UC San Diego
UC San Diego Electronic Theses and Dissertations

Title
Leaders, Audiences, and the Use of Force

Permalink
https://escholarship.org/uc/item/6rc4h16b

Author
Merrell, Brandon Cole

Publication Date
2019

Peer reviewed|Thesis/dissertation
UNIVERSITY OF CALIFORNIA SAN DIEGO

Leaders, Audiences, and the Use of Force

A dissertation submitted in partial satisfaction of the requirements for the degree Doctor of Philosophy in Political Science by

Brandon Cole Merrell

Committee in charge:
Professor David A. Lake, Co-Chair
Professor Branislav Slantchev, Co-Chair
Professor Eli Berman
Professor Lauren Prather
Professor Philip Roeder

2019
The dissertation of Brandon Cole Merrell is approved, and it is acceptable in quality and form for publication on microfilm and electronically:

Co-Chair

Co-Chair

University of California San Diego

2019
Political writers have established it as a maxim, that in contriving any system of government... every man out to be supposed to be a knave and to have no other end, in all his actions, than his private interest. By this interest we must govern him, and, by means of it, make him, notwithstanding his insatiable avarice and ambition, cooperate to public good

—David Hume, *Essays: Moral, Political, and Literary*, 1742

Force and fraud are in war the two cardinal virtues.

—Thomas Hobbes, *Leviathan*, 1651

We know how cruel the truth often is, and we wonder whether delusion is not more consoling.

—Henri Poincaré

Have republics in practice been less addicted to war than monarchies? Are not the former administered by men as well as the latter?

—Alexander Hamilton, *Federalist No. 6*, 1787

Science is made up of so many things that appear obvious after they are explained.

—Frank Herbert, *Dune*, 1965

The symbol and the metaphor are as necessary to science as to poetry.

—Jacob Brownowski, *Science and Human Values*, 1956
# TABLE OF CONTENTS

Signature Page ................................................................. iii
Epigraph ................................................................. iv
Table of Contents ............................................................. v
List of Figures ............................................................... vii
List of Tables ................................................................. viii
Acknowledgements ........................................................... ix
Vita ................................................................. xii
Abstract of the Dissertation ................................................ xiii

1 Fighting for the People: Leaders, Audiences, and the Use of Force ................. 1
   1.1 Introduction ...................................................... 2
   1.2 Theories of Crisis Escalation .................................... 5
       1.2.1 Unitary Explanations for War ......................... 5
       1.2.2 Domestic and Individual Accounts of Crisis Behavior ... 6
   1.3 The Agency Dilemma ............................................... 8
       1.3.1 Public Military Optimism ................................. 10
       1.3.2 Barriers to Information Transfers ...................... 12
       1.3.3 Divergence from Previous Domestic War Theories .... 17
   1.4 Cross-National Evidence ............................................ 19
       1.4.1 Setup, Data, and Variables ............................. 19
       1.4.2 Expectations and Results ............................... 21
   1.5 Historical Episodes .................................................. 25
       1.5.1 The War of the Pacific .................................... 26
       1.5.2 The Fall of France ...................................... 33
       1.5.3 Escalation in Vietnam ................................... 39
   1.6 Conclusion ............................................................. 46

2 The Secrecy Gambit: Clandestine Power Shifts and Preventive Conflict .............. 48
   2.1 Introduction .......................................................... 49
   2.2 Signaling, Secrecy, and Preventive Action .......................... 53
   2.3 The Secret Development Model .................................... 56
       2.3.1 Model Analysis ........................................... 61
   2.4 Discussion and Implications ........................................ 66
       2.4.1 Power Shifts and Preventive Attack .................... 66
       2.4.2 The Risks of Signaling ................................... 67
## LIST OF FIGURES

<table>
<thead>
<tr>
<th>Figure</th>
<th>Description</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>Figure 1.1</td>
<td>Bargaining with Domestic Constraints</td>
<td>13</td>
</tr>
<tr>
<td>Figure 1.2</td>
<td>Bargaining with Multiple Audiences</td>
<td>16</td>
</tr>
<tr>
<td>Figure 1.3</td>
<td>Bargaining when $F$ and $L$ Initially Agree</td>
<td>18</td>
</tr>
<tr>
<td>Figure 1.4</td>
<td>Estimated Effect of a Power Deficit (Polity Score : Power Deficit Dummy)</td>
<td>24</td>
</tr>
<tr>
<td>Figure 1.5</td>
<td>Estimated Effect of a Power Deficit (Election Dummy : Power Deficit Dummy)</td>
<td>24</td>
</tr>
<tr>
<td>Figure 1.6</td>
<td>Probability of Settlement Conditional on Size of Power Disadvantage and Polity Score</td>
<td>25</td>
</tr>
<tr>
<td>Figure 2.1</td>
<td>Secret Development Model, Period 1: Announcement or Concealment</td>
<td>57</td>
</tr>
<tr>
<td>Figure 2.2</td>
<td>Secret Development Model, Period 2: Development and Potential Exposure</td>
<td>58</td>
</tr>
<tr>
<td>Figure 2.3</td>
<td>Probability of disputes conditional on Power Shifts and Secret Alliances</td>
<td>84</td>
</tr>
<tr>
<td>Figure 3.1</td>
<td>Baseline Model of Insurgency and Governance</td>
<td>103</td>
</tr>
</tbody>
</table>
## LIST OF TABLES

<table>
<thead>
<tr>
<th>Table</th>
<th>Title</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>Table 1.1</td>
<td>Dispute Settlement, Democracy, and Relative Strength</td>
<td>23</td>
</tr>
<tr>
<td>Table 2.1</td>
<td>Model Payoffs</td>
<td>60</td>
</tr>
<tr>
<td>Table 2.2</td>
<td>Secret Alliances, Power Shifts, and MIDs</td>
<td>83</td>
</tr>
<tr>
<td>Table 2.3</td>
<td>Robustness and Alternative Specifications</td>
<td>86</td>
</tr>
<tr>
<td>Table 3.1</td>
<td>Model Payoffs</td>
<td>106</td>
</tr>
</tbody>
</table>
ACKNOWLEDGEMENTS

In his poem *The Prelude*, Wordsworth wondered how anyone could hope to “parcel out [their] intellect by geometric rules” or to identify the diverse array of experiences, circumstances, friends, and mentors on whom our insights and ideas depend. Of all the tasks I have faced during graduate school, putting into appropriate words my overwhelming appreciation for the advice and support of the following individuals is perhaps the one for which I feel most inadequate.

Although I am grateful to everyone who has shaped my experience thus far, I would above all like to thank my committee members. My co-chairs, David Lake and Branislav Slantchev, deserve particular credit for transforming my way of thinking about both politics and social science. As advisors, they offered generous guidance throughout this process, helped me locate what I hope are small kernels of truth from the garden of weeds, and—quite importantly—knew not only how to provide me with the freedom and intellectual independence I often sought, but also how to dispense with unwarranted praise and demand improvements when my arguments required major surgery. My other committee members, Eli Berman, Lauren Prather, and Phil Roeder, also share tremendous credit for my intellectual growth. Eli’s genuine enthusiasm for engaging with work that spans disciplinary lines and crosses methodological boundaries taught me to celebrate diverse approaches to scientific inquiry. Lauren was the optimal role model: generous with her time, incisive with her feedback, and always the first to respond with the encouragement and empathy that I suspect only another (relatively) early career scholar can provide. Finally, Phil has contributed, perhaps more than anyone, to my overall perspective of what political science scholarship can entail. From my first morning of graduate school, he has encouraged me to avoid speaking to a narrow and parochial audience—a temptation to which I am admittedly drawn—while at the same time reminding me to temper my remarks in instances where my framing exceeds my evidence. More than anything, all of my committee members provided a steadying influence during the times when I was most frustrated with my own work. My research will stand forever on their shoulders.
Alexei Abrahams, with whom I collaborated on much of this research, is an ideal coauthor, colleague, and friend. He is generous, reliable, good-natured, creative, and quick with a joke at any moment. I am tremendously fortunate to have made his acquaintance through David and Eli, and I look forward to continuing our research in the future.

More broadly, my work has benefited from discussions, questions, encouragement, and feedback from other friends, colleagues, and discussants, including Jeffrey Aulgur, Bernd Beber, Lawrence Broz, Tyler Bowen, Charity Butcher, Kevin Calderwood, Benjamin Campbell, Taylor Carlson, Robert Carroll, Alex Clott, Christina Cottiero, Jesse Driscoll, Mike Duda, Songying Fang, Thomas Flaherty, Mike Findley, Richard Flores, Erik Gartzke, Peter Gourevitch, Ron Gurantz, Emilie Hafner-Burton, Steph Haggard, Dotan Haim, Michael Joseph, Daniel Kent, Colin Krainin, Andrew Leber, Melissa Lee, Sichen Li, David Lindsey, Carmela Lutmar, Nathan Mariano, Kelly Matsush, Blake McMahon, James Morrow, Cécile Mouly, Matthew Nanes, Megumi Naoi, Richard Nielsen, Gerard Padró i Miquel, Kristy Pathakis, Luke Perez, Greg Phillips, Anthony Pierucci, Jeffry Ritter, Luke Sanford, Tom Schally, Christina Schneider, Robert Schub, Cameron Sells, Renard Sexton, Jake Shapiro, Ike Sharpless, Liesel Spangler, Art Stein, Zachary Steinert-Threlkeld, Barbara Walter, Alex Weisiger, David Wiens, Cathy Wu, Abigail Vaughn, Jack Zhang, and many others. I also appreciate financial support from both the UC San Diego Department of Political Science and the Institute on Global Conflict and Cooperation.

Finally, I thank Sherri Kaiser, Bill Nicolay, Glenn Prince, and Todd Graham for their early and continuing influence on my academic pursuits, along with my other close friends and family members who motivated and supported me throughout this process.

All chapters of this dissertation are being prepared for submission for publication of the material. The dissertation author was the primary investigator and author of this material.

 Chapters 2 and 3 of this dissertation are coauthored with Abrahams, Alexei S. The dissertation author was a primary author of these chapters and has permission to use the material contained therein.
VITA

<table>
<thead>
<tr>
<th>Year</th>
<th>Degree and Field</th>
<th>Institution</th>
</tr>
</thead>
<tbody>
<tr>
<td>2011</td>
<td>Bachelor of Arts, Political Science and Economics</td>
<td>Southern Illinois University, Carbondale</td>
</tr>
<tr>
<td>2015</td>
<td>Master of Arts, Political Science</td>
<td>University of California San Diego</td>
</tr>
<tr>
<td>2019</td>
<td>Doctor of Philosophy, Political Science</td>
<td>University of California San Diego</td>
</tr>
</tbody>
</table>
ABSTRACT OF THE DISSERTATION

Leaders, Audiences, and the Use of Force

by

Brandon Cole Merrell

Doctor of Philosophy in Political Science

University of California San Diego, 2019

Professor David A. Lake, Co-Chair
Professor Branislav Slantchev, Co-Chair

I examine three strategic tradeoffs between political accountability and violence. First, I argue that leaders who are accountable to their citizens may face domestic pressure to participate in risky and costly wars. Second, I show how countries can benefit from concealing their development of new military technology, whereas war can result when such efforts become publicly known. Finally, I demonstrate that citizens in fragile and under-institutionalized societies can deter their governments from engaging in abusive behavior, but these constraints on sovereign authority may come at the risk of domestic violence or instability. I support the findings with a combination of analytic theory, quantitative data, and qualitative case studies.
1 Fighting for the People: Leaders, Audiences, and the Use of Force

Abstract: Why do leaders wage wars they cannot hope to win? I demonstrate that political leaders sometimes engage in military conflict that they believe is costly and counterproductive because they would face domestic backlash if they instead pursued a peaceful settlement. In short, leaders possess private information about the costliness and riskiness of war and confront a series of strategic difficulties and disincentives to sharing this information with citizens. As a result, citizens may remain naively optimistic about the desirability of using military force. In these circumstances, domestic institutions that hold leaders accountable to their constituents can encourage rather than deter leaders from behaving aggressively. I provide two forms of empirical support for the theory. First, I examine territorial transfers that occurred between 1816 and 2014 and show that elected leaders consistently fight—and ultimately lose—asymmetric wars that autocrats avoid. Second, I provide qualitative evidence from several historical crises. The results challenge the prevailing view that democratic institutions encourage leaders to exercise discretion. Instead, domestic constraints can systematically compel accountable officials to fight riskier, costlier, and more lopsided wars than their unconstrained peers.
1.1 Introduction

In the spring of 1940, German armies rolled across Europe. As his soldiers marched into Denmark, Norway, Belgium, the Netherlands, and Luxembourg, Hitler presented the political leaders of each successive country with a choice: if they surrendered quickly, incumbent governments would retain their titles and a large degree of administrative influence throughout the subsequent occupation. If they resisted, however, Berlin would forcibly depose the leadership and install a new, more cooperative regime. As Hollander (2017, p. 46) summarizes “Germany had given each of these of these countries an offer it could not refuse.” Despite the overwhelming odds, each country but Denmark—even tiny Luxembourg—refused the offer anyway.

Why did these governments choose to wage wars they could not hope to win? Several of our most prominent theories of conflict are unconvincing. Surely optimism, for example, was not the motivating factor. None of the leaders could genuinely hope to win a war against Germany, nor did any expect that fighting would allow them to reveal unexpected strength and thereby obtain more favorable settlements in future negotiations. A second proposal—that the leaders were unwilling to commit to peace because they anticipated large shifts in the balance of power—is equally unsatisfying. Germany could already credibly threaten regime change and severe punishment in the states it conquered, so the defending countries should have viewed further tilts in favor of Berlin as largely immaterial. Finally, it seems unrealistic that the leaders believed their constituencies would be better served by war than peace.

After all, fighting would not only trigger substantial casualties in the short term but would also cause the Germans to impose significant punishments during the upcoming occupation. Why, then, did the various

---

1Countries may, for example, fail to reach settlements when they are overly optimistic about their respective prospects in war and are either unable or unwilling to share their private information with one another. See Fearon 1995, Morrow 1989, Slantchev 2003, Slantchev and Tarar 2011, Fey and Ramsay 2011, and Lindsey 2019.

2Actors may decline to compromise if they believe the distribution of power will change significantly in the future. See Fearon 1995, Powell 2006, Leventoglu and Slantchev 2007, Krainin 2017, and Merrell and Abrahams 2019.

3Although canonical models assume that fighting is costly, states may tolerate war when the price of sustaining peace is exorbitantly high—for example due to the necessity of participating in an arms race or the need to service debt obligations incurred during combat. See, for example, Coe 2012, and Slantchev 2012b.
governments decide to mount a futile defense?

In contrast to these existing explanations, I argue that the rationale for war is often grounded in domestic politics. Across a wide range of cases and circumstances—ranging from the failed defense of Europe in 1940 to the escalation of American involvement in Vietnam two and a half decades later—leaders sometimes choose to participate in aggressive military behavior for fear that doing otherwise will permanently jeopardize their domestic reputation or political standing at home. Put simply, my central finding is that leaders often engage in or escalate military operations in order to satisfy the demands of their political constituents. Indeed, even in Denmark, the only country that accepted the Reich’s offer, officials considered mounting a symbolic defense to safeguard public opinion even as German troops swarmed the capital.4

I develop the theory using an analytic stylization of a military crisis in which government leaders possess private information about the costs and risks of escalation. Because this information is not publicly available, the leader’s constituents may under some circumstances harbor naïvely optimistic beliefs about the desirability of using force. Indeed, citizens may even threaten to penalize leaders who pursue peaceful settlements. In these situations, leaders who seek to avoid violence face a dilemma: although they could attempt to dispel their constituents’ martial enthusiasm by sharing sobering information about the true costs of war, in doing so they would run two risks. First, a signal may not fully attenuate their constituents’ support for war. If voters misinterpret the message as evidence that the leader is unskilled in military affairs or dovish, citizens may react by penalizing the leader or installing an alternative who would escalate with greater intensity, thereby negating the predecessor’s effort to avoid conflict. On the other hand, a leader who successfully reveals the country’s weakness to a domestic audience may inadvertently also share the information with international observers. Such disclosures could jeopardize the nation’s bargaining position by emboldening the opponent to demand costlier concessions than they might otherwise attempt to extract or, alternatively, by dissuading potential

4 See Dethlefsen (1996, p. 25) and Merrell (2019a).
allies from offering support. These risks motivate leaders to conceal their true pessimism about war. Instead, leaders will sometimes pursue escalation despite private knowledge that accepting a settlement would leave the country better off.

To demonstrate the plausibility of the mechanism, I provide two forms of empirical support. First, I analyze international territorial exchanges that occurred between 1816 and 2014. I find evidence of a systematic relationship between regime type and crisis behavior in asymmetric conflicts. Among states that face territorial demands from a relatively strong opponents, democracies are less likely to concede territory peacefully than are autocracies. Instead, democratic leaders consistently fight—and ultimately lose—lopsided wars that autocratic countries are able to avoid. I complement the data by providing additional evidence from several historical cases in which the behavior and personal beliefs of state executives mirror the predictions of the theory.

My results yield several important theoretical and policy implications. First, whereas a long research tradition argues that “accountability to the public can restrain the war-making proclivities of leaders,” this project demonstrates that public optimism for war can also motivate peace-loving leaders to reject viable settlements and engage in counterproductive escalation. As a result, theoretical and empirical studies of conflict will remain incomplete until they account for the preferences of influential domestic constituencies. Second, whereas prevailing research argues that democratic leaders are “are highly selective... [and] prefer to negotiate when they do not anticipate military success,” I show that leaders who are accountable to the public are under some circumstances more likely to fight futile wars than their autocratic counterparts. Third, the theory contrasts with popular conceptions of domestic “audience costs.” Whereas conventional theories suggest that leaders can obtain bargaining advantages by “activating” domestic hard-liners, I show that leaders sometimes seek to suppress domestic enthusiasm for war but struggle to fully pacify

---

6 Bueno de Mesquita et al. 2003, p. 236.
their constituents. Finally, the results suggest important implications regarding counter-terrorism policy. If domestic opinion can compel leaders to escalate conflicts, democracies may present appealing targets for violent groups attempting to provoke draconian government behavior. As a result, countries may be better able to deter terrorism by developing institutions that inhibit leaders from unnecessary retaliation.

1.2 Theories of Crisis Escalation

1.2.1 Unitary Explanations for War

Fearon (1995) demonstrated the conceptual rewards of interpreting crisis diplomacy as a bargaining process. According to this stylization, war occurs when participants fail to either locate or implement acceptable settlements. In the quarter-century that followed Fearon’s insight, researchers largely consolidated to three distinct explanations for bargaining failure. First, fighting can occur when each participant in a dispute is privately optimistic about its own military prospects. Because states have incentives to bluff or otherwise misrepresent their capabilities, engaging in several rounds of hostilities may be the only means of dispelling each country’s optimism. The process of fighting facilitates settlements by revealing information about each opponent’s capabilities, thereby causing the combatants’ initially disjoint expectations to converge.

Second, countries may fail to strike bargains when they cannot convince one another that they will abide by the terms of a potential settlement. This type of commitment failure is

---

8See, for example, Lake 2002, Kydd and Walter 2006, Carter 2016.
10An earnest refusal to accept a peace deal—and therefore demonstrated willingness to commit oneself to battle—can credibly reveal the resolve of either side in ways that pre-war bombast cannot (Slantchev 2003). Likewise, a favorable outcome in the midst of fighting can attest to the relative strength of the winning belligerent and dispel the loser’s inflated perception of its own abilities.
11After observing unwelcome results, an opponent may be willing to accept a less favorable settlement than they would have tolerated in the absence of fighting. If the increase in the value of the possible settlement exceeds the costs incurred while fighting, the war was worthwhile.
particularly likely when one disputant anticipates that the balance of power will shift significantly and rapidly in favor of its opponent.  

Because a large power shift would leave the declining country relatively vulnerable to its rival, a state that foresees such a transition may choose to launch a preventive war before the shift in power begins in earnest. Although preventive wars are both costly and risky, within a range of specific conditions it becomes rational for a declining state to gamble on a short and decisive conflict in hopes of avoiding the inevitable sacrifices they would be forced to accept if their adversary successfully gained power.

A third rational, unitary explanation for war relates to the expected payoffs of conflict and peace. Although waging war is inherently costly, countries may also pay a steep price for prolonging the peace. Consider a military arms race between two countries who dispute a common border. Each participate must continually divert resources to the development of new weaponry that will deter the opponent from challenging the status quo division of territory. If the cost of producing military matériel is too high, it may be worthwhile for the states to wage a decisive war rather than permanently endure the burden of an arms race. Similarly, states who borrow to fund their military expenditures may find that their debts are unsustainable in the absence of spoils acquired while fighting, whereas the harms of defeat are dampened by the ability to repudiate their debts. In these cases, states do not fail to locate viable settlements but are instead forced to fight because no such settlements exist.

1.2.2 Domestic and Individual Accounts of Crisis Behavior

In addition to rationalist, unitary-state explanations, a variety of domestic mechanisms may facilitate the onset of war. Several authors, for example, focus on the personal characteristics of

---


14 Slantchev 2012b.
national leaders. McDermott (2001) and Horowitz, Stam, and Ellis (2015) each argue that leaders differ in their tolerance for risk and that this variation can explain their willingness to deploy troops. Similarly, Saunders (2011) suggests that leaders hold differing beliefs about the level and source of potential threats that are posed by international adversaries; she argues that leaders’ individual perceptions therefore influence the intervention strategies they select.

A second set of theories emphasize that leaders can acquire private benefits from fighting that are not available to the broader public. Although many such theories focus on a leader’s personal financial ties to the defense industry, the benefits may also relate to a decision-maker’s personal security. A leader may attempt to thwart a potential coup by initiating an international conflict and deploying likely coup participants to the front lines, thereby removing them from the leader’s proximity. Leaders may also engage in “diversionary” conflicts abroad in order to distract public attention from problems at home. In the most extreme case, a leader who expects to be ousted from office may “gamble for resurrection.” According to this theory, a leader may extend a conflict that constituents oppose in hopes that a favorable outcome will eventually convince the public that the leader is competent and should continue to retain office.

Other theories propose that domestic institutions rather than individual preferences can shape a leader’s willingness to wage war. Kriner (2010) argues that leaders who face an opposition legislature are more cautious about intervening abroad because they fear that political opponents will criticize the use of military force. Bueno de Mesquita et al. (2003) likewise contend that a leader’s willingness to participate in costly conflict is conditioned by the size of that leader’s “winning coalition,” which they define as the proportion of constituents the leader must satisfy in

---

15 Kertzer (2017) likewise shows that variation in risk preferences can explain why some individuals display more resolve than others, while Cohen (2016) and Saunders (2017) attribute leaders’ decision-making behavior to their personal experience serving in military or executive roles.

16 Chiozza and Goemans 2011.

17 Levy 1988; Richards et al. 1993; Mitchell and Prins 2004; Oakes 2006; Tarar 2006; Oakes 2012; Haun 2015 similarly argues that the leaders of weak states seek to fight foreign conflicts because doing so creates an impression that they are militarily strong and thereby deters domestic insurgents who may otherwise attempt coups.


19 Note that Kriner (2010) dismisses the possibility that deploying troops may be politically popular, as well as the possibility that the opposition party may face constraints in its capacity to criticize the incumbent.
order to retain office. The authors argue that democratically-elected leaders should be deterred from engaging in risky conflicts because, if such conflicts fail, elected leaders will face difficulty appeasing a diffuse winning coalition with private goods. In contrast, autocrats should be more risk-acceptant when entering crises because they can more easily target and compensate a small group of disenchanted supporters if the escalation backfires.

Finally, a broad literature argues that leaders can “activate” domestic audience costs in order to tie their hands during crises and access the strategic benefits of increased commitment.\textsuperscript{20} Note, however, that in each of these theories—with the exception of the audience cost mechanism—public dissatisfaction with the leader either precedes or is unrelated to the crisis. Moreover, the leader attempts to restore public confidence by securing a favorable result from war rather than with her decision to participate in the first place. In the following section, I develop a theory in which a leader’s decisions reflect not only the leader’s private beliefs about the outcome of a war, but also the public’s preferences regarding war initiation itself.

1.3 The Agency Dilemma

Leaders function as the agents of their constituents. In both democracies and autocracies, citizens and elites attempt to select representatives whose policies will reflect the desires of their supporters. Unfortunately for the interested parties, once in office public officials also face incentives to select policies that diverge from the interests of their constituents, either because they seek to secure private benefits that are unavailable to the constituency as a whole, or, alternatively, because they hope to insulate themselves from risks to which their supporters are exposed. To minimize this moral hazard problem, citizens can impose a system of incentives and punishments

\textsuperscript{20}See, for example, Fearon 1994, Smith 1998, Schultz 2001b, Baum 2004, Slantchev 2006, Tomz 2007, and Kurizaki and Whang 2015. Note that constituents function less as a source of pressure on the leader and more as a reactive mechanism that the leader can manipulate to her advantage. Likewise, the mechanism does not suggest that the public inherently supports aggressive military action; indeed, a leader must “activate” such sentiment by persuading the public to support a hardline stance.
that discourage elected officials from deviating from the wishes of their constituents. For example, citizens may observe the policies their representatives enact—and the outcomes that result from those policies—and then reelect or dismiss the officials on the basis of those outcomes.\textsuperscript{21}

Policy divergences between leaders and their constituents, however, are not always rooted in the leader’s own self-interest. Well-intentioned and sociotropic public servants may possess private information about the state of the world or the implications of potential policies that convinces them that their constituents’ preferred strategies are misguided. In this case the same mechanisms that reduce moral hazard instead create perverse incentives for public pandering. A leader who anticipates punishment if she deviates from her constituents’ instructions may opt to appease her constituents by selecting a popular policy even if she privately believes that the policy will harm those constituents in the long run.\textsuperscript{22}

This tradeoff—wherein tightening the reins to reduce moral hazard diminishes an official’s ability to exercise discretion or draw upon private knowledge while crafting policy—is particularly acute in the context of wartime decision making. Relative to the general public, elected officials possess significant informational advantages in areas relating to national security. Leaders often have access to classified information about the state’s own capabilities, estimates of enemy strength, strategic and tactical plans, appraisal’s of the adversary’s likely negotiating behavior, information about the behavior of potential allies, and many other details that may be denied to members of the broader public and even other government officials. Moreover, during military crises leaders are often unable to disclose relevant information without either jeopardizing their country’s strategic position or opening themselves to domestic criticism.\textsuperscript{23} Finally, the public is both uniquely attentive to government decision making during the opening phases of international conflict and is also poorly equipped to evaluate the net effects of security policies as time elapses.

\textsuperscript{21}Downs and Rocke 1994.
\textsuperscript{22}Canes-Wrone 2001, Bueno de Mesquita and Dickson 2007, Bas 2012, Schneider and Slantchev 2018.
\textsuperscript{23}Merrell 2019b.
impassioned citizenry lays the groundwork for the selection of suboptimal policies.

1.3.1 Public Military Optimism

Theories of public attitudes toward war often suggest that citizens are more dovish than their leaders. Although small interest groups may hold hawkish preferences, the public at large is thought to acknowledge the costs of war;\textsuperscript{24} to disapprove of conflict with other democracies;\textsuperscript{25} and to react unfavorably to increases in the amount, duration, or intensity of conflict-related casualties.\textsuperscript{26} As Morgan and Campbell 1991, p. 189 summarize, “the key feature of democracy is government by the people and... the people, who must bear the costs of war, are usually unwilling to fight.” As a result, researchers typically argue that public sentiment acts as a useful constraint upon belligerent executives who might otherwise initiate controversial and costly wars.\textsuperscript{27} Because “American military operations require public support,”\textsuperscript{28} public intolerance for military adventurism forces electorally-vulnerable leaders to behave more cautiously than their autocratic peers.\textsuperscript{29} Overall, these assumptions cultivated a belief that democratic institutions had a pacifying influence on crisis behavior. According to Moravcsik (1997, p. 531), liberal democracies are unlikely to provoke wars “because influence is placed in the hands of those who must expend blood and treasure.” Baum and Potter (2015, p. 45) echo the sentiment, claiming that “public scrutiny may, under at least some circumstances, deter leaders from using military force by disproportionately raising the expected costs of doing so.” Likewise, Caverley (2014, p. 9) describes a “remarkable consensus” within political science, “that when democracy ‘works’ a

\textsuperscript{24}Doyle 1986
\textsuperscript{25}Tomz and Weeks 2013
\textsuperscript{26}Aldrich et al. 2006; Baum and Potter 2008.
\textsuperscript{27}See, prominently, Lake (1992), although note also that the prevailing view contrasts with earlier work that found public opinion fickle and its influence potentially damaging. See, for example, Lippmann (1955), Lindsey and Lake 2014, and even Alexander Hamilton, who wondered in Federalist 6 whether “republics [have] in practice been less addicted to war than monarchies?” and concluded that “The cries of the nation... have, upon various occasions, dragged their monarchs into war, or continued them in it, contrary to their inclinations, and sometimes contrary to the real interests of the State” (quoted in Hamilton, Madison, and Jay 2008).
\textsuperscript{28}Klarevas 2002, p. 419.
\textsuperscript{29}Chiozza and Goemans 2003; Filson and Werner 2004.
moderate, effective foreign policy results.”

In contrast to prevailing theories of public pacifism, the historical record suggests that voters can exhibit either hawkish preferences or naïve optimism regarding the use of force. Significant proportions of Americans advocated swift intervention in Afghanistan following the attacks of September 11, 2001, and against Japan in the wake of the Pearl Harbor bombing. Similarly, polls conducted prior to the 2003 invasion of Iraq report that 72% of Americans supported a potential invasion even though as few as 13-37% could locate Iraq on a map. Such instances of optimism about war have earned minimal attention from researchers because they pose no political dilemma when coupled with the assumption that leaders are inherently bellicose. After all, when citizens and leaders jointly align in favor of military action it is not surprising that war will result. However, an alternative relationship between citizens and leaders remains largely unaddressed: under certain conditions, citizens are either more optimistic than their leaders about the use of military force or, alternatively, more hesitant to support a settlement. Voters’ support for escalation should exceed that of their leaders whenever available information leads the public to overestimate the threat posed by an adversary or underestimate the likely costs of conflict.

