offers few policy guidelines applicable in the Vietnamese setting. It is fallacious to equate the two, and the sooner this fallacy can be laid to rest the better it will be for policy-maker and critic alike." Yet U.S. military leaders, presumably unaware of these claims or else choosing to ignore them, tried to replicate the British army's success by adopting many of the same strategies, at the cost of substantial American and Vietnamese blood and treasure.

The results in this paper also have important implications for understanding conflict dynamics and outcomes in counterinsurgencies. In particular, the results suggest that only a certain type of organization is likely to perform most effectively at counterinsurgency: namely, the Atlas type, which has the highest commitment level in a conflict. This suggests that political and military leaders evaluating whether to engage in counterinsurgency operations would be well served in carefully examining how committed they are, to include how committed their selectorate is. If policymakers identify less than the utmost level of commitment, they may find that avoiding such a conflict may be better in the long run. Instead, if they do have an Atlas-type commitment horizon, then adopting the organizational features that encourage the military and political leaders to work in a dialectic approach to adaptation as a coordinated organization will help them achieve the greatest overall effectiveness in a counterinsurgency.

3.12 Appendix

To illustrate the model, consider the following simple example. The model contains eight agents: the political leader (Agent 1), the senior military leader (Agent 8), two operational-level military leaders (Agents 6 and 7), and four tactical-level subordinate units (Agents 2, 3, 4, and 5). The hierarchical interaction structure is illustrated in Figure 3.6.

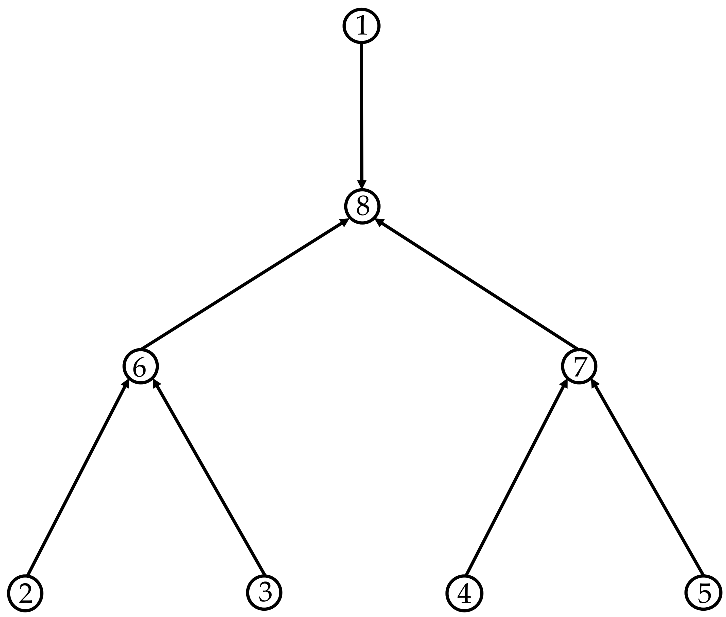


Figure 3.6: Information flow between agents in the example

Initialization. Suppose that $N = 5$, $K = 2$, and MILITARY AUTONOMY $= 4$. For simplicity, suppose also that the probability of a political and a military experiment is one. Additionally, suppose that the initial doctrine is $D = (0, 0, 1, 0, 1)$, and for illustration purposes that each decision D_i is dependent upon the two decisions adjacent to the right in D, where any dependencies to the right of the last decision loop back to the beginning of the string.[18] Therefore, D_1 depends upon D_2 and D_3, and D_4 depends upon D_5 and D_1, and so on. The complete set of possible

[18] In the simulations, the dependencies are randomly assigned.

	(0,0,0)	(0,0,1)	(0,1,0)	(0,1,1)	(1,0,0)	(1,0,1)	(1,1,0)	(1,1,1)
$f_{1,2,3}$	0.06	0.02	0.16	0.18	0.06	0.03	0.15	0.06
$f_{2,3,4}$	0.08	0.09	0.08	0.14	0.09	0.09	0.15	0.09
$f_{3,4,5}$	0.02	0.15	0.19	0.10	0.07	0.01	0.10	0.13
$f_{4,5,1}$	0.06	0.16	0.10	0.16	0.05	0.06	0.07	0.18
$f_{5,1,2}$	0.19	0.19	0.09	0.18	0.14	0.10	0.18	0.01

Table 3.5: Model example fitness levels

fitness levels in the fitness space for this example are listed in Table 3.5, which are randomly assigned. The entire doctrine's fitness level is computed as $F(D) = \sum_i f_i(D)$, the sum of all the individual fitness levels for each of the decisions. Therefore, from Table 3.5, the initial doctrine's fitness level is $F(D) = f_{1,2,3}(0,0,1) + f_{2,3,4}(0,1,0) + f_{3,4,5}(1,0,1) + f_{4,5,1}(0,1,0) + f_{5,1,2}(1,0,0) = 0.02 + 0.08 + 0.01 + 0.10 + 0.14 = 0.35$.

Step 1. Because MILITARY AUTONOMY $= 4$, the political leader and the subordinate units are each assigned a single decision to evaluate. Suppose the political leader, Agent 1, is assigned D_1, and the tactical agents, Agents 2 through 5, are assigned D_2 through D_5, respectively. Therefore, given the existing status quo doctrine, Agent 1 begins with a decision value of $D_1 = 0$. Holding constant the interdependent decisions (i.e., $D_2 = 0$ and $D_3 = 1$), Agent 1 then evaluates whether randomly changing a value in its decision subset (in this case, just a single decision) will improve its subset fitness level. Comparing the two possibilities, we see that $f_{1,2,3}(0,0,1) = 0.02$, whereas $f_{1,2,3}(1,0,1) = 0.03$. Consequently, Agent 1 will change the political requirements in the next iteration so that $D_1 = 1$.

Step 2. Tactical-level agents 2 through 5 evaluate their decision in the same manner as Agent 1. However, whereas the political leader's evaluation is a doctrinal requirement (i.e., it is implemented in the next iteration with certainty), tactical-level agents' evaluations are only recommendations to their direct superior commander. Agent 2 finds that $f_{2,3,4}(0,1,0) < f_{2,3,4}(1,1,0)$, and therefore recommends $D_2 = 1$. Agent 3 finds that $f_{3,4,5}(1,0,1) < f_{3,4,5}(0,0,1)$, and recommends $D_3 = 0$. Agent 4, finding that $f_{4,5,1}(0,1,0) > f_{4,5,1}(1,1,0)$, recommends no change, $D_4 = 0$, and Agent 5 finds that $f_{5,1,2}(1,0,0) < f_{5,1,2}(0,0,0)$, and therefore recommends $D_5 = 0$.

Step 3. Operational-level Agent 6 considers Agent 2 and 3's recommendations and evaluates which composite decision of (D_2, D_3) to recommend. Therefore, Agent 6 evaluates the status quo fitness level of $f_{2,3,4}(0,1,0) + f_{3,4,5}(1,0,1) = 0.08 + 0.01 = 0.09$, and also fitness levels of Agent 2's recommendation for D_2 and the D_3 status quo $f_{2,3,4}(1,1,0) + f_{3,4,5}(1,0,1) = 0.15 + 0.01 = 0.16$, the D_2 status quo[19] and Agent 3's recommendation $f_{2,3,4}(0,0,0) + f_{3,4,5}(0,0,1) = 0.08 + 0.15 = 0.23$, and both Agent 2 and 3's recommendations for D_2 and D_3 together, $f_{2,3,4}(1,0,0) + f_{3,4,5}(0,0,1) = 0.09 + 0.15 = 0.24$. In this case, Agent 6 decides to recommend that $(D_2, D_3) = (1,0)$. Agent 7 does the same comparisons for Agent 4 and 5's recommendations and decides to recommend that $(D_4, D_5) = (0,0)$.

Step 4. The senior commander, Agent 8, then evaluates Agent 6 and 7's recommendations to decide how to implement the military requirements in the doctrine at the strategic level. Agent 8 compares the fitness for the status quo $f_{2,3,4,5}(0,1,0,1)$, the fitness of Agent 6's recommendation for (D_2, D_3) plus the status quo for (D_4, D_5), i.e., $f_{2,3,4,5}(1,0,0,1)$; the fitness of the status quo for (D_2, D_3) plus Agent 7's recommendation for (D_4, D_5), i.e., $f_{2,3,4,5}(0,1,0,0)$; and the fitness of Agent 6 and 7's combined recommendations, i.e., $f_{2,3,4,5}(1,0,0,0)$. The result of these comparisons is that Agent 8 decides to implement military operational requirements $(D_2, D_3, D_4, D_5) = (1,0,0,1)$. Thus the senior commander implements Agent 6's recommendations for (D_2, D_3) but *rejects* Agent 7's recommendations, preferring to retain the status quo values for (D_4, D_5) instead.

Step 5. Combining the decisions, the doctrine changes to $D = (D_1, D_2, D_3, D_4, D_5) = (1,1,0,0,1)$ with $F(D) = 0.70$, and the process repeats from Step 1 until the final iteration. An equilibrium will occur when no agent prefers to make changes.

[19]Note that in this case, although the status quo value for D_2 is unchanged, its fitness level changes due to the interdependent change in D_3.

CHAPTER 4

Projecting Credibility: Alliance Commitments and Foreign Policy Pursuits

> Opportunism and flexibility . . . are military rather than civic virtues.
>
> —Sun Tzu, *The Art of War* (Sun Tzu, 1983)

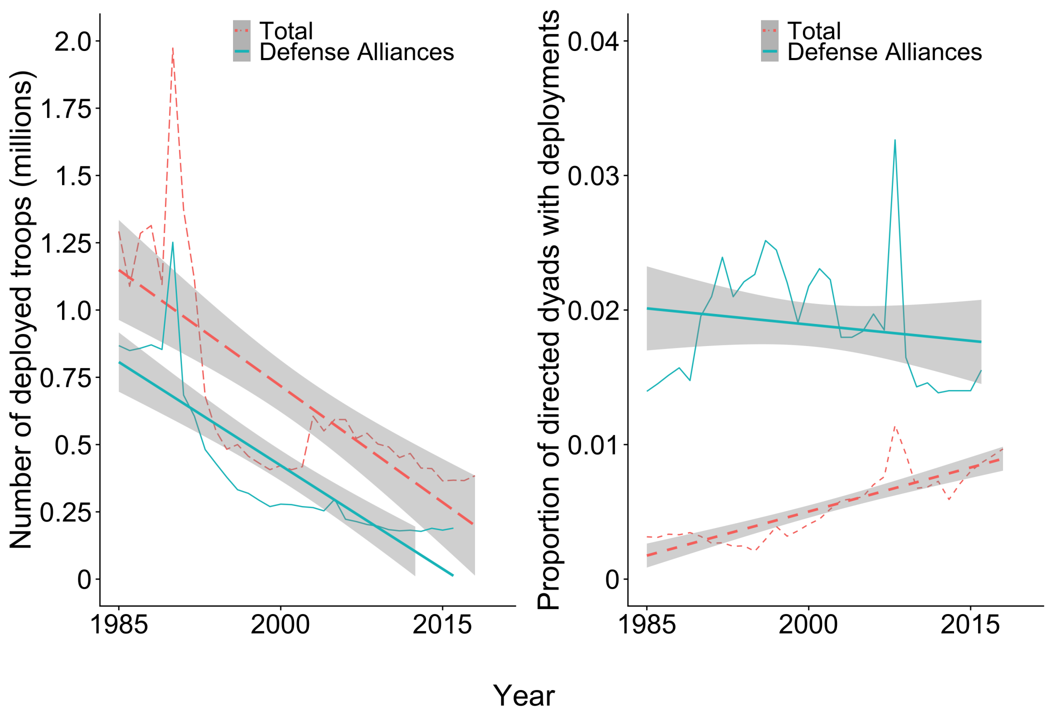
Force projection and military troop deployments are a common method states use to pursue their foreign policy interests, respond to international crises, and signal commitments to their allies. In order to demonstrate that an alliance commitment is credible, states must issue costly signals—i.e., information that communicates that a state will use force if a "red line" is crossed, and which is costly enough that only a credible actor would issue the signal. Positioning high levels of lethal force in proximity to a region of interest is one common way to issue a costly signal to deter or compel a threatening actor. Historically, long-standing troops or extended troop buildups have often been necessary to signal a deterrent or compellent threat's credibility against an adversary. For example, during the latter part of the Cold War, the U.S. stationed roughly 250,000 troops in West Germany alone, which served to signal U.S. commitment to defending NATO countries in order to deter a Soviet invasion of western Europe.[1]

However, since the end of the Cold War, states have deployed fewer troops overseas, but to more countries. Figure 4.1 illustrates these relationships.[2] The left panel plots the total annual number of troops deployed both overall and within defense alliances, and the right panel plots the proportion of dyads with deployments in the same fashion. As Figure 4.1 shows, there is a declining trend in troop levels and deployments to allies, coinciding with an increasing trend in deployments overall. What explains this phenomenon? A common explanation is that with the end of the Cold War and the rise of transnational terrorism, the threat of interstate war has declined whereas the threat of conflict from non-state actors has increased, changing the way states pursue their foreign policy interests. This may partially explain states' force projection behavior, but it is not entirely satisfying. Despite the declining threat of interstate war, alliances still remain an important feature of the international system, as states still have a long-term concern about potential hostility arising from other states. A more compelling explanation for the trends in force projection behavior must be able to take into account not only the changes in the security environment, but also the continued importance of alliances as a way to deter militarized disputes. Furthermore, another important puzzle arises from these dynamics: given the changes in the way states pursue their foreign policy interests, can they continue to credibly signal their commitments if they deploy fewer troops to their allies, and if so, how?

In this paper, I argue that states have altered the way they signal their de-

[1]My focus in this paper is on defense alliances, rather than those that are primarily offensive or serve as non-aggression pacts, because costly public signals such as troop deployments are much more important in demonstrating credibility for defense alliances. In contrast, deployments are not likely to be used in purely non-aggression pacts, where the commitment is to *not* use force, or strictly offensive alliances, which tend to be short-lived due to an imminent conflict and where secrecy is often preferred. For example, the most recent empirical example of an offensive alliance was in 1956, when France, Israel, and the U.K. signed a secret agreement shortly before initiating the Suez Crisis (Leeds, Ritter, Mitchell et al., 2002).

[2]These plots exclude deployments for United Nations and Multinational Forces and Observers peacekeeping missions.



Note: Linear trends of smoothed conditional means and standard error bounds are included. The dashed (red) line is the total level, and the solid (blue-green) line is the level within defense alliances.

Figure 4.1: Annual Deployment Trends, 1985–2018

fensive alliance commitments and pursue foreign interests as a consequence of technological advances that have increased force projection capabilities. Specifically, one of the primary mechanisms driving troop deployment levels is the speed and amount of force with which a state can respond to an international crisis. Greater response speed and volume can serve as a substitute for large quantities of standing, foreign-based troops. Increased response capability is the direct result of technological improvements in military logistics, which affects military capabilities such as strategic airlift and aircraft carriers. This mechanism of mobility capacity therefore has a direct effect on the credibility of signals sent by states within defensive alliance agreements and alters states' behavior in pursuit of foreign policy interests.

Alliances use a mixture of tying-hands and sunk cost signals, where tying-hands signals are sometimes, but not always, costly enough by themselves to credibly signal commitment (Fearon, 1997). Tying-hands signals—e.g., committing publicly to defense treaties, making public statements and speeches—generate audience costs, i.e., the costs that would be incurred *ex post* by a public audience punishing an actor that backed out of a commitment. These costs can occur through a loss in international reputation or being voted out of office at the next election, for example. This is in contrast to sinking costs—e.g., with troop deployments and overseas bases—which are paid *ex ante* regardless of whether they are ever used, making them inherently more costly. Fearon notes, however, that an empirical puzzle exists based on this theory: sunk costs convey full commitment, but we still observe only partial commitment in some cases—i.e., alliances lacking significant sunk cost signaling. Therefore, either audience costs are sufficient to signal credibility in these cases, or else there is something else going on that allows partial commitment to be sustained in equilibrium.[3]

If audience costs alone were sufficient to signal credibility, we would expect little change over time in states' capabilities to respond to international crises. If this were the case, states could simply respond to security crises with a slow build-up of their existing capabilities in order to forestall the *ex post* audience costs if they were to fail to respond, much as coalition forces took several months to build up forces after Iraqi forces invaded Kuwait in 1990 as a signal of their commitment to defend Saudi Arabia and compel Iraqi forces to withdraw from Kuwait. However, states do invest in their ability to respond quickly to crises. Therefore, I argue that it is the development of rapid mobility capabilities which underlies the partial commitment cases, and it is these capabilities that are missing from the existing explanation for alliance commitment signaling.

The deployment trends in Figure 4.1 can therefore be explained by changes in force projection capacity. As states' mobility and force projection

[3]Alternatively, it is possible that the system is not in equilibrium when there is partial commitment. However, I follow Fearon (1997) and assume that partial commitment is possible in equilibrium.

capacity increases, the need to deploy large quantities of troops in order to signal credible commitment goes down, leading to fewer deployed troops among existing alliance partnerships. Additionally, fewer deployments to allies create opportunities both to respond to emerging crises and to invest in additional non-binding security relationships—e.g., paying partners for geographic access, or providing training and short-term security rather than promising long-term protection. Moreover, the perceived threat environment in the post-Cold War order has led to a more diffuse constellation of actors than the bipolar divisions of the Cold War. Consequently, states have an incentive to diversify their partnerships to respond to these threats. Many of these partners also prefer more limited security ties rather than strong defense pacts (Ciorciari, 2010). Therefore, as force projection capacity increases, the tendency to increase the number of "noncommittal" deployments—i.e., deployments in which tying-hands signals of commitment are less likely to occur—goes up, which allows states to send and withdraw troops to security partners as a situation demands, rather than committing to the partners' long-term defense. This also help explain the puzzle of why, in recent years, there has been a growing number of limited bilateral defense agreements accompanied by nearly no new mutual defense alliance agreements (Kinne, 2018).

Furthermore, increased force projection capacity has important effects on states' tendencies toward militarized disputes. Increased force projection capacity helps deter militarized disputes against allied states. The decrease in foreign troops on allies' territories does not make allies more vulnerable to disputes, as the force projecting states' promise to quickly respond to disputes is perceived as credible by potential aggressors. At the same time, however, as states deploy to more *non-allied* locations globally, they are more likely to use military capabilities for coercive diplomacy. This increase a state's chances of initiating a militarized interstate dispute (MID). Consequently, increased force projection capacity is a double-edged sword: it helps deter disputes against allies, but it increases the likelihood a state will initiate disputes elsewhere.

Thus, in sum, the consequences of increased force projection capacity are: 1) a decreased level of troops deployed to allies and consequently a greater reliance on tying-hands signals of credible commitments in alliances; 2) an increased propensity to deploy troops for crisis responses and limited security partnerships without committing credibly to these partners' long-term defense through costly tying-hands signaling; and 3) a lower likelihood that allies will be the targets of disputes, but a higher likelihood that force projecting states will initiate militarized disputes.

4.1 Alliances and force projection

The existing literature on alliances has little to say about the impact that technological advances in force projection capacity and mobility speed have on states' alliance commitments and international deployments. Among the research that deals with alliance commitments, theoretical explanations emphasize the importance of costly signals for alliance credibility (Fearon, 1997; Morrow, 2000), as well as some of the theoretical implications of different types of signals in international crises (Slantchev, 2005; Tarar, 2013; Tarar and Leventoğlu, 2012). On the other hand, the existing literature on force projection does not say much about alliances either. Previous studies mostly attempt to explain the effects that geographic distance has on power projection through the concept of the loss-of-strength gradient (Blechman and Kaplan, 1978; Boulding, 1962; Markowitz and Fariss, 2013; Mearsheimer, 2001; Webb, 2007), whereby states become less powerful as they travel farther from their borders. The two strands of literature separately addressing alliances and force projection share a common framework for understanding states' military power. However, missing from the literature is a direct connection between the capacity for states to project force and the effect that this capacity has on states' alliance commitments.

The idea that alliances serve as a general deterrent threat can be linked to Schelling's argument that the threat of military force can coerce an actor to change its behavior based on the anticipation of such violence; therefore, the power to hurt is most useful when it is held in reserve, and an effective alliance should reduce the likelihood that an allied actor will be the target of aggressive behavior from non-allied actors (Schelling, 1966). Huth (1988) distinguishes between different types of deterrence in an alliance according to whether they are of a general form, or whether they are more immediate in nature. General deterrence involves the existence an ongoing adversarial relationship between states where the threat of war is not immediate, which encompasses the purpose of alliance commitments. In contrast, immediate deterrence arises when a potential attacker actively considers the use of force, and the defender state responds in a direct manner by threatening retaliation in order to prevent the attacker from carrying out its plan. This latter form of deterrence is thus tied to MIDs and international crises.