The existence of pessimistic leaders and optimistic citizens raises the possibility that institutions designed to enhance leader accountability may encourage those leaders to executive the desires of an aggressive domestic audience. When bellicose constituents are willing to penalize officials for exercising or advocating restraint, public officials may feel compelled to pander to their constituents by engaging in wars that they privately believe are less desirable than otherwise viable settlements. This is possible when three minimal conditions hold. First, the leader and

---

30 Several recent working papers also document hawkish sentiment on behalf of the public. See, for example, Merrell 2019b, Fang et al. 2017, and Kreps, Saunders, and Schultz 2018.
31 Consider the case of Colombian voters, who rejected a settlement that their government forged with the FARC, or of Danish citizens who rebelled against their own government when it cooperated with Nazi occupiers during World War II (Merrell 2019a).
32 Caverley (2014) contends that the median voter may support interventionism because the costs of conflict are disproportionately borne by a minority of the population.
the public must diverge on the expected payoff of fighting, with the public more optimistic than
the leader and therefore unwilling to tolerate a settlement the leader would privately accept.
Second, there must be a low probability that public support for an aggressive strategy will rapidly
evaporate once such action is underway. In other words, if public support for war will vanish
rapidly once the true costs of fighting are revealed, then so too may public support evaporate
for leaders who blunder into unnecessary wars. Finally, the leader must face some obstacle
or disincentive that prevents her from going public with information that could temper public
hawkishness.

1.3.2 Barriers to Information Transfers

Consider a bargaining interaction that closely resembles that advanced by Fearon (1995), in which
a leader \( L \) and a foreign adversary \( F \) compete for control of a continuously divisible good
represented by the unit interval \([0, 1]\). Division of the good can occur in two ways. The first is by
mutual agreement to impose a particular division \( x \), in which case \( L \) obtains \( x \) while \( F \) obtains
\( 1 - x \). Alternatively, the two sides can fight a costly all-or-nothing war if they are unable to agree
on a potential division.

Using her knowledge of each country’s military equipment and personnel, the type of
combat that could occur, the terrain around which fighting would take place, and various other
factors, \( L \) identifies \( L_o \) as the expected outcome of war if one should occur. She further identifies
\( c_L \) as the expected cost that her country would incur while fighting. Given these values, \( L \) would
accept any settlement \( x \) such that \( x > L - c_L \). In a departure from Fearon, assume that \( L \) is
accountable to a domestic actor \( D \) who lacks access to the same information as \( L \) and who is
therefore overly optimistic about the expected outcome or costs of fighting. As a result of its
optimism, \( D \) calculates an expected war outcome \( D_o > L_o \) and expected cost of combat \( c_D \leq c_L \).
These values cause \( D \) to prefer settlement to war only when \( x > D - c_D > L - c_L \).

To depict \( L \)’s accountability to \( D \), I further assume that \( D \) can penalize \( L \) for accepting
divisions that fall outside of the range that $D$ would prefer. In practical terms, such penalties could include physical violence, financial harm, removal from office, political opposition to $L$’s preferred domestic policies, or even deliberate replacement with a more aggressive alternative. If the intensity or likelihood of $D$’s punishment scales with the difference between the division $x$ that $L$ accepts and the minimal division $x = D_o - c_D$, then $L$ should be increasingly unwilling to accept settlements the further they diverge from $D$’s desires. Figure 1.1 below illustrates this relationship graphically.

**Figure 1.1: Bargaining with Domestic Constraints**

### Domestic Signaling Challenges

Within this framework, potential settlements are only problematic when they fall in the blue area of disagreement between $L$ and $D$. For example, if the foreign adversary, $F$, insisted on division $x < L - c_L$, both the leader and domestic group would prefer to reject this proposal. Similarly, if $F$ offered a settlement $x > D - c_D$, the leader would accept and the domestic group would approve of her decision. However, proposals in the region between $L - c_L$ and $D - c_D$ cause disagreement, with $L$ preferring to accept and $D$ preferring to reject such proposals.

Consider the result if $L$ accepts a settlement in the controversial region. In this case, war does not occur and $D$ remains ignorant of the true outcome and costs that would have been produced through combat. As such, $D$’s optimism regarding the use of military force may not be diminished and $D$ may seek to penalize $L$ for exhibiting military restraint. To avoid this penalty, $L$ may attempt to temper $D$’s optimism about the expected outcome and/or costs of war by sharing
informative information while the crisis is ongoing but before settlement or conflict occurs. However, $L$ may face several challenges that restrict her from fully persuading $D$ that a settlement is optimal. First, $D$ may simply be insufficiently responsive to new information about the likely costs of conflict. In other words, even an earnest effort to share information with domestic hawks may fall on deaf ears or fail to motivate a shift in deeply entrenched opinions. Merrell (2019b) provides experimental support for this possibility: in a survey experiment assessing support for military action, respondents’ support for troop deployments and drone strikes did not significantly deteriorate after exposure to information that such military actions may be highly costly. Historical evidence also suggests that leaders fear that domestic hawks are difficult to persuade. As I discuss later in this paper, Lyndon Johnson believed that his three presidential predecessors had thoroughly primed Americans to worry about the threat of communism. As such, Johnson doubted that he could convince voters that American support for South Vietnam was unnecessary or risky.\footnote{Skowronek 1997, p. 343-344.}

Leaders may face difficulty tempering domestic optimism even when audiences are highly attuned to public messages. Signals from leaders do not exist in a vacuum. Instead, leaders must compete for attention with opposing politicians, members of the media, and other actors who may advocate military aggression rather than settlement.\footnote{For example, Baum and Groeling (2005) show that opposition party criticism of incumbent politicians is prevalent before and in the immediate aftermath of domestic “rally” events, while Baum and Groeling (2009) show that the media tends to overrepresent in-party criticism of a leader’s decisions while underreporting supportive rhetoric.} Domestic audiences may struggle to discern which of these sources can provide accurate information about the desirability of war or settlement.\footnote{Kahneman and Renshon (2009) argue that citizens’ psychological biases can make individuals more receptive to and more easily persuaded by hawkish arguments than less aggressive messages, while Ashworth and Shotts (2010) show that voters apply an asymmetric burden of proof to incumbents based on whether those leaders pursue popular or unpopular policies.}

Two addition barriers to information revelation relate to the domestic group’s perception of $L$’s inherent characteristics. Although the stylized depiction assumes that $L$ can perfectly

\begin{itemize}
  \item[33]Skowronek 1997, p. 343-344.
  \item[34]For example, Baum and Groeling (2005) show that opposition party criticism of incumbent politicians is prevalent before and in the immediate aftermath of domestic “rally” events, while Baum and Groeling (2009) show that the media tends to overrepresent in-party criticism of a leader’s decisions while underreporting supportive rhetoric.
  \item[35]Kahneman and Renshon (2009) argue that citizens’ psychological biases can make individuals more receptive to and more easily persuaded by hawkish arguments than less aggressive messages, while Ashworth and Shotts (2010) show that voters apply an asymmetric burden of proof to incumbents based on whether those leaders pursue popular or unpopular policies.
\end{itemize}
calculate the expected costs and outcome of war, in reality leaders can only form estimates. Domestic audiences may doubt a leader’s military knowledge and may therefore conclude that the leader’s prediction about the likely payoff of combat is incorrect. Alternatively, $D$ may suspect that different types of leaders exist who are distinguished by their personal sensitivity to the costs of war. Put another way, some leaders may be innately dovish and therefore highly sensitive to costs incurred while fighting, while others are hawkish or relatively insensitive to casualties and other costs associated with war. If optimistic domestic audiences believe that $L$ is extremely dovish, they may conclude that her reluctance to fight stems from personal aversion to costs that the nation as a whole will not bear.

Leaders who anticipate such challenges may avoid signaling attempts altogether. This is particularly true if the leader fears that by advocating restraint or revealing her relative pessimism about the expected outcome of war she may provoke domestic backlash.\textsuperscript{36} For example, prior to his assassination President Kennedy debated whether to reduce American involvement in Vietnam despite public support for an expansion of the conflict. As he privately remarked to an aide, “If I tried to pull out completely now from Vietnam we would have another Joe McCarthy red scare on our hands.”\textsuperscript{37} Kennedy’s successor, Lyndon Johnson, similarly feared that the public would respond to an American draw-down in Southeast Asia by electing his more radical and hawkish opponent, Barry Goldwater. To prevent the election of a successor who would blunder into an even more costly conflict, Johnson underplayed his own opposition to escalation in the buildup to the election and instead pursued a moderately aggressive policy he hoped would appease hawkish audiences—particularly following the Gulf of Tonkin incident.

\textsuperscript{36}Merrell (2019b) finds that leaders who refrain from military escalation and face criticism for such inaction face significant losses of respondent support relative to those who pursue escalation, even in the face of similar criticism.\textsuperscript{37}Quoted in Gardner (1995, p. 72).
Mixing Signals with Multiple Audiences

An additional set of signaling challenges emerge when one considers the existence of $F$, the foreign adversary. Let $F$ form his own predictions about the expected outcome of war and the expected costs of combat, $F_o$ and $c_F$, respectively. Because $F$ obtains $1 - x$ in the event of peaceful settlement, $F$ would accept any settlement $x$ such that $x < F + c_F$. Figure 1.2 provides an example of this bargaining interaction in which both $D$ and $F$ are privately optimistic about the likely results of war. Thus, $F$ would accept only divisions in which $x$ is relatively small, while $D$ would accept only those divisions in which $x$ is quite large.

![Figure 1.2: Bargaining with Multiple Audiences](image)

With this distribution of initial beliefs, $L$ faces the challenge of facilitating a convergence in expectations for $F$ and $D$. Although $L$ may be able to accomplish this goal by convincing the other actors that the likely costs of war, $c_F$ and $c_D$ are much larger than either initially assumes, the realities of the conflict may be such that sending such a message is unrealistic. Instead, $L$ may need to convince $F$ that $L$’s military is stronger than $F$ perceives while simultaneously persuading $D$ that $L$’s military is weaker than $D$ believes. Sending both messages simultaneously and convincingly may present a difficult challenge for $L$.

Finally, $L$ may face difficulty creating convergence even when $L$ and $F$’s initial beliefs enable them to easily locate potential settlements. Figure 1.3 illustrates such a scenario. In this case, the leader and the foreign adversary would both agree to and division such that $L - c_L < x < F_o + c_F$. However, $D$ remains optimistic and refuse any settlement in which $x < D - c_D$. In this case, $L$ need not send mixed messages to $F$ and $D$. However, $L$’s willingness
to attempt convergence depends on two issues. The first is the efficiency with which $L$ can dispel $D$’s optimism, and the latter is the size of the bargaining range that remains after $L_o$ and $F_o$ converge.

The bottom half of Figure 1.3 illustrates the problem: in this case, $L$ has exerted effort to signal the true expected outcome of war. As a result, $F$ has updated its belief $F_o$ to converge with $L_o$. However, in this case $D$ updated its beliefs about $D_o$ less substantially. Although $L$ may continue to share information in hopes of further removing $D$’s optimism, the act of sharing information credibly may force $L$ to incur a cost. If the cost of facilitating convergence between $D$ and $F$ exceeds the size of the bargaining range—in other words, $(F_o + c_F) - (L_o - c_L)$, then $L$ will prefer to reject a settlement and pursue war rather than attempt further signaling. The costs of signaling, combined with the risk of any penalties $D$ could impose if convergence does not succeed, may motivate $L$ to forgo signaling attempts and to pursue costly wars even though $L$ recognizes that viable settlements exist.

1.3.3 Divergence from Previous Domestic War Theories

It is useful to distinguish the agency dilemma from alternative domestic war phenomena. First, domestic audiences are known to “rally ’round the flag” by raising their support for a government in the early stages of conflict.\(^38\) However, such research often refers to surges in overall patriotism as opposed to increases in support for the leader herself. Moreover, the agency dilemma mechanisms differ from traditional rallying effects in that I expect leaders to receive benefits only conditional on escalation. Rather than rallying around leaders regardless of how they react during crises, constituents will only support leaders whose escalatory choices coincide with prevailing optimism.

Second, “diversionary wars” occur when leaders initiate foreign conflicts to distract from

\(^38\)Lee 1977.
political problems at home.\textsuperscript{39} By identifying a foreign rival, the leader can create a scapegoat for domestic actors’ frustration. My mechanism, on the other hand, does not rely on the preexistence of domestic disturbances, nor does it require that the leader construct a threat where none exists. Instead, domestic antipathy or fear of an international rival precedes the leader’s interest in violence and this sentiment compels the leader to engage in escalatory military action.

Third, “gambling for resurrection” occurs when leaders escalate or prolong conflicts in hope that a lucky outcome will convince constituents that the choice to escalate was well-founded.\textsuperscript{40} In contrast, I argue that leaders enter into conflicts that they know are inferior to settlements and must hope that the costliness of such conflicts will not be quickly revealed once they occur.

Finally, audience cost theories depict circumstances in which leaders can gain a bargaining

\textsuperscript{39}Levy 1988; Richards et al. 1993; Tarar 2006.
\textsuperscript{40}Downs and Rocke 1994. See also Chiozza and Goemans 2011, who discuss “gambling for survival.”
advantage by “activating” domestic hawkishness.\footnote{Fearon 1994; Slantchev 2006; Tomz 2007.} As Crisman-Cox and Gibilisco (2018), “[audience] costs serve as a commitment device... coercing [an] opponent to back down.” Schultz (2001a) similarly states that “democratic states have managed to use threats very effectively because... the ability to signal resolve in a credible manner bestows important advantages.” My theory differs significantly from this behavior in that I do not expect leaders to deliberately cultivate hawkish sentiments. Rather, they would prefer to disabuse citizens of their enthusiasm for war but are inhibited from doing so.

1.4 Cross-National Evidence

1.4.1 Setup, Data, and Variables

The theory suggests that leaders may be punished not only for failing to win wars, but also for failing to attempt wars that their constituents believe should have been fought. Rather than attempt the difficult, risky, and costly task of persuading a domestic audience that settlement is optimal, leaders may instead engage in violence that they know is counterproductive for the country as a whole. This theoretical framework yields two testable predictions. First, leaders who are accountable to optimistic domestic groups should be less likely to settle and more likely to escalate crises than leaders who are not similarly accountable. Second, leaders who are accountable to optimistic constituents should perform less successfully in conflicts conditional on escalation.

Testing these predictions is complicated by the fact that international crises result from strategic selection: the set of countries that are targeted during crises may differ from the set of states that are never targeted. To address this concern, I use data from the Tir et al. (1998) Territorial Change (v5) dataset, which encompasses all international territorial changes that occurred between 1816 and 2014 that involved at least one nation-state. In general, territories
change hands when one state is coerced or compelled to offer a concession to another. However, countries sometimes offer territory freely. As such, I further restrict my analysis to the subset of observations in which territory was exchanged as a result of conquest, annexation, or the presence of a threat. Because my sample includes only cases in which (A) one country levied a territorial claim against another, and (B) the claim eventually succeeded, I am able to at least partially reduce heterogeneity related to strategic selection of target states on behalf of the claim-initiator.

My dependent variable in the analysis is whether the conceded territory was exchanged peacefully. When concessions occur without fighting, the variable assumes a value of “1,” even if an implicit threat of force existed. In cases where fighting occurred prior to the change of territory, the variable takes a value of “0.” Because the dependent variable is dichotomous, I use logistic regression throughout my analysis.

The independent variable of interest is an interaction between two terms. The first term is the polity score of the conceding state, which characterizes the degree to which a country is either an autocracy or democracy. The second term is dichotomous. It assumes a value of “1” when the conceding state is less powerful than the country to whom it grants territory, as determined by comparing the two countries’ “Composite Index of National Capabilities” (CINC) scores in the year that the concession occurred. Henceforth, I refer to this as a “power deficit.” In each model, I control for several additional variables that may affect a country’s willingness to fight rather than settle. These include the size of the contested territory in square kilometers, whether or not the territory was contiguous with the main body of the targeted state, and the population of the disputed territory.

42The polity variable ranges from -10 to 10, with higher values reflecting greater levels of democracy (see Marshall and Jaggers 2002). In other specifications, I use binary indicators of whether a country (1) engages in executive elections or (2) engages in parliamentary elections. Results are generally consistent across these specifications.

43The CINC index is a composite figure derived from a country’s population, urban population, iron and steel production, energy consumption, military personnel, and military expenditure (see Singer, Bremer, and Stuckey 1972). In other models, I compare military spending and the number of active military personnel rather than CINC scores as proxies of relative military power.
1.4.2 Expectations and Results

Because aggressors levy territorial claims strategically and the dataset is composed of observations in which territory ultimately changed hands, I expect to find no relationship between a power deficit and the likelihood of a settlement in autocratic states, where leaders make sober decisions about whether to accept the new division of territory or to wage war. However, I do not expect this to hold true in democracies, where leaders are subject to influence from their constituents. Among democratic states, I expect that leaders whose countries face power deficits will be more likely to fight than those who do not face such an imbalance of power.

My explanation for this prediction is that when aggressors target relatively weak democracies, they issue demands that are calibrated based on the observable balance of power but which do not account for potentially hawkish preferences among the domestic population of the target state. In other words, aggressor countries may issue territorial demands that appear likely to succeed based on an objective reading of each country’s relative power, but against democracies a subset of those demands will ultimately fail because constituents within the targeted country are naively optimistic and refuse to tolerate the requisite concessions.

Table 1 displays results from the baseline model as well as versions that include several control variables. The results are consistent with my expectations. Across each model, the interaction between a country’s polity score and its relative strength is significantly associated

---

44 When a country issues a credible demand for territory, it calibrates the size of its demand based on its expectation that the opponent will acquiesce as well as the expected cost of a conflict if the opponent chooses to fight. Because fighting is more costly than settling, aggressor states should issue the largest possible demand that they think an opponent would accept with satisfactory probability. Unless risk acceptance or the availability of information about the opponent’s capability are correlated with relative power, relative power should not be associated with the target state’s likelihood of settling rather than fighting.

45 I assume countries can observe an adversary’s military capability more easily than the enemy’s latent public opinion.

46 For evidence that voters resist settlements in territorial disputes, see Zellman (2018) and Fang et al. 2017. Note that this assumes aggressors who issue territorial demands would suffer a cost for scaling back those demands should they realize that the likelihood of conflict is greater than anticipated. Absent such a cost, the aggressor could simply fail to act upon the threat it issued once it became apparent that resistance was likely (see, similarly, Ramsay 2017).

47 As anticipated, states are less likely to peacefully concede contiguous territories and highly populated regions than alternative regions.
with the likelihood that a country concedes territory peacefully. For easier interpretation of the interaction, Figure 1.4 depicts the estimated marginal effect of a power deficit on the likelihood of settlement across various Polity scores for the targeted state. In autocracies, a power deficit is not significantly associated with a change in the likelihood that a territorial dispute will end peacefully. However, democracies who face power deficits are more likely to fight before conceding territory compared to democracies that do not face such an imbalance in relative power. Figure 1.5 provides a graphical depiction of a similar interaction, but with the continuous Polity score replaced with a dichotomous “Executive Elections” variable.\(^{48}\) Finally, Figure 1.6 provides predicted probabilities of settlement conditional on whether the targeted state is democratic or autocratic and the size of the defender’s power deficit, measured as the difference between the aggressor and defender’s CINC scores. Autocracies become more likely to settle as their power deficit increases, but the same is not true of democracies. This relationship contrasts with a wide literature that argues (1) democratic leaders should exercise caution when escalating conflicts, and (2) that autocrats should be more averse to making concessions than elected officials.\(^{49}\) Instead, the results suggest democratic leaders feel compelled participate in conflicts that they subsequently lose, but which their autocratic counterparts are able to avoid.

\(^{48}\) Results are similar if I instead use a dichotomous “Democracy” variable that takes a value of 1 when the defender’s Polity Score exceeds 6.

\(^{49}\) See, for example, Bueno de Mesquita et al. (2003), Debs and Goemans (2010), Horowitz, Stam, and Ellis (2015), and Reiter and Stam (2002).
<table>
<thead>
<tr>
<th></th>
<th>(1)</th>
<th>(2)</th>
<th>(3)</th>
<th>(4)</th>
<th>(5)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Polity × Power Interaction</td>
<td>-0.114***</td>
<td>-0.088**</td>
<td>-0.116***</td>
<td>-0.105**</td>
<td>-0.082*</td>
</tr>
<tr>
<td></td>
<td>(0.042)</td>
<td>(0.043)</td>
<td>(0.042)</td>
<td>(0.043)</td>
<td>(0.044)</td>
</tr>
<tr>
<td>Loser’s Polity Score</td>
<td>0.109***</td>
<td>0.087**</td>
<td>0.109***</td>
<td>0.093***</td>
<td>0.073*</td>
</tr>
<tr>
<td></td>
<td>(0.032)</td>
<td>(0.034)</td>
<td>(0.032)</td>
<td>(0.036)</td>
<td>(0.038)</td>
</tr>
<tr>
<td>Less Powerful</td>
<td>-0.317</td>
<td>-0.296</td>
<td>-0.308</td>
<td>-0.256</td>
<td>-0.229</td>
</tr>
<tr>
<td></td>
<td>(0.305)</td>
<td>(0.313)</td>
<td>(0.306)</td>
<td>(0.308)</td>
<td>(0.317)</td>
</tr>
<tr>
<td>Contiguous Territory</td>
<td>-0.679*</td>
<td></td>
<td></td>
<td></td>
<td>-0.682*</td>
</tr>
<tr>
<td></td>
<td>(0.373)</td>
<td></td>
<td></td>
<td></td>
<td>(0.376)</td>
</tr>
<tr>
<td>Area (Sq. Kilo.)</td>
<td></td>
<td>-0.0005</td>
<td></td>
<td>-0.001</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>(0.001)</td>
<td></td>
<td>(0.001)</td>
<td></td>
</tr>
<tr>
<td>Joint Democracy</td>
<td></td>
<td></td>
<td></td>
<td>0.364</td>
<td>0.246</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>(0.509)</td>
<td>(0.516)</td>
</tr>
<tr>
<td>Constant</td>
<td>1.478***</td>
<td>1.892***</td>
<td>1.608***</td>
<td>1.370***</td>
<td>1.990***</td>
</tr>
<tr>
<td></td>
<td>(0.241)</td>
<td>(0.345)</td>
<td>(0.317)</td>
<td>(0.262)</td>
<td>(0.427)</td>
</tr>
<tr>
<td>Observations</td>
<td>298</td>
<td>283</td>
<td>298</td>
<td>292</td>
<td>277</td>
</tr>
</tbody>
</table>

Notes: * p < 0.1, ** p < 0.05, *** p < 0.01
**Figure 1.4**: Estimated Effect of a Power Deficit (Polity Score : Power Deficit Dummy). Figure 1.4 plots the marginal effect of the defender suffering a *power deficit* across a range of *Polity scores* for the defender. The shaded region depicts 95% confidence intervals.

**Figure 1.5**: Estimated Effect of a Power Deficit (Election Dummy : Power Deficit Dummy). Figure 1.5 plots the marginal effect of the defender suffering a *power deficit* when the defender does not (0) or does (1) exhibit *executive elections*. The vertical lines depict 95% confidence intervals.
Figure 1.6: Probability of Settlement Conditional on Size of Power Disadvantage and Polity Score. Figure 1.6 plots the predicted probability a territorial dispute ends in settlement rather than war, given the size of a defender’s disadvantage in relative CINC score and whether the defender is an autocracy or democracy. Shaded regions depict 95% confidence intervals.

1.5 Historical Episodes

The analysis of territorial settlements identifies a relationship between military behavior and the domestic constraints that leaders face. However, the pattern may result from mechanisms other than public pressure. Alternatively, public optimism about war may be endogenous to leader behavior, as in “audience cost” models where leaders deliberately activate hawkish opinion during crises. To address these concerns, I supplement the statistical results with three historical analyses of executive behavior. Across each of these examples, citizens were more optimistic about the payoffs of military escalation than was their head of state. Furthermore, none of the leaders deliberately cultivated public optimism; instead, they frequently lamented their inability to sway public opinion in favor of deescalation or settlement. Finally, each leader eventually adopted more aggressive military behavior than he would likely have attempted in the absence of public pressure. As such, the cases demonstrate that leaders sometimes believe that their constituents will penalize leaders who decline to escalate.
Although the cases are not intended to “test” the mechanism at work in the theory, they may enhance our confidence that the cross-national patterns we observe are plausibly generated by the agency dilemma mechanism. Moreover, they suggest that the mechanism may apply across a broad range of difficult circumstances. For example, the behavior of the Chilean, Bolivian, and Peruvian presidents in the War of the Pacific illustrates that public pressure can simultaneously influence the decisions of several opponents at once, forcing each to participate in a war that none desire. Second, the hesitation of French leaders to surrender to Nazi Germany in 1940 demonstrates that politicians may feel pressure to persist in costly and unnecessary wars even when their constituents are familiar with the costs of war and are relatively pacifistic as a result. Finally, U.S. President Lyndon Johnson secured an overwhelming electoral victory shortly before increasing American military presence in Vietnam. The fact that he succumbed to political pressure and engaged in costly and inefficient escalation demonstrates both that the theoretical mechanism applies to offensive wars fought abroad as well as those intended to protect home soil and, moreover, that even very popular leaders may escalate when they believe failure to do so could jeopardize their domestic agenda.

1.5.1 The War of the Pacific

Background to the Crisis

The Atacama Desert runs along the western edge of South America from roughly 21° to 27° south latitude. The driest region on earth, the Atacama offers little trace of life. As a result, when Bolivia and Chile obtained their independence from the Spanish Empire in the early 1800s they wasted no time squabbling over where precisely to draw their border in the desert. Only after guano and nitrate deposits were discovered in 1840—and when their use as fertilizers was publicized the following year—did either country recognize the inherent value of the region. In 1842, as investors clamored for mining rights, the Chilean government officially defined its
northern border for the first time, declaring a boundary line of 23° south. The proposal stretched several hundred kilometers into territory that Bolivians considered their own. However, because La Paz was not prepared to contest the Chilean claim militarily, its complaints were largely ignored in Santiago until 1864, when the two countries agreed to set a new border at 24° south but to divide equally all duties from extraction conducted between 23° and 25° south.

Bolivia and Chile revisited the issue of the Atacama again in 1874. They agreed that Santiago would relinquish its claim to all territory north of 24° south and, in exchange, Bolivia would impose a twenty-five year moratorium on any tax increases that could affect Chilean firms. Three years later, in the summer of 1877, a tremendous tidal wave struck the Pacific coast. Antofagasta, the capital of Bolivia’s Atacama province, was particularly hard-hit by the disaster. In response, the town’s municipal council imposed a small property tax as well as an emergency export tax of roughly ten centavos per 100 pounds of nitrates. The Chilean company most severely affected, the Compañía de Salitres y Ferrocarril (CSFA), refused to pay its dues, citing the tax moratorium that was agreed in 1874. The Bolivian legislature, however, denied the CSFA’s appeal, arguing that the federal government was legally prohibited from invalidating municipal tax laws.

Chile’s ambassador in La Paz, Pedro Nolasco Videla, announced that Bolivia’s actions constituted an abrogation of their 1874 agreement and warned that in response Chile might reclaim the region between the 23rd and 24th parallels. The statement, however, was largely bluster—even the CSFA confided that they would prefer to avoid any conflict that might disrupt the nitrate trade. So eager was the company for a restoration of normalcy that they offered to make a voluntary payment of 1,600 pesos per year to the local government, a sum thought sufficient to cover a large proportion of Antofagasta’s reconstruction costs. Even so, Videla’s bluff succeeded in persuading La Paz to rethink the issue of taxation. In the wake of the Chilean threat, Bolivian Foreign Minister Manuel Ignacio Salvatierra announced that although the federal government

---

50 Farcau 2000, p. 40.
could not formally nullify the tax, they could assure Chile that it would not be collected.

The Fierro-Sar ratea Treaty and Domestic Outcry in Chile

Even in the midst of the taxation dispute with Bolivia, Chile’s diplomatic attention was focused elsewhere. To the east, the Argentine Republic was rapidly expanding its navy. Longtime rivals, the two countries disputed ownership of the Strait of Magellan and Patagonia. Seeking a permanent solution to the conflict, President Aníbal Pinto dispatched a history professor, Diego Barros Arana, to negotiate with Buenos Aires. Although Pinto had instructed Arana to offer Patagonia to the Argentines in exchange for the Strait, Arana instead surrendered the former while securing only joint custody of the latter. When Pinto received the news, he asked to reopen negotiations. Buenos Aires agreed and this time dispatched their own delegate to negotiate with Pinto directly. Their choice, Manuel Bilbao, authored a series of articles in Santiago’s major newspaper, El Ferrocarril, that thoroughly disparaged the Chilean people for their poor negotiating skills. In response, thousands of Chilean citizens rioted in the capital, destroying a statue erected in honor of Argentina and urging the president to reject territorial compromise.