In order for an alliance to be effective, the promise by one actor to come to another's aid if attacked must be deemed credible (Morrow, 2000). Consequently, signaling such credibility becomes essential for creating an effective alliance. Costly signals are the primary mechanism theorized for establishing credible commitments in alliances, as opposed to "cheap talk," which is a low-cost signal that therefore lacks credibility. Fearon (1997) argues that leaders can generate costly signals through two possible channels. First, leaders can tie their own hands through publicly observable commitments, which generates audience costs that the leaders will suffer *ex post* if they fail to follow through with their commitments. Alternatively, leaders can sink costs in particular ways in order to convey credible resolve *ex ante*, since a less committed actor would not pay such costs.

Following Fearon, subsequent research has explored the theoretical implications of different signals in the context of international crises. For example, military mobilization can be viewed as both a tying-hands and a sunk cost signal in situations where a crisis occurs and faces possible escalation, since it increases the probability of winning should a conflict begin and it requires costs that are paid regardless of the outcome (Slantchev, 2005). Similarly, power shifts caused by mobilization may result in wars linked to commitment problems (Tarar, 2013). On the other hand, audience costs can reduce the likelihood of war, even when the generated audience costs are not particularly high (Tarar and Leventoğlu, 2012). Notably, although these approaches have considered the impact of costly signaling in immediate crisis situations, less attention has been paid to understanding costly signaling in grand strategy and general deterrence.

Empirical studies have also uncovered various effects of alliance credibility that can be connected to the costly signals argument. For example, defensive alliances reduce the likelihood that allies will be targets of MIDs, whereas offensive or nonintervention alliances increase the likelihood that a challenger will initiate a dispute (Leeds, 2003). Similarly, defense pacts in which both high capability and high credibility exist lead to a decreased likelihood that one of the members in the agreement will be the target of a militarized dispute. Moreover, defense pacts that are designed to include more peacetime military coordination also lead to a lower likelihood of disputes targeted at member states (Johnson, Leeds, and Wu, 2015).

Other studies have also explored the issue of alliance politics from the perspective of partner states, i.e., states that host foreign troops on their soil. At the local level, hosting foreign troops can sometimes create frictions between the populace and the deployed troops, especially when the threat against the host country is less salient. For example, Japan has long dealt with local tensions arising from U.S. troops on Okinawa, where over 20,000 U.S. troops are stationed (Denyer and Kashiwagi, 2018). These local tensions can develop into domestic pressures to reduce the size of the foreign footprint. Due to these domestic political preferences, as well as a desire to maintain policy autonomy and avoid undue dependency, some states may therefore choose to limit the extent to which they provide access to allies and security partners (Ciorciari, 2010).

Distinct from the alliance literature, force projection is considered an important element of both immediate and general deterrence: the more a state can project its power outside its own borders, the more threats it can deter from other actors (Blechman and Kaplan, 1978). Existing research on force projection has primarily focused on understanding the effect of geographic distance on power projection. Scholars generally assume a loss-of-strength gradient in force projection (Boulding, 1962), in which power weakens with geographic distance, and especially so across large bodies of water (Mearsheimer, 2001). The extent to which the loss-of-strength gradient matters as force projection ability increases, however, is disputable. Markowitz and Fariss (2013), for example, show that as the cost of projecting force declines through improvements in technology and gains in economic power, states will project power at greater distances and at more frequent rates. However, Webb (2007) has argued that any advantages that a state has in projecting power are only temporary in nature, and therefore the loss-of-strength gradient will continue to matter for deterring and responding to threats as technology evolves.

Among the research that considers troop deployments explicitly, existing studies tend to evaluate the consequences of troop deployments, rather than their causes. Furthermore, many of these studies limit their focus exclusively to U.S. troop deployments, which raises concerns of sample bias. For example, recent studies have explored how troop deployments affect host states' economic development (Kane, 2012), foreign policy (Machain, 2013), defense spending (Allen, VanDusky-Allen, and Flynn, 2016), human rights violations (Bell, Clay, and Machain, 2017), financial and political stability (Aklin and Kern, 2019; Brathwaite and Kucik, 2018), foreign direct investment (Biglaiser and DeRouen, 2007), trade (Biglaiser and DeRouen, 2009), and public attitudes about foreign troops (Flynn, Machain, and Stoyan, 2019). While the consequences of such deployments are certainly important to study, evaluating the determinants of troop deployments can extend this literature by providing a more comprehensive understanding of the force projection causal pathway.

The strands of literature separately addressing alliances and force projection share a common framework for understanding states' military power. However, these research agendas have tended to operate in parallel to each other. Consequently, missing from the current literature is a direct connection between the capacity for states to project force and the effect that this capacity has on states' alliance commitments. Such an approach can shed light on how alliance dynamics have shifted over time, and explain the consequences for the way states respond to international crises and engage in other foreign policy pursuits.

4.2 Consequences of increased force projection capacity

Logistics

A military can only operate for a sustained period if it is properly supplied. Logistics have thus been a critical underlying factor in the success or failure of many military campaigns throughout history (Van Creveld, 2004). Therefore, improved logistical capabilities can significantly increase a military's ability to operate successfully. Although force projection capabilities can come in many different forms, in this paper I focus specifically on the effect of airlift and aircraft carrier capabilities. These elements of "dynamic force employment" (United States Department of Defense, 2018) have freed up many previous logistical constraints and allow greater flexibility in responding to crises. Airlift capabilities are an important feature in moving equipment and troops in a rapid manner, when minimizing crisis response time is desired. In a similar way, aircraft carriers provide blue-water navies with the capability to project force to nearly anywhere on the globe on short notice. Recently, for example, the U.S. deployed a carrier strike group to the Persian Gulf to project force in the region in response to tensions with Iran (The Economist, 2019).

The capacity to project force through airlift capabilities has shifted dramatically over the last several decades. For example, the Lockheed C-5 Galaxy currently used by the U.S. can fly at three and a half times the airspeed of the WWII-era Douglas C-47 Skytrain. Moreover, since the C-5 is capable of being refueled mid-air, its range is effectively unlimited, and is constrained only by the crew's physiological limits to continue flying without rest. At its cruise speed, the C-47 could fly for 10 hours to a range of 1,600 miles, roughly the distance between Los Angeles and Memphis. In the same amount of time, the C-5 can travel 5,710 miles, or the approximate distance from Los Angeles to Berlin—while also carrying 34 times as much cargo. The substantially increased capacity to transport supplies and troops, at greater speed and distance, illustrates how decreased logistical constraints can grant states a greater flexibility to respond to crises.

Similarly, naval capabilities among the major powers have risen dramatically over time, which have also resulted in increased mobilization capacity. As naval propulsion technology has evolved, ships have shifted from being coal-powered to oil-powered, which has decreased the overall need for bases and increased traveling efficiency (Harkavy, 1999). Moreover, the advent of nuclear-powered aircraft carriers has created a level of flexibility previously considered impossible to achieve: such carriers are able to operate without refueling for years, significantly freeing up logistical constraints. Nuclear-powered aircraft carriers can therefore respond more quickly to a crisis, remain in a location longer, and also have more space to carry fuel for aircraft, while also having substantially larger capacity to carry aircraft and military personnel (Spencer and Spring, 2007). Given that aircraft carriers also operate in international waters, they can travel to particular regions and loiter in a chosen location with relative freedom of movement, thereby reducing the constraints imposed by territorial sovereignty.

The evolution of mobility capabilities has had a profound impact on military logistical requirements, providing more capable states the ability to respond militarily to international crises with greater flexibility than ever before. For example, in 2015 Russia deployed 4,000 troops to Syria after the Syrian government requested its aid against rebel groups in the Syrian civil war. Similarly, a large coalition of states each deployed hundreds of troops to Iraq to fight the Islamic State, including Australia, Canada, Denmark, Finland, France, Germany, Hungary, Italy, the Netherlands, New Zealand, Norway, Spain, the U.K. and the U.S. Whereas troop buildups in previous decades might take months or even years to accomplish, these latter deployments typically occurred within days or weeks of initial political discussions about possible military action.

Alliance commitments

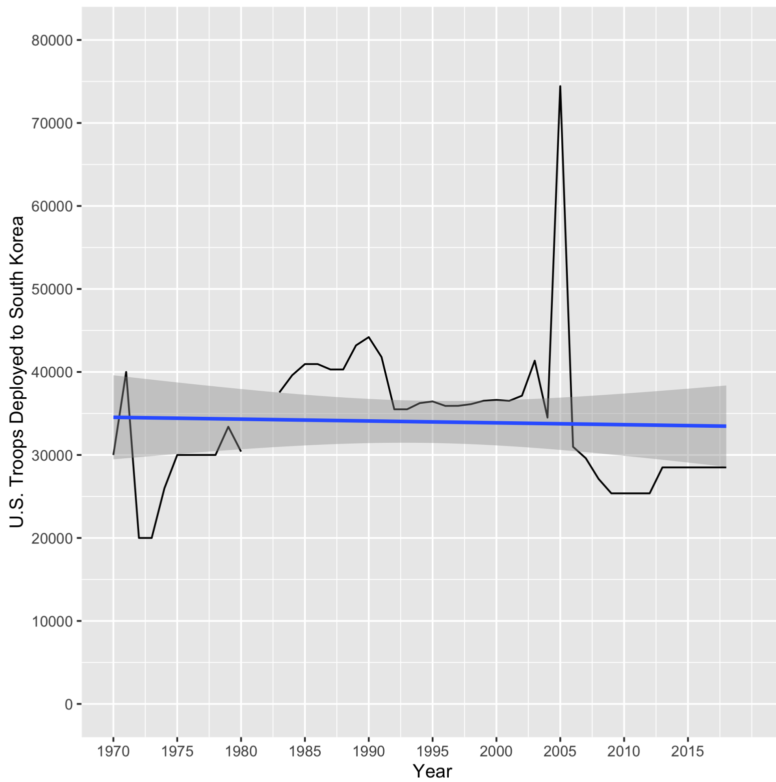
One of the primary ways in which states signal their commitment in an alliance is by deploying troops into the allied country. The presence of military forces on foreign soil is a high-visibility signal of the state's commitment to uphold its promise to protect the host country from military threats. For example, the Mutual Defense Treaty Between the United States and Republic of Korea, which entered into force in 1954 following the Korean War armistice, paved the way for the continual presence of U.S. military forces in South Korea. Yet not all defense agreements are upheld by the presence of foreign troops, and therefore audience costs may provide an alternative credible signal of resolve (Fearon, 1997). However, it is unclear when and why audience costs can be sufficient in some alliances but not others. All else equal, one would expect that a change in troop deployment levels should be directly related to a change in alliance commitments, since it indicates a shift in arrangements between states. Yet this understanding of costly signaling in incomplete.

When signaling commitment in an alliance, part of the agreement's credibility hinges on the speed at which members can respond to a potential threat in an effective manner when faced with an immediate deterrence situation. Immediate deterrence is most successful when a potential attacker determines that the costs of an attack are high and the probability of successful attack is low. Therefore, successful deterrence in a militarized dispute hinges in large part on the balance of forces between sides involved in a dispute (Huth, 1988). Moreover, military strategies generally tend to favor seizing the initiative through rapid and decisive use of force (Betts, 1991; Snyder, 1984; Van Evera, 1984). Consequently, one of the primary reasons why troops are stationed overseas is because maintaining constant force presence allows a rapid response to a crisis. Consider again the presence of U.S. troops in South Korea: since Seoul is only 35 miles from the demilitarized zone (DMZ) dividing the Korean peninsula between North and

South, a North Korean surprise invasion across the DMZ could occur so swiftly that, without a constant presence of U.S. troops, South Korea would likely suffer significantly greater costs. However, by stationing U.S. troops in South Korea, the U.S. signals its resolve to swiftly defend South Korea from the threat of North Korean attack, which has helped deter a major conflict between the two countries. The presence of troops on the Korean peninsula assures a military balance that would decrease the chances of a successful North Korean surprise offensive without very high costs. To illustrate this relationship, see Figure 4.2, which plots the number of troops the U.S. has deployed to South Korea from 1970 to 2018. This figure shows that deployment levels have remained very consistent with only a slight decline over time, which fits with the argument I present here.[4]

From a logistical standpoint, improved mobility capacity means fewer foreign bases are needed to store supplies, including fuel, food, weapons, and other military necessities. Therefore, as mobility response increases, we would expect to see fewer overseas bases and lower overall troop deployments. However, although the supply chain becomes more efficient, states may still find it in their interest for certain preexisting foreign bases to remain as logistical hubs from which to provide supplies to regional countries allied without foreign bases. For example, the U.S. Department of Defense has developed a basing strategy in recent years to minimize costs while maximizing rapid response capabilities by designating locations as either a Main Operating Base (MOB), a Forward Operating Site (FOS), or a Cooperative Security Location (CSL). The MOB acts as the logistical hub with permanently stationed troops and substantial infrastructure, and includes bases such as Ramstein in Germany, Kadena in Japan, Camp Humphreys

[4]Although data is missing for 1981–1982, the trend is still consistent over time. Additionally, the spike in troop levels in 2005 is due to a temporary increase in the number of U.S. Navy forces deployed to South Korea. *The Military Balance* (International Institute for Strategic Studies, 1970-2018), the source for this data, does not typically specify the reason for deployments, so it is unclear why this spike occurred, and most likely is a feature of when the numbers were recorded that year. For example, is possible that the 2005 spike was due to the large humanitarian assistance, including U.S. naval forces, sent in response to the tsunami that struck southeast Asia in December 2004.



Note: Linear trend of smoothed conditional means and standard error bounds are included.

Figure 4.2: Annual U.S. Deployment to South Korea, 1970–2018

in South Korea, and Al Udeid in Qatar. These bases are typically located in countries where a defense alliance agreement exists and provide support for an entire continent or geographical region. Contrasted with a MOB is the FOS, which is an expandable facility with limited military presence and possibly prepositioned equipment, such as the Sembawang port facility in Singapore and Soto Cano in Honduras. Finally, a CSL is a facility with little or no permanent presence, but which provides contingency access to an area and serves as a location for "security cooperation activities"—in other words, an area operated by an allied country but which provides access to U.S. forces in the event of an international crisis (Harkavy, 1999).

When a threat is particularly high and enduring, permanent troop deployments may still be the most effective form of deterrence for an alliance. Yet there are many cases where the threat of an attack is less salient, in which case the constant presence of troops becomes less necessary. Given the development of logistical capabilities over time and the concurrent greater speed on mobilizing in response to a crisis, states that might otherwise have used deployments to signal their credibility in an alliance can instead use rapid mobilization "on demand" as crises occur. This leads to my first hypothesis:

> *H1: Force projection capacity serves as a substitute for deployments when signaling credibility in alliances. As force projection capacity increases, deployment levels in alliances will tend to decrease.*

4.2.1 Foreign policy pursuits and militarized disputes

As states pursue security, they can face an arms-versus-allies—or equivalently, internal-versus-external balancing—tradeoff (Morrow, 1993). This tradeoff depends on the extent to which states prefer autonomy to relying upon other states for security. As states increase the capacity to project power beyond their borders, they gain increased autonomy to pursue their own interests and respond more flexibly to crises. Additionally, the reduction of forces committed to allies results in additional forces available for other deployments, if and when states perceive sufficient threats from other regions. Furthermore, as states increase the capacity to project power beyond their borders, they will more frequently project power (Markowitz and Fariss, 2013). Because states have increased autonomy with these capabilities, they will tend to place greater emphasis on limiting their commitments for new partnerships, rather than expanding into new alliance commitments. However, the increased autonomy accrued with force projection capabilities also allows states to complement existing alliance partnerships through the shift in signaling strategies as previously described. This highlights an important caveat to the theory of the internal-versus-external balancing tradeoff: in terms of force projection capabilities, increased autonomy can serve to complement existing alliances, rather than substitute for them.

In addition to increased autonomy with increased force projection capabilities, changes in the perceived threat environment over the last three decades have also changed states' incentives in the way they build security relationships. Power projecting states have an incentive to establish relationships that enable them to respond to a diffuse array of threats without abandoning their allies. While I do not directly evaluate the types of security partnerships that states establish, recent studies have shown that very few new post-Cold War alliances have been forged, whereas there has been a large increase in new security cooperation through bilateral, limited security partnerships (Kinne, 2018). Increased force projection capacity can help explain this phenomenon, as many developing countries have demonstrated a preference for more limited, flexible security ties with major powers (Ciorciari, 2010). Furthermore, many smaller states may have an incentive to partner with larger states if it is tied to financial, economic, or security gains.

At the same time, more limited security partnerships are also in the interest of major powers, who often have many other allies to which they have committed themselves for defense, or who wish to address security issues where they would prefer to avoid long-term commitments. Therefore, in order to avoid the prospect of over-promising on commitments, states with higher force projection capacity and their security partners are choosing to establish modest, non-committal ties rather than strong defense pacts. For example, France, Russia, the U.K. and the U.S. have each recently expanded their foreign military presence in several African countries by deploying troops for training programs, counter-extremist operations, and antiterrorism cooperation (Blaise, Schmitt, and Gall, 2019; O'Mahony, 2018; Schmitt, 2019; Taylor, 2017). Similarly, China recently constructed its first overseas military base in Djibouti, which houses several hundred troops and serves as a strategic naval port for the Horn of Africa (Jacobs and Perlez, 2017). The partner states in these agreements, for their part, have a number of reasons to provide access and accept foreign troop deployments, such as for financial gain, or to secure economic or military aid to deal with local threats. This leads to my second hypothesis:

> *H2: As force projection capacity increases, the likelihood of deploying troops to new locations increases.*

Finally, increased force projection capacity has important effects on states' tendencies toward militarized disputes. If force projection capacity serves as a substitute for alliance deployments as hypothesized in *H1*, then states with greater force projection capabilities will still be able to deter threats against their allies. In other words, the decrease in foreign troops on allies' territories does not make allies more vulnerable to disputes, as the force projecting states' promise to quickly respond to disputes is perceived as credible by potential aggressors. Consequently, I expect that increased force projection capacity helps deter militarized disputes against allied states. At the same time, however, if states deploy to more *non-allied* locations globally as hypothesized in *H2*, then they may be more likely to use military capabilities for coercive diplomacy. Consequently, I expect that increased force projection capacity will increase a state's chances of initiating a militarized interstate dispute (MID). Thus, I expect that increased force projection capacity is a double-edged sword: it helps deter disputes against allies, but it increases the likelihood a state will initiate disputes elsewhere. This leads to my third hypothesis:

> *H3: As force projection capacity increases, the likelihood that an allied state is the target of a militarized dispute goes down, whereas the overall likelihood of initiating a militarized dispute goes up.*

4.3 Data

My primary explanatory variable, MOBILITY INDEX$_{it}$, measures the force projection capacity that state i possesses in year t, using specific air and naval capabilities as a proxy for overall force projection capacity. To create this index, I collected data on each state's airlift capabilities and the types of aircraft carriers that it possessed in a given year from *The Military Balance* (International Institute for Strategic Studies, 1970-2018).[5] A state-year's mobility index is computed as the average of two values: the product of aircraft capacity values and the number of aircraft, summed over all aircraft types, and the product of aircraft carrier capacity values and the number of carriers, summed over all carrier types. Prior to taking the average, I re-scaled the separate measures by the maximum and minimum values for the data so that they are each in the range $[0, 1]$.[6]

The primary dependent variable, DEPLOYMENT SIZE$_{ijt}$, measures the number of troops deployed from state i to within the geographic borders of state j in year t, which I also collected from *The Military Balance* (International Institute for Strategic Studies, 1970-2018). Because this paper focuses on a state's force projection and its relation to alliances and other security agreements, I exclude deployments for United Nations or Multinational Force and Observers peacekeeping missions, as they are not directly related to force projection capacity or alliance commitments. To evaluate the probability of deployment, I create an indicator DEPLOYMENT$_{ijt}$ which equals 1 if troops were deployed from i to j in year t, and 0 if no troops were deployed.

[5] I collected the number and type of each mobility aircraft in a state's air force for each year, then recorded specifications for each aircraft's cargo capacity (pounds), cruise speed (knots), and un-refueled range (nautical miles) from various online sources. In the same manner, I also recorded each aircraft carrier's displacement (pounds), speed (knots) and un-refueled range (nautical miles). Each individual aircraft and aircraft carrier's mobility value is the product of their three respective specifications.

[6] Due to the uneven distribution of capabilities, many states have very small mobility index values. Therefore, for the remainder of the paper, I limit my focus for state i (the potential "force projector") to those states in the 60th percentile and above for MOBILITY INDEX.The distribution of MOBILITY INDEX has a large tail of states with almost no force projection capacity, and which have never deployed troops, making their state i directed-dyadic observations less relevant for understanding force projection behavior. Directed dyads for state j (the potential host of force projection) include all other states in the international system in year t. As a robustness check, I ran all statistical tests without the cutoff and the results remained unchanged.