Despite public outrage, Pinto acknowledged that his country could ill afford to fight a war. In early December he proposed a deal whereby Chile and Argentina would share custody of the Strait until an international arbitrator could settle the dispute once and for all. The Chilean public was incensed. Members of the press denounced the “miserable policy” and predicted that the legislature would reject a document that “so shamelessly betrays Chile.”51 Nevertheless, two days after Pinto signed the agreement, the upper house of the Chilean congress followed his lead. The lower house, more inclined to follow the sentiments of the public and already jockeying for position an the upcoming congressional election, was more reluctant. They still had not agreed to the peace treaty six weeks later when word arrived of a new crisis involving Bolivia.52

Escalation

When Bolivian president Hilarión Daza learned that Pinto had adopted a soft line in territorial disputes with Argentina, he assumed Pinto would be equally easy to bully on the issue of Antofagasta. Shortly after the Fierro-Sarratea Treaty was signed in Santiago, Daza declared that the ten-centavo tax that his government had assured would never be collected was not only reinstated but also applied retroactively.\(^{53}\) Unfortunately for Daza, Pinto called his bluff by deploying an ironclad, the *Blanco Encalada*, to the Antofagasta harbor and ordering the remainder of the Chilean fleet to mobilize for war. Rather than back off, Daza doubled down and declared that the CSFA’s contract was now void and that the company’s property would be auctioned off in mid-February.\(^{54}\)

In Santiago, the people demanded action. Already frustrated with Pinto’s concessions to Argentina, Chileans would not tolerate a similar outcome with Bolivia. They argued that national honor was at stake. The Chilean newspaper *El Taller* warned that if Chile would appear a “nation of shameless imbeciles” and would sacrifice the respect of the continent if Pinto accepted Bolivia’s actions.\(^{55}\) In a letter to Pinto, Interior Minister Antonio Varas summarized the public sentiment when he remarked that rioters were “marching beneath my window with an enthusiasm which I have not witnessed in my life. Either we occupy Antofagasta or they [the war opponents] will kill you and me.”\(^{56}\) Pinto privately held substantial reservations, but eventually he was persuaded to act. Four days before the proposed sale of CSFA property, Ambassador Videla reiterated his claim from the previous autumn: by implementing a new tax, Bolivia had violated the 1874 moratorium. Unless Daza repealed the tax or agreed to international arbitration within forty-eight hours, Chile would feel justified in reoccupying all territory south of the 23rd parallel.

When Daza refused, Videla requested his passports and severed diplomatic ties. Two days later,

---

\(^{53}\) Farcau 2000, p. 41.
\(^{54}\) Sater 1986, p. 5
\(^{55}\) Sater 1986, p. 9.
\(^{56}\) Sater 2007, p. 40.
on the day of the auction, two hundred Chilean troops occupied Antofagasta, though they allowed Bolivian officials to retreat peacefully to Cobija.\textsuperscript{57} Pinto’s troops arrived in the nick of time. As one of his deputies asked, “Who knows what action the public would have taken if the government had delayed one day more in occupying the littoral?”\textsuperscript{58}

Word of Antofagasta’s capture soon reached Daza. Curiously, the president waited more than a week to respond—allegedly because he did not want to distract from ongoing Carnival celebrations.\textsuperscript{59} However, when eight to ten thousand protestors massed in the capital, demanding weapons with which to oust the Chileans, Daza declared that Chile’s actions had imposed a “state of war” between the two countries.\textsuperscript{60} Even this statement, however, did not constitute an official declaration of hostilities, nor did it commit Bolivia to a specific response. Instead, Daza appealed to President Mariano Prado of Peru for a means of extricating himself from the situation.

Prado was acutely aware of his own country’s limited capacity for conflict, lamenting to the Bolivian foreign minister that “Peru has no navy, has no army, has no money; it has nothing for a war.”\textsuperscript{61} Rather than announce military support for Bolivia, Prado offered to help Bolivia negotiate a peace agreement. With Daza’s permission, Prado dispatched an emissary to Santiago, José Antonio Lavalle, with instructions to convince Pinto to accept a reinstatement of conditions that existed before the Antofagasta tax was imposed.\textsuperscript{62}

Unfortunately, when Lavalle arrived in Chile, President Pinto rejected the terms. Although the Chilean leader conceded that he personally preferred to accept the proposal, he also insisted the Chilean public would not tolerate an agreement that would restore Bolivian control of the Atacama. Instead, Pinto offered to restrict nitrate exports from his newly-acquired territory, thereby giving Peru a regional monopoly on the product—an economic outcome that amounted to a significant gain for Peru relative to pre-crisis conditions. Lavalle responded that the Pe-

\textsuperscript{57} Farcau 2000, p. 42.
\textsuperscript{58} Sater 1986, p. 16.
\textsuperscript{59} Farcau 2000, p. 42.
\textsuperscript{60} Sater 2007, p. 28.
\textsuperscript{61} Sater 2007, p. 36.
\textsuperscript{62} Farcau 2000, p. 42.
ruvian public would force the government to refuse these terms and to support their Bolivian allies. Observing that each actor’s capacity to compromise was constrained by forces beyond their control, the Chilean Foreign Minister summarized the situation by lamenting that, “Moral victories... will satisfy no one. The war might truly be a calamity, but we will have to endure it.”

As negotiations failed, Bolivia declared war on Chile; Santiago reciprocated two weeks later. In Peru, Prado also succumbed to the force of public opinion. Just as Lavalle predicted, it became clear that failure to aid Bolivia would “arouse the most intense indignation.” According to an American visitor in Lima, Prado announced his decision to wage war when “a furious mob appeared before the doors of the municipal palace and demanded [Prado’s] intentions... [and] Prado saw he must renounce Chile or lose his life.”

**Discussion and Alternatives**

What, then, was the most proximate cause of the war? By the time the first bullets were fired, none of the combatants were confident that they would win. Likewise, all recognized that the costs of fighting vastly outweighed any benefits that their country may reap. Nor did the leaders initiate conflict in order to enrich themselves personally. Finally, there is no evidence that the Chilean, Bolivian, or Peruvian governments deliberately activated domestic audiences in order to gain a bargaining advantage. Instead, each head of state was trapped by a hawkish domestic population that he felt unable to appease.

At its heart, the war occurred because Daza misjudged how Pinto would respond to provocation. After observing that Pinto was willing to sacrifice territory to avoid armed conflict

---

63 The Chilean Ambassador in Lima received a similar response when he presented the offer directly to President Prado (Farcau 2000, p. 42).  
64 Sater 1986, p. 12.  
65 Sater 2007, p. 42.  
67 A vocal minority of investors in Chile so thoroughly feared the economic consequences of war that they offered President Pinto a personal bribe of two million pesos to forgo fighting and to reinstate the terms of the 1874 agreement (Sater 2007, p. 38).
with Argentina, Daza manufactured his own crisis in hopes that Pinto would offer similar concessions in the Atacama. In some sense, Daza was correct: Pinto personally believed the border territory was unworthy of fighting over; he would happily have offered concessions rather than risk a costly military defense. Unfortunately, Daza also erred by underestimating the extent to which hawkish Chilean opinion would inhibit Pinto from offering the concessions he personally endorsed. Although Pinto explained these domestic constraints to his enemies once the crisis began, by that stage war was unavoidable and each leader was locked into a conflict they preferred to avoid.

Is it possible that Pinto exaggerated the militaristic preferences of his constituents as a negotiating tactic? Was he bluffing in hopes of forcing Peru and Bolivia to back down? Such an interpretation would appear a misreading of available evidence. The Chilean press depicted Pinto’s decision as a forced choice: they claimed that the people were deeply concerned with the country’s national dignity, an asset that “no government would be sufficiently strong or audacious to compromise without being torn apart and thrown from the Moneda like one throws garbage into the street.”\textsuperscript{68} The American ambassador in Santiago expressed a similar opinion, noting, “It is doubtful, indeed, if the administration could have taken another course and sustained itself.”\textsuperscript{69} Perhaps Bolivian envoy, José Antonio de Lavalle summarized best when he wrote, “It was impossible, completely impossible [for Pinto] to arrive at a peaceful solution, although Pinto’s government would have been disposed to go to any lengths to avoid this end... if [the dispute] had been resolved peacefully, Pinto would have been violently overthrown and the war would still have taken place.”\textsuperscript{70}

\textsuperscript{69} Sater 2007, p. 40.
\textsuperscript{70} Lavalle 1994, p. 62, as quoted in Chiozza and Goemans 2011.
1.5.2 The Fall of France

Background to the Dispute

Germany launched its western invasion of Belgium, the Netherlands, and France on May 10, 1940. Five days later, Winston Churchill awoke to a telephone call from French prime minister Paul Reynaud, who reported that German tanks and armored vehicles had broken the front near Sedan and that his country had “been defeated... we are beaten; we have lost.” The Dutch army surrendered the following morning, by which time Reynaud, acutely aware that not “a single corps of soldiers” stood between the German armies and the French capital, was likewise debating whether to order his government to evacuate Paris or merely sue for peace. In the end, Reynaud chose a third option: he postponed a withdrawal from the capital and also refrained from approaching his German adversaries about armistice terms. Instead, the French government prolonged the conflict for more than a month at the price of roughly 85,000 French lives.

By what calculus did the French prime minister choose to persist with his futile defense? Surely optimism was not a determining factor. Maurice Gamelin, commander-in-chief of the French Armed forces at the onset of the invasion, concluded by the eve of May 15 that counterattacks were impossible and that continued fighting would lead only to the “destruction of the French armies.” There is also no evidence the French hoped that continued fighting would enable them to obtain a more favorable settlement in future negotiations. In a meeting on May 25, President Albert Lebrun argued that prolonged fighting would diminish French military capabilities and therefore the “government’s freedom [to negotiate].” Gamelin’s replacement, Maxime Weygand, likewise hoped that France would secure a peace deal “while the Allies still

---

71 Churchill 1949, p. 20.
73 La Gorce 1988, p. 496. Shepperd 1990, p. 88 lists French casualties as 90,000 killed, 200,000 wounded, and another 1.9 million missing or captured.
74 Jackson 2004, p. 10. Gamelin was replaced three days later by Maxime Weygand, who subsequently raised the possibility of surrender on May 25.
held some cards in their hand.”\textsuperscript{76} The “commitment problem” explanation for war is similarly unconvincing in this situation. The theory argues that leaders may fail to commit to peace when they anticipate that an adversary will experience a rapid increase in power, but Germany could already credibly threaten to occupy the French homeland and replace the French government; further power shifts in favor of Berlin were largely immaterial. Finally, it seems unrealistic that French leaders believed their constituents would be better served by sustained war rather than peace. After all, prolonged fighting would not only trigger substantial casualties in the short term but could also create internal turmoil if no government or army remained to prevent anarchy following the German conquest.\textsuperscript{77}

**Concerns About Public Opinion**

In contrast to the unitary-state explanations, I argue that the French leadership delayed their surrender because they feared that quick capitulation would permanently jeopardize their political reputation with French citizens. As early as May 15, deputy premier Camille Chautemps, worried that the cabinet’s withdrawal from Paris would provoke “adverse public reaction, which would interpret the government’s departure as desertion.”\textsuperscript{78} Likewise, reports from May 27 suggest that Reynaud “considered the indefinite prolongation of hostilities as chimerical,” but felt “publicly committed” to continuing the war against Hitler.\textsuperscript{79} British Ambassador Ronald Campbell similarly noted his belief that “there was not a single Frenchman [in government] who did not feel, even if he would not admit it, that France was beaten,”\textsuperscript{80} and that the “forces in favor of surrender” were sufficiently strong that such an outcome “may come more quickly than we expect.”\textsuperscript{81}

Why did Reynaud feel compelled to persist in a hopeless defense? After all, the French

\textsuperscript{76}Jackson 2004, p. 132. On May 24 Weygand similarly noted that France should get “out of the ordeal which is is undergoing” if she was ever to “rise again” (Jackson 2004, p. 104).
\textsuperscript{77}Weygand was among the most significant proponents of this view. See Jackson 2004, p. 132
\textsuperscript{78}Gates 1981, p. 214.
\textsuperscript{79}Gates 1981, p. 155.
\textsuperscript{80}Gates 1981, p. 155.
\textsuperscript{81}Baxter 2006, p. 191.
population had suffered tremendous costs during the Great War. Given their knowledge of the
costliness of fighting, one might reasonably expect French citizens prefer that their government
avoid unnecessary combat rather than mount a prolonged resistance. However, as Adamthwaite
(1995, p. 169) notes, “pacifism waned after 1936.” Hucker (2007, p. 3) similarly clarifies that the
particular form of pacifism normally attributed to French citizens in the interwar period “did not
induce the defeatism” with which it is so often associated. Indeed, an opinion poll conducted
in October 1938 showed that although 57% of the French population supported appeasement at
Munich, a further 70% of respondents favored resisting additional German demands.82

In the time since Munich, government officials and members of the press also inadvert-
tently stoked the public’s desire to defend French territory. In December 1938, prime minister
Edouard Daladier announced in a speech that “France will not cede an inch of territory” to Italian
irredentists in Corsica.83 At the same time, prominent newspaper pundits argued in favor of
national defense. As Pierre-Antoine Cousteau asserted, when “our possessions are targeted, the
peace of Munich is not a precedent.”84 Finally, officials and the media cultivated the sentiment
that the country could wage its upcoming war successfully. General Weygand announced in a
speech at Lille on July 14, 1939, that “the French army is a more effective force than at any other
time in its history; it possesses equipment and fortifications of first class quality, excellent morale,
and a remarkable high command.”85 Likewise, officially sanctioned films included statements
such as “[France] is capable of facing all attacks and all challenges.”86 As Hucker (2007, p.
20) summarizes, “representations of French opinion in early 1939 demonstrated that the French
people were prepared, if necessary, to forcibly resist unreasonable demands... encouraging the

---

82 Even during the Munich negotiations, former French prime minister Pierre-Étienne Flandin remarked that although public opinion “is more likely to be in the direction of non-intervention than that of intervention,” several influential groups “are leading us into this war [and] are determined to push us into it.” (Quoted in Hucker 2007, p. 13)

83 Hucker (2007, p. 18). Daladier followed the speech with a widely-publicized tour of the island the following month.

84 Hucker (2007, p. 19). Other newspapers drew attention to the remilitarization campaign then underway, writing headlines such as, “Is France resigned to die or does she have the will to live?” (Hucker (2007, p. 18)).

85 Jackson 2004, p. 10.

86 Cited in Laborie 2001, p. 119.
French government to pursue a foreign policy of firmness rather than capitulation.”

**Blame Shifting and Eventual Capitulation**

On June 5, the day following the final evacuation of Dunkirk, the Germans renewed their attack by pushing southward. General Weygand by this point believed “the military situation to be irreparable.” He warned Reynaud that a final rupture of French defenses could occur at any moment and insisted that although the army would “continue to resist, if the Council [so] orders… the ending of hostilities must be considered soon.” Reynaud replied that although Weygand was offering “extremely competent advice about the military sphere… the question of continuing the war was a political matter.” Churchill, who had visited his French allies earlier in the day, further recalled that Marshal Pétain “had quite made up his mind that peace must be made. He believed that France was being systematically destroyed by the Germans, and that it was his duty to save the rest of the country from this fate,” but that Pétain was ashamed to present the argument to Reynaud. Instead, it would be nearly two weeks before the French government pursued an armistice and a further week until peace terms were signed.

We can attribute this delay in seeking peace to disagreement over who should take public responsibility for the armistice. On June 12, Weygand, supported by “virtually all the senior army commanders,” announced to the Council of Ministers that the war was irretrievably lost and that it was essential for the French government to seek an armistice. However, several members of the Council, including Chautemps, rejected the proposal “because public opinion was not yet prepared for it.” Reynaud’s detractors would likewise claim after the war that “he believed an armistice to be inevitable and that the stands he took, the speeches he made, the orders he gave

---

87 Jackson 2004, p. 123 similarly argues that Daladier’s enthusiasm for intervention in Finland was motivated by concerns about those in the “opposing camp who felt he was prosecuting the war insufficiently energetically.”


89 Gilbert 2000, p. 145.


91 Churchill and Cook 2013, p. 290.


were all... mere posturing intended for public consumption” (Gates 1981, p. 190). Of particular concern to Reynaud was that soliciting an armistice would “shift the responsibility for the defeat [from the military] to the politicians.”94 Instead, the prime minister proposed an alternative: Weygand, as commander-in-chief of the French military, should declare a ceasefire while the Council relocated to North Africa to maintain at least the appearance of resistance. Weygand refused, declaring that he would “never agree to bring such disgrace on the flags of the French army,” and claiming that Reynaud was merely trying to deflect responsibility for defeat away from the Cabinet.95

In the end, Reynaud chose to resign rather than concede.96 On June 16 he was replaced by Marshal Pétain, who sought to open peace negotiations with Germany, an act for which Pétain and other members of the Vichy regime would find themselves on trial after the war. Others in the cabinet, notably Charles de Gaulle, moved abroad and became popular rallying points for a ‘Free France’ both during and after the war.97 On paper the German peace terms were surprisingly lenient: France would continue to exist as a sovereign state, would maintain jurisdiction of its overseas territories, and could even maintain small local military units to ensure domestic order.98

The Anglo-French Alliance

Although this paper argues that French politicians hesitated to settle with Germany because they feared that doing so would cost them public support, one alternative possibility merits discussion. Some argue that the French leadership worried that rapid capitulation would trigger backlash not from French citizens but rather from the British allies on whose fortunes France would rely for liberation. Indeed, the two countries had reached an agreement on March 28 that neither would

---

94 Jackson 2004, p. 104.
96 According to testimony taken after the war from three former ministers—two supporters of Reynaud, the other an opponent—on the day of his resignation there was still a slight majority in the Council opposed to an armistice (Jackson 2004, p. 139).
97 As Jackson (2004, p. 142) quips, Reynaud “missed the chance to be de Gaulle,” an error for which he “never forgave himself.”
98 Gilbert 2000, p. 149.
sign a peace treaty with Germany without the other ally’s consent.

Nevertheless, there are several reasons to doubt that maintaining Allied support was Reynaud’s primary motivation when refusing to settle. First, the French leaders openly expressed their military pessimism to their British allies throughout the engagement. If Reynaud was concerned about losing British support, he would instead have feigned commitment to the war effort. Second, Churchill on several occasions indicated that the British could provide very little military assistance to France. 99 French leaders not only acknowledged that British capabilities were constrained, they further believed that Britain would quickly succumb to Germany, thereby negating any hopes that the alliance would pay off long-term. 100 Third, the French politicians ought not have felt obliged to abide by their commitments against unilateral peace. When Reynaud broached the subject of French capitulation, Churchill reportedly instructed his colleague that “If it is thought best for France in her agony that her Army should capitulate, let there be no hesitation on our account.” 101 Finally, when total German victory over the French seemed imminent, the British offered France an opportunity to form a political union that would permanently bind the two states. If prioritizing the alliance with Britain in hopes of achieving long-term victory was essential to the French, they should have accepted this offer rather than capitulate to Germany. Instead, they rejected it. 102 Thus, it seems the French cabinet was more concerned with the need to appease domestic audiences than the desire to reassure their British allies.

99 When Weygand requested British air reinforcements, arguing that “Now is the decisive moment” and that it was “wrong to keep any squadrons back in England,” Churchill refused, replying that “This is not the decisive point and this is not the decisive moment. That moment will come when Hitler hurls his Luftwaffe against Great Britain” (Churchill 1949, p. 147).

100 Jackson 2004, p. 103.


102 Pétain famously quipped that such an agreement would be akin to “fusion with a corpse” (Shlaim 1974, p. 53).
1.5.3 Escalation in Vietnam

Background and Preferences

When Lyndon Johnson assumed the U.S. presidency, he opposed the expansion of American military operations in Vietnam. As vice president, Johnson drafted a prophetic memo to President Kennedy in which he described the risks of deploying U.S. combat forces: “We had better remember the experience of the French who wound up with several hundred thousand men in Vietnam and were still unable to [succeed]... Before we take any such plunge we had better be sure we are prepared to become bogged down chasing irregulars and guerrillas over the rice fields and jungles of Southeast Asia while our principal enemies China and the Soviet Union stand outside the fray and husband their strength.”

After Kennedy’s assassination, Johnson continued to resist calls to deploy troops or increase bombing operations in Vietnam. In April 1964, he lamented that the military was “trying to get me in a war over there... I turned them down three times last week.” The following month, the president expressed his private reservations to McGeorge Bundy, explaining that with Vietnam it “looks to me that we’re getting into another Korea... I don’t see what we can hope to get out of this.”

Pressure to escalate military operations came not only from Johnson’s military advisors but also from the public at large. American voters, primed by three successive administrations to consider Southeast Asia a national security priority, supported American efforts to secure the region against the communist threat. Johnson was acutely attuned to such foreign policy hawks, who he referred to as the “great lurking monster” of American politics.

---

103 Quoted in Warner (1994). Although there is considerable debate over whether Kennedy would eventually have withdrawn from Vietnam, evidence suggests that he worried about the political consequences of doing so. He intimated privately to Senator Mike Mansfield, “I can’t do it [withdraw] until 1965—after I’m re-elected” (Asprey and Asprey 1994, p. 761-762).


105 McMaster and Williams 1997, p. 325.

106 Herring (1995, p. 134). Johnson knew from personal experience how politically damaging hawkish critiques could cause for a president. As a freshman senator, Johnson had criticized Truman for rejecting the Joint Chiefs’ recommendations to increase U.S. air power in Korea, asserting that “all [the administration’s] effort is seemingly directed toward staying out of the war we are already in” (McMaster and Williams 1997, p. 52).
observed that the “Goldwater crowd” of war-hawks was “more numerous, more powerful, and more dangerous than the fleabite professors,” and General William Westmoreland similarly admitted in an interview that Johnson was substantially more concerned with appeasing hawkish public opinion than he was with the anti-war movement. Johnson himself acknowledged that because “1964 was an election year” he would be forced to “take some action to show that his administration was on top of the situation” in Vietnam.

**Defusing the War Hawks**

Despite his concerns about hawkish pressure, Johnson initially followed his personal preferences and positioned himself as the ‘peace’ candidate in contrast to Barry Goldwater in the 1964 election. During the campaign, Johnson sought to reduce American support for the intensification of violence in Southeast Asia. When General Westmoreland recommended that the Administration adopt a “people-to-people program, to get the American people... some emotional attachment to the South Vietnamese,” Johnson shot down the idea for fear that if Americans became emotionally aroused the “hawks might take over control.” As Press Secretary Bill Moyers claimed, the administration’s conclusion was that public debate on Vietnam should be kept at “as low a level as possible.” The president hoped that public hawkishness could be kept at a low simmer; he feared that if the topic gained a foothold in public discourse the resulting attitude would create irresistible pressures for escalation. Secretary of State Dean Rusk further explained that the administration deliberately avoided “military parades through the cities [and] beautiful movie stars selling out war bonds... we felt that in a nuclear world it is just too dangerous for an entire people to get too angry and we deliberately played this down.”

---

107 Herring 1995, p. 140.
109 Stempel 1965, p. 221, as quoted in Caverley 2014.
110 Charlton and Moncrieff 1978, p. 137.
112 Charlton and Moncrieff 1978, p. 115. Likewise, when Johnson learned that General Curtis LeMay was considering retirement—at which point the general planned to openly criticize the administration’s policy in Vietnam,
Johnson’s attempt to defray public concern with Vietnam was also motivated by his fear that the topic would distract from his domestic agenda. The president sought to push his Great Society legislation through Congress as quickly as possible after election; if Vietnam became a contentious topic it could divert congressional attention or create political divisions that would be difficult to bridge. Johnson viewed the Great Society as his opportunity to create a lasting political legacy of reform. Although the president personally objected to the escalation of American involvement, the possibility that he would lose, as he put it, “the woman I really loved” for “that bitch of a war on the other side of the world” was intolerable. From his installment in office until the spring of 1968, LBJ continually fretted that public criticism that he was ‘not doing enough’ on Vietnam would undermine his legislative goals. He therefore sought to downplay the conflict in Southeast Asia and plotted a course that would safeguard his political capital and insulate him from criticism.

**Escalatory Incentives**

In the late summer and autumn of 1964, a series of incidents tested Johnson’s capacity to maintain the peace. The first was the so-called “Gulf of Tonkin Incident” of August 2, 1964, in which the USS Maddox exchanged gun fire with several North Vietnamese torpedo boats. The president was convinced the incident occurred in response to U.S. covert operations in the gulf and decided to downplay the event to deflect calls for retaliation. However, two days later McNamara received word that both the Maddox and the USS Turner Joy were being followed by North Vietnamese vessels and were preparing for an attack. This time, Johnson’s advisors warned the president that a military response was essential in order to deny Goldwater an opportunity to “accuse him of vacillating or being an indecisive leader.”

---

113 McMaster and Williams 1997, p. 194.
114 Herring 1995, p. 130.
response that would appease the Republicans, they “might take such an action that would... put the administration in a position where we had to do things we thought would be very unwise, that might involve bringing in the Chinese or offending somebody else.” To ward off this possibility, Johnson ordered a series of retaliatory air strikes and called on Congress to pass a resolution giving him the authority to “take all necessary measures in support of freedom and in defense of peace in southeast Asia.” When Congress adopted the Gulf of Tonkin Resolution on August 7, 1964, Johnson’s popularity surged.

Although Ambassador Maxwell Taylor and the Joint Chiefs of Staff advocated a sustained bombing campaign in the wake of the Gulf of Tonkin incident, Johnson was reluctant to increase U.S. operations. Instead, he approved only the resumption of U.S. naval patrols as well as aerial reconnaissance, maritime raids, and leaflet drops. In a message to Taylor, the president declared that he would not be drawn into a war against North Vietnam merely because “our own people are careless or imprudent.” Five days after the destroyer patrols in the Gulf of Tonkin were resumed, Johnson received word of another skirmish between U.S. forces and Vietnamese patrol boats. Urged to authorize a new round of retaliatory attacks, Johnson once again deflected: “Hell... those dumb, stupid sailors were probably just shooting at flying fish.” Rather than commit himself to a bombing campaign, he asked Ambassador Taylor to draft more optimistic assessments of the situation in Vietnam, once again hoping to appease hawks and downplay the desirability of escalation. In a campaign event in Manchester, New Hampshire at the end of September, the president reminded voters of the need to be “very cautious and careful” and noted that his administration would go on the offense against Vietnam “only as a last resort.”

---

117 Some allege that Johnson and McNamara deliberately manufactured the second Gulf of Tonkin incident in order to justify their plans for interventionism, but this account is suspicious. If Johnson sought a reason to retaliate, he could have done so following the first incident. Instead, he waited to respond until McNamara received intelligence—later revealed to be faulty—that a second military exchange had taken place.
121 McMaster and Williams 1997, p. 151.
122 Johnson 1971.
Finally, on November 1, 1964, just days before the election, the North Vietnamese successfully conducted a mortar attack on Bien Hoa air base in which 27 U.S. aircraft were damaged or destroyed. When Ambassador Taylor asked the president to consider retaliation, Johnson once again declined, though he first asked Special Assistant Bill Moyers to inquire with pollsters whether “failure to respond to this attack immediately will be taken by the voters as a sign of weakness.”\textsuperscript{123} He continued to advocate an approach of relative restraint, optimistic that the sorties he approved following the Gulf of Tonkin incident were sufficient to allay public criticism for the time being.

**Protecting the Domestic Agenda**

After his election, Johnson focused his effort on pushing through his Great Society legislation. Although public pressure to escalate in Vietnam was rising, Johnson had reason to doubt that aerial bombing would yield reliable results. On October 5, 1964, George Ball delivered a memo to the president that criticized current U.S. policies. According to Ball, that there was little evidence that even a substantial air campaign could convince Hanoi to “permanently abandon its aggressive tendencies.”\textsuperscript{124} Indeed, the memo argued that escalatory tactics might inspire the North to reciprocate, forcing the U.S. to deploy ground troops and creating a costly spiral from which the Johnson would be unable to extricate himself. By January 27, 1965, Bundy and McNamara had similar concerns. Although the current policy of limited involvement had temporarily appeased the hawks, it would eventually lead to “defeat and an invitation to get out in humiliating circumstances.”\textsuperscript{125} Bundy and McNamara noted that the president now faced a choice: he could either use unrestricted military power to appease the hawks or begin the process of draw-down and withdrawal “with no major addition to our present military risks.”\textsuperscript{126}

Administration opinions on the dilemma were split. Ball noted the enormous costs and

\textsuperscript{123} McMaster and Williams 1997, p. 174.
\textsuperscript{124} McMaster and Williams 1997, p. 166-167.
\textsuperscript{125} McMaster and Williams 1997, p. 214.
\textsuperscript{126} McMaster and Williams 1997, p. 215.
minimal benefits of the war, and encouraged Johnson to exit the conflict as gracefully as possible. Vice President Humphrey similarly urged Johnson to “cut his losses in Vietnam,” arguing that the president’s sweeping victory in November granted him a mandate to ignore Republican critics who preferred escalation. Unfortunately, appeasing Republican senators remained a high priority for Johnson, who believed he needed their support to ensure the adoption of his Great Society Legislation. Bundy and McNamara were also concerned with appeasing the hawks. “You need Vietnam to save your administration,” they wrote. “If we lose because we have withheld our military power, you will be blamed, and nothing can undo the damage.” Bundy reiterated this claim when, on February 7, he returned from a visit to Vietnam and cautioned that the current limited intervention campaign was ineffective. As such, he recommended that the president select a new policy that would “damp down the charge that we did not do all we could have done.” Although under an escalated policy “U.S. casualties would be higher,” this price would be politically “cheap” compared to the cost of withdrawal or defeat. Confiding to Carl Rowan that “Just between you and me, all I want to do is bloody their noses a little bit” Johnson approved “Rolling Thunder,” an eight-week air campaign against the North Vietnamese.

Although he was willing to approve a bombing campaign to appease the war hawks and avoid criticism as a dove, Johnson still doubted the desirability of committing U.S. troops. He granted a request from General Westmoreland for roughly 1,500 soldiers to defend the air base at Da Nang, but when the Joint Chiefs recommended that Johnson deploy a full 90,000 troops to Vietnam, the president objected and delivered only five thousand men. Johnson likewise sought to draw down the aerial bombing, even though he worried that doing so would cost him public support. A Gallup poll from April 1965 found that only 21% of respondents thought that

---

127 Skowronek 1997, p. 343.
130 Gardner (1995, p. 169). Polls from the period suggest that there was very little public enthusiasm for either negotiation or withdrawal. In a Gallup poll gathered before Johnson announced the expanded air campaign, 67% of respondents thought the U.S. should continue its present efforts in Vietnam, 47% of respondents thought that the U.S. should persist even if doing so risked nuclear war. Only 20% of those surveyed preferred withdrawal.
the U.S. should stop its bombing; 59% recommended that it continue. Despite these figures, Johnson agreed to a week-long hiatus that he hoped would encourage the North to negotiate. During the break, the president fretted to McNamara, “the public has never wanted us to stop the bombing... we don’t want to [stop] too long else we lose our base of support.” When the ceasefire ended without successful negotiations, pressure mounted for more aggressive military action. In another poll from June, 47% of respondents supported “sending more troops to defend South Vietnam,” and a further 19% recommended that the U.S. maintain current troop deployments; only 11% of respondents preferred to take most troops out.

Finally, on June 5, Johnson assembled his primary advisors, including Ball, Bundy, McNamara, and Rusk for a decisive policy meeting about Vietnam. The questions the president posed in the meeting reveal his caution, uncertainty, and political motivations. Johnson acknowledged that his advisers had “no plan for victory militarily or diplomatically,” but also that he “shudder[ed] to think what all ’em [in the public] would say” if he chose to withdraw. McGeorge Bundy similarly recalled that despite the president’s reservations about escalating the war, “his unspoken object was to protect his legislative program.” In the end, the least costly course of action when judged by immediate domestic politics was to deploy additional troops, and in July 1965 Johnson announced that he was increasing U.S. combat strength to 125,000 personnel, committing the U.S. to a conflict he never intended to fight and had no plan to win. As Herring (1995) summarizes, “Johnson’s inability to wage war in cold blood produced what appears on the surface a great anomaly—one of the shrewdest politicians of the twentieth century committing a form of political suicide by taking the nation into a war he would have preferred not to fight.”

---

1.6 Conclusion

The finding that domestic political pressure can motivate leaders to initiate crises has important implications for international relations theory. First, although others have argued that domestic politics can influence conflict behavior, to my knowledge this paper is the first to argue that latent public enthusiasm for war can directly compel leaders to engage in combat. In other popular theories, the public is not initially reluctant to settle. For example, when leaders activate “audience costs” they deliberately cultivate hawkish public preferences in order to gain an international bargaining advantage. In contrast, the “agency dilemma” pertains to cases in which leaders consistently sought to suppress public enthusiasm for fighting but were unable to sway public sentiment sufficiently. As such, this project highlights the fact that future analyses of war must account for the foreign policy preferences of the political constituents within each belligerent nation as well as the rational interests of the unitary state.

Second, I show that democracies in some cases appear more willing than autocracies to participate in costly and futile military campaigns. This finding contrasts with the prevailing view that “democracies are not eager to pursue wars they do not expect to win” and that democratic leaders “are highly selective; they prefer to negotiate when they do not anticipate military success.” Identifying conditions in which hawkish constituents can compel leaders to engage in inefficient fighting should also yield important policy implications given our emerging understanding of the “provocation” strategies that belligerent groups often pursue. If citizens demand retaliation or escalation in the wake of violent episodes, then democratic states may present appealing targets for groups who aim to provoke a draconian response. As a result, states may be better able to dissuade transnational violence by tying their hands in ways that would prevent significant retaliation rather than by enhancing their capacity to respond with force.

The project therefore suggests several directions for future work. First, researchers should

---

135 See, for example, Fearon 1994, Haynes 2012, Kurizaki and Whang 2015.
136 Bueno de Mesquita et al. 2003, p. 236. See also Reiter and Stam 2002.
137 Carter 2016; Kydd and Walter 2006.
investigate specific conditions in which leaders are particularly likely to modify their crisis decision-making in response to public opinion.\textsuperscript{138} This may include the role of term limits, electoral competition and the proximity of political challengers, the ease with which foreign policy failures can be attributed, as well as domestic pressures that autocrats face.\textsuperscript{139} Models that allow leaders to select private or public negotiations, conditional on the preferences of their constituents, may also be worthwhile.\textsuperscript{140} A third branch of research should analyze environments and attributes that make foreign policy a particularly salient public issue, including the domestic economic environment, military conscription policies, and female enfranchisement.\textsuperscript{141} Finally, researchers should investigate whether constituent optimism can promote violent outcomes or encourage ‘revolutionary momentum’ in situations involving domestic violence, such as military coups and civil movements.\textsuperscript{142}

All chapters of this dissertation are being prepared for submission for publication of the material. The dissertation author was the primary investigator and author of this material.

\textsuperscript{140}For similar work, see Kirpichevsky and Lipsy 2018, Carson and Yarhi-Milo 2017, and Carnegie and Carson 2018.
\textsuperscript{141}Shaver (2015), for example, shows that audiences are particularly optimistic when employment is high, while Trager et al. (2018) suggest in a recent conference paper that women are less optimistic about the use of force and that female voting participation is related to a decline in violence by democratic states.
\textsuperscript{142}See, for example, Abrahams and Merrell 2019.
2 The Secrecy Gambit: Clandestine Power Shifts and Preventive Conflict

Abstract: Under what conditions should rising states reveal or conceal their military capabilities? We present a formal model in which military announcements reveal information both about a state’s current capabilities as well as its potential development trajectory. The results suggest that several common conclusions about crisis behavior should be qualified. First, we identify conditions in which states will coerce their adversaries by signaling strength, but also circumstances in which they will attempt to conceal their military capabilities or even eschew opportunities for growth altogether. The model also clarifies two forms of preventive war that researchers often conflate: wars of discovery and wars of suspicion. Finally, we show that the possibility of covert activity compels all states to take costly actions to reassure suspicious adversaries. While this reassurance tax is part of the ‘gambit’ played by ambitious states, it imposes a burden on those that cannot pursue clandestine development. We support the theory with historical data on secret alliances and show that the results generate novel predictions for research on military arming, allying, and counterinsurgency.
2.1 Introduction

Public signals are part and parcel of international politics. During crises, negotiations, and conflicts, strong actors make costly moves to demonstrate their abilities, announce their intentions, and motivate observers to either cooperate or concede. A broad body of research describes how credible signals of strength enable states to realize greater success in war and secure more favorable outcomes in peace.¹ Despite these benefits, however, actors frequently choose to avoid public signaling, opting instead to conceal their capabilities and obscure their strengths. Militaries, for instance, often keep new weapons system covert and advanced technologies secret for years after completion.² Allied countries likewise forge clandestine agreements and conceal the extent of their cooperation, even if announcing their ties might pay dividends at the bargaining table.³ Even individuals can engage in secretive behavior: whereas some citizens signal their support for rebel or government forces during civil conflicts, others strive to conceal the extent of their collaboration or allegiance.

This paper introduces a new framework for evaluating how political actors choose between signaling and secrecy. The answer exposes and resolves an important but under-acknowledged tension between prevailing theories of deterrence, crisis bargaining, and war. According to one view, states should deter hostile rivals by signaling strength and resolve.⁴ Clear and credible military demonstrations allow opponents to reconcile conflicting beliefs, identify feasible bargains, and avoid wars that might occur if information remained asymmetric.⁵ Nevertheless, signaling is no panacea. A contrasting literature argues that military demonstrations can provoke war rather than resolve it: a country that discovers its opponent is rapidly increasing in power may doubt the

²Prominent American examples include the F-117 Night Hawk, which was revealed publicly only when its production run was nearing completion, many years after it had achieved operating capability.
³The “Alliance Treaty Obligations and Provisions” dataset (Leeds et al. 2002) suggests that roughly 20% of alliances that existed between 1815 and 1956 were forged in secret and were concealed from non-members until after expiration. More recently, powerful states often train and clandestinely equip military proxies or sub-state allies (see Alpher 2015 and Carson 2018).
opponent’s commitment to peace and initiate a preventive attack on the basis of this concern.\textsuperscript{6} States that seek to maximize their fortunes and minimize the overall risk of war must therefore tread a fine line. To deter hostile rivals they must develop and display military assets that attest to their strengths and capabilities while avoiding demonstrations that could provoke preventive attacks by opponents who infer that the balance of power is shifting too rapidly.

We assess how actors manage this dilemma by modeling their decision to develop and announce military technologies—such as arms or alliances—that could bolster future bargaining power. Relative to previous models, our theory combines several innovations. First, we allow military demonstrations to occur at intermediate stages of development. As a result, such demonstrations convey new information both about a country’s existing power as well as its development trajectory. Traditional arming models abstract from this dilemma by focusing largely on the acquisition of nuclear weapons. Although they are of obvious importance in international relations, such weapons are also unusual insofar as they confer military benefits only once a state’s development process is fully complete. Theories tailored around the nuclear case mirror this logic, providing the arming side with a single significant boost in power when it crosses the final development threshold. Such models therefore emphasize the importance of secrecy, giving states have little incentive to reveal their development progress at intermediate stages.\textsuperscript{7}

This paper focuses instead on situations in which new military demonstrations could reveal two pieces of information: (1) the signaling country is already stronger than its adversaries believed, and (2) the signaler is poised to secure additional power in the future.\textsuperscript{8} Modeling the dual consequences of military announcements allows us to better capture the strategic dilemmas


\textsuperscript{7}See, for example, Kydd 2000, Baliga and Sjöström 2008, Debs and Monteiro 2014 and Bas and Coe 2016.

\textsuperscript{8}Even nuclear tests raise questions about how quickly a rising country will obtain second-strike capability, thermonuclear weapons, improved delivery systems, miniaturized devices, etc. Indeed, prior to the Soviet Union’s development of thermonuclear weapons several American military advisors argued that the U.S. should exploit its nuclear advantage by attacking the relatively weak Soviet Union. The likelihood of war decreased when the Soviet Union’s own nuclear tests convinced the Americans that war would be costlier than they previously anticipated.
embedded in a state’s decision to reveal or announce a wide array of new technologies. For example, Chinese demonstrations of advanced weapons systems, such as anti-ship missiles and stealth aircraft, foment unease among American officials not only because such technologies augment China’s existing capabilities but also because the advancements suggest an overall development trajectory that could allow China to achieve genuine military parity with the U.S. in the years ahead.\textsuperscript{9} Likewise, North Korea’s Hwasong-14 missile tests in July 2017 indicated not only that the regime could now launch low-payload devices against targets in Alaska or Hawaii, but also that Pyongyang’s missile development was progressing with sufficient speed that areas throughout the continental U.S. could soon become vulnerable as well.

By analyzing how states manage their countervailing incentives to hide and announce new weapons, relationships, or technologies, our theory draws attention to several under-appreciated aspects of crisis behavior and generates novel predictions for empirical research. We contribute to growing literatures on military signaling and secrecy by identifying conditions in which states will either advertise or attempt to conceal their emerging capabilities.\textsuperscript{10} Whereas a long research tradition emphasizes the benefits of signaling strength, recent work on military arming shows that countries can benefit from an opponent’s uncertainty regarding imminent shifts in military power.\textsuperscript{11} These conclusions stand in tension when military signals convey information about both immediate and future capabilities. We reconcile the contrasting findings by modeling uncertainty as endogenous to the potentially-rising country’s decision of whether to announce or conceal the full extent of its current capabilities. Our results demonstrate that this choice is an important element of a state’s strategic arsenal. In some circumstances, countries will deliberately cultivate uncertainty by keeping existing military assets covert. States may also forgo risky development pathways that lie within their grasp, fearing that such efforts could result in premature exposure

\textsuperscript{9}Allison 2017.
\textsuperscript{11}For the benefits of signaling, see Banks 1990. For strategic concealment, see several of the preceding citations along with Debs and Monteiro 2014, Bas and Coe 2016, Bas and Schub 2016.
and war. However, the pursuit of secrecy and the avoidance of development are not universal, and under other conditions countries can safely reveal that they anticipate rapid growth in the near future.

Our second contribution is to demonstrate that the possibility of covert development imposes an externality on non-developers. In short, when countries cannot discern whether their opponents are poised to rapidly develop, they must treat all adversaries with suspicion—including those who lack either the interest or ability to pursue clandestine activity. Just as law-abiding citizens must sometimes make costly behavioral adjustments to avoid or reduce criminal profiling by suspicious authorities, non-developing countries must offer bargaining concessions to assuage the suspicions of uncertain rivals. This reassurance tax constitutes a significant cost that states could avoid if their capacities for clandestine activity were credibly diminished or they inhabited a world in which secret development was impossible. We argue that empirical research on actors under suspicion of illicit or sensitive activity—ranging from weapons proliferation by rogue states to civilian cooperation with counterinsurgent groups—would benefit from acknowledging this dynamic.

Finally, we distinguish between two types of preventive action that arise from secrecy, which we term wars of suspicion and wars of discovery. The former can occur when a country strongly suspects that an opponent will attempt military development in the future. If the price of reassurance exceeds the opponent’s capacity to pay, no credible concession will dissuade the suspicious side from preventive action. When the level of suspicion falls below this threshold, peace is sustainable in the short-term, but opponents may gamble by pursuing development in the future. Wars of discovery occur when these development attempts are exposed prematurely. The results suggest that the threat of a “large and rapid power shift” is neither a necessary nor a sufficient condition for preventive war. Preventive attacks may occur in the absence of such an

12 The prediction that countries will sometimes deliberately avoid pursuing potent military improvements or opt out of powerful alliances should motivate researchers to reconsider the assumption that observable patterns in allying and arming reflect states’ universal thirst for military power (Mearsheimer 2001 and Grieco 1988).

imminent shift, as long as opponents are highly suspicious that a transition will occur. Likewise, even ongoing power shifts need not provoke war if they can be successfully concealed. By highlighting the difference between these mechanisms, our theory establishes a new framework for analyzing preventive conflicts and yields advice for policymakers who are suspicious of their adversaries.

2.2 Signaling, Secrecy, and Preventive Action

Whether they arise from the acquisition of territory, the formation of alliances, or the development of new weapon technologies, public shifts in military power enable states to coerce their adversaries. When war occurs, strong countries are more likely than weak ones not only to win but also to impose heavy costs on their enemies. Opponents who observe that their adversaries have gained strength should therefore become less eager to fight and more willing to settle peacefully, even at the cost of significant concessions. An important but easily overlooked nuance within this framework is that countries do not obtain negotiating benefits merely by acquiring power. Although advantages in military strength may prove useful if fighting occurs, augmented capabilities only endow a state with coercive leverage in pre-war negotiations insofar as opponents recognize or believe that such capabilities exist. A country whose rivals mistakenly perceive that it is strong may, for example, be better able to coerce its adversaries during crisis negotiations than a genuinely strong country whose enemies believe that it is weak. In summary, success in pre-war bargaining relies on the perception of strength, whereas success in war itself may hinges on whether countries genuinely possess military power when called upon to use it.

Because war and bargaining outcomes rely, respectively, on the acquisition and appearance of strength, countries face incentives not only to pursue improvements in military power but also to demonstrate credibly the gains they obtain. A wide theoretical literature identifies mechanisms

---

14Countries can coerce their enemies without engaging in combat, as long as rivals believe that combat would occur if negotiations failed.
through which genuinely strong states can separate themselves from weaker imposters.\textsuperscript{15} Research on this issue has produced two general conclusions. The first is that credible signals of strength allow strong actors to obtain better bargaining outcomes.\textsuperscript{16} The second is that public signals help countries avoid conflict. Speaking loosely, the rationale is as follows: although sufficiently large shifts in the distribution of power may force states to renegotiate the status quo, this process can occur through compromise rather than conflict. Fighting should occur only when disagreement persists regarding the distribution of power and the extent to which the status quo must be revised in order to satisfy either opponent.\textsuperscript{17} Public demonstrations of military capacity help countries avoid war by enabling them to reconcile their conflicting beliefs about relative military strength and then identify appropriate bargains.

Unfortunately, signals of strength are risky as well as informative. Although states can dispel their opponents’ optimism by credibly demonstrating military power, ‘preventive war’ theorists argue that such demonstrations may sometimes provoke war rather than resolve it. Because a rising state can use its growing coercive power to extract concessions from an adversary, a country who discovers that its opponent is rapidly gaining strength may doubt the opponent’s commitment to existing agreements and may initiate a preventive attack on the basis of this concern. By initiating war before a power transition is complete, the declining country can enter the conflict on the best possible terms and, if victorious, avoid the sacrifices it would be forced to offer an emerging power that completed its rise unchallenged. Today, preventive war is among the most widely invoked concepts in discussions of international conflict.\textsuperscript{18} Indeed, A.J.P. Taylor (1954, p. 166) famously claimed that “Every war between the Great Powers... started as a preventive war.”

\textsuperscript{16}Morrow 1989, Banks 1990.  
strength can deter conflict, in other situations they may provoke preventive attacks. How can growing states manage the tradeoff between deterrence and provocation? When will the fear of preventive war dissuade a state from announcing its capabilities or attempting further development?

One means of answering these questions is with a formal model that explicitly links a country’s current military capabilities with its overall development trajectory and allows a country to choose between concealing or announcing their existing capabilities. Because the decision to arm and the decision to signal are often examined separately—or within the context of nuclear weapons, which provide only limited benefits until development concludes—existing research largely overlooks the dual risks and rewards of sending military signals. Theorists have identified numerous mechanisms that enable countries to credibly signal existing power, but relatively little research examines how or why states should conceal existing capabilities. Slantchev (2010) and Lindsey (2015) represent two exceptions to this rule; each author analyzes a scenario in which a country has an incentive to “feign weakness.” However, their models are limited to situations in which war is either imminent or ongoing and deterrence attempts are therefore of limited value. Likewise, although Baliga and Sjöström (2008), Debs and Monteiro (2014), and Bas and Coe (2016) depict scenarios in which opponents benefit by concealing ongoing development, they focus on cases in which development is unrelated to perceptions of existing power and confers benefits only upon completion; as a result, states lack incentives to signal ongoing progress. To fully analyze the signaling dilemma, a model should allow states to either announce or conceal military assets that provide new information about both existing and future capabilities. In addition, efforts to keep conceal technology or development should fail when clandestine activities leak or are inadvertently exposed. The model we present hereafter incorporates each of these possibilities.
2.3 The Secret Development Model

To analyze how countries decide whether to announce or conceal military power, we present a two-period game in which two risk-neutral players, \( R \) (the rising actor) and \( S \) (the suspicious actor), contest a continuously divisible good we represent as the interval \([0,1]\).\(^{19}\)

At the beginning of the game, Nature determines whether \( R \) is “Normal” or “Powerful.” \( R \)’s type determines her probability of victory in a war against \( S \). Following convention, we model war as a costly lottery that determines which player will obtain sole control of the contested good. If war occurs, both players suffer a cost for participating, \( C_i \in (0, 1] \), with \( i \in \{R, S\} \), and the loser of the contest leaves empty-handed. If \( R \) is a normal type, she expects to prevail against \( S \) with probability \( p \), but when \( R \) is powerful she enjoys a larger probability of victory: \( p + \pi \). \( R \)’s type is private information. Although \( S \) knows \( R \) is powerful with probability \( \sigma \), \( S \) remains unaware of \( R \)’s actual type in this stage unless \( R \) chooses to “announce” her power.\(^{20}\) By allowing \( S \) to confront uncertainty about \( R \)’s type, we depict a world in which an adversary may possess clandestine technology, covert equipment, or latent partnerships that raise its overall capabilities beyond what opponents can estimate from observables alone.

After \( R \) learns her type and chooses whether to announce it, the players engage in ultimatum bargaining.\(^{21}\) \( R \) proposes a division of the contested good, which we denote \( x_i \in [0, 1] \). \( S \) can reject this proposal by initiating an all-or-nothing war as explained above. If war occurs, the game ends. Alternatively, \( S \) can accept the proposal, in which case \( R \) receives the value she demanded, \((x_i)\), \( S \) obtains the complement, \((1 - x_i)\), and play proceeds to a second period. We list the overall payoffs associated with each outcome in Appendix B.\(^{22}\)

\(^{19}\)Throughout the paper we refer to \( R \) using feminine pronouns and to \( S \) using masculine pronouns.

\(^{20}\)We deliberately allow \( R \)’s announcement to be both credible and costless because this establishes the most difficult condition for secrecy. If signaling strength was costly or imperfectly credible, it would be easier for us to identify equilibria in which actors pursue secrecy. Our goal is instead to identify conditions in which rational players may avoid signaling even when demonstrations of power are free and perfectly informative. For similar reasons, the current model does not allow weak types to “bluff” by making false signals of strength.

\(^{21}\)Fearon 1995, Powell 1999, etc.

\(^{22}\)Following convention, we discount players’ second-period payoffs by \( \delta \in [0, 1] \).
In the second period, \( R \)’s type determines the range of military development strategies available to her. By “development” we refer to a process that could significantly augment \( R \)’s military power upon completion: the construction of vehicles and weaponry, the expansion of armed forces, the formalization of a military alliance, or the pursuit of weapons of mass destruction. In all such cases, a country’s type—i.e., normal or powerful— influences the level of development it can access in the short term. A powerful country with greater industrial equipment and scientific knowledge can produce larger numbers of more advanced vehicles, aircraft, and munitions compared to a country that begins with either fewer factories or less research and development experience. Similarly, a state equipped with a large population as well as powerful economic and logistical infrastructure can recruit, train, and deploy larger numbers of soldiers than could an opponent that lacked these capabilities. Alliances may also be forged more easily by states that enjoy strong relations with potential partners than by others that have alienated themselves from their neighbors. Finally, our logic may apply to the scenario that exists after a country’s initial development of nuclear weapons. States that have already developed rudimentary
or low-yield nuclear devices are better poised to acquire thermonuclear weapons or to increase the size of their stockpile compared to states that are earlier in the process of nuclear arming.

![Figure 2.2: Secret Development Model, Period 2: Development and Potential Exposure](image)

We depict this process in the model by restricting a normal-type \( R \) to “normal development,” or the marginal level of arming that a non-powerful country could accomplish within a single time period. When \( R \) is powerful she enjoys the additional option of attempting “high development,” which represents an investment in arming that, once complete, would boost her military capabilities beyond that which can be accomplished by a normal player within the same timeframe. Both types of arming result in improvements in \( R \)’s probability of victory in war against \( S \). If \( R \) completes the process of normal development, her probability of victory increases by \( D \), whereas if she successfully achieves high development her probability of victory increases instead by \( \Delta \). Because high development entails a larger improvement in capabilities than normal
development, we assume that $\Delta > D \geq 0$, while the maximum probability of victory remains capped at $p + \pi + \Delta \leq 1$.

Although potentially fruitful, high development is also costlier and riskier than normal development. We depict these characteristics in two ways. First, if $R$ pursues high development she must pay a cost, $K \in [0, 1]$. In addition, we allow the process of high development to begin clandestinely, but assume that $R$’s effort may be “exposed” with probability $\epsilon$, which is common knowledge. This risk represents the collective chance that ongoing development may be discovered via espionage, may leak via unauthorized or accidental sharing of clandestine information, or could be prematurely revealed through other means. If exposure occurs, information about $R$’s type and development choice becomes common knowledge. In response to such exposure, $S$ may initiate preventive war by attacking $R$, in which case we assume the “high development” process remains incomplete and the probability of victory for each side depends purely on $R$’s type. Alternatively, $S$ may respond to $R$’s exposure by eschewing war and allowing $R$ to continue the development process.

If $R$ either pursues normal development or, alternatively, pursues high development while avoiding exposure or war with $S$, the two players once again engage in ultimatum bargaining. As before, $R$ makes an offer $x_i$ that $S$ may either reject or accept, resulting in either war or peace. However, two events occur before bargaining occurs. First, the process development that $R$ initiated at the beginning of this time period reaches maturity, so that $R$’s probability of victory in

---

23 For simplicity, we set the cost of normal development to zero for all types of $R$, but our results are substantively consistent if we instead assume that normal development also costs $K_N$, with $0 < K_N < K$.

24 Results are consistent if $R$ can instead pursue high development publicly after revealing power in the first period.

25 We can alternatively assume that the development process is partially complete, so $R$ obtains an improvement in his probability of victory that ranges in size between $[0, \Delta)$ depending on the degree of development $R$ achieves prior to exposure. This change does not eliminate the equilibria we identify in the following section.

26 In the version of the model we present here, $S$ can only interrupt ongoing development by initiating conflict if exposure occurs. We choose to model the interaction in this way because the structure better reflects the empirical decision-making process of a suspicious actor. In the opening period of our game, $S$ is uncertain whether his adversary is powerful. To ward against the possibility of a fait accompli in the future, he initiates preventive conflict in the first period against adversaries who are unable to reassure him with sufficiently generous bargaining offers. Once this screening process is complete, $S$ allows his adversaries to proceed unless the new exposure of information about an ongoing development process forces him to reconsider preventive action.
war is now raised by either \( D \) or \( \Delta \) depending on the form of development she pursued. Second, all information about \( R \)'s type becomes common knowledge, so that \( R \)'s true probability of victory is known to both players. This reflects the fact that countries routinely engage in successful military demonstrations in order to depict their capabilities once the development process is complete. Although countries sometimes face difficulty signaling their abilities, we omit this analysis from this model for the sake of conceptual clarity and to remain consistent with our depiction of first-period signaling as perfectly credible and costless.\(^27\)

### Table 2.1: Model Payoffs.

<table>
<thead>
<tr>
<th>Outcome</th>
<th>Player R’s Payoff</th>
<th>Outcome</th>
<th>Player S’s Payoff</th>
</tr>
</thead>
<tbody>
<tr>
<td>( O_1 ):</td>
<td>( p - C_R + \delta(p) )</td>
<td>( O_1 ):</td>
<td>( 1 - p - C_S + \delta(1 - p) )</td>
</tr>
<tr>
<td>( O_2 ):</td>
<td>( x_1 + \delta(p + D - C_R) )</td>
<td>( O_2 ):</td>
<td>( (1 - x_1) + \delta(1 - p - D - C_S) )</td>
</tr>
<tr>
<td>( O_3 ):</td>
<td>( x_1 + \delta(x_2) )</td>
<td>( O_3 ):</td>
<td>( (1 - x_1) + \delta(1 - x_2) )</td>
</tr>
<tr>
<td>( O_4 ):</td>
<td>( p + \pi - C_R + \delta(p + \pi) )</td>
<td>( O_4 ):</td>
<td>( 1 - p - \pi - C_S + \delta(1 - p - \pi) )</td>
</tr>
<tr>
<td>( O_5 ):</td>
<td>( x_3 + \delta(p + \pi + D - C_R) )</td>
<td>( O_5 ):</td>
<td>( (1 - x_3) + \delta(1 - p - \pi - D - C_S) )</td>
</tr>
<tr>
<td>( O_6 ):</td>
<td>( x_3 + \delta(x_4) )</td>
<td>( O_6 ):</td>
<td>( (1 - x_3) + \delta(1 - x_4) )</td>
</tr>
<tr>
<td>( O_7 ):</td>
<td>( x_3 + \delta(p + \pi + \Delta - C_R - K) )</td>
<td>( O_7 ):</td>
<td>( (1 - x_3) + \delta(1 - p - \pi - \Delta - C_S) )</td>
</tr>
<tr>
<td>( O_8 ):</td>
<td>( x_3 + \delta(x_5 - K) )</td>
<td>( O_8 ):</td>
<td>( (1 - x_3) + \delta(1 - x_3) )</td>
</tr>
<tr>
<td>( O_9 ):</td>
<td>( x_3 + \delta(p + \pi - C_R - K) )</td>
<td>( O_9 ):</td>
<td>( (1 - x_3) + \delta(1 - p - \pi - C_S) )</td>
</tr>
<tr>
<td>( O_{10} ):</td>
<td>( x_3 + \delta(p + \pi + \Delta - C_R - K) )</td>
<td>( O_{10} ):</td>
<td>( (1 - x_3) + \delta(1 - p - \pi - \Delta - C_S) )</td>
</tr>
<tr>
<td>( O_{11} ):</td>
<td>( x_3 + \delta(x_6 - K) )</td>
<td>( O_{11} ):</td>
<td>( (1 - x_3) + \delta(1 - x_6) )</td>
</tr>
<tr>
<td>( O_{12} ):</td>
<td>( p + \pi - C_R + \delta(p + \pi) )</td>
<td>( O_{12} ):</td>
<td>( 1 - p - \pi - C_S + \delta(1 - p - \pi) )</td>
</tr>
<tr>
<td>( O_{13} ):</td>
<td>( x_7 + \delta(p + \pi + D - C_R) )</td>
<td>( O_{13} ):</td>
<td>( (1 - x_7) + \delta(1 - x_8) )</td>
</tr>
<tr>
<td>( O_{14} ):</td>
<td>( x_7 + \delta(x_8) )</td>
<td>( O_{14} ):</td>
<td>( (1 - x_7) + \delta(1 - p - \pi - D - C_S) )</td>
</tr>
<tr>
<td>( O_{15} ):</td>
<td>( x_7 + \delta(p + \pi + \Delta - C_R - K) )</td>
<td>( O_{15} ):</td>
<td>( (1 - x_7) + \delta(1 - p - \pi - \Delta - C_S) )</td>
</tr>
<tr>
<td>( O_{16} ):</td>
<td>( x_7 + \delta(x_9 - K) )</td>
<td>( O_{16} ):</td>
<td>( (1 - x_7) + \delta(1 - x_9) )</td>
</tr>
<tr>
<td>( O_{17} ):</td>
<td>( x_7 + \delta(p + \pi - C_R - K) )</td>
<td>( O_{17} ):</td>
<td>( (1 - x_7) + \delta(1 - p - \pi - C_S) )</td>
</tr>
<tr>
<td>( O_{18} ):</td>
<td>( x_7 + \delta(p + \pi + \Delta - C_R - K) )</td>
<td>( O_{18} ):</td>
<td>( (1 - x_7) + \delta(1 - p - \pi - \Delta - C_S) )</td>
</tr>
<tr>
<td>( O_{19} ):</td>
<td>( x_7 + \delta(x_{10} - K) )</td>
<td>( O_{19} ):</td>
<td>( (1 - x_7) + \delta(1 - x_{10}) )</td>
</tr>
</tbody>
</table>

\(^{27}\)An alternative setup in which \( R \) must pay a modest cost to demonstrate its power before the final bargaining stage would not eliminate the equilibria we identify in the following section.
2.3.1 Model Analysis

We show that Perfect Bayes Equilibria (henceforth, equilibria or PBE) exist that yield five distinct behaviors: (1) demonstrations of power, in which the rising player (R) announces her power and derives a bargaining benefit in the first period; (2) wars of discovery, in which the suspicious player (S) initiates a preventive attack after becoming aware of ongoing development by R; (3) fait accompli, in which R conceals her power and completes high development without exposure, thereby gaining a second-period bargaining advantage; (4) wars of suspicion, in which S launches a preventive attack because he suspects that R is pursuing a fait accompli; and, (5) strategic restraint, in which R eschews high development to reduce the risk of conflict. We omit discussion of equilibria that rely on mixed-strategies in favor of those in which both R and S adopt pure strategies.