I also include a number of additional explanatory variables in the data set to account for additional influences on deployment likelihood and the number of troops deployed. Because overall capabilities such as military expenditures and number of military personnel are important in understanding deployment tendencies, I include the Composite Index of National Capabilities (Singer, Bremer, and Stuckey, 1972), CINC_{it}, to measure state i's capabilities in year t, as well as a measure for the proportion of i's capabilities compared to the total dyadic capabilities, $\text{CINC}_{it}/(\text{CINC}_{it}+\text{CINC}_{jt})$, which captures relative military imbalance and possible dependence within the dyad.[7] Similarly, because greater trade dependence may influence alliance commitments, I measure total dyadic trade as a proportion of i's total trade with $\text{TRADE}_{ijt}/\text{TOTAL TRADE}_{it}$ (Barbieri, Keshk, and Pollins, 2008). I also include a measure for political regime type to control for differences in domestic preferences with POLITY SCORE_{it} (Marshall, Gurr, and Jaggers, 2014). To account for recent historical events that may influence deployments, I include three variables: $\text{TERRORIST FATALITIES LAST 2 YEARS}_{jt}$, the total number of fatalities due to terrorist attacks in state j during the previous two years (National Consortium for the Study of Terrorism and Responses to Terrorism (START), 2018); $\text{COMMON THREATS LAST 2 YEARS}_{ijt}$, the total number of militarized disputes in which states i and j were on the same side in the previous two years, and $\text{HOSTILE MILITARIZED DISPUTES LAST 2 YEARS}_{ijt}$,[8] the total number of hostile militarized disputes in which states i and j were on opposing sides in the previous two years (Palmer, D'Orazio, Kenwick et al., 2020).[9] Differences in foreign policy preferences are included via Bailey, Strezhnev, and Voeten (2017)'s measure of absolute distance between state i's and j's ideal point estimates from UN General

[7]Including state i's CINC score and the dyadic proportion of CINC does not create collinearity in the model, because the former is a monadic measure and the latter is dyadic. The Pearson correlation coefficient between the two variables in the full dataset is 0.27, and in the alliance-only subset is 0.29. For further comparison between MOBILITY INDEX and CINC, see the appendix.

[8]A MID is considered "hostile" if there is a use of force or it leads to war.

[9]Common threats and hostile MIDs are entirely uncorrelated in the data: the Pearson correlation coefficient for both the full and dataset and for alliances only is 0.00.

Assembly voting, FOREIGN POLICY DIFFERENCE$_{ijt}$.

To code whether states have a defense agreement, where an alliance member promises active military support if the other alliance member's sovereignty or territory is attacked, I include DEFENSE ALLIANCE$_{ijt}$ from Leeds, Ritter, Mitchell et al. (2002). Defensive alliances include multilateral treaties such as NATO, the Rio Pact, and the Warsaw Pact, and bilateral treaties such as between China and North Korea, Kenya and Ethiopia, and France and Gabon. I exclude non-defensive alliances, such as those that are primarily either offensive or non-aggression pacts, because they do not relate to signaling commitments to deter threats. Examples of excluded alliance treaties are the Paris Charter and the non-aggression pact between India and the Philippines.

I also measure geographic distance between states, DISTANCE$_{ij}$, as the great circle distance between their capital cities (Gleditsch, 2008), because greater distance may increase the need for deployments to signal alliance commitments. Finally, to control for the effects of dyadic colonial history and the Cold War, I add indicators COLONIAL HISTORY$_{ij}$ (Hensel, 2018), which equals 1 if both states in a dyad were previously part of a colonial relationship and 0 otherwise, and COLD WAR$_t$, which equals 1 if the year is prior to 1990, and 0 otherwise.

A summary of the data used in the analysis is provided in Table 4.1. Although I collected mobility and deployment data starting in 1970, due to gaps in reporting annual data in *The Military Balance* prior to 1984, I exclude these years from my analysis.[10] Additionally, since the current MID dataset runs through 2010, I limit my focus in the empirical analysis to the period from 1985 to 2010.[11]

[10]For example, *The Military Balance* does not list many Soviet Union, U.K., or U.S. troop levels deployed to other countries until 1984.

[11]Since missing data in the interior of a period under study can cause problems in time-series-cross-section data (Beck and Katz, 2011), my strategy for missing data was the following: first, I interpolated missing values for MOBILITY INDEX$_{it}$ and DEPLOYMENT SIZE$_{ijt}$, and carried forward the last observed value for POLITY SCORE$_{it}$ and POLITY SCORE$_{jt}$. Then, for any states where either MOBILITY INDEX$_{it}$, DEPLOYMENT$_{ijt}$, or POLITY SCORE$_{it}$ was missing, I removed any directed dyad in which they were the potential force projector

Variable	*Mean*	*SD*	*Median*	*Min*	*Max*
DEPLOYMENT SIZE$_{ijt}$	78.90	3,118.08	0	0	541,400
DEPLOYMENT$_{ijt}$	0.01	0.11	0	0	1
MOBILITY INDEX$_{it}$	0.03	0.12	0.003	0.0001	0.95
CINC$_{it}$	0.02	0.03	0.007	0.0003	0.21
POLITY SCORE$_{it}$	2.89	7.39	7	-10	10
CINC$_{it}$/(CINC$_{it}$+CINC$_{jt}$)	0.76	0.28	0.89	0.001	1
TRADE$_{ijt}$/TOTAL TRADE$_{it}$	0.006	0.03	0.0001	0	0.96
TERRORIST FATALITIES LAST 2 YEARS$_{jt}$	37.74	212.30	0	0	6,667
COMMON THREATS LAST 2 YEARS$_{ijt}$	0.03	0.27	0	0	7
HOSTILE MILITARIZED DISPUTES LAST 2 YEARS$_{ijt}$	0.005	0.10	0	0	4
FOREIGN POLICY DIFFERENCE$_{ijt}$	1.14	0.87	0.97	0	4.89
MILITARIZED DISPUTE$_{ijt}$	0.006	0.08	0	0	1
ALLY j MILITARIZED DISPUTE TARGET$_{ijt}$	0.19	0.39	0	0	1
DEFENSE ALLIANCE$_{ijt}$	0.09	0.28	0	0	1
DISTANCE$_{ij}$	4,882.62	2,758.52	4,721	0	12,420
COLONIAL HISTORY$_{ij}$	0.02	0.15	0	0	1
COLD WAR$_{t}$	0.17	0.37	0	0	1
YEAR	1997.91	7.38	1998	1985	2010

Note: $N = 242{,}422$ directed dyads.

Table 4.1: Descriptive Statistics

Additionally, since the U.S. was the dominant global military power throughout the time period analyzed in the data, it is possible that the data is primarily observing U.S. behavior, which could skew the results. To address concerns about the uneven distribution of states observed in the data, I provide a summary of State A observations for the U.S. in Table 4.2. This table shows that U.S. deployments do account for a relatively large proportion of the overall deployments both within alliances and throughout the international system. This is not surprising: we would expect the U.S. to deploy more because greater military power tends to increase force projection behavior. More importantly, however, in both instances these proportions are less than half of the total deployments. Therefore, it does not appear that the data is predominantly accounting for U.S. behavior, as over half of the deployments in alliances—and nearly three quarters of deployments

(state i) from the data. For any states where DISTANCE$_{ij}$ data was missing, I completely removed them from the data. Finally, I imputed remaining missing values—on POLITY SCORE$_{jt}$, TRADE$_{ijt}$/TOTAL TRADE$_{it}$, and FOREIGN POLICY DIFFERENCE$_{ijt}$—using Amelia II (Honaker, King, and Blackwell, 2015).

	Prop. U.S. Observations	*Prop. U.S. Deployments*
Alliances only	0.08	0.46
All observations	0.02	0.27

Table 4.2: Proportion of U.S. Observations and Deployments

overall—are from non-U.S. states. As an additional robustness check, I conducted all of the main empirical models in the following section without the U.S. and show that the results do not change substantially. These results are available in the appendix.

4.4 Data Analysis

I model the determinants of alliance deployment levels with an autoregressive distributed lag model (ADL) as follows:

$$Y_{ijt} = \theta Y_{ijt-1} + \beta_0 X_{it} + \beta_1 X_{it-1} + \beta_2 X_{it-2} + \gamma Z_{ijt} + \eta_{ij} + \tau_t + \epsilon_{ijt} \quad (4.1)$$

Because the model in Eq. 4.1 focuses on alliance deployment levels, I focus strictly on those directed dyads for which an alliance exists in year t in the data. The lagged dependent variable is included on the right-hand side of the equation to account for the fact that deployment levels from i to j in year t are likely influenced by past deployment levels in the same direction. Furthermore, the main explanatory variable X_{it} (MOBILITY INDEX$_{it}$) is lagged twice for both theoretical and data-driven reasons: theoretically, an increase in force projection capacity might not have a contemporaneous effect on deployment levels, but will instead be lagged over time. Furthermore, based on the data, I lag the variable by two years in order to remove serial correlation from the data, which I tested for with a Lagrange Multiplier test.

In addition to the main explanatory variable, I include a number of additional covariates Z_{ijt} to account for other influences on deployment levels, including state i's CINC score, polity score, dyadic CINC dependence, and dyadic trade dependence; state j's deployment levels to state i; dyadic foreign policy differences; the number of common threats in the past two years within the dyad; dyadic distance; dyadic colonial history; a Cold War indicator; and a yearly time trend to account for technological changes over time. In the fixed effects version of the model, I include fixed effects for directed dyad η_{ij} and time τ_t to control for unobserved directed-dyad-specific and year-specific heterogeneity, respectively.[12] Finally, I also include an i.i.d. error term, ϵ_{ijt}.

To test the second hypothesis, I consider the determinants of deployment likelihood for all directed dyads, and not just those within defensive alliances. To do so, I use a logistic regression model to test the probability of a new deployment in year t, using the same modeling structure as in Eq. 4.1 for logistic regression, except dropping the lagged independent variables and adding time splines for years since last deployment.[13] In order to remove the contaminating effects of long deployment spells, I limit my analysis to the onset of a new deployment (Beck, Katz, and Tucker, 1998), with repeats possible for directed dyads in non-consecutive years.

For the third hypothesis, I evaluate the determinants of MID likelihood for all directed dyads, and for those only within defensive alliances. I again use a logistic regression model and test the probability of a MID in year t. With the defense alliance data, I evaluate the probability that state j will be the target of a MID by a state other than i. With the full data, I evaluate the probability that state i and j will be in an opposing MID in year t. As with the deployment likelihood model, I use time splines for years since last MID for the respective dependent variable.

[12] Although including fixed effects with a lagged dependent variable is known to create "Nickell bias" of order $1/T$ in the parameter estimates (Nickell, 1981), this is not a major concern for my data because the time period is relatively long ($T = 26$), and thus the bias will be small. I follow the recommendation of Beck and Katz (2011) in modeling the ADL model with fixed effects given $T > 20$.

[13] I exclude distributed lags in this model because I already account for time dependence with the cubic time spline.

4.4.1 Results

Table 4.3 reports the determinants of alliance deployment levels. The first two models report the estimated effect of force projection capacity on deployment size, without fixed effects (Model 1) and with fixed effects (Model 2). Models 3 and 4 incorporate additional covariates in the same fashion. In each of the models, the long-term effect of MOBILITY INDEX on DEPLOYMENT SIZE can be calculated by $(\sum \beta)/(1-\theta)$ where the β values are the coefficients for MOBILITY INDEX$_{it}$ and its lags, and θ is the coefficient for DEPLOYMENT SIZE$_{it-1}$, as defined in Eq. 4.1. For ease of interpretation, I also report the calculated values of these long-term effects in Table 4.3.

In all but the first model of Table 4.3, the long-run effect of MOBILITY INDEX is negative, indicating that an increase in force projection capacity is associated with a decrease in deployment levels to allies. However, the magnitude of this effect differs according to whether or not fixed effects are included. Yet because the models without fixed effects contain a substantial amount of serial correlation as indicated by the results of the LM test statistic—which is compared to a χ^2 statistic with 1 d.f.—whereas the models with fixed effects do not, Models 2 and 4 provide more reliable estimates for the estimate of the long-run effect.

The coefficients' signs switch from positive to negative in the second-year lag for each of the models in Table 4.3, which can be cause for concern that the distributed lags have high multicollinearity. While I determined the ADL model's lag length according to a Lagrange Multiplier test in order to remove serial correlation in the error term, this does not resolve multicollinearity concerns. However, additional diagnostic checks using Bayesian Model Averaging, and available in the appendix, support the conclusion that the opposite signs are genuine and not due to multicollinearity. While it is unclear why the trend is not persistently negative, one possible explanation is that within the first year following an increase in force projection capacity, the increase has an upward pressure on a state's tendency to deploy troops. However, when looking at the long-run effect, this ini-

	Dependent variable:			
	LOG(DEPLOYMENT SIZE$_{ijt}$)			
	(1)	(2)	(3)	(4)
LOG(DEPLOYMENT SIZE$_{ijt-1}$)	0.94***	0.73***	0.93***	0.73***
	(0.002)	(0.006)	(0.002)	(0.006)
SIDE i COVARIATES				
MOBILITY INDEX$_{it}$	0.98***	0.30	0.90***	0.17
	(0.26)	(0.29)	(0.27)	(0.29)
MOBILITY INDEX$_{it-1}$	3.42***	3.18***	3.38***	3.14***
	(0.34)	(0.33)	(0.34)	(0.33)
MOBILITY INDEX$_{it-2}$	$-$4.33***	$-$3.73***	$-$4.31***	$-$3.81***
	(0.31)	(0.33)	(0.31)	(0.33)
CINC$_{it}$			0.94***	7.31***
			(0.30)	(1.17)
POLITY SCORE$_{it}$			$-$0.00009	$-$0.002
			(0.0005)	(0.002)
CINC$_{it}$/(CINC$_{it}$+CINC$_{jt}$)			0.01	$-$0.15
			(0.01)	(0.15)
TRADE$_{ijt}$/TOTAL TRADE$_{it}$			0.20***	$-$0.07
			(0.07)	(0.21)
SIDE j COVARIATE				
LOG(DEPLOYMENT SIZE)$_{jit}$			0.01***	0.01
			(0.003)	(0.009)
DYADIC COVARIATES				
FOREIGN POLICY DIFFERENCE$_{ijt}$			$-$0.02***	0.02
			(0.007)	(0.02)
COMMON THREATS LAST 2 YEARS$_{ijt}$			0.03***	0.02***
			(0.005)	(0.006)
LOG(DISTANCE$_{ij}$)			0.01***	
			(0.004)	
COLONIAL HISTORY$_{ij}$			0.10***	
			(0.009)	
TIME COVARIATES				
COLD WAR$_t$			0.02*	
			(0.009)	
YEAR TREND$_t$			0.0002	
			(0.0004)	
MOBILITY INDEX (LONG RUN EFFECT)	1.17	$-$0.94	$-$0.51	$-$1.84
Year fixed effects		✓		✓
Directed dyad fixed effects		✓		✓
Observations	19,605	19,605	19,605	19,605
Clusters	1,079	1,079	1,079	1,079
Years	26	26	26	26
LM autocorrelation test statistic	139.55	2.58	133.60	2.48

Notes: Panel corrected standard errors in parentheses. *p<0.1; **p<0.05; ***p<0.01

Table 4.3: Determinants of Alliance Deployment Levels, 1985–2010

tial upward pressure is reversed by the overall tendency to deploy fewer troops, where the magnitude of the long-term negative coefficient exceeds and counteracts the short-term increase.

An additional modeling challenge to Eq. 4.1 is possible endogeneity between deployment levels in alliances and force projection capacity: in particular, technological investments and force posture policy decisions might predict force projection capacity, rather than (or in addition to) the other way around. To deal with this threat to causal inference, I employ instrumental variable estimation, using state *i*'s past ruling party political ideology in the government as instruments for MOBILITY INDEX$_{it}$. The appendix discusses the IV estimation strategy in detail, and shows that the results in Table 4.3 still hold when accounting for the possibility of endogeneity.

Next, to test the second hypothesis, Table 4.4 reports the determinants of deployment onset probability for all directed dyads. I include several additional covariates that might explain deployment probability: state *j*'s domestic regime type, the logged number of terrorist fatalities in *j* during the last two years, and the number of hostile militarized disputes between *i* and *j* in the last two years. I also include a binary indicator for whether the states share an alliance commitment to control for deployments within alliances, and I add a cubic spline for the time since *i*'s last deployment to *j*.

The model in Table 4.4 offer clear empirical evidence for the second hypothesis. As MOBILITY INDEX increases, the likelihood of deployment onset also increases. Furthermore, because the unit of analysis is a directed-dyad-year, an increase in the likelihood of deployment onset means that either states are deploying to new locations, or new states are deploying troops, or both. Consequently, when the results from the first hypothesis are also considered, we can conclude that as force projection capacity increases, states are deploying fewer troops, but to more countries.

What effect does a global force posture have on deterrence? To evaluate my third hypothesis, I test three models with different dependent variables. Table 4.5 reports the results of these tests. In the first model in Table 4.5, I evaluate only states in defense alliances. The dependent variable in this

	Dependent variable:
	PR(DEPLOYMENT$_{ijt}$)
SIDE *i* COVARIATES	
MOBILITY INDEX$_{it}$	3.11***
	(0.34)
CINC$_{it}$	9.37***
	(1.97)
POLITY SCORE$_{it}$	0.07***
	(0.01)
CINC$_{it}$/(CINC$_{it}$+CINC$_{jt}$)	1.42***
	(0.29)
TRADE$_{ijt}$/TOTAL TRADE$_{it}$	0.93
	(1.74)
SIDE *j* COVARIATES	
DEPLOYMENT$_{jit}$	1.14***
	(0.37)
POLITY SCORE$_{jt}$	−0.04***
	(0.008)
LOG(TERRORIST FATALITIES LAST 2 YEARS$_{jt}$)	0.23***
	(0.02)
DYADIC COVARIATES	
FOREIGN POLICY DIFFERENCE$_{ijt}$	−0.22***
	(0.05)
COMMON THREATS LAST 2 YEARS$_{ijt}$	0.42***
	(0.08)
HOSTILE MILITARIZED DISPUTES LAST 2 YEARS$_{ijt}$	0.47***
	(0.14)
DEFENSE ALLIANCE$_{ijt}$	−0.11
	(0.15)
LOG(DISTANCE$_{ij}$)	−0.82***
	(0.06)
COLONIAL HISTORY$_{ij}$	0.69***
	(0.21)
TIME COVARIATES	
COLD WAR$_{t}$	0.68***
	(0.13)
TIME SINCE LAST DEPLOYMENT SPLINE$_{ijt}$	0.01**
	(0.008)
Observations	236,274

Note: Robust standard errors in parentheses clustered by directed dyad. *p<0.1; **p<0.05; ***p<0.01

Table 4.4: Determinants of Deployment Onset Probability, 1985–2010

model is the probability that state j will be the target of a MID from a state other than i. In the second model, I use the full dataset, and the dependent variable is the probability that state i will initiate a MID with state j. Finally, the third model is the reverse of the second, so that the dependent variable is the probability that state i will be the target of a MID with state j. I also include most of the same covariates as in Table 4.4, though I leave out trade, deployment, terrorist fatalities, colonial history, and Cold War covariates, as there is no theoretical reason for these variables to be associated with MID likelihood.

	Dependent variable:		
	Pr(Ally j MID target$_{ijt}$)	Pr(MID$_{ijt}$)	Pr(MID$_{jit}$)
	(1)	(2)	(3)
Side i covariates			
Mobility index$_{it}$	-2.21^{***}	0.50^{**}	0.08
	(0.31)	(0.20)	(0.28)
CINC$_{it}$	11.90^{***}	19.62^{***}	23.05^{***}
	(1.59)	(0.80)	(1.04)
Polity score$_{it}$	-0.05^{***}	0.01^{***}	0.02^{***}
	(0.004)	(0.004)	(0.005)
CINC$_{it}$/(CINC$_{it}$+CINC$_{jt}$)	-2.29^{***}	-2.70^{***}	-3.53^{***}
	(0.06)	(0.10)	(0.14)
Side j covariates			
Polity score$_{jt}$	0.03^{***}	-0.03^{***}	-0.03^{***}
	(0.004)	(0.004)	(0.005)
Dyadic covariates			
Foreign policy difference$_{ijt}$	0.44^{***}	0.23^{***}	0.24^{***}
	(0.03)	(0.03)	(0.04)
Defense alliance$_{ijt}$		-0.97^{***}	-0.91^{***}
		(0.09)	(0.14)
log(Distance$_{ij}$)	-0.02	-1.53^{***}	-1.45^{***}
	(0.02)	(0.03)	(0.04)
Time covariates			
Time since j last MID target spline$_{ijt}$	-0.05^{***}		
	(0.003)		
Time since last dyadic MID spline$_{ijt}$		0.002	-0.02^{***}
		(0.004)	(0.005)
Observations	19,605	242,422	242,422

Note: Robust standard errors in parentheses clustered by directed dyad. *p<0.1; **p<0.05; ***p<0.01

Table 4.5: Determinants of Militarized Dispute Probability, 1985–2010

The models in Table 4.5 provide strong empirical support for the third hypothesis. First, states with higher force projection capacity can still effectively deter disputes against their allies. Additionally, it is worth noting that for the first model, DISTANCE_{ij} has no measurable effect on i's ability to deter disputes against allied j. Therefore it appears that with the capacity to project force overseas, allies that are geographically distant are still credible from the view of potentially hostile actors.

However, states with higher force projection capacity are also more likely to initiate disputes overall. Although according to Model 3 in Table 4.5 there is no significant effect of force projection capacity on the likelihood of being targeted for a dispute, in the appendix I show that this is likely a feature of an uneven distribution of power between the U.S. and other states. Instead, the U.S. appears to more likely to be the target of a dispute as its force projection capacity increases, whereas other states appear less likely to be the target of a dispute. While these results are driven by the U.S. in the data, it is may simply be the effect of major power status rather than a feature that is unique to the U.S. as an actor.

4.5 Alliance burden sharing and network autocorrelation

A potential limitation to the model in Eq. 4.1 is that it assumes each directed dyad is independent of other directed dyads. However, each state has potentially dozens of directed dyads within its alliance network, all of which can have an effect on each other through their network connectivity. Thus, for example, state i's deployment levels to j depend on i's deployment levels to all of its other connections since it has a finite number of troops it can deploy. U.S. troops stationed on the Korean peninsula, for example, affect on the number of troops the U.S. can deploy to the Middle East. Likewise, alliance burden sharing can affect states' decisions on how to posture their forces: i's deployment levels to j can also depend on other states' deployment levels to j, especially if they are part of the same alliance agreement. For example, when the U.S. deployed troops to West Germany during the

Cold War, it meant that other NATO states did not have to deploy as many troops to the same area in order to signal NATO credibility. Although I account for spatial dependence in the first set of models for Eq. 4.1 by using panel-corrected standard errors (Beck and Katz, 1995), I also explore a spatial error model (SEM) that explicitly takes these relationships into account as follows (Anselin, 1988; Beck, Gleditsch, and Beardsley, 2006):

$$Y_{ijt} = \theta Y_{ijt-1} + \beta_0 X_{it} + \beta_1 X_{it-1} + \beta_2 X_{it-2} + \gamma Z_{ijt} + \eta_{ij} + \tau_t + \epsilon_{ijt} + \lambda w_{it} \epsilon \tag{4.2}$$

When $\lambda = 0$, the model in Eq. 4.2 is identical to the OLS model in Eq. 4.1. Thus, the SEM explicitly accounts for unobserved spatial dependence among directed dyads by parameterizing the spatial correlation, λ, in the error term.[14] Additionally, w_{it} is the ith row of connectivity matrix W for year t, where $w_{ijt} > 0$ when i and j are connected in year t, and 0 otherwise. Thus W is an $NT \times NT$ block-diagonal matrix, with connectivity for each year blocked along the diagonal and zeroes elsewhere.

To construct W, I converted the alliance network structure for all defense alliances in the ATOP dataset (Leeds, Ritter, Mitchell et al., 2002) from a binary graph to a binary edge graph, in order to accommodate the data's directed-dyadic structure. An edge graph, $E(G)$, takes the edges from G and treats them as nodes in $E(G)$. Two nodes in $E(G)$ are then joined if they share a common node in G (Bondy and Murty, 1976). Using the edge graph as the connectivity matrix, I then row standardized W with equal weights for all non-zero values in w_{it}.

To account for alliance network autocorrelation, I model a spatial error model incorporating connectivity matrix W. Table 4.6 reports the results of this model. Comparing the results in Table 4.6 to Model 4 in Table 4.3, it is clear from the coefficient and standard error for λ that although positive spatial correlation exists, it does not substantially change the estimates

[14] I also consider a spatial autoregressive model (SAR) in the appendix. However, diagnostic tests also available in the appendix indicate that the SEM is a better model for my data than the SAR.

	Dependent variable:
	log(Deployment size$_{ijt}$)
log(Deployment size$_{ijt-1}$)	0.74***
	(0.005)
Side i covariates	
Mobility index$_{it}$	0.15
	(0.40)
Mobility index$_{it-1}$	3.02***
	(0.45)
Mobility index$_{it-2}$	-3.71^{***}
	(0.39)
CINC$_{it}$	5.04***
	(1.26)
Polity score$_{it}$	-0.002
	(0.002)
CINC$_{it}$/(CINC$_{it}$+CINC$_{jt}$)	-0.14
	(0.13)
Trade$_{ijt}$/Total trade$_{it}$	0.06
	(0.18)
Side j covariate	
log(Deployment size)$_{jit}$	-0.007
	(0.007)
Dyadic covariates	
Foreign policy difference$_{ijt}$	0.006
	(0.02)
Common threats last 2 years$_{ijt}$	0.02**
	(0.008)
Spatial correlation covariate	
λ	0.51***
	(0.02)
Mobility index (long run effect)	-2.16
Year fixed effects	✓
Directed dyad fixed effects	✓
Observations	19,605
LM autocorrelation test statistic	5.78

Notes: Asymptotic adjusted standard errors in parentheses. *p<0.1; **p<0.05; ***p<0.01

Table 4.6: Spatial Error Model of Alliance Deployment Levels, 1985–2010

for the long-run effect of MOBILITY INDEX on DEPLOYMENT SIZE. Consequently, spatial dependence does not appear to bias the estimates for the main variables of interest from the original OLS regression. However, the SEM does reduce the magnitude of $CINC_{it}$ without substantially changing the magnitude of the long-run effect for MOBILITY INDEX, further demonstrating that CINC is measuring a different aspect of capabilities than force projection capacity.[15]

4.6 Conclusion

The empirical evidence in this paper demonstrates that an increase in force projection capacity has a depressing effect on deployment levels. States tend to rely less on costly signaling in their alliance commitments as their force projection capacity increases, instead relying more on tying hands signaling in order underwrite the credibility of their alliances. As Fearon (1997) notes, states would prefer to rely on audience costs since they are only paid if a crisis occurs. My results provide evidence of this tendency. However, audience costs alone are generally not sufficient in demonstrating credible commitments. Instead, audience costs signals are supported by the capacity to project force with high speed and volume.

Therefore, the primary theoretical contribution that this study makes is that tying-hands signaling in alliances without significant sunk-cost signaling does not operate purely through audience costs. Instead, credibility is also signaled by the specific capabilities that states develop, which allows states to quickly respond to a crisis situation. One of the key capabilities in this regard is the level of mobility speed and capacity, which I measure by creating an annual index for airlift and aircraft carrier capabilities.

More broadly, this paper also adds to our understanding of states' se-

[15]The Lagrange Multiplier test for the SEM indicates that some spatial correlation remains in the error term, and therefore the spatial weights matrix did not fully capture spatial dependence. However, based on the value for the test statistic in Table 4.6, the amount of remaining spatial correlation is small and therefore should not have a substantial impact on the reported results.

curity pursuits. The arms-versus-allies—or equivalently, internal-versus-external balancing—tradeoff expressed in the literature is useful for understanding the extent to which states prefer autonomy versus relying on other states for security (Morrow, 1993). However, rather than being substitutes for each other, I show that force projection capabilities can increase autonomy while also complementing alliances through a shift in signaling strategies. Furthermore, these dynamics apply to both symmetric and asymmetric alliance relationships (Morrow, 1991). Additionally, the more flexible, smaller deployments also give a weaker partner an opportunity to engage in security cooperating while preserving a high degree of autonomy.

Moreover, the results in this paper have important implications for the changing nature of deterrence. With increased force projection capacity, the distinction between immediate deterrence and general deterrence blurs: both occur through flexible force projection, which is more prevalent when there is a high capacity to project force. This increased level of flexibility to respond to crises has an additional effect on states' force projection behavior: the cost savings from fewer deployments to allies create new opportunities to pursue foreign policy interests in other areas without needing to generate strong commitment signals. Consequently, increased force projection capacity increases the propensity for states to generate new deployments into other countries. These can occur through both crisis responses and limited security partnerships, such as training foreign military forces or conducting combined antiterrorism operations. At the same time, however, as states deploy to more *non-allied* locations globally, they are more likely to use military capabilities for coercive diplomacy. Thus, increased force projection capacity acts a double-edged sword: it helps deter disputes against allies, but it increases the likelihood a state will initiate disputes elsewhere.

The findings in this paper suggest a number of areas for future research. For example, while I control for regime type in my analysis, I do not explicitly evaluate the differences in signaling across political systems. Instead, the models I use assume there is no difference in motivations for signaling based on regime type. However, since a lack of sunk-cost signaling in alliances indicates that audience cost signals are therefore more important in signaling credibility, are democratic regimes more likely to use rapid mobility as a substitute for troop deployments? Conversely, are autocratic regimes more likely to continue to rely on deployments to signal credibility? Similarly, do autocracies that develop rapid mobility capabilities tend to have fewer defense alliances, and if so, given their tendency to initiate more militarized disputes, do they therefore use such capabilities in a more offensive manner, or are such capabilities used primarily for immediate deterrence?

This paper also has several important policy implications for the changing international security environment. For example, it helps explain what policymakers can expect from China as it expands its force projection capacity, both in terms of airlift capabilities and aircraft carriers. Over the last two decades, China has expanded its force projection capacity, which is now greater than that of both France and the U.K., as Figure 4.3 shows. China launched its first aircraft carrier in 2011, followed by a second carrier in 2017 that will soon be fully commissioned for missions (Buckley, 2017). China has also announced plans to construct four additional carriers by 2035 (Chan and Rui, 2019). With such a swift expansion, rather than establish defensive alliances, China is more likely to pursue limited security partnerships with like-minded states. Thus as China continues to increase its capacity to project force, policymakers can expect further Chinese involvement in foreign security partnerships where its interests lie. Doing so will provide China with greater flexibility in pursuit of its foreign policy interests, but will also limit its ability to deter or compel other states from meddling in its interests in the long run, thereby increasing the chance that China will become involved in militarized disputes.

Additionally, there are important policy implications for the NATO alliance, which has recently faced challenges to its long-run sustainability and legitimacy. For example, the U.S. has been evaluating the costs of maintaining the current 35,000 troops deployed to Germany (Hudson, Sonne, DeYoung et al., 2018). These challenges have raised alarms in Europe about the

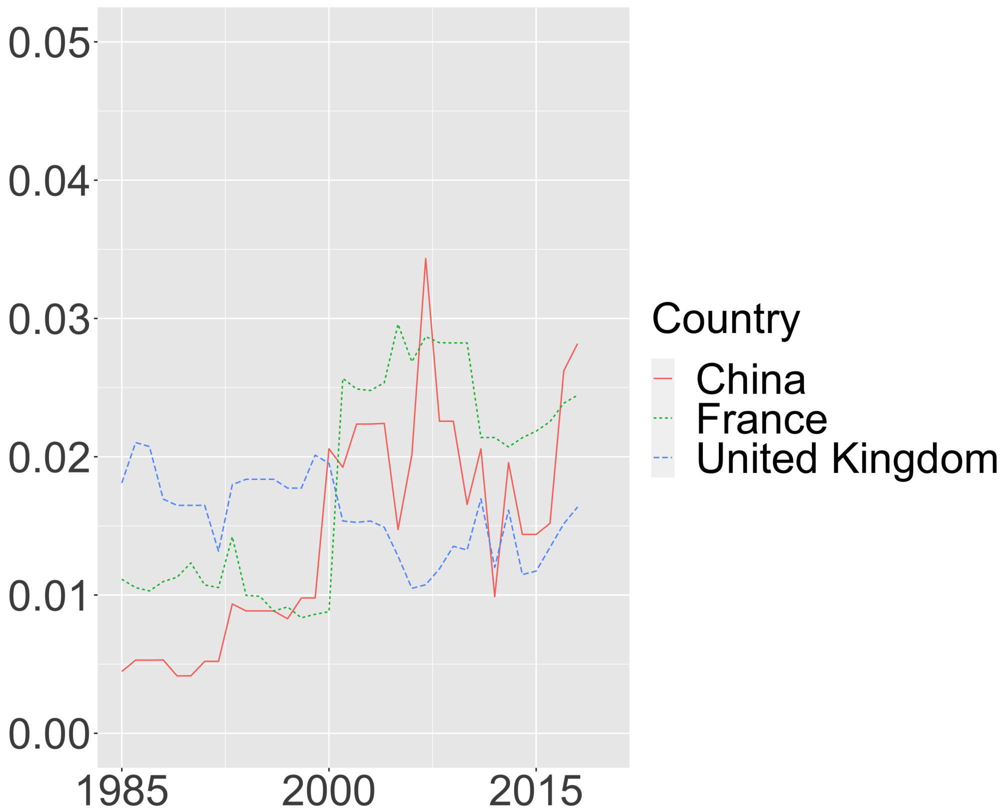


Figure 4.3: China, France, and U.K. MOBILITY INDEX, 1985–2018

strength of NATO credibility if the U.S. were to withdraw some its troops from the region. However, this paper provides reassurance that the alliance can still remain credible so long as U.S. political leaders maintain both their public commitment to the alliance and the military's force projection capacity, even if troops were to withdraw from Germany or other parts of the European continent.

There are also cautionary implications for policymakers based on the findings presented here. While developing greater force projection capabilities may reduce infrastructure costs of overseas bases, there is a risk of over-promising within alliances when too many of them are underwritten by force projection capacity without any substantial sunk costs. Force projection is a limited resource, and being spread too thin can come at great cost to alliance credibility, and ultimately reputation, if the likelihood of threats increases and not enough resources are available to respond to them all. Thus policymakers must still consider the expected costs of stationing troops versus the expected threat likelihood and the cost of responding to crises with ad hoc, flexible deployments.

4.7 Appendix

4.7.1 Mobility Index vs. CINC Comparison

One might suspect that as a state's overall capabilities increase, its force projection capabilities will also increase. The Pearson correlation coefficient for mobility capacity and the Composite Index of National Capabilities (CINC) (Singer, Bremer, and Stuckey, 1972), often used by political scientists to measure national power, is 0.65, indicating moderately strong positive correlation between the two measures. However, a closer evaluation reveals some striking differences between the two indices. For example, Figure 4.4 plots CINC and MOBILITY INDEX values over time for several of the most capable states according to these indices. While the two measures have different scales,[16] the plots are lined up in relative terms to each other for easier comparison.

Several items are worth briefly noting about the plots in Figure 4.4. First, whereas some states such as China and India have seen large increases in relative capabilities over time according to CINC, neither of these states' force projection capabilities have experienced any apparent increase, relative or absolute, according to MOBILITY INDEX. Moreover, because both states have relatively high CINC capabilities even in 1985, one might assume they each have some relatively high force projection capacity. However, MOBILITY INDEX suggests that this is a poor assumption. Thus, CINC may overestimate force projection capacity. In contrast, the U.S. has remained steady in relative capabilities by CINC standards, but has increased in both relative and overall force projection capacity according to MOBILITY INDEX measures. This suggests that CINC underestimates force projection capacity in other cases. In total, all of these differences indicate that MOBILITY INDEX is capturing a different aspect of national capabilities than CINC.

[16] CINC represents an annual proportion of world total capabilities, so that relative national power is time independent, whereas MOBILITY INDEX is scaled relative to its maximum value. I do not use an annual proportion because this removes the time-trending effect of improvements in force projection capacity, whereas I want to explicitly measure the effects of these changes over time.

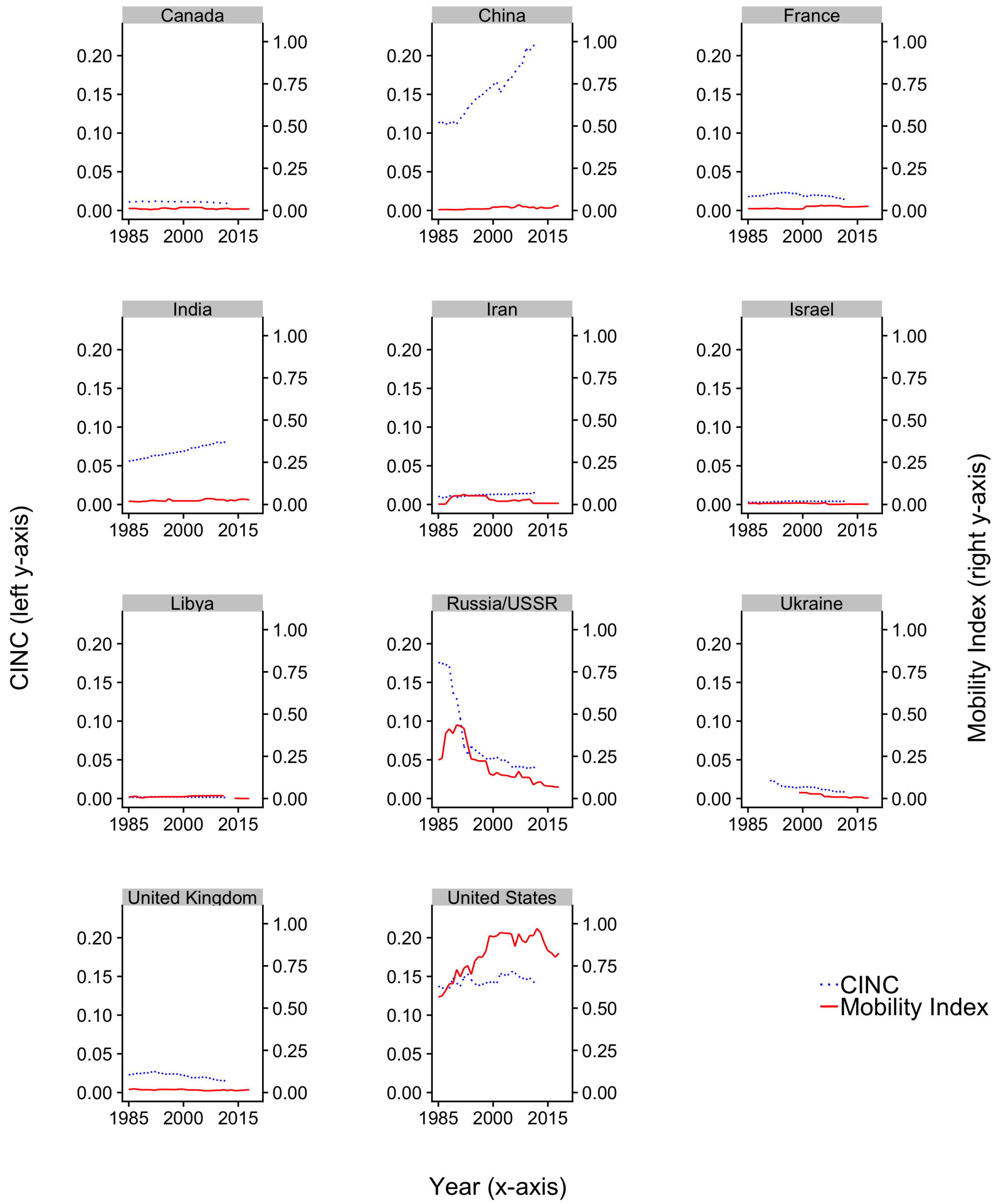


Figure 4.4: Sample Comparison of CINC vs. MOBILITY INDEX, 1985–2018

4.7.2 Autoregressive distributed lag model diagnostics

Because distributed lags can be highly correlated in a regression model, I ran diagnostics on the results in Table 4.3 to evaluate whether the alternating signs on the lag coefficients might indicate multicollinearity. To do so, I used Bayesian Model Averaging (BMA), which checks variable selection across models and accounts for uncertainty by averaging over the best models.