Fait Accompli, Wars of Discovery, and the Risk of Exposure

We begin by considering the players’ behaviors in the second period. First, notice that because all information about relative power becomes public prior to the last round of bargaining, the game will always end peacefully if play proceeds to that step.\(^{28}\) The fact that the outcome will be peaceful, however, does not imply that R’s development is irrelevant. Rather, the level of R’s development determines the size of the division she can extract during negotiations. R can successfully demand a larger share after completing high development as opposed to low development. The act of completing high military development and then using the associated improvements in power to gain bargaining leverage over an adversary constitutes a fait accompli.

The rewards of a successful fait accompli are offset by the fact that war will occur if

\(^{28}\)Because player S incurs a cost for fighting, he can only credibly reject proposals when the portion of the contested good he would obtain by accepting, \(1 - x_i\), is smaller than his war payoff. Player R’s cost of war is likewise non-zero, and because both players agree about the expected outcome of war, she strictly prefers to make an offer that S will accept rather than reject. As a result, R will always demand \(x_i\) such that S’s acceptance payoff \(\geq\) his war payoff, and S will accept. For further discussion of why players will achieve agreement rather than costly rejection when engaging in ultimatum bargaining in the final period of a complete information game, see Fearon 1995.
high development is exposed prior to completion. To observe this, notice that if exposure occurs player $S$ has two options. First, he can attack, which yields an expected payoff of $1 - p - \pi - C_S$. Alternatively, $S$ can refrain from attacking. Because the latter case would force $S$ to negotiate with $R$ after she completes her development, $S$ would obtain a payoff of only $1 - p - \pi - C_S - \Delta$. He therefore prefers to attack whenever $R$’s attempt at high development is exposed. We refer to preventive attacks that result from the premature exposure of ongoing development as *wars of discovery*.

We now identify conditions in which $R$ will risk a war of discovery in hopes of accomplishing a fait accompli. $R$’s continuation payoff from choosing normal development when powerful is $p + \pi + D + C_S$. In contrast, $R$’s expected continuation payoff from high development depends on her probability of exposure and reduces to $p + \pi + \Delta + C_S - K - \epsilon(C_R + C_S + \Delta)$. R will therefore pursue high development when the risk of exposure is sufficiently low, or, more precisely, when $\epsilon < \frac{\Delta - K - D}{\Delta + C_R + C_S}$. We refer to this condition as the *Low Development Threshold (LDT)*. When the inequality is reversed and $\epsilon$ exceeds the LDT, the risk of exposure is high enough that even a powerful-type $R$ would instead pursue strategi c restraint by choosing normal development in the second period. Notice that high development is more likely when the costs of war ($C_R$ and $C_S$), the cost of high development ($K$), and the benefits of normal development ($D$) are each small.

**Announcement, Reassurance, and Wars of Suspicion**

The previous section describes how the risk of premature exposure ($\epsilon$) influences whether $R$ would attempt or avoid high development in the second period. Player $S$, however, can take action in the first period to prevent the game from reaching this stage. In particular, if $S$ suspects that $R$ is powerful and is poised to pursue high development, $S$ may initiate a preventive war by rejecting

\[\text{This matches the result in Debs and Monteiro (2014) that preventive wars occur when adversaries are certain of imminent power shifts.}\]

\[\text{We calculate } R\text{’s payoffs by comparing } \epsilon(\text{preventive war}) + (1 - \epsilon)(\text{peace after a second round of bargaining}), \text{ which is equivalent to } \epsilon(p + \pi - C_R - K) + (1 - \epsilon)(p + \pi + \Delta + C_S - K).\]

62
R’s proposals during the first round of bargaining. Although risky, such a war would end the game and deny R the opportunity for further growth, thereby ensuring a fait accompli does not occur.

Under what conditions would S take preventive military action in the first period? When could R reassure or compensate S in order to dissuade him from such attacks? We answer these questions by identifying the minimum bargaining share \((1 - x_i)\) that S would accept rather than reject. We can characterize S’s acceptance threshold by considering two parameters that usefully divide the parameter space into relevant regions. To begin, because \(\epsilon\) is common knowledge S knows whether the LDT has been crossed and can anticipate whether R would attempt high development in the second period if given the opportunity. S, however, may not know R’s type unless R has chosen to reveal it. We therefore let \(\hat{\sigma} \in [0, 1]\) depict S’s belief that R is powerful at each decision node.

First consider the region of the parameter space in which \(\epsilon > LDT\) and the risk of exposure is sufficiently high that R will always pursue low development. Moreover, as we explain below, R will always announce when she is powerful and—regardless of whether she is powerful or normal—will propose a division that induces S to forgo preventive violence. To understand the logic, we can compare S’s rejection thresholds in the first bargaining period depending on whether he observes an announcement of power from R. When S is uncertain whether R is powerful, he will reject any bargaining proposals in which \(x_i > p + \hat{\sigma}(\pi) + C_S - \delta(D + C_S)\). Because the size of the division R can extract \((x_i)\) increases with \(\hat{\sigma}\), R can maximize her bargaining leverage with a Demonstration of Power that removes S’s uncertainty and causes \(\hat{\sigma}\) to assume a value of 1. After making this announcement, R will propose the largest division \((x_i)\) that S would accept. This move induces S to avoid preventive war and enables R to obtain a larger payoff than she could expect to achieve through fighting. Because a powerful R will always announce her type if

\[ 31 \text{ S anticipates the following payoff from rejecting R's proposal and initiating preventive war: } 1 - x_i + \delta(1 - p - \pi - D - C_S). \text{ In contrast, when S is uncertain of R's type he anticipates a payoff of } 1 - x_i + \delta(1 - p - D - C_S) - (\hat{\sigma} \times \delta)(\pi). \text{ If he accepts R's proposal, S's rejection payoff exceeds his acceptance payoff when the condition in the text holds.} \]

\[ 32 \text{ R's payoff from announcing her power and proposing the maximum division } (x_i) \text{ that S would accept is } p + \pi + C_S + \delta(p + \pi). \text{ In contrast, R's expected payoff if war occurs at this stage is } p + \pi - C_R + \delta(p + \pi). \text{ The former value exceeds the latter when } C_R + C_S > 0, \text{ which is always true.} \]
\( \varepsilon > LDT \), when player \( S \) fails to observe such an announcement he will instead conclude that \( R \) is normal and will raise his bargaining expectations accordingly. Nevertheless, even a normal-type \( R \) will propose a division that induces acceptance from \( S \).\(^{33}\) In summary, when the risk of exposure is sufficiently high, \( R \) will forgo high development, will announce her existing power if given the opportunity, and will avoid preventive war by proposing a division that \( S \) will accept.

Now consider the region of the parameter space in which \( \varepsilon > LDT \) and a powerful \( R \) would attempt high development. Within this region, announcements of power by \( R \) could provoke war by alerting \( S \) to the possibility that \( R \) will pursue a large power shift in the future. To ward against the possibility of preventive attack, \( R \) may opt to conceal her power in the first period—in effect, maintaining temporary secrecy in order to pursue a fait accompli. Regardless of whether she is genuinely powerful, \( R \) must also reassure \( S \) by offering a relatively generous bargaining division that accounts for \( S \)'s suspicion that \( R \) may secretly be strong (\( \hat{\sigma} \)). Unfortunately for \( R \), however, \( S \)'s level of suspicion is sometimes so high that he will reject any proposal that \( R \) can credibly offer. In these circumstances, \( S \) initiates a War of Suspicion in order to remove any possibility of a fait accompli—even though such an option may not in actuality be within \( R \)'s grasp.

To understand the intuition for these results, we once again begin by comparing \( S \)'s rejection thresholds depending on whether \( R \) announces power. If \( S \) is certain \( R \) is powerful, he rejects any bargaining proposal in which \( x_i > p + \pi + C_S - \delta[C_S + \Delta(1 - \varepsilon)] \).\(^{34}\) In contrast, when \( S \) is uncertain about \( R \)'s type, he rejects proposals in which \( x_i > p + C_S + \hat{\sigma}(\pi) - \delta(D + C_S) + (\delta \times \hat{\sigma})[D - \Delta(1 - \varepsilon)] \).\(^{35}\) Although the perception of status quo power continues to provide \( R \) with bargaining benefits, as characterized by the term \( \hat{\sigma}(\pi) \) in the preceding inequality, these rewards are offset by the threat that \( R \) could achieve a large boost in power in the second period and then

---

\(^{33}\)When \( R \) is a normal-type, her payoff from proposing the maximum division \( (x_i) \) that \( S \) would accept is \( p + C_S + \delta(p) \), whereas \( R \)'s expected payoff from war is \( p - C_R + \delta(p) \). Because \( C_S + C_R > 0 \), the former always exceeds the latter.

\(^{34}\)\( S \)'s expected payoff from fighting if \( R \) will pursue high development and he knows \( R \) is powerful is \( 1 - p - \pi - C_S + \delta(1 + \pi - p) \), while his expected payoff from accepting \( R \)'s offer is \( 1 - x_i + \delta(1 - p - \pi - \Delta - C_S) + (\varepsilon \times \delta \times \Delta) \). \( S \)'s expected war payoff exceeds his expected peace payoff when the condition in the text holds.

\(^{35}\)In this case, \( S \)'s expected payoff from fighting is \( 1 - p - C_S + \delta(1 - p) - \delta[\pi + (\delta \times \pi)] \), whereas his expected payoff from accepting \( R \)'s offer is \( 1 - x_i + \delta(1 - p - D - C_S) + (\delta \times \delta)[D - \Delta(1 - \varepsilon) - \pi] \).
force \( S \) to yield significant concessions: \((\delta \times \hat{\sigma})[D - \Delta(1 - \varepsilon) - \pi]\).\(^{36}\) As a result, \( R \)’s behavior differs from the low-development region of the parameter space, where \( R \)’s bargaining leverage increased with \( \hat{\sigma} \) and \( R \) therefore always prefers to announce her power. In the high-development region, the opposite result holds: \( R \) can secure a more favorable first-period agreement from \( S \) when she does not reveal her strength.\(^{37}\) If she announced her capacity for future growth, \( R \) would be forced to compensate \( S \) by requesting a smaller share of the contested good \((x_i)\) than she could claim if \( S \) remained uncertain about \( R \)’s type or if \( R \) could commit to normal development. This requirement gives \( R \) an incentive to conceal her existing power and to mimic the behavior of a normal type in order to prolong \( S \)’s uncertainty about \( R \)’s development trajectory.

Unfortunately for \( R \), there are circumstances in which \( R \) cannot propose a sufficiently generous division to reassure \( S \) and avoid preventive conflict. In the opening round of bargaining, the minimum amount \( R \) can claim is \( x_i = 0 \), as this would entail forfeiting the entirety of the contested good to player \( S \). Nevertheless, when \( S \) is highly suspicious even this amount may not be sufficient to satisfy \( S \).\(^{38}\) In this case, even the most generous feasible offer from \( R \) would still lead to a preventive War of Suspicion in which \( S \) initiates preventive attacks despite his uncertainty about \( R \)’s current and future capabilities. In some cases, these wars of suspicion will be waged correctly against secretly-powerful actors who are concealing their abilities, but in other cases they will target normal-type states that are unable to signal their relative weakness.

Even when the players can avoid wars of suspicion, one final aspect of their interaction merits consideration. When \( S \) is uncertain about \( R \)’s type and remains concerned that \( R \) may develop significantly in the second period, \( R \) must provide reassurance in the form of a smaller initial bargaining offer than she could make if \( S \) was not suspicious. When \( R \) is powerful and is concealing this power, she expects to recoup these losses after engaging in high development in

---

\(^{36}\)Because \( D < \Delta(1 - \varepsilon) \) when \( R \) will pursue high development, \( \hat{\sigma}(D - \Delta(1 - \varepsilon)) \) is negative.

\(^{37}\)The size of the division \( R \)’s could claim after announcing power would only exceed the division she could claim while \( S \) remained uncertain if \( \hat{\sigma} > 1 \), which is impossible.

\(^{38}\)\( S \) will require a first-period offer that exceeds the entire contested good when \( \hat{\sigma} > \frac{\delta(D + C_S - p - C_S)}{(x + \delta\Delta(1 - \varepsilon))} \).
the second period and, as a result, will always attempt to find a peace deal in the opening stage.\textsuperscript{39} When \( R \) is normal, however, concessions made in the first round of bargaining are unrecoverable. In effect, when \( R \) is normal she suffers a Reassurance Tax simply because she inhabits a world in which secret development is possible.\textsuperscript{40} By reassurance tax, we refer to the decrease in the size of the first-period division \( R \) can peacefully propose when \( S \) is uncertain of \( R \)’s type, compared to what \( R \) could otherwise obtain if she could clear away \( S \)’s suspicion by demonstrating that she was a normal type who would not enact high development in the future.\textsuperscript{41}

\subsection*{2.4 Discussion and Implications}

\subsubsection*{2.4.1 Power Shifts and Preventive Attack}

Several enduring political questions ask how actors respond to shifts in the balance of power. Robust literatures explore whether new alliances deter or provoke conflict,\textsuperscript{42} how actors choose between arming or allying in response to threats,\textsuperscript{43} and why countries sometimes accommodate emerging rivals but in other cases initiate war.\textsuperscript{44} Central to all these discussions is the assumption that power shifts may provoke adversaries to adopt preventive behavior or even engage in preventive attacks.\textsuperscript{45} In recent decades, formal theorists have identified specific mechanisms

\begin{align*}
\text{When } R \text{ is powerful, her payoff from inducing acceptance from } S \text{ exceeds her payoff from war when } \hat{\sigma} > (|\hat{\epsilon} - \hat{\delta}| - 1)(C_S + C_R) + \hat{\delta}(K) - 1 \text{ The right hand side of the inequality is negative, so the condition is always satisfied.}
\end{align*}

\textsuperscript{40}\text{The reassurance tax is positive when } \hat{\sigma} > 0 \text{, which is always true.}

\textsuperscript{41}\text{In some cases, the cost of reassurance is itself so large that a normal-type } R \text{ would prefer to fight rather than offer the concessions necessary to induce peaceful agreement from } S. \text{ More precisely, when she is normal, } R \text{'s war payoff exceeds her peace payoff when } S \text{ is uncertain when: } \hat{\sigma} > \frac{C_S + C_R}{\hat{\sigma}^2 - \hat{\epsilon}^2 (1 - \hat{\epsilon})}. \text{ In these circumstances, each type of } R \text{ will propose the maximum division } (x_i) \text{ that } S \text{ would accept if } S \text{ knew } R \text{'s true type. The difference between the offers made by the powerful and normal types of player } R \text{ allow } S \text{ to discern between types, thereby enabling each state to avoid war.}


\textsuperscript{43}\text{Morrow 1993, Glaser 2004, Monteiro and Debs 2014, Yarhi-Milo, Lanoaszka, and Cooper 2016.}


\textsuperscript{45}\text{Thucydides 1954, p. 1.23 famously attributed the cause of the Peloponnesian War to "the growth of Athenian power and the fear which this caused in Sparta." Other prominent discussions of power shifts and war include Gilpin 1981, Levy 1987, Walter 1997, Copeland 2000, Trachtenberg 2007, and Bell and Johnson 2015.}
through which such wars can arise. Fearon (1995) introduced an analytic stylization of preventive war as a result of impending shifts in power, Powell (2004, 2006) identified a general condition by which “large and rapid” power shifts should guarantee conflict under complete information, and Krainin (2017) generalized this result to also include slower, long-term shifts.\footnote{See also Powell 1999, Leventoğlu and Slantchev 2007, and Schub 2017.} Across all such models, preventive war occurs when one state knows that its enemy is poised to achieve a significant increase in military power in the future. Because this power shift would enable the rising state to extract costly bargaining concessions from its opponents, countries on the cusp of decline initiate risky preventive action to stop the shift from taking place.

In contrast to Powell, our model demonstrates that “large and rapid power shifts” are neither a necessary nor sufficient cause of preventive war. Preventive attacks can occur even in the absence of genuine arming as long as an opponent suspects that military development will occur in the future. Furthermore, even the presence of a genuine power shift may not cause war if an opponent is unaware or uncertain that the shift will occur.\footnote{See Debs and Monteiro 2014 and Bas and Coe 2016 for similar results.} These results produce important implications for studies of military arming, power shifts, and preventive conflict. Researchers should not assume that large swings in the balance of power will consistently predict preventive conflict, both because potential power shifts may not be apparent to adversaries until they are publicly revealed and also because suspicious adversaries may launch preventive attacks even when the distribution of power will remain fixed. Instead, empirical research on preventive attacks should account for adversaries’ beliefs about the likelihood of shifts in the balance of power.

\subsection*{2.4.2 The Risks of Signaling}

International crises can end in two ways: the first is when one actor loses the capacity to resist—as when its military forces are exhausted through battle—and the second is when at least one actor agrees to concede the disputed stakes. The player that grants such a concession must believe the
terms of a potential peace deal are preferable to the payoff she would receive if she allowed the crisis to continue or escalate. Conventional wisdom therefore suggests that strong countries have incentives to signal their military strength and resolve. Credible and convincing signals should persuade opponents that the strong country is unwilling to grant large concessions, so the onus for compromise rests with the weaker actor, which should expect to perform poorly if negotiations fail and war begins.  

This model demonstrates that countries also face incentives to forgo signaling—and to eschew development—in appropriate circumstances. Although states should eagerly announce their military development when they can do so without provoking conflict, signaling strength may not be useful when the opposite is true. These findings contribute to our understanding of the risks and benefits of signaling. Whereas previous work recognizes the risks of signaling when war is already imminent—Slantchev (2010), for example, observes that signaling can enable opponents to prepare better for war, while Lindsey (2015) shows that signals may allow adversaries to respond with better tactics during war—our model demonstrates that sharing information can directly provoke war where none would otherwise occur.

Finally, because our model identifies conditions in which secrecy is preferable as well as those in which announcements can be made safely, our results may help explain variation in the effort states direct toward concealing nuclear weapons development. Previous work identifies nuclear development as the optimal circumstance in which states should prioritize secrecy rather than alert their adversaries to the imminent completion of a weapons system. Nevertheless, countries exert varying levels of effort toward such secrecy. Consider the case of North Korea, which made little effort to conceal its interest in nuclear weapons development early 2000s, going so far as to withdraw from the Treaty on the Non-Proliferation of Nuclear Weapons (NPT), insist to American diplomats that it already possessed weapons, and expel several IAEA inspectors. In contrast to many other potential proliferators, Pyongyang already maintained relatively robust

---

48Banks (1990), Fearon (1995), etc.
49Debs and Monteiro (2014), Bas and Coe (2016).
conventional military capabilities—a factor that limited the size of the power shift that could occur even if the state obtained nuclear weapons. Countries with smaller conventional militaries or for whom nuclear weapons would cause a larger shift in military capability relative to key adversaries may need to devote greater effort to maintaining secrecy.

2.4.3 Wars of Discovery and Wars of Suspicion

Our model clarifies the logic of preventive war by showing that such wars occur through two distinct mechanisms. Wars of discovery can occur when clandestine activities are prematurely exposed. Consider the Soviet behaviors that provoked the 1962 Cuban Missile Crisis. Concerned about a potential U.S. invasion of Cuba, Soviet Premier Nikita Khrushchev decided to deploy weapons on the island in hopes that such weapons would equalize “what the West likes to call the ‘balance of power.’”\(^50\) To avoid provoking American preventive actions, the Soviets sought to keep the extent of their military relationship with Cuba a secret the power shift was complete.\(^51\) Indeed, when Cuban leaders proposed that the two countries announce their partnership in order to establish immediate deterrent benefits, Khrushchev refused the request, promising instead to reveal the full extent of Soviet-Cuban military cooperation in a fait accompli once all deployments were operational..\(^52\) The Soviets’ secrecy gambit backfired on October 14, when American U-2 reconnaissance identified offensive missile sites in San Cristobal. Although Khrushchev eventually chose to withdraw the missiles, their discovery nearly provoked a war of discovery with the United States.\(^53\)

---
\(^{50}\)Quoted in George and Smoke (1974, p. 462). President Kennedy’s statements reflect a similar sentiment, including his lament that “The Soviet move had been undertaken so swiftly, so secretly, and with so much deliberate deception... that it represented a provocative change in the delicate status quo” (Quoted in Lebow 2000, p. 15).

\(^{51}\)Several Soviet personnel argued that exposure could provoke war (see Lebow and Stein 1995 and Lebow 2000).

\(^{52}\)Hansen 2002.

\(^{53}\)Our model abstracts from analyzing how the leaders averted war in this case. It is possible that when Khrushchev initially opted to attempt secret development he underestimated the costs associated with a war of exposure. When Soviet missile deployment began, President Kennedy had not yet publicly pledged to prevent Cuba from obtaining offensive military capabilities. As a result, Khrushchev might have assumed that exposure could result in American preventive action that fell short of nuclear conflict. As the crisis elapsed, a series of unauthorized incidents coupled with Kennedy’s behavior led Khrushchev to believe that widespread nuclear engagement was more likely.
On the other hand, *wars of suspicion* arise when threatened states suspect their adversaries are or will conceal significant development activity—regardless of whether such activity actually exists. One prominent example of such suspicion relates to the United States’ invasion of Iraq in 2003. Substantial debate exists in the literature as to why the United States concluded that Iraq possessed WMD and, more importantly, why Iraq was unable to quell American suspicions. Although other researchers blame the United States for its failure to gather accurate intelligence about Iraqi WMD development, their explanation raises the question of why Iraq was unable to “clear the air.” As President George W. Bush lamented in his memoirs, “If Saddam [Hussein] didn’t have WMD, why wouldn’t he just prove it to the inspectors?” Our model provides a formal explanation for both Hussein’s failure to provide information and the United States’ choice to invade: the strategic environment inhibited Iraq from sharing information that would have reassured the U.S.

In the language of our model, American suspicion ($\hat{\sigma}$) that Iraq was developing WMD was extremely high in the prelude to the invasion. The United States developed a perception throughout the 1990s that “Iraq would never be forthcoming, and that if it was blocking access to the UN [inspectors], then it must have something to hide.” In effect, even though Iraq was not developing WMD, the state was subject to a reassurance tax because it inhabited a security environment in which clandestine activity was possible. To complicate matters further, Hussein, was initially unwilling to offer convincing evidence that he had dismantled his nuclear program for

---

55 See Kaufmann (2004), Flibbert (2006), Lake (2010), and Debs and Monteiro (2014) for arguments that the U.S. failed to optimally gather and process information.
56 Bush 2010, p. 269. Note that although there is still much debate about whether members of the Bush administration maintained additional interest in the conflict, the overriding question is why Iraq did not demonstrate that the United States’ publicly-stated rationale for war was built on flawed estimations of Iraqi WMD production.
57 This corresponds to Lake’s (2010) conclusion that Hussein chose not to reveal his lack of WMD because by doing so he would have incurred steep domestic costs and been constrained from deterring other opponents.
58 Duelfer and Dyson 2011, p. 97. Following the terrorist attacks of September 11, 2001, members of the U.S. intelligence community also became acutely aware that they lacked the capabilities to detect all potential security threats in a timely manner (Debs and Monteiro 2014). As a result, the Bush administration adopted its “one percent” doctrine, according to which it treated even a one percent chance that Iraq could develop nuclear weapons as an unacceptable risk (Lake 2010).
fear that doing so would also alert his domestic and regional opponents to his military weakness.\textsuperscript{59}

Even when Hussein’s priorities shifted and he allowed United Nations (UN) inspectors to return in November 2002, the United States remained suspicious that Iraq was concealing additional capabilities and was on a trajectory to obtain WMD. In short, the largest credible concession that Hussein could make in terms of inspections still could not reassure the United States. In the end, the Bush administration remained suspicious of Iraqi development and the U.S. embarked upon a war of suspicion.

By identifying this distinction between wars of discovery and wars of suspicion, we make three contributions. First, we bring formal models of preventive war into better harmony with historical and qualitative accounts of conflict. Whereas canonical formalizations depict wars of discovery in which actors become aware of ongoing or imminent power shifts, historians often describe wars of suspicion in which states remain uncertain but suspicious of their rivals’ developments. Our model not only draws attention to both causal logics but also identifies conditions in which each category of preventive war is likely to occur.

Second, distinguishing between these separate mechanisms allows us to clarify the means by which states may attempt to avoid preventive war. Countries can eliminate the risk of wars of discovery by forgoing development—after all, development cannot be exposed if no development has occurred. However, countries cannot always eliminate the threat of wars of suspicion: when player $S$’s suspicion level exceeds $\delta(D+C_S)-p-C_S\left(\frac{\delta(D+C_S)-p-C_S}{(\pi+\delta(D-D(1-e)))}\right)$, player $R$ will lack any credible means of demonstrating that she has eschewed secret development and war is unavoidable.

Finally, acknowledging the difference between discovery and suspicion improves our understanding of how uncertainty relates to preventive conflict, an issue much debated in recent work. Whereas Debs and Monteiro (2014, p. 2) claim that “when power shifts are endogenous...\textsuperscript{59} As Gordon and Trainor (2006, p. 63) explain, Hussein’s “top priority was protecting his government against potential coups and internal threats... Iran, an adversary with whom he had fought a bloody eight-year war, was next on the list of dangers.” The Iraqi leader appears to have believed that maintaining ambiguity over his WMD arsenal would simultaneously quell domestic unrest and deter attacks from Tehran. Duelfer (2004, p. 32) likewise argues that “This led to a difficult balancing act between the need to disarm to achieve sanctions relief while at the same time retaining a strategic deterrent. The Regime never resolved the contradiction inherent in this approach.”
preventive war requires uncertainty,” their view contrasts with Krainin (2017, p. 106), who argues that “incomplete information is not necessary to cause war using the logic of commitment problems.” Our model shows that uncertainty yields different effects across each category of preventive conflict. Wars of suspicion are indeed caused by uncertainty: adversaries could avoid conflict if only $R$ could credibly prove that it was not developing. However, wars of discovery are not caused by uncertainty but rather by the removal thereof: if player $S$ remained uncertain about $R$’s ongoing development, no discovery would occur and $S$ would refrain from fighting.

2.4.4 Empirical Generality and the Reassurance Tax

Throughout this paper we refer to “war,” “countries,” and “military development” in order to facilitate intuition. However, the mechanisms we identify should apply to interactions beyond full-scale interstate wars. For example, although we discuss “preventive war,” the action could describe any activity—violent or otherwise—in which risky or costly actions by one actor can thwart the relative gains of another. A country that discovers an adversary’s secret research lab could, for example, attempt to destroy the lab or its employees with airstrikes, covert sabotage, cyberattacks, etc. None of these actions constitute “war” in popular parlance, but all fit the mould of our model by depicting costly and risky actions that can prevent the rise of an adversary.

Likewise, the players in our model could include any strategic actors that are locked in bilateral conflict, ranging from states and insurgents to businesses and labor unions. Consider the example of businesses, which face the decision of whether to patent the technologies they develop. A firm that submits a patent can lock in a small flow of benefits in the form of licensing fees that competitors must pay if they adopt similar technology. However, patents also sometimes expose development pathways by making technological solutions public to attentive competitors. In some cases, the information contained within a public patent can allow competing firms to catch up in the research and development race in ways that counteract the revenue gained via licensing. As a result, businesses sometimes choose to delay or even forgo patent filings to avoid giving
the competition an opportunity to react.  

Similarly, when a labor union discovers pro-business lobbies are secretly advancing anti-union legislation, the union may take preventive action by carrying out strikes or counter-lobbying to quash the bill. Although such forms of competition are not war in the colloquial sense, and “legislation” is not the same as “military development,” the situation parallels the strategic calculus laid bare by our model.

The wide applicability of our model is particularly helpful for making intuitive sense of one of our key discoveries, the reassurance tax. The option of secrecy, as noted earlier, is a boon to actors wishing to pursue ambitious power shifts, but a burden to those happy with the status quo. To see this latter point, imagine a citizen living in a surveillance state, where citizens’ actions are constantly monitored for signs of radicalization or criminality. Under such circumstances, citizens must curtail their behavior in various ways in order to avoid drawing suspicion. If the government is for some reason suspicious of a particular religious group, for example, citizens may avoid converting to that religion or associating with its members. Current members may likewise choose to eschew outward signals of devoutness either in public or on social media for fear of being profiled. We think of these self-imposed curtailments as compensation paid by citizens to reassure suspicious authorities.

The same concept also applies to the burgeoning research on civilian wartime informants. In this literature, a government seeks to crush an insurgency embedded in a civilian population, but requires strategic information from the civilians on the whereabouts of insurgent weapons caches or hideouts. Models in this literature have so far treated the interaction as a one-off game: the civilian chooses whether or not to share with the government information about

---

60 The WD-40 company famously chose not to patent its eponymous product “in order to avoid having to disclose the ingredients publicly” (Martin 2009). Likewise, Elon Musk refuses to patent technologies developed by his company SpaceX, arguing that “Our primary long-term competition is in China. If we published patents, it would be farcical, because the Chinese would just use them as a recipe book” (Anderson 2012.)

61 Consider the actions of crypto-Muslim Moriscos who hid their identity by practicing taqiyya during and after the Spanish inquisition (Harvey 2005), the behavior of Jews who claimed false identities in occupied Europe, and the experience of targeted civilians in Vietnam, Colombia, Mosul, and other recent or contemporary conflict settings.

62 Condra and Shapiro 2012, Shapiro and Weidmann 2015.
the insurgents, and the game ends thereafter.\textsuperscript{63} In reality, of course, intelligence agencies cultivate informants to provide a flow of information over time. As a repeated game, secrecy becomes paramount, since the insurgents have a reason and opportunity to engage in prevention by killing or turning informants that may collaborate with the state.

The importance of secrecy in these situations is sufficiently intuitive that previous researchers have assumed that formalizing the interaction would provide limited intellectual returns. Our model, however, reveals surprising results related to reassurance: because both non-informants live under the same weight of suspicion as informants, they must take costly steps to reassure insurgents that they are not cooperating with the government. Such steps may range from avoiding places frequented by government officials to refusing government services for fear of them being viewed as compensation, or even actively helping the insurgency. In all cases this implies a polarizing effect, where civilians are forced to take sides in ways that may either help or hinder the government. Technological shocks that facilitate secrecy, such as the introduction of mobile phones or counterinsurgency hotlines, raise the \textit{reassurance tax} and exacerbate civilian polarization.\textsuperscript{64} Empirically, this polarization is also problematic for difference-in-differences (DID) studies that hope to capture the effect of civilian informants on outcomes such as the incidence of attacks or government casualties. In particular, wherever informing behavior is observed, ‘reassurance’ activity also exists and DID estimates will conflate these two effects.

\section*{2.5 Secret Alliances and Wars of Suspicion}

We provide empirical support for the existence of \textit{Wars of Suspicion} using historical data on secret military alliances, which serve as a useful test for our argument both theoretically and empirically. First, military alliances match our theoretical assumptions by capturing the dual signals that military arming can send. Prevailing theories of alliance formation focus largely

\textsuperscript{63}Berman, Shapiro, and Felter 2011.

\textsuperscript{64}See Shapiro and Weidmann (2015) for an example of such a shock.
on the first half of this dualism, emphasizing that alliances enable states to credibly signal their collective power. In short, by forming and announcing a military alliance, two countries reveal their intention to cooperate with each other on matters of policy or defense. Opponents that observe the alliance and consider it genuine should react as though the allies have aggregated their capabilities. Even if the aggregation is inefficient, the allied coalition should enjoy greater deterrent capability and bargaining leverage than its individual members could each claim prior to the announcement.

Despite these deterrent and bargaining benefits, alliance announcements may also alert opponents to the development trajectory of an allied coalition. As we discuss above, the literature on preventive war argues that a threshold exists in military arming whereby declining states may pursue preventive wars when they anticipate sufficiently large and rapid shifts in the balance of power. By announcing an alliance, a state whose individual development trajectory otherwise fell within acceptable bounds might signal that its development may instead occur at a much higher rate, thereby motivating opponents to initiate preventive conflict so as to thwart the imminent power shift. The desire to avoid preventive conflicts may motivate rising states to conceal military alliances until their development has progressed sufficiently that it may be safely revealed.

Consider, for example, the clandestine military cooperation between Germany and the Soviet Union in the European interwar period. Following Germany’s defeat in WWI, the victors placed strict limits on the size and scope of the German military, limiting membership in the German military to 100,000 soldiers and explicitly forbidding the country from procuring submarines, aircraft, or armored vehicles. Germany bypassed these regulations in part by forming a secret agreement with the Soviet Union whereby the Reichswehr dispatched advisors to train young Soviet officers in exchange for access to hidden military bases and manufacturing plants inside Russia. Throughout the 1920s and early 1930s, thousands of German scientists and engi-

66Because alliances are costly to initiate and maintain, only states that are committed to the alliance terms and who expect to gain from military cooperation should opt to ally.
neers worked within the Soviet Union testing military prototypes, developing chemical weapons, and laying the groundwork for the mass production of new German weapons and equipment. In the early phases, the two countries sought to conceal their activity out of concern that the western allies would undertake preventive action if the scale of German rearmament became public knowledge. Indeed, when an emissary from Germany was briefly detained while ferrying documents pertaining to the military cooperation, he worried that the great war would “reignite if the allies discovered what I [was] carrying.” By 1933, however, these fears had dissipated, and Hitler concluded that his country’s development had progressed far enough that he could publicly acknowledge rearmament without provoking the allies into preventive action.

The German-Soviet cooperation is far from the only case in which states concealed their military cooperation. Secret alliances are also a useful test case for our theory because of their long existence as a common feature of international politics. For example, Grosek (2007) identifies 593 secret treaties that existed between 1521 and 2000. Although more limited in temporal coverage than Grosek’s qualitative approach, the Alliance Treaty Obligations and Provisions (ATOP) dataset facilitates the systematic study of secret military alliances that have existed since the end of the Napoleonic wars. Based on the ATOP data, at minimum roughly 20 percent of all alliances formed between 1815 and 1956 were secret in nature.

Other researchers have used this data to provide support for alternative theories of conflict escalation. Most prominently, Bas and Schub (2016) argued that the existence of secret alliances could provide support for the theory that wars sometimes result from mutual optimism. According to their theory, states that participate in secret alliances possess greater military capabilities than their adversaries can anticipate. Opponents that lack accurate information about such alliances

---

67 Johnson 2016.
69 Bas and Schub (2016). Whether additional clandestine alliances existed and whether secret alliances remain prevalent today is difficult to assess—although such agreements might have declined in frequency, they may also exist while remaining hidden from researchers. See also Ritter 2003.
70 On mutual optimism, see Morrow 1989; Fearon 1995; Powell 1999; Slantchev 2003; Slantchev and Tarar 2011; Fey and Ramsay 2011; and Lindsey 2019.
liances will remain “optimistic” insofar as they overestimate their own likelihood of victory or underestimate the likely costs of a potential conflict against a secretly-allied foe. Applying their theory to dyadic interstate data between 1816 and 1923, Bas and Schub (2016) find an association between the presence of secret alliances and the likelihood of Multilateral Interstate Disputes (MIDs).

Because theories of mutual optimism relate to the status quo distribution of power—as opposed to how the distribution of power might shift in the future—Bas and Schub (2016) did not account for the possibility of power shifts in their empirical analysis. In contrast, our theory depicts a scenario that includes both information asymmetries about existing power as well as potential commitment problems related to shifts in future power. As a result, we argue that the relationship between secret alliances and conflict may not stem purely from mutual optimism. For example, if war would not occur if both sides were fully informed about one another’s capabilities, one wonders why secretly-allied states would not avert inefficient wars by revealing their alliances on the eve of conflict. In other words, if alliance participants genuinely possess hidden military strength as a result of their collaboration, disclosing the alliance should induce an optimistic opponent to return to the bargaining table, thereby allowing both sides to avoid the costs of fighting.

Bas and Schub (2016) provide two possible explanations for this behavior. One focuses on domestic factors. In this view, leaders continue to conceal their alliances when publicly announcing the alliance would force the leader to incur domestic or international political costs. This mechanism requires a questionable reading of history. Many of the alliances in the data were signed by autocratic states during a time period in which citizens lacked significant political influence and were only weakly attuned to international relations—conditions in which leaders should be insulated from the costs of domestic backlash. The authors’ second explanation.

71 Audience-based explanations struggle to explain why states would conceal alliances during negotiations but adhere to them once conflict begins. If constituents can deter leaders from even announcing alliances, how could leaders fulfill the terms of their alliances if called upon to do so? Deploying troops on behalf of an unpopular ally should pose a larger political liability than merely announcing one’s support. For accounts of secret cooperation that
for why states continue to keep their alliances secret on the eve of conflict relates to Slantchev (2010)’s “Feigning Weakness.” In this case, the allies believe (1) the announce of the alliance may not deter the adversary from initiating conflict, and (2) the announcement would reduce the allies’ tactical advantages on the battlefield—by, for example, alerting the enemy to the possibility that the war would be fought on multiple fronts. This possibility, however, abstracts away from the theory of mutual optimism by arguing that war would still occur even if the allied side announced its alliance and revealed accurate information about the expected outcome of war.

We resolve these concerns by demonstrating that the empirical relationship between secret alliances and conflict may not result from mutual optimism but rather from wars of suspicion. In this view, adversaries initiate wars against countries they suspect may be secretly allied. If opponents’ suspicions perform better than random chance—in other words, if they are more likely to pursue wars of suspicion against secret allies than against states that are not secretly allied, an association between secret alliances and conflict should exist in the data. Moreover, unlike the mutual optimism explanation, our theory regarding wars of suspicion provides a justification for why states forge secret alliances in the first place: as in the example of Germany and the Soviet Union, alliance members choose to conceal their cooperation out of concern that revealing the alliance would provoke preventive attacks from adversaries who realize that the allies may experience more rapid military growth than would otherwise be possible.

How can we distinguish empirically between these differing explanations for the association between secret alliances and war? Whereas the mutual optimism account suggests that

---

72 See also Lindsey (2015).
73 The empirical support we offer throughout this section is also consistent with our wars of discovery mechanism. In this case, the MIDs we observe in the data might have occurred because opponents discovered evidence of secret alliances forged between rapidly-growing states. Unfortunately, we cannot adjudicate between this possibility and our war of suspicion mechanism, because we lack the evidence necessary to provide a reliable measure of whether alliances remained perfectly clandestine in the immediate prelude to each MID. Note, however, that the although possibility of premature exposure is consistent with our overall theory, it is problematic for the mutual optimism explanation: alliances should not contribute to private optimism if they become public knowledge prior to a dispute. MIDs that occur in the wake of public announcement or discovery must therefore result from an alternative mechanism.
such alliances should consistently raise the probability of conflict within all dyads in which they exist, the war of suspicion theory is conditional. According to our theory, maintaining alliance secrecy is most important when allies are acquiring military power relative to a rival state. As such, opposing states should be most concerned about the potential existence of secret alliances—and most willing to attempt wars of suspicion—when they confront adversaries that are experiencing improvements in military power. In contrast, countries should only rarely launch wars of suspicion against secretly-allied adversaries that are not growing in relative power.

2.5.1 Research Design

To test for the existence of a conditional relationship between shifts in military power, secret alliances, and the onset of conflict, we begin by replicating the empirical setup in Bas and Schub (2016). The unit of analysis throughout our tests is the non-direct dyad year, using only politically relevant dyads in which states are contiguous, are separated by less than 400 miles of water, or at least one is a major power. We further restrict our analysis to observations that fall between 1816 and 1923 because, as we discuss above, secret alliances almost entirely disappear from observable data outside of this range.\(^{74}\)

To assess whether each member of a dyad participated in a secret alliance with a third party, we used data from ATOP, which codes whether any or all alliance provisions are secret. Throughout this section, we code alliances as secret only if all of their provisions are secret. To account for the possibility that rational actors acknowledge the possibility that their adversaries may participate in secret alliances, Bas and Schub (2016) develop a measure that incorporates each actor’s prior belief about the existence of such alliances. They base this prior on the average rate of secret alliances in the dataset and the share of the world’s military capabilities that those

\(^{74}\)Because secret alliances are unobserved outside of this range, including subsequent observations would violate the positivity assumption that treated and control units exist across strata (see Petersen et al. 2012). The only secret alliance to appear in the ATOP data post 1923 is the 1956 alliance between the United Kingdom, France, and Israel that preceded the Suez Crisis.
alliances possessed. Thus, the variable takes a value of “1” if and when a country within a dyad participates in one or more secret alliances that render the country more powerful than its adversary would anticipate if such alliances were randomly dispersed across countries. In contrast, the measure takes a value of “0” when members of a dyad participate in no secret alliances or, alternatively, are participants in a smaller or weaker set of secret alliances than their adversaries’ baseline beliefs would suggest. For the sake of simplicity, we refer to this measure as Secret Alliances throughout our analysis.

To account for the conditional relationship between Secret Alliances and shifts in the balance of power, we created a second variable, Power Shift, which characterizes the change in the two states’ relative military capabilities over a recent time period. To measure the states’ military capabilities, we used CINC scores (Singer, Bremer, and Stuckey 1972), a composite measure of a state’s population totals, industrial output, military personnel, and defense expenditure that, although crude, is widely used in the conflict literature as a proxy for military power during the time period of our data. Our Power Shift variable measures the change in the weaker state’s share of the two states’ combined CINC scores that occurred over the preceding three, five, or ten years. When evaluating our hypothesis, our explanatory variable of interest is the interaction between the Secret Alliances variable and the Power Shift variable. We predict that Secret Alliances should be associated with significantly higher likelihoods of conflict only when they are accompanied by large Power Shifts.

Our primary dependent variable is Militarized Interstate Dispute (MID) onset, a binary measure that assumes a value of “1” in observations where one dyad member threatens to use force, initiates a display of force, or actually uses force up to and including war. Limiting the outcome variable to a binary measure in which force is actually used yields similar results.

We include a series of control variables to account for confounding factors that are prominently associated with MIDs in existing literature. Because previous research suggests that

---

75 This measure is further offset by the possibility that an adversary is itself party to a secret alliance.
conflict may be more likely to occur between states with similar capabilities,\textsuperscript{77} we include a *Relative Capabilities* variable that measures the weaker state’s current share of total dyadic capabilities using CINC scores, with .5 representing power parity between the two sides. Geographically proximate states engage in conflict at higher rates than distant states,\textsuperscript{78} so we include a binary *Contiguity* variable that indicates whether the states in a dyad share a border or are separated by less than 400 miles of water.\textsuperscript{79} An extensive literature on the democratic peace suggests that conflict should occur less frequently in dyads where both states are democracies compared to dyads where one or both states are non-democracies.\textsuperscript{80} We therefore include a *Joint Democracy* variable that indicates whether the Polity IV scores of both states in a dyad exceed six.\textsuperscript{81} States that are themselves allied are less likely to initiate disputes with each other, so we also include an *Allied Dyad* indicator for whether the dyad members share a formal alliance.\textsuperscript{82} Finally, we include polynomial measures of the time elapsed since the dyad’s most recent MID: *Peace Year, Peace Years*\textsuperscript{2}, and *Peace Years*\textsuperscript{3} so as to account for temporal dependence in the data.\textsuperscript{83}

### 2.5.2 Results and Robustness

We estimate dichotomous outcome models using logistic regression with standard errors clustered by dyad. We find consistent support for our hypothesis across various model specifications. Table 2.2 presents results using several distinct duration periods for our *Power Shifts* variable. Model 1 measures the change in the weaker state’s relative capabilities over the preceding year, while successive models increase the duration of the power shift to a maximum of five years.\textsuperscript{84}

\textsuperscript{77}Reed 2003.


\textsuperscript{79}Because the dataset is restricted to politically-relevant dyads, the only non-contiguous states are major powers in observations where they are not adjacent or proximate to the other member of their dyad.


\textsuperscript{81}Marshall and Jaggers 2002

\textsuperscript{82}Gibler and Sarkees 2002.

\textsuperscript{83}Carter and Signorino 2010.

\textsuperscript{84}Although existing literature on preventive wars often focuses on what Powell (1999) referred to as “large and rapid” power shifts (emphasis added), other research suggests that states may feel threatened by shifts in power that occur more gradually or over longer time periods. See Krainin (2017) for a recent formalization of one such
Notice first that across all specifications that main effect of *Power Shifts* is either null or negatively associated with the occurrence of a MID at conventional significance levels, in line with our expectation that when states engage in visible power shifts—such as those related to substantial changes in CINC scores—they should do so at sufficiently conservative levels so as not to provoke preventive conflict. Likewise, across all specifications, the main effect of *Secret Alliances* is positively associated with MID occurrence. This result is consistent with Bas and Schub (2016), but is also consistent with our interpretation that opponents are able to target their suspicion against states that they believe are likely candidates for secret alliances.

Most importantly, notice that the existence of *Secret Alliances* within a dyad is associated with a *larger* increase in the probability of a MID when a power shift is also occurring. To facilitate interpretation of the interaction between *Power Shifts* and *Secret Alliances*, Figure 2.3 plots the predicted probability of a MID at different levels of each constitutive variable. When the weaker state within a dyad has experienced only a small power shift over the preceding five years, the existence of a secret alliance does not significantly increase the probability of an international dispute relative to the alternative case in which secret alliances do not exist. Once again, this is consistent with our prediction: even if an opponent suspects that a *Secret Alliance* may exist, this suspicion is insufficient to motivate the opponent to initiate preventive conflict because power is observably shifting between the two states only gradually. In contrast, on the right hand side of the figure, when observable swings in power are already relatively large, the added suspicion that *unobservable* shifts in power are also occurring via the potential presence of secret alliances is often enough to motivate opponents to attempt preventive conflict. Indeed, in this area of the figure we observe that secret alliances are associated with substantially increased probabilities of disputes relative to the baseline scenario.\(^{85}\)

---

85 With the exception of *Joint Democracy*, the direction and significance of our control variables are also consistent with our predictions across all specifications.
Table 2.2: Secret Alliances, Power Shifts, and MIDs. Models use logistic regression with non-directed politically relevant dyad years as the unit of analysis. Dyad-clustered standard errors are shown in parentheses. Results for higher-level Peace Years are not shown.

<table>
<thead>
<tr>
<th></th>
<th>(1) 1 year</th>
<th>(2) 2 years</th>
<th>(3) 3 years</th>
<th>(4) 4 years</th>
<th>(5) 5 years</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Secret Alliances × Power Shift</strong></td>
<td>18.64***</td>
<td>17.68***</td>
<td>15.36**</td>
<td>14.65***</td>
<td>13.91***</td>
</tr>
<tr>
<td></td>
<td>(5.537)</td>
<td>(5.032)</td>
<td>(5.010)</td>
<td>(4.276)</td>
<td>(4.197)</td>
</tr>
<tr>
<td><strong>Secret Alliance</strong></td>
<td>0.438**</td>
<td>0.353*</td>
<td>0.311*</td>
<td>0.219</td>
<td>0.187</td>
</tr>
<tr>
<td></td>
<td>(0.140)</td>
<td>(0.156)</td>
<td>(0.156)</td>
<td>(0.155)</td>
<td>(0.164)</td>
</tr>
<tr>
<td><strong>Power Shift</strong></td>
<td>-4.188</td>
<td>-4.900**</td>
<td>-3.005</td>
<td>-1.861</td>
<td>-1.876</td>
</tr>
<tr>
<td></td>
<td>(2.262)</td>
<td>(1.622)</td>
<td>(1.579)</td>
<td>(1.407)</td>
<td>(1.270)</td>
</tr>
<tr>
<td><strong>Relative Capabilities</strong></td>
<td>1.697**</td>
<td>1.797**</td>
<td>1.637**</td>
<td>1.476*</td>
<td>1.486*</td>
</tr>
<tr>
<td></td>
<td>(0.573)</td>
<td>(0.586)</td>
<td>(0.580)</td>
<td>(0.581)</td>
<td>(0.599)</td>
</tr>
<tr>
<td><strong>Contiguity</strong></td>
<td>0.793***</td>
<td>0.763***</td>
<td>0.739***</td>
<td>0.675***</td>
<td>0.684***</td>
</tr>
<tr>
<td></td>
<td>(0.182)</td>
<td>(0.182)</td>
<td>(0.181)</td>
<td>(0.181)</td>
<td>(0.182)</td>
</tr>
<tr>
<td><strong>Joint Democracy</strong></td>
<td>-0.0809</td>
<td>0.0148</td>
<td>0.106</td>
<td>0.134</td>
<td>0.227</td>
</tr>
<tr>
<td></td>
<td>(0.258)</td>
<td>(0.244)</td>
<td>(0.232)</td>
<td>(0.227)</td>
<td>(0.220)</td>
</tr>
<tr>
<td><strong>Allied Dyad</strong></td>
<td>-0.715***</td>
<td>-0.654**</td>
<td>-0.627**</td>
<td>-0.606**</td>
<td>-0.600**</td>
</tr>
<tr>
<td></td>
<td>(0.217)</td>
<td>(0.214)</td>
<td>(0.214)</td>
<td>(0.214)</td>
<td>(0.220)</td>
</tr>
<tr>
<td><strong>Peace Years</strong></td>
<td>-0.154***</td>
<td>-0.161***</td>
<td>-0.172**</td>
<td>-0.178***</td>
<td>-0.181***</td>
</tr>
<tr>
<td></td>
<td>(0.0155)</td>
<td>(0.0163)</td>
<td>(0.0171)</td>
<td>(0.0174)</td>
<td>(0.0174)</td>
</tr>
<tr>
<td><strong>Constant</strong></td>
<td>-3.117***</td>
<td>-3.039***</td>
<td>-2.902***</td>
<td>-2.757***</td>
<td>-2.731***</td>
</tr>
<tr>
<td></td>
<td>(0.164)</td>
<td>(0.178)</td>
<td>(0.185)</td>
<td>(0.188)</td>
<td>(0.195)</td>
</tr>
<tr>
<td><strong>Observations</strong></td>
<td>24056</td>
<td>23318</td>
<td>22666</td>
<td>22044</td>
<td>21455</td>
</tr>
</tbody>
</table>

*Notes: * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$
To further support our theory, we present several additional models designed to address alternative theories that are otherwise consistent with our initial empirical results. Table 2.3 displays the results of these alternative specifications. The first possibility we consider relates to strategic selection. Perhaps states form secret alliances when they anticipate imminent attack. Although announcing these alliances might reduce the risk of conflict by deterring potential adversaries, states may prefer to keep the alliances secret in order to reap tactical advantages once war begins in earnest, as we alluded to earlier in our discussion of Slantchev (2010) and Lindsey (2015). If this were the case and if power shifts directly increase the probability of conflict, then the creation of secret alliances in the prelude to such conflicts might explain the relationship we observe. To address this possibility, Model (6) drops all observations in which at least one state signed a secret alliance that year or in the previous two years. The results remain consistent.

\[\text{Note that neither of these theories were designed with secret alliances in mind, though Slantchev (2010) acknowledges their potential applicability. As such, neither theory provides a clear explanation for why alliances formed immediately in the prelude to conflict would remain credible. If attack is certain, a third party who forms an alliance with a targeted state will be forced to suffer the costs of conflict if it honors the alliance, whereas it could conceivably avoid these costs by remaining neutral.}\]
with our hypothesis. Because the secret alliances that remain in this sample were forged at least three years prior to the initiation of a dispute, it is harder to believe that they were crafted by allies who were wary of imminent attack.

A second possibility is that rising states form secret alliances in hopes of carrying out attacks against their adversaries. \(^{87}\) If such allies consistently delay their intended attacks while their capabilities mature, this behavior may produce the pattern we observe in the results—either in place of or alongside our War of Suspicion mechanism. Model (7) therefore limits our measure of Secret Alliances to include only those agreements that are purely defensive in nature. Despite this restriction, the results remain consistent with our predictions. Although not shown in the table, our results are also consistent when we combine the restrictions of models (6) and (7), thereby limiting our observations to cases in which new alliances were not formed in the three years preceding a conflict and limiting our Secret Alliances variable to measure purely defensive alliances. That we find support for our hypothesized relationship even among relatively lengthy defensive alliances increases our skepticism of the mutual optimism explanation for such conflicts. If states did not form defensive alliances out of concern of imminent attack by an adversary, it is curious why they would make no effort to reveal their alliance once conflict appeared likely. Revealing such an agreement should lower the opponent’s expectation of victory or raising its anticipated costs for fighting, thereby reducing the opponent’s willingness to fight and allowing the allies to avoid war. If wars nevertheless occur in these circumstances, states must harbor and incentive not to reveal their alliances even when conflict appears likely. Our theory helps to explain this puzzle: revealing an alliance might intensify rather than reduce the opponent’s incentive to attack by revealing definitively that two states that previously appeared independent are not only allied but on a rapid development trajectory. In these circumstances, allies would prefer to keep the alliance secret and risk a war of suspicion rather than reveal the alliance and abandon the possibility of peace entirely.

\(^{87}\) For a discussion of secret offensive alliances, see Ritter (2003).
Table 2.3: Robustness and Alternative Specifications. Models use logistic regression with non-directed politically relevant dyad years as the unit of analysis and dyad-clustered standard errors in parentheses. Results for higher-level Peace Years are not shown.

<table>
<thead>
<tr>
<th></th>
<th>(6) Non-Recent</th>
<th>(7) Defensive</th>
<th>(8) Actual Force</th>
<th>(9) Continuous</th>
</tr>
</thead>
<tbody>
<tr>
<td>Secret Alliances × Power Shift</td>
<td>13.96*</td>
<td>10.41*</td>
<td>10.92*</td>
<td>6.752*</td>
</tr>
<tr>
<td></td>
<td>(5.977)</td>
<td>(5.215)</td>
<td>(5.494)</td>
<td>(3.316)</td>
</tr>
<tr>
<td>Secret Alliance</td>
<td>-0.0677</td>
<td>0.153</td>
<td>0.192</td>
<td>0.338</td>
</tr>
<tr>
<td></td>
<td>(0.376)</td>
<td>(0.167)</td>
<td>(0.217)</td>
<td>(0.207)</td>
</tr>
<tr>
<td>Power Shift (5 years)</td>
<td>-1.088</td>
<td>-0.784</td>
<td>-2.331</td>
<td>-4.923*</td>
</tr>
<tr>
<td></td>
<td>(1.187)</td>
<td>(1.216)</td>
<td>(1.664)</td>
<td>(2.192)</td>
</tr>
<tr>
<td>Relative Capabilities</td>
<td>1.453</td>
<td>1.695**</td>
<td>1.113</td>
<td>1.782**</td>
</tr>
<tr>
<td></td>
<td>(0.869)</td>
<td>(0.560)</td>
<td>(0.696)</td>
<td>(0.564)</td>
</tr>
<tr>
<td>Contiguity</td>
<td>0.655*</td>
<td>0.665***</td>
<td>0.726***</td>
<td>0.725***</td>
</tr>
<tr>
<td></td>
<td>(0.258)</td>
<td>(0.180)</td>
<td>(0.205)</td>
<td>(0.189)</td>
</tr>
<tr>
<td>Joint Democracy</td>
<td>0.462</td>
<td>0.268</td>
<td>0.232</td>
<td>0.248</td>
</tr>
<tr>
<td></td>
<td>(0.534)</td>
<td>(0.221)</td>
<td>(0.248)</td>
<td>(0.225)</td>
</tr>
<tr>
<td>Allied Dyad</td>
<td>-0.894**</td>
<td>-0.616**</td>
<td>-0.844**</td>
<td>-0.593**</td>
</tr>
<tr>
<td></td>
<td>(0.335)</td>
<td>(0.223)</td>
<td>(0.262)</td>
<td>(0.221)</td>
</tr>
<tr>
<td>Peace Years</td>
<td>-0.203***</td>
<td>-0.181***</td>
<td>-0.234***</td>
<td>-0.180***</td>
</tr>
<tr>
<td></td>
<td>(0.0045)</td>
<td>(0.0174)</td>
<td>(0.0221)</td>
<td>(0.0174)</td>
</tr>
<tr>
<td>Constant</td>
<td>-2.710***</td>
<td>-2.732***</td>
<td>-2.885***</td>
<td>-3.047***</td>
</tr>
<tr>
<td></td>
<td>(0.274)</td>
<td>(0.194)</td>
<td>(0.221)</td>
<td>(0.261)</td>
</tr>
<tr>
<td>Observations</td>
<td>13643</td>
<td>21455</td>
<td>21455</td>
<td>21455</td>
</tr>
</tbody>
</table>

*p < 0.05, **p < 0.01, ***p < 0.001
Our last two checks relate to the coding of our outcome and treatment variables. One concern is that our theory predicts preventive military action rather than merely the advent of a military dispute. As such, model (8) adopts an alternative dichotomous outcome variable in that assumes a value of “1” only if a dispute escalates to the use of military force. With this setup, the predicted relationship remains consistent with our hypothesis. The final alternative specification relates to the binary nature of our Secret Alliance variable. Although Bas and Schub use the same dichotomous measure in their article, they also construct a continuous variable that indicates the degree to which the secret alliances present in a dyad are collectively stronger (values >1) or weaker (values <1) than opponents would assume if such alliances existed only at baseline levels. In model (9) we substitute this continuous measure of alliance strength for our original dichotomous measure and find that the results continue to align with our predictions.

Overall the relationship between Secret Alliances, Power Shifts, and international conflict remains strong and consistent across a wide range of model specifications that address alternative theoretical explanations. Although secret alliances are far from the only example of clandestine military activities that could yield significant power shifts and provoke Wars of Suspicion by concerned adversaries, we believe these results provide plausible support for our theory as well as preliminary evidence to support the empirical commonality of the mechanisms we outline.

2.6 Conclusion

This article presents several major claims. First, the ability to announce or conceal military development is an important strategic tool. Whereas previous research suggests that signals of strength enable states to access deterrent and bargaining benefits, we show that such displays can also provoke conflict. As a result, countries have incentives to hide their capabilities even when signaling is credible and costless. Similarly, states should not always pursue the largest long-term improvements in military capability. To minimize the risk of preventive war, countries
may rationally avoid forming alliances with strong partners or developing potent military tech-
nologies—even if such options are free to implement. Researchers should revisit theoretical and
empirical research that assumes increases in military power are either universally desirable or are
consistently associated with favorable conflict and bargaining outcomes.