BMA is useful for identifying which variables are the most relevant predictors for the model. Using this technique, the variables with high posterior probabilities are most likely useful predictors in the model. While I am less concerned with which variables to select in the model, I am concerned with whether BMA suggests that the distributed lags for MOBILITY INDEX are useful predictors or not. In Tables 4.7 through 4.10, I report the best BMA-based model selected for each of the four models in Table 4.3. In all four tables, the best models all indicate that the distributed lags are useful predictors, and moreover, that the coefficient on the second lag switches from positive to negative. This provides support for the conclusion that these lags are meaningfully explaining the outcome, and are not providing spurious coefficients due to multicollinearity.

	p!=0	EV	SD	Best Model
LOG(DEPLOYMENT SIZE$_{ijt-1}$)	100.0	0.94	0.002	0.94
MOBILITY INDEX$_{it}$	80.8	0.79	0.46	0.98
MOBILITY INDEX$_{it-1}$	100.0	3.56	0.43	3.42
MOBILITY INDEX$_{it-2}$	100.0	-4.28	0.30	-4.33
Posterior probability				0.81

Table 4.7: Bayesian Model Averaging for Model 1, Table 4.3

4.7.3 Network autocorrelation diagnostics

The following diagnostic tests were performed to evaluate spatial dependence in the data using network connectivity matrix W, as described in the

	p!=0	EV	SD	Best Model
LOG(DEPLOYMENT SIZE$_{ijt-1}$)	100	0.73	0.005	0.73
MOBILITY INDEX$_{it}$	0	0.00	0.00	.
MOBILITY INDEX$_{it-1}$	100	3.39	0.27	3.39
MOBILITY INDEX$_{it-2}$	100	-3.64	0.26	-3.64
Posterior probability				1

Table 4.8: Bayesian Model Averaging for Model 2, Table 4.3

	p!=0	EV	SD	Best Model
LOG(DEPLOYMENT SIZE$_{ijt-1}$)	100.0	0.93	0.002	0.93
MOBILITY INDEX$_{it}$	67.7	0.64	0.50	0.86
MOBILITY INDEX$_{it-1}$	100.0	3.58	0.42	3.37
MOBILITY INDEX$_{it-2}$	100.0	-4.25	0.31	-4.31
Additional covariates	...	...	...	...
Posterior probability				0.34

Table 4.9: Bayesian Model Averaging for Model 3, Table 4.3

	p!=0	EV	SD	Best Model
LOG(DEPLOYMENT SIZE$_{ijt-1}$)	100.0	0.73	0.005	0.73
MOBILITY INDEX$_{it}$	0.0	0.00	0.00	.
MOBILITY INDEX$_{it-1}$	100.0	3.26	0.27	3.26
MOBILITY INDEX$_{it-2}$	100.0	-3.73	0.26	-3.73
Additional covariates	...	...	...	...
Posterior probability				0.73

Table 4.10: Bayesian Model Averaging for Model 4, Table 4.3

main body of the paper.

In Table 4.11, I conducted a Moran's I diagnostic test (Moran, 1950), which provides a coefficient to evaluate the global level of spatial autocorrelation in the residuals of the estimated model. The observed Moran's I is compared to the expected value of Moran's I under the null hypothesis of no spatial autocorrelation, which is computed as $-\frac{1}{N-1}$. As indicated in Table 4.11, we can reject the null hypothesis and conclude that there is spatial autocorrelation in the model.

Diagnostic Test	*Test Statistic*	*p-value*
Moran's I	32.88	$< 2.2^{-16}$
Observed Moran's I	*Expectation*	*p-value*
0.05	-0.002	2^{-6}

Table 4.11: Moran's I Test for Spatial Dependence

Table 4.12 reports the results of Lagrange Multiplier (LM) diagnostics for spatial autocorrelation in the residuals of the estimated model, comparing the spatial error model (SEM, "ERR") and the spatial autoregression model (SAR, "LAG"). LMERR is the simple test statistic for the LM test for error dependence, LMLAG is the same but for the LM test for a missing spatially lagged dependent variable, RLMERR and RLMLAG are variants of these two tests that are robust to the possible presence of the other, and SARMA is a portmanteau test that is equivalent to LMERR + RLMLAG. In Table 4.12, all of the tests are statistically significant, meaning that there is spatial autocorrelation and both the SAR and SEM may each be viable models. However, looking at their robust counterparts, RLMLAG, has a smaller test statistic, suggesting that the SEM may likely be the better model to capture the spatial autocorrelation.

Next, Table 4.13 reports the results of the SAR, in case it is actually the better model to capture spatial autocorrelation. Substantively, these results are not markedly different from the SEM model: the long-run effect for MOBILITY INDEX is -1.63, and the spatial lag covariate indicates positive spatial autocorrelation. However, with the spatial lag covariate and the spatial weights matrix modeled explicitly in the model, the marginal effects of each covariate are somewhat less clear than a standard OLS model: not only are there direct effects, which are reflected in the values of the covariates in

Diagnostic Test	*Test Statistic*	*p-value*
LMERR	992.50	$< 2.2^{-16}$
LMLAG	228.08	$< 2.2^{-16}$
RLMERR	777.74	$< 2.2^{-16}$
RLMLAG	13.33	0.0003
SARMA	1005.80	$< 2.2^{-16}$

Table 4.12: Lagrange Multiplier Diagnostics for Spatial Dependence

Table 4.13, but there are indirect network effects as well. To explore these effects in more detail, Table 4.14 provides direct, indirect, and total impact measures for each of the covariates.

Finally, Table 4.15 compares the post-regression LM tests for both the SAR and SEM models, as well as each of these models' values for the Akaike information criterion (AIC) estimator (Akaike, 1973). The LM test for the SAR model indicates that there is still a large amount of spatial autocorrelation in the model's residuals, whereas the LM test for the SEM indicates that nearly all spatial autocorrelation is accounted for. Additionally, the AIC for the SEM is lower than for the SAR, which is preferred. These two post-regression diagnostics provide further support for the conclusion that the SEM is the more appropriate model for spatial autoregression with the data.

4.7.4 Instrumental variable regression diagnostics

The following diagnostic tests were performed to evaluate instrumental variable regression (IV) as described in section 4.7.6 below.

In Table 4.16, I report the results of the Wu-Hausman test (Nakamura and Nakamura, 1981). This test evaluates whether IV is as consistent as OLS, where the null hypothesis is that IV is equally consistent to OLS. If we fail to reject the null, then OLS is preferable to IV. However, rejecting the null hypothesis, which we can do based on the results in Table 4.16, means

	Dependent variable:
	LOG(DEPLOYMENT SIZE$_{ijt}$)
LOG(DEPLOYMENT SIZE$_{ijt-1}$)	0.73^{***}
	(0.005)
SIDE i COVARIATES	
MOBILITY INDEX$_{it}$	0.24
	(0.29)
MOBILITY INDEX$_{it-1}$	2.77^{***}
	(0.32)
MOBILITY INDEX$_{it-2}$	-3.45^{***}
	(0.28)
CINC$_{it}$	4.44^{***}
	(1.03)
POLITY SCORE$_{it}$	-0.001
	(0.002)
CINC$_{it}$/(CINC$_{it}$+CINC$_{jt}$)	-0.10
	(0.12)
TRADE$_{ijt}$/TOTAL TRADE$_{it}$	0.04
	(0.18)
SIDE j COVARIATE	
LOG(DEPLOYMENT SIZE)$_{jit}$	0.001
	(0.007)
DYADIC COVARIATES	
FOREIGN POLICY DIFFERENCE$_{ijt}$	0.02
	(0.01)
COMMON THREATS LAST 2 YEARS$_{ijt}$	0.006
	(0.006)
SPATIAL LAG COVARIATE	
ρ	0.24^{***}
	(0.02)
Observations	19,605
Directed dyads	1,079
Years	26

Note: Standard errors in parentheses. *p<0.1; **p<0.05; ***p<0.01

Table 4.13: Spatial Autoregression Model

Variable	*Direct*	*Indirect*	*Total*
LOG(DEPLOYMENT SIZE$_{ijt-1}$)	0.73	0.23	0.96
MOBILITY INDEX$_{it}$	0.24	0.07	0.31
MOBILITY INDEX$_{it-1}$	2.77	0.88	3.65
MOBILITY INDEX$_{it-2}$	−3.46	−1.09	−4.55
CINC$_{it}$	4.45	1.41	5.86
POLITY SCORE$_{it}$	−0.0005	−0.0002	−0.0007
CINC$_{it}$/(CINC$_{it}$+CINC$_{jt}$)	−0.11	−0.03	−0.14
TRADE$_{ijt}$/TOTAL TRADE$_{it}$	0.05	0.01	0.06
LOG(DEPLOYMENT SIZE)$_{jit}$	0.001	0.0005	0.002
FOREIGN POLICY DIFFERENCE$_{ijt}$	0.02	0.005	0.02
COMMON THREATS LAST 2 YEARS$_{ijt}$	0.006	0.002	0.008

Table 4.14: SAR impact measures

Diagnostic Test	*Test Statistic*	*p-value*
LMERR$_{SAR}$	320.99	2.2^{-16}
LMERR$_{SEM}$	5.78	0.02
Model	*AIC*	*Log likelihood*
SAR	24,513	−11,139.57
SEM	24,186	−10,976.15

Table 4.15: Comparisons between SAR and SEM models

therefore that IV is more consistent than OLS.

Finally, the Sargan test evaluates overidentifying restrictions, thereby indicating whether or not the instruments are exogenous (Sargan, 1958). The null hypothesis for this test is that the instruments are exogeneous. Based on the results in Table 4.16, we fail to reject the null hypothesis and can conclude that the instruments are valid.

4.7.5 Is U.S. behavior driving the results?

Because the U.S. was the dominant global military power throughout the time period analyzed in the data, it is possible that U.S. behavior is driving the results. To address this possibility, I re-ran each of the main models

Diagnostic Test	*df1*	*df2*	*Test Statistic*	*p-value*
WU-HAUSMAN	3	11,799	3.87	0.009
SARGAN	3		1.21	0.75

Table 4.16: Diagnostics for Instrumental Variables

for the three hypotheses in the main text with the U.S. excluded from the dataset. The results are listed in Tables 4.17, 4.18, and 4.19, respectively. The substantive results for the first two hypotheses remain the same as the results in the main text, supporting the argument that U.S. behavior is not driving the results in a unique way. For the third hypothesis in Table 4.19, the results are slightly different but should be interpreted carefully. Based on this table, it appears that for states other than the U.S., increased force projection capacity still decreases the probability that an ally will be the target of a militarized dispute and increases the probability that the state will initiate militarized disputes overall. However, increased force projection lowers the probability of being the target of a militarized dispute overall, which is different from the main results that include the U.S. Therefore, at most, we can conclude that increased force projection increases the likelihood that the U.S. will be the target of militarized disputes overall—if we believe that the U.S. is an outlier case. However, this is also a biased sample because it systematically excludes all the observations for the most powerful force projecting state, which has played a meaningful role in deterring threats and maintaining stability in the international system. Consequently, a more conservative (and generalizable) conclusion from these results would be that increased force projection capacity may increase the likelihood that the dominant power will be the target of militarized disputes, while it decreases the likelihood of being the target of militarized disputes for less powerful states.

4.7.6 Tests for endogeneity

An additional modeling challenge to Eq. 4.1 is possible endogeneity between deployment levels in alliances and force projection capacity: in particular, technological investments and force posture policy decisions might predict force projection capacity, rather than (or in addition to) the other way around. To deal with this threat to causal inference, I employ instrumental variable estimation, using state *i*'s past ruling party political ideology in the government as instruments for MOBILITY INDEX$_{it}$. To code political ideology, I use the Database of Political Institutions (Cruz, Keefer, and Scartascini, 2018), which provides a categorical coding (left, center, right) for the executive and opposition parties' ideological orientation.[17] I expect past ruling party ideology, conditional on other covariates, to be a strong predictor of military spending and thus force projection capacity, with right-leaning parties likely to spend more on the military than left-leaning parties. To illustrate this relationship, Table 4.20 reports the results of a first-stage regression with MOBILITY INDEX$_{it}$ regressed on PARTY IDEOLOGY$_{it-2}$, and likewise for each of the other lagged instruments, which I explain in more detail below. Because PARTY IDEOLOGY is a categorical variable, the results are reported relative to the baseline, which is a right-leaning government. The results in Table 4.20 illustrate that, as ideology moves from right to left, there is a negative relationship between party ideology and force projection capacity.

I address the possibility that MOBILITY INDEX is endogenous through instrumental variable estimation, using past domestic political ideology as an instrument. Table 4.21 reports the results of this estimation. About one-third of the observations from the original results are no longer available because of missing data for domestic political regime ideology. Additionally, the standard errors are larger, though this can occur due to weaker in-

[17] I used the following rules to code state *i*'s ruling party political ideology: 1) if the executive is not competitively elected, then use the executive's ideology; otherwise, if the executive is competitively elected and 2a) the opposition's vote share in the legislature is less than or equal to 50%, then use the executive's ideology; 2b) the opposition's vote share in the legislature is greater than 50%, then use the opposition's ideology.

struments. However, even with fewer observations and potentially weaker instruments, the results in Table 4.21 are similar to those in Models 4 and 5 in Table 4.3, including the estimate for the long-run effect of MOBILITY INDEX, providing further empirical support for the first hypothesis.

According to the first-stage F-statistics reported in Table 4.21, the instruments appear strong enough to obtain the expected outcome. Additionally, because the coding for domestic political ideology is categorical, the endogenous variables are over-identified, which allows me to test instrument exogeneity. Based on diagnostic checks (available in the appendix), the results of the Sargan test for instrument exogeneity indicate that all of the instruments are valid and exogenous. Therefore, if we are willing to accept that the instrumental variable assumptions are valid, the instruments satisfy the exclusion restriction and we can conclude that the results in Table 4.21 provide consistent estimates.

Another possibility is that the instruments are not perfectly exogenous, but are "plausibly exogenous" to the extent that the exclusion restriction is not perfectly met, but is close enough that we can still infer causality. To evaluate whether the instruments are plausibly exogenous, I use Conley, Hansen, and Rossi (2012)'s technique of relaxing the IV exclusion restriction by defining a parameter, γ, that reflects the degree to which the instruments may diverge from true exogeneity. By removing the direct effect of the instrument on the dependent variable at different values for γ and then evaluating the coefficients for the endogenous parameters in an IV regression, I can determine how large γ would need to be for the coefficient estimates to be insignificant.

Figure 4.5 plots the range of possible coefficient values across γ for the instantaneous effect of MOBILITY INDEX in the IV regression. If we have a prior belief that the direct effect of the instrument on deployment levels is negative—i.e., that left-leaning ruling parties have a depressing effect on troop deployment levels, compared to right-leaning parties—then only the range of values where γ is negative are relevant. Given this belief, Figure 4.5 therefore shows that the direct (negative) effect of the instrument, γ, must be approximately -0.027 in order for the instantaneous effect of MOBILITY INDEX on deployment levels to be zero. Substantively, this means that the direct effect of left-leaning governments must decrease deployment levels by 2.7% in order for the instantaneous effect of force projection capacity to be removed.

Furthermore, in Figure 4.6 I plot the range of possible values across γ for the long-run effect of MOBILITY INDEX. The long-run effect becomes *stronger* as we relax the exclusion restriction by making γ more negative. This result provides strong evidence that I have identified the correct causal direction for the main argument that force projection capacity will tend to decrease deployment levels. The long-run effect is removed only when γ is equal to approximately 0.04, meaning that left-leaning governments would need to *increase* deployment levels by 4% in order for the long-run effect of force projection capacity to be removed.

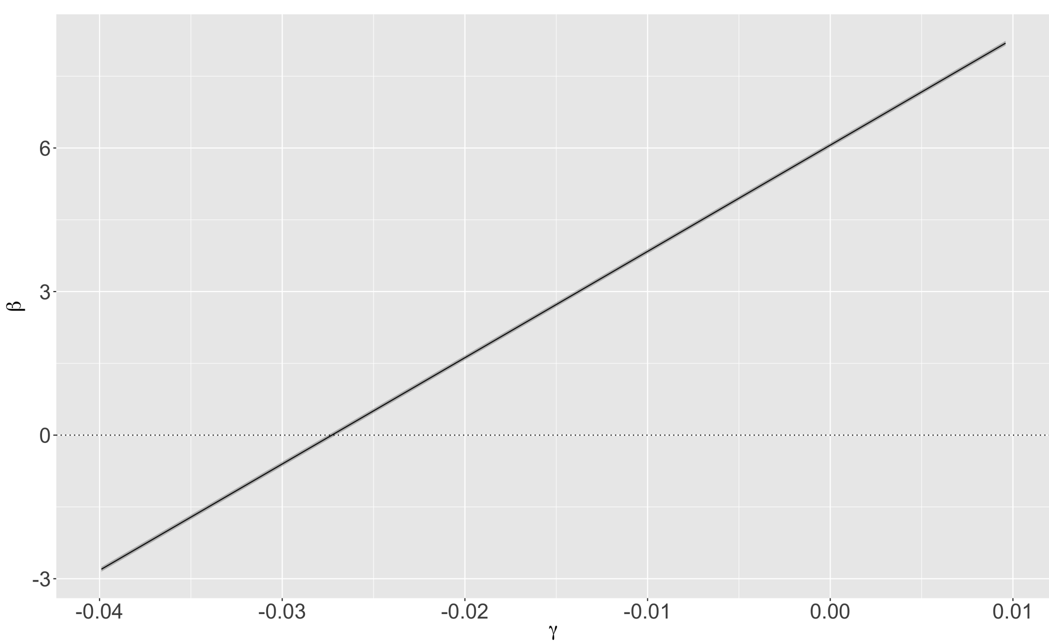


Figure 4.5: Range of possible coefficient values across γ for MOBILITY INDEX in the IV regression with 95% confidence interval.

I also explored other approaches to causal inference in case political

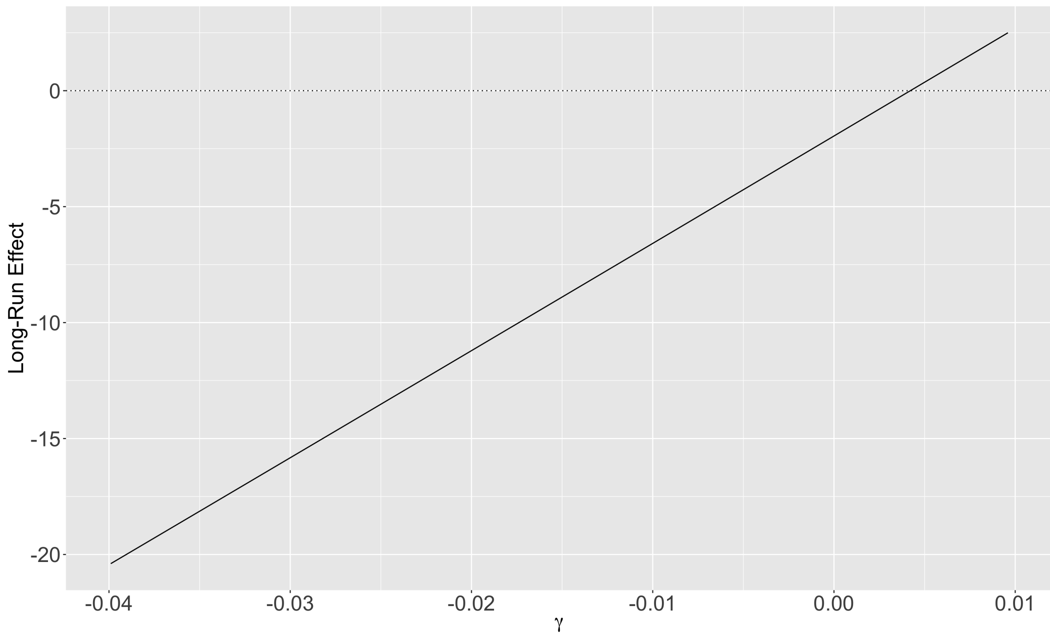


Figure 4.6: Range of possible long-run effect values of MOBILITY INDEX across γ in the IV regression.

party ideology is not a valid instrument. One approach I considered was to use commodity shocks as an instrument, in line with previous research that has also explored commodity shocks as an identification strategy (e.g., Bazzi and Blattman, 2014; Carreri and Dube, 2017). A key difficulty in using commodity shocks as an instrument for MOBILITY INDEX, however, is that cross-nationally, states respond to shocks in different ways. Depending on the domestic economy, political preferences about how a government budget is allocated, and other domestic variables that I do not measure, states will vary in the way they respond to shocks. Moreover, I considered only looking at specific commodities such as steel, iron, and other metals that might affect production costs for aircraft and naval capabilities. However, this distinction poses separate challenges because these shocks would only correlate with the production of new capabilities, but it would have no measurable impact on capabilities that have already been built or purchased and are operational. Separately, I also considered using covariate balancing generalized propensity score (CBGPS) matching (Fong, Hazlett, and Imai, 2018). However, this method has separate limitations that make estimating causal effects in my model particularly challenging. In particular, due to the time-series-cross-section data structure, matching methods such as CBGPS do not account for the large number of repeated measurements on the same units. While other methods are being developed to deal with this issue (e.g., Imai, Kim, and Wang, 2019), currently they only allow for a binary treatment variable, whereas MOBILITY INDEX is continuous, making these other methods less ideal for my purposes.

	Dependent variable:			
	LOG(DEPLOYMENT SIZE$_{ijt}$)			
	(1)	(2)	(3)	(4)
LOG(DEPLOYMENT SIZE$_{ijt-1}$)	0.93***	0.72***	0.92***	0.72***
	(0.002)	(0.006)	(0.003)	(0.006)
MOBILITY INDEX$_{it}$	6.42***	1.53**	5.72***	1.08
	(0.60)	(0.71)	(0.63)	(0.70)
MOBILITY INDEX$_{it-1}$	−7.45***	−5.60***	−7.08***	−5.25***
	(0.82)	(0.73)	(0.81)	(0.71)
MOBILITY INDEX$_{it-2}$	0.71***	0.0004	0.36	0.36
	(0.47)	(0.54)	(0.48)	(0.56)
CINC$_{it}$			1.70***	4.83***
			(0.32)	(1.33)
POLITY SCORE$_{it}$			−0.0001	−0.001
			(0.0004)	(0.002)
CINC$_{it}$/(CINC$_{it}$+CINC$_{jt}$)			0.009	−0.05
			(0.01)	(0.12)
TRADE$_{ijt}$/TOTAL TRADE$_{it}$			0.25**	0.08
			(0.10)	(0.21)
LOG(DEPLOYMENT SIZE)$_{jit}$			0.004	−0.002
			(0.004)	(0.009)
FOREIGN POLICY DIFFERENCE$_{ijt}$			0.04***	0.02
			(0.008)	(0.01)
COMMON THREATS LAST 2 YEARS$_{ijt}$			0.02***	0.02***
			(0.005)	(0.007)
LOG(DISTANCE$_{ij}$)			0.006	
			(0.004)	
COLONIAL HISTORY$_{ij}$			0.13***	
			(0.02)	
COLD WAR$_t$			−0.006	
			(0.008)	
YEAR TREND$_t$			−0.0008**	
			(0.0004)	
MOBILITY INDEX (LONG RUN EFFECT)	−4.56	−14.45	−12.18	−13.42
Year fixed effects		✓		✓
Directed dyad fixed effects		✓		✓
Observations	17,584	17,584	17,584	17,584
Clusters	987	987	987	987
Years	26	26	26	26
LM autocorrelation test statistic	49.52	4.73	47.34	6.32

Notes: Panel corrected standard errors in parentheses. *p<0.1; **p<0.05; ***p<0.01

Table 4.17: Determinants of Alliance Deployment Levels (excluding U.S.), 1985–2010

	Dependent variable:
	PR(DEPLOYMENT$_{ijt}$)
MOBILITY INDEX$_{it}$	3.10***
	(0.47)
CINC$_{it}$	10.10***
	(2.05)
POLITY SCORE$_{it}$	0.07***
	(0.01)
CINC$_{it}$/(CINC$_{it}$+CINC$_{jt}$)	1.74***
	(0.33)
TRADE$_{ijt}$/TOTAL TRADE$_{it}$	2.24
	(1.39)
DEPLOYMENT$_{jit}$	0.49
	(0.51)
POLITY SCORE$_{jt}$	−0.06***
	(0.009)
LOG(TERRORIST FATALITIES LAST 2 YEARS$_{jt}$)	0.31***
	(0.02)
FOREIGN POLICY DIFFERENCE$_{ijt}$	−0.17**
	(0.07)
COMMON THREATS LAST 2 YEARS$_{ijt}$	0.62***
	(0.09)
HOSTILE MILITARIZED DISPUTES LAST 2 YEARS$_{ijt}$	0.38**
	(0.17)
DEFENSE ALLIANCE$_{ijt}$	−0.32
	(0.22)
LOG(DISTANCE$_{ij}$)	−0.92***
	(0.06)
COLONIAL HISTORY$_{ij}$	0.62***
	(0.24)
COLD WAR$_{t}$	0.49***
	(0.13)
TIME SINCE LAST DEPLOYMENT SPLINE$_{ijt}$	−0.01
	(0.008)
Observations	231,026

Note: Robust standard errors in parentheses clustered by directed dyad. *p<0.1; **p<0.05; ***p<0.01

Table 4.18: Determinants of Deployment Onset Probability (excluding U.S.), 1985–2010

	Dependent variable:		
	Pr(Ally j MID target$_{ijt}$)	Pr(MID$_{ijt}$)	
	(1)	(2)	(3)
Mobility index$_{it}$	-4.75^{***}	0.89^{*}	-3.55^{***}
	(1.18)	(0.47)	(0.97)
CINC$_{it}$	14.19^{***}	18.67^{***}	23.87^{***}
	(2.19)	(0.83)	(1.11)
Polity score$_{it}$	-0.04^{***}	0.01^{***}	0.02^{***}
	(0.004)	(0.004)	(0.006)
CINC$_{it}$/(CINC$_{it}$+CINC$_{jt}$)	-1.82^{***}	-2.47^{***}	-3.27^{***}
	(0.07)	(0.10)	(0.14)
Polity score$_{jt}$	0.02^{***}	-0.03^{***}	-0.03^{***}
	(0.005)	(0.004)	(0.005)
Foreign policy difference$_{ijt}$	0.31^{***}	0.07^{**}	0.05
	(0.06)	(0.03)	(0.04)
Defense alliance$_{ijt}$		-1.10^{***}	-1.14^{***}
		(0.11)	(0.16)
log(Distance$_{ij}$)	-0.17^{***}	-1.62^{***}	-1.57^{***}
	(0.03)	(0.03)	(0.04)
Time since j last MID target spline$_{ijt}$	-0.04^{***}		
	(0.003)		
Time since last dyadic MID spline$_{ijt}$		0.001	-0.02^{***}
		(0.004)	(0.005)
Observations	17,584	236,446	236,446
Note: Robust standard errors in parentheses clustered by directed dyad.			*p<0.1; **p<0.05; ***p<0.01

Table 4.19: Determinants of Militarized Dispute Probability (excluding U.S.), 1985–2010

	Dependent variable:		
	MOBILITY INDEX$_{it}$	MOBILITY INDEX$_{it-1}$	MOBILITY INDEX$_{it-2}$
	(1)	(2)	(3)
PARTY IDEOLOGY (CENTER)$_{it-2}$	0.005***		
	(0.002)		
PARTY IDEOLOGY (LEFT)$_{it-2}$	-0.01^{***}		
	(0.002)		
PARTY IDEOLOGY (CENTER)$_{it-3}$		0.003	
		(0.002)	
PARTY IDEOLOGY (LEFT)$_{it-3}$		-0.01^{***}	
		(0.002)	
PARTY IDEOLOGY (CENTER)$_{it-4}$			-0.0002
			(0.002)
PARTY IDEOLOGY (LEFT)$_{it-4}$			-0.015^{***}
			(0.002)
Additional covariates	✓	✓	✓
Year fixed effects	✓	✓	✓
Directed dyad fixed effects	✓	✓	✓
Note: Panel corrected standard errors in parentheses.			*p<0.1; **p<0.05; ***p<0.01

Table 4.20: First-Stage Estimation

	Dependent variable:
	LOG(DEPLOYMENT SIZE$_{ijt}$)
LOG(DEPLOYMENT SIZE$_{ijt-1}$)	0.75***
	(0.008)
SIDE *i* COVARIATES	
MOBILITY INDEX$_{it}$	6.06**
	(2.95)
MOBILITY INDEX$_{it-1}$	-6.13^{*}
	(3.43)
MOBILITY INDEX$_{it-2}$	-0.41
	(2.48)
CINC$_{it}$	9.55
	(6.61)
POLITY SCORE$_{it}$	-0.006
	(0.005)
CINC$_{it}$/(CINC$_{it}$+CINC$_{jt}$)	-0.46
	(0.37)
TRADE$_{ijt}$/TOTAL TRADE$_{it}$	-0.27
	(0.24)
SIDE *j* COVARIATE	
LOG(DEPLOYMENT SIZE)$_{jit}$	0.02
	(0.01)
DYADIC COVARIATES	
FOREIGN POLICY DIFFERENCE$_{ijt}$	0.04
	(0.03)
COMMON THREATS LAST 2 YEARS$_{ijt}$	0.02***
	(0.008)
MOBILITY INDEX (LONG RUN EFFECT)	-1.96
Year fixed effects	✓
Directed dyad fixed effects	✓
Observations	12,654
Clusters	816
Years	26
PARTY IDEOLOGY$_{it-2}$ F-statistic	103.63
PARTY IDEOLOGY$_{it-3}$ F-statistic	126.00
PARTY IDEOLOGY$_{it-4}$ F-statistic	154.91

Note: Panel corrected standard errors in parentheses. *p<0.1; **p<0.05; ***p<0.01

Table 4.21: Instrumental Variables Estimation

CHAPTER 5

Conclusion

> We stood all alone a year ago, and to many countries it seemed that our account was closed, we were finished . . . But instead our country stood in the gap. There was no flinching and no thought of giving in; and by what seemed almost a miracle to those outside these Islands, though we ourselves never doubted it, we now find ourselves in a position where I say that we can be sure that we have only to persevere to conquer.
>
> —Winston Churchill
> London, UK, October 29, 1941
> (Churchill, 1941)

This dissertation explores the causes and consequences of military commitment problems in order to answer the following question: in dynamic environments where conditions can change quickly and substantially over time, why do military commitment problems arise, how do they impact states' ability to deter threats and fight wars, and how can states overcome them? To answer this question, I investigate three factors that are necessary for military organizations to be effective in dynamic situations: resolve, adaptation, and flexibility.

In Chapters 2 through 4, I focus on a particular factor within a given context to evaluate military effectiveness. These chapters provide a demonstration of how resolve, adaptation, and flexibility contribute to military effectiveness in the context of conventional conflict, unconventional conflict, and international security, respectively.

Each of these three factors is related to a different commitment problem that military forces often have to confront. First, high signals of resolve within an army can make the commitment to fight credible, such that commanders and soldiers believe fighting in combat is their best option, rather than fleeing or surrendering. Second, a high level of commitment to a conflict by political leaders can create better conditions for the military to adapt to novel situations and improve their doctrines. Finally, high flexibility through improved force projection capabilities can make security commitments to other states credible, as it allows military forces to respond to crises more quickly and efficiently.

5.1 Summary of findings

Chapter 2: "Until the Bitter End? The Diffusion of Surrender Across Battles"

Why do some armies fight until the bitter end, but others collapse and surrender? Existing research has highlighted the importance of battlefield resolve for the onset, conduct, and outcome of war, but has left these life-and-death decisions mostly unexplained. We know little about why battle-level surrender occurs, and why it stops.

In this chapter, I argue that surrender emerges from a collective-action problem: success in battle requires that soldiers choose to fight as a unit rather than flee, but individual decisions to fight depend on whether soldiers expect their comrades to do the same, all else equal. As a result, surrender becomes contagious across battles because soldiers take cues from what other soldiers did when they were in a similar position. Conversely, where no recent precedent exists, mass surrender is unlikely.[1]

I find empirical support for this claim using a new data set of conventional battles in all interstate wars from 1939 to 2011. The results offer several contributions to research on interstate conflict. I demonstrate that battlefield surrender can be contagious because of a collective-action problem within military organizations. Success in battle requires that soldiers fight as a cohesive unit, but individual decisions to fight depend on whether soldiers expect their comrades to do the same. As troops learn of past decisions to surrender within their own army, they lose confidence in their unit's resolve, and decide to flee rather than fight. This pattern is particularly acute if the expected costs of surrender are also low—either because troops believe the opponent will treat prisoners well, or because senior officers have recently surrendered, shaking the credibility of threats to punish desertion and surrender by the rank and file.

In addition to diffusion, I examined several alternative explanations of surrender. I found tentative, if mixed, support for a few factors that might affect the parameters of the collective action model—like international law, principal-agent problems, and offensive advantages. However, I found no evidence that surrender depends on political regime type, recruitment methods, or relative national power. Although data limitations prevent us from directly testing several other explanations—mutual surveillance, discipline, and ideology—I sought to at least account for them econometrically, through combatant and war fixed effects. I also demonstrated that it is information specifically on past surrender, rather than military effectiveness generally, that drives soldiers' decisions.

The determinants of surrender are theoretically important because these life-and-death choices tend to resonate well beyond individual battles. Although previous research suggests that combatants acquire information

[1]Military organizations can devise ways to try to manipulate soldiers' incentives, which can create some variation in the way collective action is observed empirically. The Soviet use of "blocking detachments" (Lyall and Zhukov, 2020), for example, may have helped prevent more soldiers from surrendering in the Winter of 1941/2 even though many Soviet soldiers had surrendered a few months before.

about war-fighting resolve through battle outcomes, scholars often treat resolve as an exogenous cause of war termination. In the approach taken in this chapter, by contrast, battlefield resolve is of primary theoretical interest. If wars are a continuation of political bargaining, reconciling informational asymmetries through the use of force, then understanding the mechanisms and processes influencing battlefield resolve is crucial for explaining and predicting bargaining outcomes. This chapter's results illustrate that wartime resolve does not depend solely on political leaders' assessment of probabilistic battlefield outcomes. Instead, military officers and their troops are the primary actors mutually influencing each other's behavior. Because soldiers' choices in future battles depend on precedents set by others in the past, it is these cascading battlefield decisions that ultimately shape and constrain leaders' choices. South Vietnam's final rout in 1975 illustrates how surrender can cascade across an army; likewise, German resolve to fight until the bitter end in 1945 illustrates how the decision to continue fighting also has its own "cascading" effect by signaling a decision to continue fighting.

Chapter 3: "Adapting Counterinsurgency Doctrine in the Shadow of the Future"

Modern insurgencies create significant evolutionary pressures on conventional militaries to improve their doctrines and overall military effectiveness. Why are some militaries more effective at evolving their counterinsurgency doctrine than others?

This chapter develops a theory which argues that complex local conditions force militaries to optimize their organizations for the commitment horizon they possess in a conflict. The commitment horizon is the extent to which the future matters for the present—i.e., how much the "shadow of the future" (Axelrod, 1984) weighs on a military's decision-making—and is related to the concept of resolve, i.e., the "firmness or steadfastness of purpose" (Kertzer, 2017). The optimization based upon commitment horizon drives militaries to take on particular organizational characteristics, such as the amount of delegation and the tolerance for experimenting with new tactics, which affect their performance in systematic ways as a result.

Using an agent-based model, I evaluate the theory by simulating different organizational characteristics and evaluating how doctrinal effectiveness changes based on different commitment horizons. Results from the simulations suggest a typology of counterinsurgent militaries according to their commitment horizon. Contrary to existing arguments that counterinsurgency effectiveness is determined by a particular set of strategies or by a fixed, preexisting military culture, the results also suggest that doctrinal effectiveness follows a U-shaped curve as the commitment horizon increases. To illustrate the logic of the model, I operationalize commitment horizon and describe the typology and its implications in several historical counterinsurgency cases: the U.S. in Iraq and Afghanistan; the British in Malaya and Iraq; and the French in Indochina and Algeria.

This chapter contributes to our understanding of the determinants of military doctrine and battlefield effectiveness in counterinsurgency operations by providing a novel approach to explaining military adaptation during wartime. I have shown that militaries waging counterinsurgencies can be categorized according to a typology defined by their commitment horizons. Depending on how large the shadow of the future looms, militaries will adopt certain organizational features to optimize their performance in a complex environment. These features—the level of military autonomy and the tolerance for political and military experimentation—directly affect the extent to which learning and doctrinal adaptation will occur, and thus overall performance. As the results of my model indicate, overall counterinsurgency effectiveness follows a U-shaped curve as the commitment horizon increases. Those militaries with a moderate commitment horizon tend to perform most poorly, as they lack both the urgency to adapt and the drive to persist and search for better solutions over time. On the other hand, the militaries that experience the greatest overall effectiveness are those that have the highest commitment level, because the organizational characteristics adopted allow the political and military levels to work in conjunction to find improvements over time. Furthermore, this chapter demonstrates that military culture—as manifested in the level of military autonomy, and the tolerance for political and military experimentation—can develop through processes within the organization itself in response to a state's commitment horizon.

Chapter 4: "Projecting Credibility: Alliance Commitments and Foreign Policy Pursuits"

How does the ability to quickly project large-scale military force—such as with transport aircraft and naval aircraft carriers—affect states' alliance commitments and the likelihood of militarized disputes with other states?

In this chapter, I argue that states with greater capacity to project force can deploy fewer troops to allies for deterrent purposes, instead combining their capacity to respond swiftly in a crisis with public words and actions—such as statements by political leaders—to demonstrate their commitment to allies. However, greater force projection capacity is a double-edged sword: although it helps deter threats against allies, it also increases a state's likelihood of initiating militarized disputes in other parts of the world.

I evaluate the effect of force projection capacity on troop deployments using a new index for force projection capacity from 1985 to 2018. The empirical evidence in this chapter demonstrates that an increase in force projection capacity has a depressing effect on deployment levels. States tend to rely less on costly signaling in their alliance commitments as their force projection capacity increases, instead relying more on tying hands signaling in order underwrite the credibility of their alliances. However, audience costs alone are generally not sufficient in demonstrating credible commitments. Instead, audience costs signals are supported by the capacity to project force with high speed and volume.

Moreover, the results in this chapter have important implications for the changing nature of deterrence. With increased force projection capacity, the distinction between immediate deterrence and general deterrence blurs: both occur through flexible force projection, which is more prevalent when there is a high capacity to project force. This increased level of flexibility to respond to crises has an additional effect on states' force projection behavior: the cost savings from fewer deployments to allies create new opportunities to pursue foreign policy interests in other areas without needing to generate strong commitment signals. Consequently, increased force projection capacity increases the propensity for states to generate new deployments into other countries. These can occur through both crisis responses and limited security partnerships, such as training foreign military forces or conducting combined antiterrorism operations. Furthermore, an increased capacity to project force increases the likelihood of initiating militarized disputes with other states, but it also works to effectively deter militarized disputes against allies.

5.2 Directions for future research

This dissertation opens several future research avenues. For example, although I have demonstrated that surrender can have a contagion effect across battles, we do not analyze how this process begins within battles, and what critical events must occur to jump-start surrender and its subsequent diffusion. While my focus in Chapter 2 is on inter-battle dynamics, a more explicit focus on intra-battle behavior is needed to understand the conditions leading to initial organizational breakdown.

Further research is also needed to understand how different political-military institutions affect whether the diffusion process occurs, or whether it can be reversed. We know little about why some military organizations can absorb losses and adapt to changing circumstances, while others are unable to recover from battles in which soldiers surrendered en masse. By disaggregating wars into battles and stepping away from the classical approach of treating the military as a unitary actor, we can better understand how collective action dynamics affect battlefield outcomes and, ultimately, decisions to initiate, continue or terminate war.

Chapter 3 opens up at least two avenues for future research. First, I do not model varying organizational structures, such as one with an uneven number of agents between vertical or horizontal levels, or one that is more networked in its connectivity between agents. Future research could extend the model to explicitly evaluate how changes in organizational design and structure affect doctrinal consensus. Second, while I focus on the role of the commitment horizon in determining counterinsurgency effectiveness, there are also implications for civil-military relations that I do not systematically explore. For example, the varying levels of military autonomy and the tolerance for political and military experimentation each have implications about the relative distribution of decision-making power between civil and military authorities and the relative openness to change, respectively. Future research could more explicitly evaluate the implications arising from different organizational arrangements, and how democratic and autocratic regimes might vary in this regard.

Additionally, the findings in Chapter 4 suggest a number of other areas for future research. For example, while I control for regime type in my analysis, I do not explicitly evaluate the differences in signaling across political systems. Instead, the models I use assume there is no difference in motivations for signaling based on regime type. However, since a lack of sunk-cost signaling in alliances indicates that audience cost signals are therefore more important in signaling credibility, are democratic regimes more likely to use rapid mobility as a substitute for troop deployments? Conversely, are autocratic regimes more likely to continue to rely on deployments to signal credibility? Similarly, do autocracies that develop rapid mobility capabilities tend to have fewer defense alliances, and if so, given their tendency to initiate more militarized disputes, do they therefore use such capabilities in a more offensive manner, or are such capabilities used primarily for immediate deterrence?

5.3 Policy implications

In addition to opening up new research avenues, this dissertation also provides several important policy implications. For example, in Chapter 2, we showed that if political authorities wish to maintain the resolve of their armies in battle, they should worry less about how dangerous a combat environment is likely to be, and more about how to address recent precedents for mass surrender.

Additionally, the results in Chapter 3 suggest that only a certain type of organization is likely to perform most effectively at counterinsurgency: namely, one which has the highest commitment level in a conflict. This suggests that political and military leaders evaluating whether to engage in counterinsurgency operations would be well served in carefully examining how committed they are, to include the commitment of the selectorate—i.e., the individuals who can choose the political leader (Bueno de Mesquita, Smith, Siverson et al., 2003). If policymakers identify less than the utmost level of commitment, they may find that avoiding such a conflict may be better in the long run. Instead, if they do have a very high commitment horizon, then adopting the organizational features that encourage the military and political leaders to work in a dialectic approach to adaptation as a coordinated organization will help them achieve the greatest overall effectiveness in a counterinsurgency.