Our second finding is that the threat of clandestine development can itself provoke
war and create significant distributional consequences. Whereas previous theoretical research
on preventive conflict focuses largely on complete-information environments where one actor
becomes aware of an imminent power shift by the opposing side, we demonstrate that the mere
suspicion of such a shift can compel an enemy to fight. Likewise, all actors under suspicion
of development must pay concessions to their opponents in order to achieve peace, even if
development has not actually occurred and will not occur in the future. Thus, the possibility that
some actors are engaged in secret behavior creates negative externalities for those who do not
participate in secret development.

Critics may complain that our stylized description of military arming abstracts too far from
the complexity of the international environment. Indeed, we acknowledge that the model does not
incorporate several realistic nuances that may yield interesting behaviors. Future researchers may
investigate how the strategic behavior may change when opponents can invest resources in either
intelligence or, alternatively, counter-arming. Identifying how states balance these options could
prove a worthwhile extension. Likewise, our model considers a strategic interaction between only
two players, whereas in reality military technologies are fungible and countries often confront
several adversaries simultaneously. How countries might behave when arming promises to deter
aggression from one opponent but risks inciting aggressive action from another opponent is an
important and intriguing question for future research. Finally, our model provides actors with a
decision to pursue or avoid a development whose scale is exogenously determined. This reflects
the fact that many power shifts are beyond states’ ability to select and that the range of available
options if often heavily constricted. Nonetheless, future work should investigate how states behave when they can endogenously determine the scale and speed of their developments. Though our model abstracts away from these and many other potential extensions, the cases suggest that it usefully describes a series of important historical episodes and, perhaps more importantly, demonstrates that behavior as seemingly diverse as civilian interactions with counterinsurgents, German and Soviet cooperation in the interwar period, and firms’ decisions to file patents are well explained within a consistent theoretical framework.

Finally, the results should inform our thinking about the relationship between information, secrecy, and war. Dominant theories of international conflict suggest that war is most likely when actors are unable to share information, but our model shows that in some circumstances strategic concealment allows states to maintain peace. Thus, whereas Blainey (1988, p. 56) lamented that war must sometimes “provide the stinging ice of reality” that eliminates optimism and enables states to settle, we show that an alternative is also true. When states are initially uncertain about the presence of an upcoming power shift, the “stinging ice of reality” can alert them to a potential threat and thereby itself provoke war.

All chapters of this dissertation are being prepared for submission for publication of the material. The dissertation author was the primary investigator and author of this material.

Chapter 2 is coauthored with Abrahams, Alexei S. The dissertation author was the primary author of this chapter and has permission to use the material contained herein.

---

88 See Krainin (2017) and Schub (2017) for examples.
3 Monopolies of Violence: Civil Insurgency and the Accountability Deficit

Abstract: Researchers of civil conflict and state-building often argue that governments should possess a “monopoly on violence” with which to enforce laws, facilitate commerce, and provide security. Monopolists, however, are notorious for predatory behavior. How can citizens simultaneously empower a sovereign to maintain order while also deterring that ruler from engaging in abuse? Using analytic theory and evidence from contemporary crises, this paper demonstrates that optimal governance can emerge when rulers are held accountable by the latent threat of insurgent violence. Although the presence of armed insurgents may provoke civil conflict, the threat of unrest can also inhibit government predation, encourage the development of institutional safeguards, and motivate rulers to extend generous political and economic guarantees to citizens. The results challenge popular conclusions regarding the benefits of democratic rule, the desirability of state consolidation, the causes of civil violence, and the optimal use of foreign intervention in civil conflicts.
3.1 Introduction

During the spring of 1989, Chinese citizens participated in anti-corruption demonstrations and pro-democracy marches in hundreds of cities across the country. In Beijing alone, more than a million protesters gathered at Tiananmen Square to demand political reforms that would lessen state authoritarianism and improve accountability. After weeks of demonstrations, the government responded. Rather than compromise, however, Party officials instead imposed martial law, denounced the demonstrators as rebels seeking to overthrow the state, and ordered the military to suppress additional protests with force.¹ Thirty years later, images of the events at Tiananmen are among the most recognizable of the 20th century, but the political and social reforms the protestors sought have yet to materialize. Far from expanding political freedoms, the Chinese government has increased press censorship, restricted academic freedom, detained political dissidents, encroached upon the rights of social minorities, and suppressed elements of civil society. Party reformers have given way to conservatives, and official figures suggest that government spending on domestic “stability maintenance” (weiwen) now outstrips funding directed toward national defense.²

By several measures, the lack of Chinese reform in the face of such events is surprising. Social scientists have long argued that economic growth, international exposure, and external pressure should contribute to political liberalization.³ Moreover, a growing literature argues that widespread social demonstrations are capable of affecting reform and inducing democratization even within oppressive and authoritarian states.⁴ China, however, is hardly unique in the degree to which empirical outcomes contrast with prevailing theories. From Palestinians advocating

¹Although casualty counts related to the operation remain in dispute, estimates of the civilian death toll range from nearly 300 to more than 10,000, with thousands of others arrested and detained.
²See Wang 2013 for additional discussion, but note that official figures may be unreliable.
³China is a commodity importer, unlike other regimes whose resistance to democratization and indifference to public opinion may reflect a “resource curse” (Auty 1994, Coxhead 2007, Ross 1999, DeMeritt and Young 2013, Ross 2015).
for human rights reforms to women protesting against discriminatory government policies or Black Lives Matter demonstrators seeking credible constraints on state behavior, civilian groups throughout the world continue to harbor concerns about undue government power, sometimes even in societies where both the franchise and the right to participate in nonviolent protests are enshrined by law. As millions of citizens once again flood the streets—this time in Hong Kong—demanding protections for their political rights, researchers and policymakers alike have asked, “under what conditions can civil demonstrations succeed in inducing state concessions? Why do political protests that appear to enjoy widespread domestic or international support often fail to achieve their goals?” Most broadly, “how can citizens secure economic and political rights from powerful regimes?”

The protest scenarios we outlined above each fall within a consistent framework. In all cases, citizens disagree with their government about the appropriate balance of two desirable goods: state power and accountability. On the one hand, citizens acknowledge that strong governments help them avoid the brutalities of life under anarchy. Because life without government is often “nasty, brutish, and short,” citizens seek to establish a sovereign whose “monopoly on violence” enables it to enforce rules, facilitate exchange, and provide security. Unfortunately, however, such sovereigns are neither universally just nor perfectly benevolent. With the power to enforce the law comes the temptation to live above it. Leaders who acquire a monopoly on authority may therefore act with little regard for those they rule, implementing policies that privilege preferred groups at the expense of others. The fundamental difficulty for vulnerable citizens, then, is that they must “must first enable the government to control the governed; and in the next place oblige it to control itself.”

We present a solution to this dilemma that challenges several popular conclusions regarding the importance of state consolidation, the desirability of democratic rule, the causes of civil violence, and the effects of foreign intervention in domestic conflicts. In short, we argue

---

5Hobbes 1651; Weber 1919.
6Madison 1788.
that citizens can resolve the governance dilemma and improve state accountability by leveraging the credible threat of insurgent violence. According to our theory, citizens, insurgents, and governments are mutual participants in a “governance market.” As consumers of governance, citizens seek from the state a package of rights, liberties, assets, and opportunities. Because these goods are costly to provide, a monopolist government will tend to under-supply them, resulting in restrictions on the political and economic freedoms that citizens desire. Citizens, who are relatively weak and poorly organized, are ill-equipped to secure more generous governance through either mass demonstrations or attempts at organized violence. Embedded among the citizens, however, are political activists and entrepreneurs of violence (“insurgents”) who enjoy a special talent for social organizing and militancy. Whether they genuinely care for citizens’ well-being or simply pander to citizens’ interests in the service of ulterior motives, these insurgents can present a credible alternative to the incumbent regime. As the government continues to disappoint citizens, the weight of public support shifts to the insurgents, emboldening them to launch a militant challenge against the state.

When the market is healthy, citizens function as influential consumers who can shift their support between the incumbent and the potential insurgents depending on the relative package of rewards that each side can credibly offer. Even when insurgents remain latent rather than active market participants, their threat of competitive entry applies pressure on the incumbent, which worries that civilians may abandon the establishment and instead support insurgent entry. To avoid this outcome, the government may attempt to head off the insurgents by developing and delivering more generous political and economic institutions than the insurgents could match. The state’s promise of reform is made credible by the possibility that citizens could shift their allegiance if the government reneges, thereby enabling the insurgents to engage in violent attacks. In sum, citizens can leverage the existence of latent insurgents to extract economic and political concessions from recalcitrant regimes—even when the insurgents themselves are predatory and self-interested.
Our framework contrasts with theories that portray government accountability as dependent on a transition to democratic rule, the existence of civil society or mass social movements, the altruistic or sociotropic preferences of influential actors, the economic self-interest of a stationary ruler, the development of social norms, or reliable external enforcement—alternatives whose strengths and shortcomings we discuss in the following section. Instead, we argue that regimes may also be constrained from predation by the threat of organized, violent dissent. Put simply, the latent threat of insurgent violence can dissuade rulers from rolling back existing concessions or attempting authoritarian reversals. Moreover, when insurgents successfully use coercive threats to elicit preemptive government reforms, the violence they threaten need never occur.

Our findings challenge the widespread belief that democracy is either a necessary or sufficient condition for accountable governance. We acknowledge that the tradeoff between power and accountability is most overt in states where institutional safeguards and robust political protections either have decayed or are yet to emerge. Nevertheless, we argue that the tradeoff persists even within well-institutionalized democratic settings. Whereas researchers sometimes argue that democracy fosters the emergence of institutional and social constraints against tyranny—ranging from constitutional checks and balances to peaceful protest groups and nonviolent social movements—we propose that even mature democratic regimes are ultimately constrained from predation by the threat that organized dissenters will impose costs to which incumbent regimes are highly sensitive.

The effectiveness of organized threats in inducing state reforms remains under-appreciated by scholars and policymakers who overlook a fundamental truth of coercion: violence is most productive when it remains unobserved. Implicit threats of insurgent violence may succeed in motivating state reforms without becoming apparent to researchers; indeed, such threats may

---

7In countries throughout the Middle East and Africa, the scale of human suffering testifies to the horrors associated with regimes that are armed with either too little coercive power or too much. In parts of Libya, Syria, Iraq, and the Democratic Republic of Congo, the collapse or fragility of the central government has left citizens vulnerable to predation from violent opportunists. Elsewhere, in places like Egypt and Bahrain, governments monopolize violence so comprehensively that they may themselves prey upon citizens with impunity.
exist in conjunction with nonviolent tactics to which researchers instead assign primary credit. Only when these efforts fail to induce reforms do insurgents feel pressure to militarize in visible ways, leading outsiders to observe violence precisely in situations where governments are most resistant to compromise. As a result, those who assess “success” and “failure” purely based on observable outcomes may be apt to over-attribute government concessions to peace movements and to under-acknowledge the role of underlying coercive threats in motivating reform. Moreover, government actors can often outmatch and out-escalate softer forms of protest, such as strikes and boycotts. Denying citizens the ultimate coercive option of supporting an organized insurgent movement may in some cases leave civilians vulnerable to systemic abuse and state predation.

Although our model suggests that across a wide range of conditions the threat of insurgent violence is itself sufficient to induce government reforms and improve civilian welfare, even a cursory reading of history suggests that insurgent movements often succumb to state prevention, may provoke state repression, or may escalate to full-blown civil violence. In the latter half of this paper, we turn to the various obstacles and “market failures” that can disrupt the process of peaceful reform. Civil violence can arise through a series of familiar channels, including repressive or preventive action by the incumbent; attempts by either the insurgent group or the state to signal strength and resolve; and undue optimism on behalf of the government, insurgents, or even the civilians themselves. In other cases, violence emerges through mechanisms that are less well-known and that stem from the interactions of several non-unitary actors, including the behavior of heterogeneous civilian populations or multiple competing insurgent groups. Because all of these mechanisms emerge from a single cohesive framework, the theory offers an adaptable, powerful, and parsimonious platform for additional research on the causes of civil violence.

Finally, whereas the baseline model treats domestic governance and civil conflict as

---

8For example, our argument suggests that although black protestors who eschewed violence throughout the U.S. civil rights movement deserve high esteem and respect for their efforts, they may nevertheless owe a portion of their success to the existence of their more radical peers who, in the words of Malcolm X, were “not handcuffed by the disarming philosophy of nonviolence.” See recent work relating to the Black Freedom Movement, including Hill 2004, Umoja 2013, Cobb 2014, and Levy 2018.
existing within an autarkic environment, the last section of the paper illustrates how the theory yields valuable predictions for the study of foreign intervention in civil conflicts. Although foreign aid and intervention may quell civil disputes when properly calibrated, even efforts that seek to cultivate stability while minimizing collateral damage can instead provoke the recidivism of violence or abuse when applied incorrectly. In particular, foreign intervention that disarms an insurgent group and relieves an incumbent of competitive pressure may harm citizens long-term by reducing the ruler’s incentive to reform. International actors who seek to improve citizens’ quality of life may therefore best achieve their goals not by eliminating insurgent actors or toppling despotic regimes but instead by ensuring that incumbent governments face latent but powerful coercive competitors.

3.2 Constraining the State

Writing in the midst of the English Civil War, Thomas Hobbes famously argued that life without government is “solitary, poor, nasty, brutish, and short.”⁹ Nothing prevents the strong from plundering the weak, the unscrupulous from preying upon the unwitting, or the dishonest from defrauding the less discerning. Anarchic conditions also suppress economic exchange. Without recourse to external enforcement, individuals face difficulty establishing durable contracts over goods and services. Governments, according to Hobbes, emerge as a natural response to these circumstances. Rather than persist under anarchy, individuals instead submit themselves to the rule of a single sovereign authority they hope will provide order, stability, and security. Even political economists who doubt this narrative of state emergence nevertheless endorse it as a justification for state consolidation, arguing that the centralization and monopolization of force are the optimal means of organizing violence.¹⁰ As North (1981, p. 24) summarized, “throughout history, individuals, given a choice between a state... and anarchy, have decided for the former.”

⁹Hobbes 1651.
¹⁰Weber 1919; Olson 1993; North, Wallis, and Weingast 2009.
Unfortunately, the benefits of state authority are neither automatic nor risk-free. Governments are inherently coercive: by design they are empowered to use violence or the threat thereof to motivate changes in civilian behavior. Nearly all government actions—even those that provide overall social benefits—make at least some individuals worse off. Moreover, states empowered to use coercion for the collective good may instead exert their authority for private gain. Throughout the world powerful rulers routinely select policies to enrich themselves or their supporters and to predare upon subjects they oppose. In 1762, Rousseau noted that the inherent difficulty in building a society is “to find a form of association which will defend and protect... the goods of each associate,” but within which each citizen, “while uniting himself with all, may still... remain as free as before.”

Classical political theorists and contemporary political economists have advanced several potential solutions for Rousseau’s puzzle of how societies might establish order while preserving liberty. At the most formative level, concerns about state predation dissipate if we assume that that governments are either inherently benevolent or, alternatively, are perfectly responsive to altruistic or other-regarding constituents. Plato, for example, proposed that “philosopher kings” might serve as benevolent dictators who ruled with the interests of society in mind. To the extent that these leaders strayed from justice, they could be induced toward good behavior by “guardians” who were themselves indoctrinated to prioritize social interests over personal ambition. In effect, Plato’s proposal merely kicks the can down the road: an optimal society is ruled by elites who harbor sociotropic interests and whose negative impulses are in turn checked by other altruistic citizens, ad infinitum. Where these sociotropic impulses originate and through what means they become widely held throughout the ruling class remain open questions.

11 Governments may in principle facilitate coordination without exercising coercion, but this role does not justify the formalized, hierarchical, and monopolistic states to which we and other political scientists devote our attention.
12 This argument that political leaders can enact benevolent policies when properly educated and guided by mandarin advisors is also commonly invoked in Chinese political thought. See Yao and Yao 2000, De Bary and Lufrano 2001, Xuetong 2013, Qin 2016.
13 Spinoza (1675) lamented that “no men are less fit to govern than theorists or philosophers” who “sing the praises of a human nature nowhere to be found... and conceive men not as they are, but as they would like them to be.” See, however, Bowles and Gintis (2011) for an alternative account.
An alternative proposal is that national regimes might be induced to behave by a supranational authority or world government that intervenes to safeguard the interests of citizens. Once again, this mechanism leaves unspecified from whence such noble motivations might arise. Why would a world sovereign, once empowered, use its influence for liberal rather than authoritarian ends? As a practical matter international relations researchers and policymakers acknowledge the inherent difficulties of establishing durable and efficient supranational bodies that challenge state sovereignty.\(^{14}\) Attempts to develop and implement such governments have consistently faced challenges—even the strongest and most successful examples, such as the European Union, remain limited in policy scope, face jurisdictional conflicts with national authorities, and are ill-equipped to prevent governments or empowered majority groups from acting against minority interests.

Researchers who study the intersection of political authority and economic development suggest a third possibility. According to this view, citizens function as a renewable resource from whom rulers extract taxes and rents. Just as the assignment of property rights can motivate individuals to avoid over-exploitation of natural resources,\(^{15}\) powerful actors who consolidate political control over a region acquire incentives to eschew high levels of taxation that would deplete the economic productivity of the population they govern.\(^{16}\) In this respect, the relationship between these sovereigns and their subjects is “similar to the relation between a parasite and its host.”\(^{17}\) Indeed, in the parlance of the literature, such governments are known as ‘stationary bandits’ who continually extract the surpluses of production from those who inhabit their lands.

Although some authors assert that these ‘bandits’ will develop legal or even constitutional protections that limit their capacity to extract,\(^{18}\) these proposals suffer from several limitations. First, only minimal protections for property rights are necessary to facilitate economic growth.

---


\(^{15}\)See, for example, Alchian and Demsetz 1973 and Hardin 1968.

\(^{16}\)Olson 1993.

\(^{17}\)Wagner 2010, p. 119.

\(^{18}\)See, for example, North and Weingast 1989 and Root 1989.
A state concerned with fostering economic development need neither engage in generous social spending nor extend significant political liberties to citizens. As such, stationary bandit explanations leave unanswered the question of how citizens secure freedoms and opportunities beyond those that are minimally conducive to growth. Second, although researchers sometimes advance the claim that “constitution[s] and other political institutions…place restrictions on the state,” this faith in institutionalization often leaves unspecified the mechanism that transforms such documents from sheafs of paper into credible guarantees. Governments routinely engage in reversion and backsliding, even over commitments as ostensibly sacrosanct as the power of the purse.

Perhaps the most popular explanation for how citizens can restrain the state’s predatory impulses relates to the role of democratic and social movements. Acemoglu and Robinson (2005) advance the theory that elites voluntarily extend the franchise to avoid revolution and that citizens subsequently use their political influence to obtain additional economic concessions from the state. This account faces its own set of challenges. First, despite popular discussion in the literature, mass “peasant revolts” rarely succeed in toppling incumbent regimes. Only the most optimistic citizens could expect to gain as much by participating in a revolt as they stand to lose by doing so. As such, few social protests involve the mass mobilization of all members of society. Instead, successful or credible challenges to state authority are often channeled through political insurgents and entrepreneurial networks who organize small groups of motivated participants. Second, Acemoglu and Robinson assume that both revolutions and enfranchisement trigger a transfer of power in which a unified citizenry gains the capacity to

---

19 Even the British slave trade was long justified on economic grounds. See Eltis 1987.
20 North and Weingast 1989, p. 805.
21 See Cox 2012. Indeed, as Ziblatt (2006, p. 312) summarizes, “the central empirical puzzle” that occupied democratic theorists in the mid-twentieth century was to explain “the democratic reversals of the interwar period… given that democracy had appeared so secure in the world’s most advanced economies.”
22 Acemoglu and Robinson assume the costs of participating in rebellion are independent from the ultimate success or failure of the movement, i.e., that failed revolutionaries face no penalty for their attempt.
23 The American Revolution was instigated more directly by the actions of the Sons of Liberty and the pamphleteering of Thomas Paine and James Otis Jr. than by mass support for rebellion; likewise, once the war began fewer than 15% of colonial residents participated in the conflict.
set policy directly. This characterization of “democratic victory” in the aftermath of revolution appears dubious when contrasted with historical realities. In particular, it overlooks the possibility that the insurgents who participate in rebellion may, upon obtaining power, fail to fulfill their promises to former supporters.24 A richer and more realistic theory would acknowledge the possibility of post-war disagreement or conflict between ordinary citizens and the leaders of the revolutionary movement.25

Finally, even if democracy is established, the system of government remains vulnerable to various familiar criticisms. Foremost among these is the peril of factionalized or oppressive majority rule. As Schmitter and Karl (1991, p. 78) summarize, “competition among factions is a necessary evil in democracies… [yet democrats] tend to disagree about the best forms and rules for governing factional competition.” Likewise, Weingast (1997, p. 246) concedes that “Democratic stability requires that citizens agree on the limits on the state that they are to defend. Such agreement is neither natural nor automatic.” Most broadly, concerns about adequate representation, participation, and authoritarian reversion are familiar to scholars of democratic politics.26 Collectively, these challenges raise serious questions about whether the institutions of democracy provide adequate guarantees for citizens who desire generous political rights and durable protections from government abuse. In the next section, we introduce an alternative mechanism through which citizens can obtain such guarantees.

24Victorious revolutionaries routinely neglect to establish democracy, and the successful overthrow of an established regime may not portend peace or stability. Thousands were massacred in the aftermath of the Haitian Revolution—a popularly-supported rebellion that nevertheless yielded an authoritarian empire rather than a democratic state. Even when “revolutions” succeed and erect democratic institutions, victorious states often face a series of subsequent challenges from unsatisfied domestic groups; in the United States these manifested in a series of revolutions, ranging from Shay’s Rebellion and the Whiskey Rebellion through the American Civil War.

25Under the theory advanced by Acemoglu and Robinson, one is left to wonder why former American slaves declined to grant themselves robust civil rights after gaining the franchise in the 1860s. The explanation most consistent with the theory is that these individuals, now empowered to set policy, simply desired no further concessions or protections.

3.3 Model

We now present a model of trilateral conflict between citizens, insurgents, and an incumbent government. Game-theoretic models, like fables, deliberately abstract from reality, clearing away true but distracting phenomena in order to highlight an essential tradeoff. In our case, we seek to identify how a potential insurgency can inhibit or assist citizens in obtaining governance concessions from a ruling regime, as well as conditions in which this process may provoke violence. The baseline, complete-information game therefore illustrates an environment in which actors could induce civil war but manage to avoid violent outcomes in equilibrium. Instead, in the main equilibrium of interest, the government is reluctant to share the perquisites of control with an insurgency and, as a result, preemptively commits itself to generous public governance so as to dissuade citizens from supporting the would-be challenger. The insurgency realizes that it lacks sufficient support to extort concessions from the government and therefore chooses not to enter the political arena. This ‘no entry’ equilibrium therefore exhibits precisely the welfare-enhancing dynamic that we believe citizens can harness in optimal circumstances: keeping the government accountable via the threat of insurgency without needing to follow through on that threat.

Following the tradition of those who use formal theory to analyze political violence, we then discuss how minor deviations from our baseline setup—including information asymmetries, commitment deficits, and frictions between citizen principals and their insurgent agents—create equilibria in which civil violence may occur instead of reform. As a paradigmatic example, the government may fail to commit itself to sufficiently good governance, either because it underestimates the level of governance the insurgents can credibly offer or because it overestimates the insurgents’ cost of entry. Insurgents subsequently enter the market, whereafter the government still has a chance to avert war by propitiating the insurgents with a sufficiently large concession. Civil war ensues if the government once again miscalculates or is overly dismissive of the

---

28 For example, the government may institute land reforms (Kapstein 2017), extend suffrage (Acemoglu and Robinson 2005), or subsidize the costs of housing and food (Bates 1981).
insurgency’s coercive capacities.

We deliberately keep the baseline model simple so as to demonstrate that minor modifications generate many familiar phenomena documented by the civil war and COIN researchers. For example, we show the model can be easily adapted to include insurgent fragmentation, heterogeneous preferences among civilians, signaling violence, preventive violence, and so on. The model therefore serves as an easily customizable platform for future theoretical work on civil governance and intrastate conflict. Nevertheless, through all of these modifications, citizens continue to confront the same basic tradeoff between government accountability and domestic stability. Likewise, areas of the parameter space remain in which the expected benefits of accountability that result from insurgent threats outweigh the accompanying risks of destabilization. The core result of our baseline model is therefore robust to a broad range of plausible modifications.

3.3.1 Baseline Setup

Our baseline game tree is displayed below (see Table 3.1 for payoffs). Following recent theoretical work on ‘hearts and minds’ counterinsurgency (COIN), 29 we consider a hypothetical country with three domestic, strategic, unitary actors: citizens, an insurgency, and a government. In keeping with the crisis bargaining tradition, we suppose there exists a finite, rivalrous resource whose total value is normalized to 1, and which is continuously divisible on the interval \([0, 1]\). 30 At the start of the game, we assume the resource is controlled by the government.

Our first departure from the literature is to assume that the government and citizens are locked in a zero-sum conflict over the finite resource. This means that if the government provides citizens \(g_1 \in [0, 1]\), then it retains for itself only \(1 - g_1\). Although this setup is typical in the formal study of interstate conflict, it may seem unusual to those who analyze domestic civil-government relations. After all, citizens rely on the government for the supply of various public goods, most

29 See, for example, the model in Berman, Shapiro, and Felter 2011.
Figure 3.1: Baseline Model of Insurgency and Governance
notably protection from external predation. The government, in turn, depends on citizens as a tax base. This mutual dependency ought to imply complementary interests and a positive-sum game: the government benefits if the citizens prosper, since this prosperity increases tax revenues.\textsuperscript{31} As we know from the labor-firm bargaining literature, however, two actors can engage in a productive and mutually beneficial relationship but nevertheless find themselves locked in a zero-sum conflict over how to split the surplus of production. Governments throughout the world enable their citizens to operate as a tax base while still denying those citizens various human rights and civil liberties. It is over these additional rights, liberties, assets, and opportunities that governments and citizens engage in zero-sum conflict. For example, neither the members of Hong Kong’s anti-extradition movement nor the protestors involved in France’s Yellow Vest Movement face economic deprivation that forces them below subsistence levels. Rather, both groups seek additional political or economic rights from their respective states. Like laborers collectivizing around union leaders in hopes of securing higher wages and benefits, citizens may need to organize behind insurgencies to secure their goals.

In our second departure from many theoretical treatments of civil war, the insurgency begins the game as a \textit{latent} actor who has not yet decided to enter the ‘governance market.’ One can conceive of this insurgency as a potential political entrepreneur that citizens may know through the community’s associational life. It may for example be a religious group, a community association, a charity, a network of business affiliates, and so on.\textsuperscript{32} If the citizens provide sufficient support to the latent actor, it can evolve into a viable challenger to the incumbent regime by taking the form of a labor union, social movement, militant group, or rebellion. Even in its latent capacity, however, the potential insurgency achieves a certain degree of legitimacy among citizens.\textsuperscript{33} We translate this into material terms via the parameter $\nu$, which represents the degree to which the insurgents would treat citizens generously if the insurgency achieved a position of

\textsuperscript{31}Olson 1993, North, Wallis, and Weingast 2009.
\textsuperscript{32}Many successful militant groups began as charities and community associations and did not conceive of themselves as latent insurgents until much later in their development (Berman 2011).
\textsuperscript{33}Roy 2013, Cammett 2014.
political authority in the future. More precisely, the insurgents’ ties to the community before entering the political arena ensure that the citizens will obtain a fraction $\nu_i$ of any concessions won by the insurgents later on. The incumbent government, meanwhile, sets an analogous commitment parameter $\nu_g$.

After observing $\nu_g$ and $\nu_i$, the insurgency decides whether or not to enter the market and challenge the government. If the insurgents choose to stay out of politics, they obtain a payoff of 0, while the government gives the citizens $g_1 = \nu_g$ and retains $1 - \nu_g$ for itself. To the outside observer, this outcome generates no dramatic or remarkable phenomena: no civil war is fought, nor peace deals brokered. Indeed, the insurgency itself, by remaining latent, is never observed—an outcome that has significantly influenced the phenomenology and normative construction of insurgency by the international community, as we discuss in the following section.

Alternatively, the insurgency can enter the political arena at cost $c_f$. In this case, the government has a chance to avert civil war by offering insurgents a take-it-or-leave-it peace deal. Under the terms of the deal, the government offers $g_2$ to the insurgents and retains $1 - g_2$ for itself. The two actors then rule as a duopoly of violence, fulfilling their credible commitments to the citizens by offering them, respectively, fractions $\nu_g(1 - g_2)$ and $\nu_i g_2$ of the finite resource. Alternatively, civil war is imminent when insurgents reject the government’s offer. Citizens decide their degree of support for the insurgency, $\sigma \in [\sigma_{\text{min}}, \sigma_{\text{max}}]$. This influences the insurgency’s probability of victory in war, which we define as $p = p'(\sigma)$. Nature then tosses an unfair coin, awarding total victory to either the insurgency or the government. Regardless of the outcome, the insurgency pays a fighting cost $c_i$ and the government pays a fighting cost $c_g$. In the course of the war, a fraction $c$ of the finite resource is also destroyed. This can be thought of as destruction of property, loss of life, loss of tourism revenues, capital flight, and so on. Of the remaining $1 - c$, the victor shares with citizens the fraction $\nu_g$ or $\nu_i$ to which it is committed.

\[34\text{See Grossman (1995) for an analysis of how this might work.}\]
Table 3.1: Model Payoffs.