Furthermore, Chapter 4 has several important policy implications for the changing international security environment. For example, it helps explain what policymakers can expect from China as it expands its force projection capacity. China launched its first aircraft carrier in 2011, followed by a second carrier in 2017 and plans to construct four additional carriers by 2035 (Chan and Rui, 2019). As China continues to increase its capacity to project force, policymakers can expect further Chinese involvement in foreign security partnerships where its interests lie. However, China will likely pursue these interests without making many, if any, strong commitments to new defensive alliances. Doing so will provide China with greater flexibility in pursuit of its foreign policy interests, but will also limit its ability to deter or compel other states from meddling in its interests in the long run, thereby increasing the chance that China will become involved in militarized disputes.

Additionally, there are important policy implications for the NATO alliance, which has recently faced challenges to its long-run sustainability and legitimacy. For example, the U.S. has been evaluating the costs of maintaining the current 35,000 troops deployed to Germany. These challenges have raised alarms in Europe about the strength of NATO credibility if the U.S. were to withdraw some its troops from the region. However, Chapter 4 provides reassurance that the alliance can still remain credible so long as U.S. political leaders maintain both their public commitment to the alliance and the military's force projection capacity, even if troops were to withdraw from Germany or other parts of the European continent (assuming the security situation with Russia does not change significantly).

Finally, there are also cautionary implications for policymakers based on the findings presented in Chapter 4. While developing greater force projection capabilities may reduce infrastructure costs of overseas bases, there is a risk of over-promising within alliances when too many of them are underwritten by force projection capacity without any substantial sunk costs. Force projection is a limited resource, and being spread too thin can come at great cost to alliance credibility, and ultimately reputation, if the likelihood of threats increases and not enough resources are available to respond to them all. Thus policymakers must still consider the expects costs of stationing troops versus the expected threat likelihood and the cost of responding to crises with ad hoc, flexible deployments.

In summary, the organizational solutions to commitment problems I consider—resolve, adaptation, and flexibility—directly influence military effectiveness, and along with other types of commitment solutions, provide a better framework for understanding military effectiveness than existing approaches. For both scholars and practitioners alike, this dissertation demonstrates that a careful study of military commitment problems can improve our evaluation of military effectiveness as well as our understanding of how states build and sustain effective military organizations.

BIBLIOGRAPHY

Adamsky, Dima, and Kjell Inge Bjergam, eds. 2012. *Contemporary Military Innovation: Between Anticipation and Adaptation*. New York, NY: Routledge.

Akaike, Hirotugu. 1973. Information Theory and an Extension of the Maximum Likelihood Principle. In *Proceedings of the Second International Symposium on Information Theory*, edited by B.N. Petrov and F. Caski, 267–281. Budapest, Hungary: Akademiai Kiado.

Akam, Simon. 2018. *The Changing of the Guard: The British Army since 9/11*. London, UK: Scribe.

Aklin, Michaël, and Andreas Kern. 2019. Moral Hazard and Financial Crises: Evidence from American Troop Deployments. *International Studies Quarterly* 63:15–29.

Allen, Michael A., Julie VanDusky-Allen, and Michael E. Flynn. 2016. The Localized and Spatial Effects of US Troop Deployments on Host-State Defense Spending. *Foreign Policy Analysis* 12 (4):674–694.

Ambler, John Steward. 1966. *The French Army in Politics 1945-1962*. Columbus, OH: Ohio State University Press.

Angeletos, George-Marios, Christian Hellwig, and Alessandro Pavan. 2007. Dynamic Global Games of Regime Change: Learning, Multiplicity, and the Timing of Attacks. *Econometrica* 75 (3):711–56.

Anselin, Luc. 1988. *Spatial Econometrics: Methods and Models*. Dordrecht, Netherlands: Kluwer.

Arreguín-Toft, Ivan. 2001. How the Weak Win Wars: A Theory of Asymmetric Conflict. *International Security* 26 (1):93–128.

Avant, Deborah D. 1993. The Institutional Sources of Military Doctrine–Hegemons in Peripheral Wars. *International Studies Quarterly* 37 (4):409–430.

———. 1994. *Political Institutions and Military Change: Lessons from Peripheral Wars*. Ithaca, NY: Cornell University Press.

Axelrod, Robert. 1984. *The Evolution of Cooperation*. New York, NY: Basic Books, revised edn.

———. 1997. *The Complexity of Cooperation: Agent-Based Models of Competition and Collaboration*. Princeton, NJ: Princeton University Press.

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

Barbieri, Katherine, Omar Keshk, and Brian Pollins. 2008. Correlates of War project trade data set codebook, v. 2.0. Available at http://correlatesofwar.org.

Barno, David, and Nora Bensahel. 2020. The Role of Doctrine. In *Adaptation under Fire: How Militaries Change in Wartime*, edited by David Barno and Nora Bensahel, chap. 2. New York, NY: Oxford University Press.

Bartov, Omer. 1992. *Hitler's Army: Soldiers, Nazis, and War in the Third Reich*. New York: Oxford University Press.

Baumann, Oliver, Jens Schmidt, and Nils Stieglitz. 2019. Effective Search in Rugged Performance Landscapes: A Review and Outlook. *Journal of Management* 45 (1):285–318.

Bazzi, Samuel, and Christopher Blattman. 2014. Economic Shocks and Conflict: Evidence from Commodity Prices. *American Economic Journal: Macroeconomics* 6 (4):1–38.

Beck, Nathaniel, Kristian Skrede Gleditsch, and Kyle Beardsley. 2006. Space is More than Geography: Using Spatial Econometrics in the Study of Political Economy. *International Studies Quarterly* 50:27–44.

Beck, Nathaniel, and Jonathan N. Katz. 1995. What to Do (and Not to Do) with Time-Series Cross-Section Data. *American Political Science Review* 89 (3):634–647.

———. 2011. Modeling Dyanmics in Time-Series-Cross-Section Political Economy Data. *Annual Review of Political Science* 14:331–352.

Beck, Nathaniel, Jonathan N. Katz, and Richard Tucker. 1998. Taking Time Seriously: Time-Series-Cross-Section Analysis with a Binary Dependent Variable. *American Journal of Political Science* 42 (4):1260–1288.

Belkin, Aaron, Michael Clark, Gulriz Gokcek, Robert Hinckley, Thomas Knecht, and Eric Patterson. 2002. When Is Strategic Bombing Effective? Domestic Legitimacy and Aerial Denial. *Security Studies* 11 (4):51–88.

Bell, Sam R., K. Chad Clay, and Carla Martinez Machain. 2017. The Effect of US Troop Deployments on Human Rights. *Journal of Conflict Resolution* 61 (10):2020–2042.

Bernstein, Serge. 1993. *The Republic of de Gaulle 1958-1969*. Cambridge, UK: Cambridge University Press.

Betts, Richard K. 1991. *Soldiers, Statesmen, and Cold War Crises*. McGill-Queen's Press-MQUP.

Biddle, Stephen. 2004. *Military Power: Explaining Victory and Defeat in Modern Battle*. Princeton, NJ: Princeton University Press.

———. 2021. *Nonstate Warfare: The Military Methods of Guerrillas, Warlords, and Militias*. Princeton, NJ: Princeton University Press.

Biddle, Stephen, Jeffrey A. Friedman, and Jacob N. Shapiro. 2012. Testing the Surge: Why Did Violence Decline in Iraq in 2007? *International Security* 37 (1):7–40.

Biddle, Stephen, and Stephen Long. 2004. Democracy and Military Effectiveness: A Deeper Look. *Journal of Conflict Resolution* 48 (4):525–46.

Biglaiser, Glen, and Jr. DeRouen, Karl. 2007. Following the Flag: Troop Deployment and U.S. Foreign Direct Investment. *International Studies Quarterly* 51:835–854.

———. 2009. The Interdependence of U.S. Troop Deployments and Trade in the Developing World. *Foreign Policy Analysis* 5 (3):247–263.

Blaise, Lilia, Eric Schmitt, and Carlotta Gall. 2019. U.S. and Tunisia Are Fighting Militants Together. Just Don't Ask Them About It. *The New York Times* March 2, 2019.

Blechman, Barry M., and Stephen S. Kaplan. 1978. *Force Without War: U.S. Armed Forces as a Political Instrument*. Brookings Institution.

Bondy, J.A., and U.S.R. Murty. 1976. *Graph Theory with Applications*. New York, NY: North Holland.

Boot, Max. 2013. *Invisible Armies: An Epic History of Guerrilla Warfare from Ancient Times to the Present*. New York, NY: Liveright.

Boulding, Kenneth E. 1962. *Conflict and Defense: A General Theory*. Harper and Row.

Brathwaite, Alex, and Jeffrey Kucik. 2018. Does the Presence of Foreign Troops Affect Stability in the Host Country? *Foreign Policy Analysis* 14 (4):536–560.

Brooks, Sarah M. 2007. When Does Diffusion Matter? Explaining the Spread of Structural Pension Reforms Across Nations. *Journal of Politics* 69 (3):701–15.

Buckley, Chris. 2017. China, Sending a Signal, Launches a Home-Built Aircraft Carrier. *The New York Times* April 25, 2017.

Bueno de Mesquita, Bruce, Alastair Smith, Randolph M. Siverson, and James D. Morrow. 2003. *The Logic of Political Survival*. Cambridge, MA: MIT Press.

Bueno de Mesquita, Ethan. 2010. Regime Change and Revolutionary Entrepreneurs. *American Political Science Review* 104 (3):446–66.

Burgett, Donald. 1999. *Seven Roads to Hell: A Screaming Eagle at Bastogne*. Novato, CA: Presidio Press.

Burton, Brian, and John Nagl. 2008. Learning as We Go: The US Army Adapts to Counterinsurgency in Iraq, July 2004–December 2006. *Small Wars & Insurgencies* 19 (3):303–327.

Byman, Daniel. 2016. 'Death Solves All Problems': The Authoritarian Model of Counterinsurgency. *Journal of Strategic Studies* 39 (1):62–93.

Carlsson, Hans, and Eric Van Damme. 1993. Global Games and Equilibrium Selection. *Econometrica* 61 (5):989–1018.

Carreri, Maria, and Oeindrila Dube. 2017. Do Natural Resources Influence Who Comes to Power, and How? *The Journal of Politics* 79 (2):502–518.

Casper, Brett Allen, and Scott A. Tyson. 2014. Military Compliance and the Efficacy of Mass Killings at Deterring Rebellion. Unpublished Manuscript. Available at SSRN 2434297.

Castillo, Jasen. 2014. *Endurance and War: The National Sources of Military Cohesion*. Stanford, CA: Stanford University Press.

Catignani, Sergio. 2012. 'Getting COIN' at the Tactical Level in Afghanistan: Reassessing Counter-Insurgency Adaptation in the British Army. *Journal of Strategic Studies* 35 (4):513–539.

Chan, Minnie, and Guo Rui. 2019. China Will Build 4 Nuclear Aircraft Carriers in Drive to Catch US Navy, Experts Say. *South China Morning Post* February 6, 2019.

Charrad, M., N. Gahhazli, V. Boiteau, and A. Niknafs. 2014. NbClust: An R Package for Determining the Relevant Number of Clusters in a Data Set. *Journal of Statistical Software* 61 (6):1–36.

Churchill, Winston. 1941. Address to Harrow School, London, UK. October 29, 1941.

Ciorciari, John D. 2010. *Limits of Alignment: Southeast Asia and the Great Powers since 1975*. Washington, DC: Georgetown University Press.

Clodfelter, Michael. 2008. *Warfare and Armed Conflicts: A Statistical Encyclopedia of Casualty and Other Figures, 1494-2007*. Jefferson, NC: McFarland & Company, 3rd edn.

Clutterbuck, Richard L. 1966. *The Long, Long War: Counterinsurgency in Malaya and Vietnam*. New York, NY: Praeger.

Cochran, Kathryn McNabb, and Stephen B. Long. 2017. Measuring Military Effectiveness: Calculating Casualty Loss-Exchange Ratios for Multilateral Wars, 1816–1990. *International Interactions* 43 (6):1019–40.

Cohen, Eliot A. 2021. Exit Strategy. *The Atlantic* April 13, 2021.

Conley, Timothy G., Christian B. Hansen, and Peter E. Rossi. 2012. Plausibly Exogenous. *Review of Economic Studies and Statistics* 94 (1):260–272.

Council on Foreign Relations. 2021. The U.S. War in Afghanistan 1999–2021. Available at https://www.cfr.org/timeline/us-war-afghanistan.

Cronin, Audrey Kurth. 2006. How al-Qaida Ends: The Decline and Demise of Terrorist Groups. *International Security* 31 (1):7–48.

Cruz, Cesi, Philip Keefer, and Carlos Scartascini. 2018. Database of Political Institutions 2017 (DPI2017). Available at https://mydata.iadb.org/Reform-Modernization-of-the-State/Database-of-Political-Institutions-2017/938i-s2bw.

Davenport, Christian. 2015. *How Social Movements Die: Repression and Demobilization of the Republic of New Africa*. New York: Cambridge University Press.

Denyer, Simon, and Akiko Kashiwagi. 2018. In Japan's Okinawa Island, U.S. Military Bases Take Center Stage in Election. *The Washington Post* September 28, 2018.

Dixon, Paul. 2009. 'Hearts and Minds'? British Counter-Insurgency from Malaya to Iraq. *Journal of Strategic Studies* 32 (3):353–381.

Dupuy, Trevor. 1984. Analysis of Factors that Have Influenced Outcomes of Battles and Wars. Vol. 1. Technical Report, CAA SR-84-6. Dunn Loring, VA: Historical Evaluation and Research Organization.

Edmond, Chris. 2013. Information Manipulation, Coordination, and Regime Change. *Review of Economic Studies* 80 (4):1422–58.

Edwards, Bob, and Sam Marullo. 1995. Organizational Mortality in a Declining Social Movement: The Demise of Peace Movement Organizations in the End of the Cold War Era. *American Sociological Review* 60 (6):908–27.

Enterline, Andrew J., Emily Stull, and Joseph Magagnoli. 2013. Reversal of Fortune? Strategy Change and Counterinsurgency Success by Foreign Powers in the Twentieth Century. *International Studies Perspectives* 14 (2):176–98.

Fall, Bernard B. 1961. *Street Without Joy: Indochina at War, 1946-54*. Harrisburg, PA: Stackpole.

Farrell, Theo. 2010. Improving in War: Military Adaptation and the British in Helmand Province, Afghanistan, 2006-09. *Journal of Strategic Studies* 33 (4):567–594.

Farrell, Theo, Frans Osinga, and James A. Russell, eds. 2013. *Military Adaptation in Afghanistan*. Stanford, CA: Stanford University Press.

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

———. 1997. Signaling Foreign Policy Interests: Tying Hands Versus Sinking Costs. *Journal of Conflict Resolution* 41 (1):68–90.

Feaver, Peter D. 1999. Civil-Military Relations. *Annual Review of Political Science* 2:211–241.

———. 2003. *Armed Servants: Agency, Oversight, and Civil-Military Relations*. Cambridge, MA: Harvard University Press.

Flynn, Michael E., Carla Martinez Machain, and Alissandra T. Stoyan. 2019. Building Trust: The Effect of U.S. Troop Deployments on Public Opinion in Peru. *International Studies Quarterly* 63 (3):742–755.

Foley, Robert T. 2012. A Case Study in Horizontal Military Innovation: The German Army, 1916-1918. *Journal of Strategic Studies* 35 (6):799–827.

Fong, Christian, Chad Hazlett, and Kosuke Imai. 2018. Covariate Balancing Propensity Score for a Continuous Treatment: Application to the Efficacy of Political Advertisements. *Annals of Applied Statistics* 12 (1):156–177.

Francisco, Ronald A. 2004. After the Massacre: Mobilization in the Wake of Harsh Repression. *Mobilization: An International Quarterly* 9 (2):107–26.

Galula, David. 1964. *Counterinsurgency Warfare: Theory and Practice*. New York, NY: Praeger.

Gates, Scott, and Ragnhild Nordås. 2016. Recruitment and Retention, and Religion in Rebel Groups: Exploring the Advantages of Extremism. Unpublished Manuscript.

George, Alexander L., and Andrew Bennett. 2005. *Case Studies and Theory Development in the Social Sciences*. Cambridge, MA: MIT Press.

Getmansky, Anna. 2013. You Can't Win If You Don't Fight: The Role of Regime Type in Counterinsurgency Outbreaks and Outcomes. *Journal of Conflict Resolution* 57 (4):709–34.

Gleditsch, Kristian Skrede. 2008. Distance Between Capital Cities [Data file]. Retrieved from http://ksgleditsch.com/data-5.html.

Goldman, Emily O., and Richard B. Andres. 1999. Systemic Effects of Military Innovation and Diffusion. *Security Studies* 8 (4):79–125.

Granovetter, Mark. 1978. Threshold Models of Collective Behavior. *American Journal of Sociology* 83 (6):1420–43.

Grauer, Ryan. 2014. Why Do Soldiers Give Up? A Self-Preservation Theory of Surrender. *Security Studies* 23 (3):622–55.

Grauer, Ryan, and Michael C. Horowitz. 2012. What Determines Military Victory? Testing the Modern System. *Security Studies* 21 (1):83–112.

Gudmundsson, Bruce I. 1995. *Stormtroop Tactics: Innovation in the German Army, 1914-1918*. Westport, CT: Praeger.

Hack, Karl. 2009. The Malayan Emergency as Counter-Insurgency Paradigm. *Journal of Strategic Studies* 32 (3):383–414.

Hammarström, Mats. 1994. The Diffusion of Military Conflict: Central and South-East Europe in 1919–20 and 1991–92. *Journal of Peace Research* 31 (3):263–80.

Hamner, Christopher H. 2011. *Enduring Battle: American Soldiers in Three Wars, 1776–1945*. Lawrence: University Press of Kansas.

Harkavy, Robert E. 1999. Long Cycle Theory and the Hegemonic Powers' Basing Networks. *Political Geography* 18:941–972.

Hart, B.H. Liddell. 1991. *Strategy*. New York, NY: Meridian, 2nd edn.

Hartigan, J.A., and M.A. Wong. 1979. A K-Means Clustering Algorithm. *Journal of the Royal Statistical Society* 28 (1):100–108.

Heggoy, Alf. 1972. *Insurgency and Counterinsurgency in Algeria*. Bloomington, IN: Indiana University Press.

Helmbold, Robert L., and Aqeel A. Kahn. 1986. Combat History Analysis Study Effort (CHASE): Progress Report for the Period August 1984-June 1985. National Technical Information Service.

Hensel, Paul R. 2018. ICOW Colonial History Data Set, version 1.1 [Data file]. Retrieved from https://www.paulhensel.org/icowcol.html.

Holmstrom, Bengt. 1982. Moral Hazard in Teams. *The Bell Journal of Economics* 13 (2):324–40.

Homer. 2002. *The Odyssey*. Fagles, Robert, tr. New York, NY: Penguin Books.

Honaker, James, Gary King, and Matthew Blackwell. 2015. *Amelia*. R package version 1.7.4.

Horne, Alistair. 1977. *A Savage War of Peace: Algeria 1954-1962*. New York, NY: New York Review of Books.

Horowitz, Michael C. 2010a. *The Diffusion of Military Power: Causes and Consequences for International Politics*. Princeton, NJ: Princeton University Press.

———. 2010b. Nonstate Actors and the Diffusion of Innovations: The Case of Suicide Terrorism. *International Organization* 64 (1):33–64.

Horowitz, Michael C., and Shira E. Pindyck. 2020. What is a Military Innovation? A Proposed Framework. Unpublished Manuscript. Available at SSRN 3504246.

Hothorn, Torsten, Kurt Hornik, and Achim Zeileis. 2006. Unbiased Recursive Partitioning: A Conditional Inference Framework. *Journal of Computational and Graphical Statistics* 15 (3):651–674.

Hudson, John, Paul Sonne, Karen DeYoung, and Josh Dawsey. 2018. U.S. Assessing Cost of Keeping Troops in Germany as Trump Battles with Europe. *The Washington Post* June 29, 2018.

Huntington, Samuel P. 1957. *The Soldier and the State: The Theory and Politics of Civil-Military Relations*. Cambridge, MA: Harvard University Press.

Huth, Paul K. 1988. Extended Deterrence and the Outbreak of War. *American Political Science Review* 82 (2):423–443.

Imai, Kosuke, In Song Kim, and Erik Wang. 2019. Matching Methods for Causal Inference with Time-Series Cross-Sectional Data. Unpublished Manuscript.

International Committee of the Red Cross. 2016. Treaties, State Parties, and Commentaries. Available at https://www.icrc.org/applic/ihl/ihl.nsf/vwTreatiesByDate.xsp.

International Institute for Strategic Studies. 1970-2018. *The Military Balance*. London, UK: The International Institute for Strategic Studies.

Jacobs, Andrew, and Jane Perlez. 2017. U.S. Wary of Its New Neighbor in Djibouti: A Chinese Naval Base. *The New York Times* February 25, 2017.

Jervis, Robert. 1978. Cooperation under the Security Dilemma. *World Politics* 30 (2):167–214.

Johnson, Douglas. 2000. Bernard Tricot: De Gaulle's Adviser Who Helped Set Algeria Free. *The Guardian* June 13, 2000.

Johnson, Jesse C., Brett Ashley Leeds, and Ahra Wu. 2015. Capability, Credibility, and Extended General Deterrence. *International Interactions* 41 (2):309–336.

Johnston, Patrick B. 2012. Does Decapitation Work? Assessing the Effectiveness of Leadership Targeting in Counterinsurgency Campaigns. *International Security* 36 (4):47–79.

Jones, Daniel M., Stuart A. Bremer, and J. David Singer. 1996. Militarized Interstate Disputes, 1816–1992: Rationale, Coding Rules, and Empirical Patterns. *Conflict Management and Peace Science* 15 (2):163–213.

Jones, Seth G. 2017. *Waging Insurgent Warfare: Lessons from the Vietcong to the Islamic State*. Oxford, UK: Oxford University Press.

Kane, Tim. 2012. Development and US Troop Deployments. *Foreign Policy Analysis* 8:255–273.

Kauffman, Stuart, and Simon Levin. 1987. Towards a General Theory of Adaptive Walks on Rugged Landscapes. *Journal of Theoretical Biology* 128 (1):11–45.

Keegan, John. 1976. *The Face of Battle*. London: Penguin.

Kelly, George Armostrong. 1965. *Lost Soldiers: The French Army and Empire in Crisis 1947-1962*. Cambridge, MA: MIT Press.

Kertzer, Joshua D. 2017. Resolve, Time, and Risk. *International Organization* 71 (S1):S109–S136.

Kier, Elizabeth. 1997. *Imagining War: French and British Military Doctrine between the Wars*. Princeton, NJ: Princeton University Press.

Kilcullen, David. 2010. *Counterinsurgency*. New York, NY: Oxford University Press.

Kim, Hyohoung, and Peter S. Bearman. 1997. The Structure and Dynamics of Movement Participation. *American Sociological Review* 62 (1):70–93.

Kinne, Brandon J. 2018. Defense Cooperation Agreements and the Emergence of a Global Security Network. *International Organization* 72 (4):799–837.

Kissinger, Henry A. 1969. The Viet Nam Negotiations. *Foreign Affairs* 47 (2):211–234.

Kollars, Nina A. 2015. War's Horizon: Soldier-Led Adapation in Iraq and Afghanistan. *Journal of Strategic Studies* 38 (4):529–553.

Koopmans, Ruud. 2004. Protest in Time and Space: The Evolution of Waves of Contention. In *The Blackwell Companion to Social Movements*, edited by David A. Snow, Sarah A. Soule, and Hanspeter Kriesi, 19–46. Malden, MA: Blackwell.

Kuran, Timur. 1991. Now Out of Never: The Element of Surprise in the East European Revolution of 1989. *World Politcs* 44 (1):7–48.

Leeds, Brett Ashley. 2003. Do Alliances Deter Aggression? The Influence of Military Alliances on the Initiation of Militarized Interstate Disputes. *American Journal of Political Science* 47 (3):427–439.

Leeds, Brett Ashley, Jeffrey M. Ritter, Sarah McLaughlin Mitchell, and Andrew G. Long. 2002. Alliance Treaty Obligations and Provisions, 1815-1944. *International Interactions* 28 (3):237–260.

Legro, Jeffrey W. 1995. *Cooperation under Fire: Anglo-German Restraint during World War II*. Ithaca, NY: Cornell University Press.

Lehmann, Todd C., and Yuri M. Zhukov. 2019. Until the Bitter End? The Diffusion of Surrender Across Battles. *International Organization* 73 (1):133–169.

Levitt, Barbara, and James G. March. 1988. Organizational Learning. *Annual Review of Sociology* 14:319–40.

Long, Austin. 2016. *The Soul of Armies: Counterinsurgency Doctrine and Military Culture in the US and UK*. Ithaca, NY: Cornell University Press.

Lyall, Jason. 2010. Do Democracies Make Inferior Counterinsurgents? Reassessing Democracy's Impact on War Outcomes and Duration. *International Organization* 64 (1):167–92.

———. 2014. Why Armies Break: Explaining Mass Desertion in Conventional War. Unpublished Manuscript. Available at SSRN 2524561.

Lyall, Jason, and Isaiah Wilson III. 2009. Rage Against the Machines: Explaining Outcomes in Counterinsurgency Wars. *International Organization* 63 (1):67–106.

Lyall, Jason, and Yuri M. Zhukov. 2020. Fratricidal Coercion and Battlefield Performance in Modern War. Unpublished Manuscript.

MacCoun, Robert J. 1993. Unit Cohesion in Military Performance. In *Sexual Orientation and US Military Personnel Policy: Policy Options and Assessment*, 283–331. Santa Monica, CA: RAND.

MacCoun, Robert J., Elizabeth Kier, and Aaron Belkin. 2006. Does Social Cohesion Determine Motivation in Combat? An Old Question with an Old Answer. *Armed Forces & Society* 32 (4):646–654.

Machain, Carla Martinez. 2013. The Effect of US Troop Deployment on Host States' Foreign Policy. *Armed Forces & Society* 39 (1):102–123.

Mack, Andrew J.R. 1975. Why Big Nations Lose Small Wars: The Politics of Asymmetric Conflict. *World Politics* 27 (2):175–200.

MacQueen, J. 1967. Some Methods for Classification and Analysis of Multivariate Observations. In *Proceedings of the Fifth Berkeley Symposium on Mathematical Statistics and Probability*, vol. I, edited by Lucien M. Le Cam and Jerzy Neyman, 281–297. University of California Press.

Macy, Michael W. 1991. Chains of Cooperation: Threshold Effects in Collective Action. *American Sociological Review* 56 (6):730–47.

Mao, Tse-Tung. 1961. *On Guerrilla Warfare*. Griffith, Samuel B., tr. New York, NY: Praeger.

March, James G. 1991. Exploration and Exploitation in Organizational Learning. *Organization Science* 2 (1):71–87.

Markowitz, Jonathan N, and Christopher J Fariss. 2013. Going the Distance: The Price of Projecting Power. *International Interactions* 39:119–143.

Marshall, Monty G., Ted Robert Gurr, and Keith Jaggers. 2014. Polity IV Project. Available at http://www.systemicpeace.org/inscrdata.html.

Marshall, S.L.A. 1947. *Men Against Fire: The Problem of Command in Future War*. New York: William Morrow.

McLauchlin, Theodore. 2010. Loyalty Strategies and Military Defection in Rebellion. *Comparative Politics* 42 (3):333–50.

———. 2015. Desertion and Collective Action in Civil Wars. *International Studies Quarterly* 59 (4):669–679.

McNeill, William H. 1982. *The Pursuit of Power: Technology, Armed Force, and Society since AD 1000*. Chicago: University of Chicago Press.

Mearsheimer, John J. 1989. Assessing the Conventional Balance: The 3:1 Rule and Its Critics. *International Security* 13 (4):54–89.

Mearsheimer, John J. 2001. *The Tragedy of Great Power Politics*. New York, NY: Norton and Company.

Merom, Gil. 2003. *How Democracies Lose Small Wars: State, Society, and the Failures of France in Algeria, Israel in Lebanon, and the United States in Vietnam*. Cambridge, UK: Cambridge University Press.

Millett, Allan R., Williamson Murray, and Kenneth H. Watman. 1988. The Effectiveness of Military Organizations. In *Military Effectiveness*, vol. I, edited by Allan R. Millett and Williamson Murray, chap. 1, 1–30. Boston, MA: Allen & Unwin.

Milner, Marc. 1984. Convoy Escorts: Tactics, Technology, and Innovation in the Roayl Canadian Navy, 1939-1943. *Military Affairs* 48 (1):19–25.

Moran, P.A.P. 1950. Notes on Continuous Stochastic Phenomena. *Biometrika* 37 (1/2):17–23.

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.

———. 1993. Arms Versus Allies: Trade-Offs in the Search for Security. *International Organization* 47 (2):207–233.

———. 2000. Alliances: Why Write Them Down? *Annual Review of Political Science* 3:63–83.

———. 2014. *Order within Anarchy: The Laws of War as an International Institution*. New York, NY: Cambridge University Press.

Most, Benjamin A., and Harvey Starr. 1980. Diffusion, Reinforcement, Geopolitics, and the Spread of War. *American Political Science Review* 74 (4):932–46.

Mumford, Andrew. 2011. *The Counter-Insurgency Myth: The British Experience of Irregular Warfare*. New York, NY: Routledge.

Murray, Williamson. 2011. *Military Adaptation in War: With Fear of Change*. Cambridge, UK: Cambridge University Press.

Muth, Jörg. 2011. *Command Culture: Officer Education in the U.S. Army and the German Armed Forces, 1901-1940, and the Consequences for World War II*. Denton, TX: University of North Texas Press.

Nagl, John A. 2002. *Learning to Eat Soup with a Knife: Counterinsurgency Lessons from Malaya and Vietnam*. Chicago, IL: University of Chicago Press.

Nakamura, Alice, and Masao Nakamura. 1981. On the Relationships Among Several Specification Error Tests Presented by Durbin, Wu, and Hausman. *Econometrica* 49 (6):1583–1588.

Nash, John F., Jr. 1950. Equilibrium points in *n*-person games. *Proceedings of the National Academy of Sciences* 36 (1):48–49.

National Consortium for the Study of Terrorism and Responses to Terrorism (START). 2018. Global Terrorism Database [Data file]. Retrieved from https://www.start.umd.edu/gtd.

Nickell, Stephen. 1981. Biases in Dynamic Models with Fixed Effects. *Econometrica* 49 (6):1417–1426.

O'Hanlon, Michael E., and Ian Livingston. 2011. Iraq Index: Tracking Variables of Reconstruction & Security in Post-Saddam Iraq. *Brookings Institution*.

O'Mahony, Jennifer. 2018. UK Troops Sent to Africa's Jihadist Belt to Fight Extremism and Migration in British 'Pivot to Sahel'. *The Telegraph* October 6, 2018.

Palmer, Glenn, Vito D'Orazio, Michael R. Kenwick, and Roseanne W. McManus. 2020. Updating the Militarized Interstate Dispute Data: A Response to Gibler, Miller, and Little. *International Studies Quarterly* 64 (2):469–475.

Pampinella, Stephen. 2015. The Effectiveness of Coercive and Persuasive Counterinsurgency Practices since 1945. *Civil Wars* 17 (4):503–26.

Paret, Peter. 1964. *French Revolutionary Warfare from Indochina to Algeria: The Analysis of a Political and Military Doctrine*. New York, NY: Praeger.

Patel, M. Hammad, Mujtaba Ahmed Abbasi, M. Saeed, and Shah Jamal Alam. 2018. A Scheme to Analyze Agent-Based Social Simulations Using Exploratory Data Mining Techniques. *Complex Adaptive Systems Modeling* 6 (1):1–17.

Paul, Christopher, Colin P. Clarke, and Beth Grill. 2010. *Victory Has a Thousand Fathers: Sources of Success in Counterinsurgency*. Santa Monica, CA: RAND Corporation.

Peters, Heidi M., and Sofia Plagakis. 2021. *Department of Defense Contractor and Troop Levels in Afghanistan and Iraq: 2007-2020*. Washington, DC: Congressional Research Service.

Porch, Douglas. 2013. *Counterinsurgency: Exposing the Myths of the New Way of War*. Cambridge, UK: Cambridge University Press.

Posen, Barry S. 1984. *The Sources of Military Doctrine: France, Britain, and Germany between the World Wars*. Ithaca, NY: Cornell University Press.

Pottier, Philippe. 2007. GCMA/GMI: A French Experience in Counterinsurgency during the French Indochina War. *Small Wars & Insurgencies* 16 (2):125–146.

Powell, Robert. 1996. Bargaining in the Shadow of Power. *Games and Economic Behavior* 15 (2):255–89.

———. 2004. Bargaining and Learning While Fighting. *American Journal of Political Science* 48 (2):344–61.

———. 2006. War as a Commitment Problem. *International Organization* 60 (1):169–203.

Ramsay, Kristopher W. 2008. Settling It on the Field: Battlefield Events and War Termination. *Journal of Conflict Resolution* 52 (6):850–79.

Reiter, Dan, and Allan C. Stam. 1997. The Soldier's Decision to Surrender: Prisoners of War and World Politics. Paper presented at the 1997 Annual Meeting of the American Political Science Association, Washington, DC.

———. 2002. *Democracies at War*. Princeton, NJ: Princeton University Press.

Roggio, Bill, and Alexandra Gutowski. 2021. Mapping Taliban Control in Afghanistan. *FDD's Long War Journal*. Foundation for Defense of Democracies.

Rosen, Stephen Peter. 1991. *Winning the Next War: Innovation and the Modern Military*. Ithaca, NY: Cornell University Press.

Russell, James A. 2010. Innovation in War: Counterinsurgency Operations in Anbar and Ninewa Provinces, Iraq, 2005-2007. *Journal of Strategic Studies* 33 (4):595–624.

———. 2011. *Innovation, Transformation, and War: Counterinsurgency Operations in Anbar and Ninewa, Iraq, 2005-2007*. Stanford, CA: Stanford University Press.

Ruttig, Thomas, and Jelena Bjelica. 2018. The State of Aid and Poverty in 2018: A New Look at Aid Effectiveness in Afghanistan. *Afghanistan Analysts Network*.

Sargan, J.D. 1958. The Estimation of Economic Relationships Using Instrumental Variables. *Econometrica* 26 (3):393–415.

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

Schelling, Thomas C. 1960. *The Strategy of Conflict*. Cambridge, MA: Harvard University Press.

———. 1966. *Arms and Influence*. New Haven, CT: Yale University Press.

———. 2006. *Micromotives and Macrobehavior*. New York, NY: W.W. Norton & Company, revised edn.

Schmitt, Eric. 2019. Russia's Military Mission Creep Advances to a New Front: Africa. *The New York Times* March 31, 2019.

Shils, Edward A., and Morris Janowitz. 1948. Cohesion and Disintegration in the Wehrmacht in World War II. *Public Opinion Quarterly* 12 (2):280–315.

Short, Anthony. 1975. *The Communist Insurrection in Malaya, 1948–1960*. London, UK: Frederick Muller.

Showalter, Dennis, ed. 2013. *The Encyclopedia of Warfare*. London: Amber Books.

Siebold, Guy L. 2007. The Essence of Military Group Cohesion. *Armed Forces & Society* 33 (2):286–295.

Siegel, David A. 2011. When Does Repression Work? Collective Action in Social Networks. *The Journal of Politics* 73 (4):993–1010.

Simmons, Beth A., and Zachary Elkins. 2004. The Globalization of Liberalization: Policy Diffusion in the International Political Economy. *American Political Science Review* 98 (1):171–89.

Singer, J. David, Stuart Bremer, and John Stuckey. 1972. Capability Distribution, Uncertainty, and Major Power War, 1820-1965. In *Peace, War, and Numbers*, edited by Bruce Russett, 19–48. Beverly Hills, CA: Sage.

Singer, Joel David. 1979. *Explaining War: Correlates of War Project*. Beverly Hills, CA: Sage.

Slantchev, Branislav L. 2003. The Principle of Convergence in Wartime Negotiations. *American Political Science Review* 97 (4):621–32.

———. 2005. Military Coercion in Interstate Crises. *American Political Science Review* 99 (4):533–547.

Smith, Alastair, and Allan C. Stam. 2004. Bargaining and the Nature of War. *Journal of Conflict Resolution* 48 (6):783–813.

Smith, Caroline, and James M. Lindsay. 2003. Rally 'Round the Flag: Opinion in the United States before and after the Iraq War. *Brookings Institution*.

Snyder, Jack. 1984. Civil-Military Relations and the Cult of the Offensive, 1914 and 1984. *International Security* 9:108–146.

Sorley, Lewis. 1999. *A Better War: The Unexamined Victories and Final Tragedy of America's Last Years in Vietnam*. New York, NY: Harcourt.

Spencer, Jack, and Baker Spring. 2007. The Advantages of Expanding the Nuclear Navy. *The Heritage Foundation* Web Memo No. 1693. Available at http://www.heritage.org/research/reports/2007/11/the-advantages-of-expanding-the-nuclear-navy.

Stam, Allan C. III. 1999. *Win, Lose, or Draw: Domestic Politics and the Crucible of War*. Ann Arbor, MI: University of Michigan Press.

Sun Tzu. 1983. *The Art of War*. Clavell, James, tr. New York, NY: Delta.

Taber, Robert. 2002. *War of the Flea: The Classic Study of Guerrilla Warfare*. Washington, DC: Brassey's.

Tarar, Ahmer. 2013. Military Mobilization and Commitment Problems. *International Interactions* 39 (3):343–366.

Tarar, Ahmer, and Bahar Leventoğlu. 2012. Limited Audience Costs in International Crises. *Journal of Conflict Resolution* 57 (6):1065–1089.

Taylor, Alan. 2017. Operation Barkhane: France's Counterterrorism Forces in Africa. *The Atlantic* October 24, 2017.

The Economist. 2019. The Brewing Conflict Between America and Iran. *The Economist* May 9, 2019.

Tilman, Robert O. 1966. The Non-Lessons of the Malayan Emergency. *Asian Survey* 6 (8):407–419.

Toft, Monica Duffy, and Yuri M. Zhukov. 2012. Denial and Punishment in the North Caucasus: Evaluating the Effectiveness of Coercive Counter-insurgency. *Journal of Peace Research* 49 (6):785–800.

Toronto, Nathan. 2005. Military Recruitment Data Set, version 2005.1. Available at https://nathantoronto.com/research.

Trinquier, Roger. 1961. *Modern Warfare: A French View of Counterinsurgency*. New York, NY: Praeger.

Ucko, David H. 2009. *The New Counterinsurgency Era: Transforming the U.S. Military for Modern Wars*. Washington, DC: Georgetown University Press.

———. 2010. The Malayan Emergency: The Legacy and Relevance of a Counter-Insurgency Success Story. *Defence Studies* 10 (1-2):13–39.

Ucko, David H., and Robert Egnell. 2013. *Counterinsurgency in Crisis: Britain and the Challenges of Modern Warfare*. New York, NY: Columbia University Press.

United States Army. 2014. *FM 3-24/MCWP 3-33.5: Insurgencies and Countering Insurgencies*. Washington, DC: Headquarters, Department of the Army.

United States Department of Defense. 2018. *Summary of the 2018 National Defense Strategy of the United States of America*.

United States Senate. 2008. *Inquiry into the Treamtent of Detainees in U.S. Custody*. 110 S. Rpt. 54. Committee on Armed Services.

Van Creveld, Martin. 2004. *Supplying War: Logistics from Wallenstein to Patton*. Cambridge University Press, 2nd edn.

Van Evera, Stephen. 1984. The Cult of the Offensive and the Origins of the First World War. *International Security* 9 (1):58–107.

———. 1998. Offense, Defense, and the Causes of War. *International Security* 22 (4):5–43.

Webb, Kieran. 2007. The Continued Importance of Geographic Distance and Boulding's Loss of Strength Gradient. *Comparative Strategy* 26 (4):295–310.

Weidmann, Nils B. 2015. Communication Networks and the Transnational Spread of Ethnic Conflict. *Journal of Peace Research* 52 (3):285–96.

Weigley, Russell F. 1988. The Political and Strategic Dimensions of Military Effectiveness. In *Military Effectiveness*, vol. III, edited by Allan R. Millett and Williamson Murray, chap. 10, 341–364. Boston, MA: Allen & Unwin.

Weisiger, Alex. 2016. Learning from the Battlefield: Information, Domestic Politics, and Interstate War Duration. *International Organization* 70 (2):347–75.

Whitlock, Craig. 2019a. The Afghanistan Papers: Built to Fail. *The Washington Post* December 9, 2019.

———. 2019b. The Afghanistan Papers: Stranded Without a Strategy. *The Washington Post* December 9, 2019.

Windrow, Martin. 2004. *The Last Valley: Dien Bien Phu and the French Defeat in Vietnam*. Cambridge, MA: Da Capo Press.

Zhukov, Yuri M., and Brandon M. Stewart. 2013. Choosing Your Neighbors: Networks of Diffusion in International Relations. *International Studies Quarterly* 57 (2):271–87.

Zisk, Kimberly Marten. 1993. *Engaging the Enemy: Organization Theory and Soviet Military Innovation, 1955-1991*. Princeton, NJ: Princeton University Press.

Zolotarev, Vladimir Antonovich. 1997. Prikazy narodnogo komissara oborony SSSR. 22 iyunya 1941 g.–1942 g. [Orders of the USSR state commissar of defense. 22 June 1941–1942]. *Russkiy arkhiv: Velikaya Otechestvennaya* 13 (2-2).