<table>
<thead>
<tr>
<th>Payoff #:</th>
<th>Outcome:</th>
<th>Government:</th>
</tr>
</thead>
<tbody>
<tr>
<td>O₁:</td>
<td>No Ins., Peace</td>
<td>1 − g₁</td>
</tr>
<tr>
<td>O₂:</td>
<td>Ins, Peace</td>
<td>(1 − g₂)(1 − v₉)</td>
</tr>
<tr>
<td>O₃:</td>
<td>Ins, Fight, Ins wins</td>
<td>−c₉</td>
</tr>
<tr>
<td>O₄:</td>
<td>Ins, Fight, Gov wins</td>
<td>(1 − c)(1 − g₃) − c₉</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Payoff #:</th>
<th>Outcome:</th>
<th>Insurgent:</th>
</tr>
</thead>
<tbody>
<tr>
<td>O₁:</td>
<td>No Ins., Peace</td>
<td>0</td>
</tr>
<tr>
<td>O₂:</td>
<td>Ins, Peace</td>
<td>(1 − vᵢ)g₂ − c₉</td>
</tr>
<tr>
<td>O₃:</td>
<td>Ins, Fight, Ins wins</td>
<td>(1 − c)(1 − i₂) − cᵢ − c₉</td>
</tr>
<tr>
<td>O₄:</td>
<td>Ins, Fight, Gov wins</td>
<td>−cᵢ − c₉</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Payoff #:</th>
<th>Outcome:</th>
<th>Citizens:</th>
</tr>
</thead>
<tbody>
<tr>
<td>O₁:</td>
<td>No Ins., Peace</td>
<td>g₁</td>
</tr>
<tr>
<td>O₂:</td>
<td>Ins, Peace</td>
<td>v₉(1 − g₂) + vᵢg₂</td>
</tr>
<tr>
<td>O₃:</td>
<td>Ins, Fight, Ins wins</td>
<td>(1 − c)i₂</td>
</tr>
<tr>
<td>O₄:</td>
<td>Ins, Fight, Gov wins</td>
<td>(1 − c)g₃</td>
</tr>
</tbody>
</table>

3.3.2 Baseline Solution

In the baseline game, finding a subgame-perfect equilibrium in which citizens benefit from the threat of insurgency is straightforward. To see this, note that if insurgents enter the market at cost c₉ and accept a peace offering of g₂ from the government, they share g₂vᵢ with the civilians while retaining for themselves a final payoff of (1 − vᵢ)g₂ − c₉. If instead the insurgents reject the peace deal and go to war, their expected payoff is p(1 − c)(1 − vᵢ) − cᵢ − c₉. Accordingly, the insurgents will accept the government’s peace offer whenever g₂ ≥ p(1 − c) − cᵢ. The government, if it desires peace, will therefore offer exactly g₂ = p(1 − c) − cᵢ to the insurgents, retaining [1 − p(1 − c)](1 − vᵢ)(1 − v₉) for itself. The government’s expected payoff from war, meanwhile, is (1 − p)(1 − c)(1 − v₉) − c₉. The government therefore always prefers peace over war with the insurgents, since [1 − p(1 − c) + cᵢ/(1 − vᵢ)](1 − v₉) > (1 − p)(1 − c)(1 − v₉) − c₉.

The civilians anticipate that insurgent entry will always lead to peace, with a payoff of [1 − p(1 − c) + cᵢ/(1 − vᵢ)]v₉ + [p(1 − c) − cᵢ/(1 − vᵢ)]vᵢ. Because the coefficients in the preceding expression
sum to one, the citizens’ peace payoff is a convex combination of the government’s and insurgents’
commitments: $(1 - \lambda_p)\nu_g + \lambda_p\nu_i$. If the insurgency does not enter the political arena, citizens
obtain a payoff of $g_1 = \nu_g$ from the government. It immediately follows that citizens are supportive
of the insurgency whenever $\nu_i > \nu_g$. In other words, citizens support insurgent entry precisely
when the insurgency is more committed towards citizens than is the incumbent government.

Knowing this, the insurgents can make an informed choice about whether or not to
challenge the government with entry. Whenever $\nu_i > \nu_g$, insurgents feel emboldened by the
prospect of citizens’ support, $\sigma = \sigma_{\text{max}}$. They enter the political arena, anticipating a peace deal
with payoff $(1 - \nu_i)(p_{\text{max}}(1 - c) - \frac{c_i}{1 - \nu_i}) - c_f$, where $p_{\text{max}} \equiv p'(\sigma_{\text{max}})$. On the other hand, when
$\nu_g \geq \nu_i$, the insurgents anticipate that citizens will not support them ($\sigma = \sigma_{\text{min}}$), and entry results
in payoff $(1 - \nu_i)(p_{\text{min}}(1 - c) - \frac{c_i}{1 - \nu_i}) - c_f$, where $p_{\text{min}} \equiv p'(\sigma_{\text{min}})$. Notice that if the insurgents’
payoff is positive in both of the above expressions, then the insurgents always decide to enter
(even if the citizens would prefer them not to do so). Similarly, if the insurgents’ payoff is negative
in both of the above expressions, then insurgents never enter (even if the citizens would prefer
insurgent entry). The interesting part of the parameter space, of course, is where the insurgency’s
payoff is positive when it enjoys the citizens’ support (when $p = p_{\text{max}}$), but negative when it does
not (when $p = p_{\text{min}}$). In this part of the parameter space, citizens have enough leverage over war
outcomes to incentivize or discourage insurgent entry, and so the citizens can give or withdraw
their support to manipulate the insurgency to enter or stay out of politics.

Finally, we can understand how the government chooses its commitment level $\nu_g$. If the
government chooses $\nu_g < \nu_i$, the citizens support insurgent entry and a peace deal is reached with a
payoff to the government of $[1 - p_{\text{max}}(1 - c) + \frac{c_i}{1 - \nu_i}](1 - \nu_g)$. Indeed, in this case, the government
should set $\nu_g = 0$ to maximize its peace payoff, obtaining $1 - p_{\text{max}}(1 - c) + \frac{c_i}{1 - \nu_i}$. Alternatively, if
the government chooses $\nu_g \geq \nu_i$, citizens dissuade the insurgency from entry and the government
obtains $1 - \nu_g$. In this case, the government should choose $\nu_g = \nu_i$ to maximize its payoff. The
government therefore prefers to deter insurgent entry whenever $1 - \nu_i > 1 - p_{\text{max}}(1 - c) + \frac{c_i}{1 - \nu_i}$,
which simplifies to $v_i < p_{\text{max}}(1 - c) - \frac{c}{\nu_i}$. Note that the righthand side of this expression is the offer $g_2$ derived earlier, which is positive in the relevant area of the parameter space. Furthermore, because marginal increases to $v_i$ cause an increase to the lefthand side and a decrease to the righthand side, the inequality is violated for sufficiently large $v_i$. There exists an inflection point $v_i^* \in (0, 1)$, therefore, such that the government raises its commitment to citizens and deters insurgent entry whenever $v_i \leq v_i^*$, while abandoning its commitment to citizens and inviting insurgent entry whenever $v_i > v_i^*$. We therefore discover two subgame-perfect equilibria: the no-entry equilibrium, where $v_i \leq v_i^*$; and the entry equilibrium, where $v_i > v_i^*$. This latter equilibrium represents an extreme in which the government is really a government only in name: it offers citizens nothing, and uses its military threat to extort and consume a significant portion of the country’s resources, before washing its hands of the responsibilities of governance.\footnote{We can interpret this equilibrium as an incumbent choosing to relinquish control of an area of territory to an insurgent group. For similar results, see both Spolaore 2014 and Acharya and Lee 2018.}

Because we are concerned with the continued provision of governance, our discussion focuses instead on the more typical no-entry equilibrium, where the insurgents offer a good outside governance option to citizens, but one that is not so generous that the government opts not to compete.

### 3.4 Discussion of the Baseline Model

In subsequent sections, we modify the game to allow for several realistic nuances, including government suppression of potential insurgencies, asymmetric information about expected war outcomes, heterogeneous citizens, fragmented insurgents, and external interveners. Before considering these extensions, however, we must first highlight several simple but under-appreciated insights about civil conflict, governance, and statebuilding that emerge from the baseline setup.
3.4.1 Why a “Violent Insurgency”?

The model does not suggest that coercive violence is the only means by which citizens can secure generous economic and political rights. For example, those who are fortunate enough to live under the benevolent leadership of altruistic rulers or the protective umbrella of reliable external enforcers can thrive even in the absence of a credible insurgent threat against the state. Unfortunately, however, these alternative conditions are difficult to achieve and are unlikely to prove reliable in the long term. Political leadership can change hands and international supporters may retreat as their own political will evaporates. Our model therefore seeks to investigate how citizens may hold a state accountable even when prevailing conditions prevent them from turning to either of the aforementioned alternatives. When citizens cannot count on leaders to act out of their own generosity or to respond to foreign pressure, coercive threats that induce state concessions from arise from elsewhere in the domestic environment.

Even domestic pressure, however, could arise through other means. Citizens that inhabit consolidated democracies may find their governments constrained by institutional checks and balances such as independent judiciaries, professional police and security services, and openly contested elections—along with the accumulated inertia of social norms and collective expectations that can render these institutions credible. Unfortunately, even well-respected social norms can decay; history abounds with situations in which governments reversed or rolled back institutional safeguards that citizens once thought inviolable. Although widespread social revolt or

---

36 See, for example, Weingast 1997. More broadly, the simplest means by which citizens can impose a cost on state officials is when credible democratic institutions provide citizens with voting rights in open elections. Unfortunately, these hallmarks of democracy do not themselves guarantee state responsiveness to minorities: popular majorities can oppose concessions to marginalized groups, and states can rescind the franchise when it threatens their interests.

37 In the United States alone, presidents Lincoln, Grant, and Franklin Roosevelt each suspended the writ of habeus corpus. Likewise, black enfranchisement was effaced by Jim Crow in the latter half of the 1800s. More recently, the current administration has rolled back civil protections for Muslims and immigrants, while state legislatures have expanded barriers to abortion. Even the American Revolution was itself in some ways a response to British encroachment on colonial home rule. In a broader context, the 2019 protests in Hong Kong occurred in a political context where over the preceding six months the state had expelled foreign journalists, banned an opposition political party and detained its members, jailed previous protestors, and threatened to erode non-extradition guarantees. Finally, executives have historically reversed political institutions even as fundamental as the legislative power of the purse (Cox 2012) and universal enfranchisement or openly contested elections, as has occurred in scenarios ranging
peaceful protests may thwart such reversals in specific circumstances, elsewhere the government may respond only to the additional coercive pressure made possibly by organized insurgency. More precisely, peaceful protests should only succeed when (1) incumbent governments are especially sensitive to the costs that unorganized demonstrators can impose through nonviolent means and (2) the state faces a low cost for offering concessions relative to the price of breaking a protest. In contrast, nonviolent demonstrations will be less effective when states are willing to either ignore such protests—because the demonstrators are ill-equipped to impose sufficient costs on the state—or when states can more easily engage in repressive actions that would quash the effectiveness of peaceful demonstrations.

Our model demonstrates that citizens may turn to insurgent groups to exert coercion against the state when softer forms of protest and political opposition are either inaccessible or under threat. Although efforts by citizens to conduct letter campaigns, public protests, and boycotts were observably associated with long-term success in, for instance, the American Civil Rights and Black Freedom movement, such efforts might have faced insurmountable challenges if they instead confronted a less conciliatory political regime. It is difficult to envision such movements meeting great success against governments akin to those of Germany in the late 1930s, modern China, various states throughout the contemporary Middle East, or other countries where rulers would face minimal costs when busting boycotts, imprisoning protestors, or censoring political speech. The latent threat of organized insurgency—though still difficult to wield, as we explain in the following section—provides the citizens with an alternative and potentially more fruitful coercive approach in these settings. Indeed, even the successes we associate with nonviolent and diffuse protest are often (1) channeled through a smaller and better-organized

from the establishment of The Four Hundred in ancient Athens to the democratic reversals that plagued the European inter-war period (Ziblatt 2006).

38Consider, for instance, an extension of the model in which citizens can confront the state directly rather than channeling their support toward an insurgency that will fight on their behalf. This new means of direct confrontation would prove preferable to citizens only in circumstances where the state was also sensitive to the costs citizens could impose relative to the cost of conflict against an insurgency and, additionally, in cases where the citizens could undertake direct action at low cost themselves. Along these lines, see Lake 1992.

“insurgency” group that does not rely on the equal participation of all disaffected citizens, and (2) are accompanied by underlying threats of insurgent violence that remain latent as a result of preemptive government reforms.\textsuperscript{40}

\subsection{The Accountability Deficit}

According to the baseline model, when a latent political actor can credibly commit to govern citizens more generously than the incumbent government ($v_i > v_g$), citizens have cause to support the entry of that actor onto the political stage. We define this difference $v_i - v_g$ as the government’s accountability deficit, which in intuitive terms represents the gap between the terms of governance that the incumbent currently offers and the proposal of its latent competitor. This concept may remind readers of the ‘sovereignty gap’ introduced in Ghani and Lockhart (2009). In that volume, the authors argue that when governments deliver less than what a sovereign ‘ought’ to provide, they create governance voids that make the country vulnerable to violent entry by domestic political entrepreneurs that may precipitate state failure. Our accountability deficit is similar, albeit with an important distinction that yields separate implications.

In our model, the government’s accountability deficit is not measured relative to an absolute norm of how governments ought to engage their citizens. Rather, the deficit is determined by the relative quality of governance to which either side can credibly commit. This distinction is important because it demonstrates that the anticipated quality of rebel governance, $v_i$, operates as a reservation wage or minimum threshold that incumbent governments must match in order to avoid civil strife. As $v_i$ rises, the incumbent faces incentives to raise its own parallel commitment, $v_g$, not because doing so is normatively desirable but rather because the incumbent seeks to reduce citizen support for the insurgent competitor. Thus, in contrast with Ghani and Lockhart (2009), who propose that a government’s failure to clear the ‘sovereignty gap’ may provoke

\textsuperscript{40}See, for example, the discussion of organized black insurgencies in the American south as described by Hill 2004, Umoja 2013, Cobb 2014, and Levy 2018.
state collapse, our model suggests that governments should respond to looming accountability deficits by initiating a process of reform—at least under the circumstances outlined in the baseline setup.  

3.4.3 **Insurgency as a Tool of Statebuilding**

This model presents the existence of insurgents as an important point of leverage that enables citizens to hold their governments accountable. The threat that civilians may support insurgents in an upcoming conflict is itself sufficient to induce an incumbent to extend durable political and economic concessions to the populace. Our mechanism contrasts with existing theories that rely instead on assumptions about the performance of democratic institutions, citizens’ capacity to challenge the state directly, or the generosity of external actors or stationary bandits. Instead, our framework requires only that the government and insurgents both acknowledge that their prospects in conflict are related—even if only slightly—to the share of civilian support they will each enjoy if escalation occurs.

Because insurgent presence can benefit civilians even when insurgents are themselves inherently predatory and oppositional to civilian interests, our mechanism challenges two polarized characterizations of insurgency that researchers and policymakers often invoke. The predominant view, represented widely in the statebuilding and counterinsurgency literature, derives from the ‘opportunism’ logic of rebellion, and casts insurgents as ‘bad actors’ who undermine peace and stability. According to this view, the international community should assist domestic governments in suppressing insurgents and restoring government control. Indeed, “after the US and other Western militaries became involved in the wars in Iraq and Afghanistan... counterinsurgency attained a lofty pedestal among policymakers and academics.”

---

41 This mechanism is consistent with concessions by popularly-constrained dictators, as in Egorov, Guriev, and Sonin 2009. North 1981 introduces a similar argument but omits the possibility that the threat of violent contestation can create productive pressure on a monopolist regime.

42 Leites and Wolf Jr 1970

43 Jones 2017
violent opportunists provides an ethical and practical pretext for intervention and suppression.

An alternative view, more widely adopted among scholars of social movements, builds on the ‘grievances’ logic of rebellion and casts insurgents in a more sympathetic light. According to this paradigm, citizens are aggrieved by government mistreatment or negligence and decide to rebel against the incumbent regime. Insurgents participate in this rebellion as heroic ‘freedom fighters’ or militants who use violence in hopes of remedying the needs of the citizens whose opinions they share and whose interests they represent. Scholars and development practitioners who adopt this view sometimes advocate supporting insurgent movements or, alternatively, encourage incumbent governments to amend their ways by adopting ‘hearts and minds’ approaches that may address citizens’ grievances and obviate violent rebellion.

According to our model, insurgents deserve neither such censure nor such praise, and ought not be characterized as inherently good or bad. We assume insurgents are violent political opportunists without any intrinsic regard for citizens. Left to their own devices, the insurgents would prefer to obtain the entirety of the divisible good and would offer nothing to the citizens. In order to defeat or extort the government, however, insurgents require citizens’ support, which they can only obtain by sufficiently internalizing citizens’ interests. Thus, although insurgents are intrinsically opportunist, citizens can manipulate insurgent behavior in order to advance the citizens’ own goals, creating the appearance of insurgent altruism. Moreover, by delegating the process of fighting to a vicarious proxy, citizens resolve the collective action problem in their conflict against the government and wield a credible threat of violence with which to extort concessions. Insurgents therefore serve as a useful tool for citizens so long as their incentives to enter or stay out of politics hinge on citizens’ support. Like a watchdog, the insurgency is useful to citizens so long as it remains tethered to their interests.

In view of our model, then, the international community can no longer assume that suppressing opportunistic insurgents in the interest of stability is welfare-enhancing for citizens.

44Gurr 1970
On the contrary, citizens sometimes benefit when their government faces a credible threat of insurgency, so long as this threat motivates the government to commit to more generous governance. Any foreign assistance that relieves the government of competitive pressure will reduce citizens’ welfare. Conversely, foreign assistance that raises the insurgents’ chances of winning can actually be helpful to citizens, so long as it is carefully calibrated to ensure that it does not decouple the insurgents’ decision-making from citizen support.

3.4.4 Empirical Inference and Latent Insurgencies

As we discuss above, the threat of insurgent entry can deter government misbehavior. However, as we discuss in the model extensions section, governments that ignore or underestimate this threat will behave negligently toward citizens. In this case, insurgents may opt to compete with the incumbent for domestic control, and divided governance or civil war can ensue, accompanied by all the destruction, violence, and inefficiencies that such outcomes typically entail.45

Our baseline model therefore suggests that we should only observe insurgency in situations where governments fail to publicly and credibly commit to levels of governance that their citizens desire. In other words, when the latent threat of insurgency succeeds in motivating the incumbent regime to enact credible reforms, the process of institutionalizing should appease citizens, obviate their willingness to support latent insurgents, and deter insurgents from entering the governance market. In this case, the threat of insurgent entry stimulates an improvement in incumbent governance quality even though the insurgents themselves remain latent, unobserved, and uncredited by researchers. Instead, we classify the country as one in which no insurgency exists, peace has prevailed, and the rulers fulfill the desires of the citizens. Alternatively, when the latent challenge fails to motivate good behavior by the incumbent, citizens support insurgency entry and a violent competitor emerges on the political scene. In this case, researchers observe insurgency, civil war, divided governance, and potentially territorial deconsolidation.

45For discussion of oligopolies of violence and their discontents, see Fukuyama 2011, Fukuyama 2014.
Because we observe insurgencies in the company of violence, we interpret them first as a symptom of a government’s incapacity to maintain a monopoly of control and, second, as a direct cause of violence that otherwise would not occur. This interpretation is misleading insofar as it overlooks the censoring problem related to insurgent entry and government reform. Insurgents exist in latent form even in the ‘no entry’ equilibrium—they are simply harder to detect empirically because they do not enter the political limelight in ways that attract attention from researchers. Indeed, Jones speculates that “most would-be insurgents may only make it to the ‘pre-insurgency’ stage” (2017). They instead remain charities, religious clubs, and other organizations within civil society. Latent or low-level insurgencies build ties to the community and solve collective action problems among themselves, but their unrealized potential to enter the political sphere as credible challengers helps keep the incumbent honest.

By measuring the impacts of insurgencies primarily under conditions of civil war and divided governance, researchers introduce a selection bias that discards examples of successful insurgent competitive pressure, thereby encouraging a conclusion that insurgencies are a scourge. Instead, our model suggests that researchers should pursue two goals. First, they should attempt to estimate the positive impacts that latent insurgencies have on governance and political institutionalization. Second, we should attribute violent outcomes not to the inherent existence of an insurgency but rather to factors that inhibit or discourage governments from undertaking institutional reforms so as to prevent the latent insurgent movement’s realization as a violent competitor.

3.5 Model Extensions and the Causes of Civil Violence

Our baseline model deliberately oversimplifies civil conflict for the sake of illuminating the concepts we introduced in the previous section. We now discuss how the model can be easily extended to mirror reality more closely. Most notably, civil strife is always successfully
averted in the baseline model, whereas in reality it is not. We begin by discussing how minor modifications to the model allow for traditional rationalist channels to violence, including but not limited to commitment deficits and information asymmetries between each of the actors. Our baseline model also assumes that each of the three strategic actors are unitary. In practice, however, citizens have heterogeneous preferences, and insurgencies are likewise notorious for fragmentation. We sketch how the model may be amended to account for these realities. Finally, though it is more properly the topic of a follow-on paper, we discuss how ‘outside players,’ namely international actors, can contribute unwittingly or deliberately to instability.

Throughout all of these possible extensions, citizens continue to face the same basic stability-accountability tradeoff. The extensions help make explicit the baseline model’s implicit risks of preventive violence, signaling violence, violence due to heterogeneity of citizens, violence due to fragmentation among insurgents, or violence due to foreign interference. All of these destabilization risks must be weighed by citizens against the anticipated benefits of achieving accountable governance. In some situations, these risks overwhelm those benefits of facilitating insurgency or even tolerating the presence of latent competitors, but in other areas of the parameter space these risks are offset by the benefits of government reform.

### 3.5.1 Preventive Violence

No government wants to be pressured into making concessions to any adversary, foreign or domestic. In our baseline model, the government’s only option for avoiding negotiations with the insurgency is to win the citizens’ hearts and minds prior to insurgent entry. In reality, however, governments have another option: they can attempt preventive action to suppress insurgents who remain in infancy and have yet to mature into credible challengers.\(^{46}\) We may interpret violence by governments against local social welfare providers, NGOs, political dissidents, activists, writers,

\(^{46}\)For formalizations of this logic in either interstate or civil conflict settings, see Powell 1999, Powell 2006, Leventoğlu and Slantchev 2007, Krainin 2017, and Merrell and Abrahams 2019.
journalists, community organizers, religious figures, or others as efforts to crush potential political competitors before they organize and militarize. Concomitant with arrests and disappearances, governments may pursue broad campaigns of censorship, propaganda, and disinformation in order to throw a wrench in the works of civil collectivization.

Several extensions of the baseline model could allow for preventive violence. To sketch perhaps the simplest extension, consider a government who can attempt prevention as an alternative to institutionalization at the beginning of the game. The preventive action, which is undertaken at cost $c_{g,\text{prev}}$, can either successfully crush the latent insurgency (probability $q$), or fan the flames with oxygen as onlookers are outraged by the government’s crackdown (probability $1-q$). We can allow this preventive action to consume a portion of the contested resource, so that only $1-c_{\text{prev}}$ remains. When prevention succeeds, the latent insurgency is eliminated, so the government faces no consequences for failing to institutionalize ($\nu_g = 0$), and can consume the entire remaining finite resource, obtaining $1-c_{\text{prev}}-c_{g,\text{prev}}$ while citizens obtain 0. If prevention fails, the game proceeds as in the baseline model, but with the finite resource reduced to $1-c_{\text{prev}}$. Since $\nu_g = 0$, there is a clear accountability deficit ($\nu_i > \nu_g$), so citizens support insurgent entry. The government’s decision to undertake preventive action therefore reduces to a calculation of whether the expected payoff of attempting prevention, $q(1-c_{\text{prev}}-c_{g,\text{prev}}) + (1-q)[(1-p_{\text{max}}(1-c)+\frac{c_i}{1-\nu_i})]$, exceeds the payoff for heading off insurgent entry: $1-\nu_i$. This reduces to an assessment of the costliness of preventive action and the likelihood of success.

Within this setup, counterinsurgent technologies provided by foreign countries, companies, or researchers that help the government undertake surveillance, arrests, detentions, extraditions, and so on, will raise the probability of successful prevention, motivate the government to suppress potential insurgents rather than enact reforms, and thereby reduce citizen welfare. Likewise, when governments can more efficiently suppress insurgencies through preventive action than via credible governance reforms, a more potent threat of insurgency—i.e., larger values of $\nu_i$ that
force the government to make larger concessions—may provoke government violence that would reduce civilian welfare rather than improve it.

Consider the behavior of Joseph Kabila’s government in Kinshasa since the signing of the February 2013 “Peace, Security, and Cooperation Framework.” According to the terms of that deal, Kabila agreed to reform Congolese institutions. In turn, the U.N. authorized the deployment of a new and more aggressive “Force Intervention Brigade” that was tasked with suppressing rebels in the east—including the M23, the most violent Congolese rebel movement. Within eight months, these measures succeeded, and the U.N.-backed Congolese army forced the M23 to disband. Although numerous hostile challengers remained, the expansion of U.N. peacekeeping efforts improved the Congolese state’s preventive capacity and raised the costs that challenger groups faced for combating Kabila directly. As a result, these changes enabled Kabila to further consolidate his political control. In the absence of viable opposition, he began to delay institutional reforms and initiated crackdowns on political freedoms. In January 2015, he jailed hundreds of non-violent political protestors, expanded a campaign of political intimidation, imposed widespread media censorship, and announced plans to delay a promised round of presidential elections that were originally scheduled for 2016.\footnote{A Congo Research Group poll found that Kabila would have earned only 7.8% of the vote if elections were held.} Only in the face of mounting international pressure and the reemergence of widespread internal violence did Kabila agree to step down from power and allow elections to occur in late December 2018.\footnote{Even the results of these elections were widely contested. Election observers report that opposition candidate Martin Fayulu received a landslide share of the vote, but Félix Tshisekedi was instead installed as president.}

### 3.5.2 Mutual Optimism and Signaling Violence

If the insurgency survives preventive violence, it decides to enter or stay out of the political arena. If it enters, we can assume it anticipates citizen support, so \( p = p_{\text{max}} \). The government can avert war by offering a peace deal \( g_2 \), which the insurgency accepts whenever \( g_2 \geq p_{\text{max}}(1 - c) - \frac{c_i}{\nu_i} \). Any disagreement over these parameter values, however, can cause
bargaining to break down. If the government believes, for example, that the insurgents’ costs of
war are higher than they really are \((\hat{c}_i > c_i)\), and offers
\[ g_2 = p_{\max}(1 - c) - \frac{\hat{c}_i}{1 - \nu_i}, \]
the insurgency will reject the offer and violence will ensue. Events will unfold similarly if the insurgency
underestimates its own costs of fighting.

The coercive actors may also be mutually optimistic about their prospects of citizen
support, as this informs their expected war payoffs. For example, if the insurgents believe the
citizens are on their side, so \( p = p_{\max} \), while the government believes the citizens are on their side,
so \( p = p_{\min} \), then the insurgency will only accept
\[ g_2 \geq p_{\max}(1 - c) - \frac{\hat{c}_i}{1 - \nu_i}, \]
while the government will only offer
\[ g_2 = p_{\min}(1 - c) - \frac{\hat{c}_i}{1 - \nu_i} < p_{\min}(1 - c) - \frac{\hat{c}_i}{1 - \nu_i}, \]
and violence will once again ensue.

To model decision making in the presence of imperfect information over these parameters,
we can follow the usual approach of introducing subjective beliefs for each actor, namely
probability distributions over parameter values. The end result of this modification is that civil
strife occurs with nonzero probability in equilibrium. In the baseline model, citizens support
insurgent entry whenever there is an accountability deficit \( \nu_i > \nu_g \). With a nonzero risk of civil
strife, the calculation amends to
\[ \nu_i - \text{risk of violence} > \nu_g. \]
The term \textit{risk of violence} is merely
the product of the probability of violence and the loss associated with that outcome. Rearranging,
we obtain \( \nu_i - \nu_g > \text{risk of violence} \), which clarifies that the accountability deficit only provokes
the citizens to support insurgent entry if the risks of civil strife are outweighed by the anticipated
benefits.

Throughout history, optimism has often motivated governments to deny preemptive
concessions to their subjects and to instead undertake costly wars. One example our readers may
find familiar is the interaction between the leaders of Great Britain, Patriot insurgents, and colonial
residents in the prelude to the American Revolutionary War. Although the British were aware of
growing frustration among citizens in the colonies, they also doubted that a large proportion of the
population would back the extremist insurgencies if a conflict escalated. Moreover, the Americans
appeared to lack sufficient weaponry, possessed neither a standing army nor a standing navy,
and have never previously coordinated on military defense. As a result, British leaders assumed their patriot opponents possessed relatively low coercive capacity and that Britain could secure a quick and relatively easy victory. These perceptions led them to resist offering concessions to appease their colonial subjects, further motivating sections of the populace to support the Patriot insurgency, which subsequently entered the political market wholesale and initiated conflict with Britain.49

3.5.3 Heterogeneous Citizens

In recent work, Berman, Shapiro, and Felter (2011) model a three-way strategic interaction between a citizens, insurgents, and a government who seeks to induce civilian cooperation. Our baseline model parallels theirs by including the assumption that the three strategic actors are unitary or homogenous. Nevertheless, this unitary actor assumption does not hold in practice, and more realistic models of civil conflict should allow for heterogeneity among citizens along with fragmentation among insurgent groups.50

Let us first consider heterogeneity among citizens. According to the hearts-and-minds COIN strategy, insurgents depend upon the support of citizens for survival and victory. If the government can win over the citizens’ hearts and minds, it can convince them to retract their support for insurgents and the insurgent movement will collapse. But the hearts-and-minds COIN strategy is a war-fighting strategy as opposed to a statebuilding strategy, and is designed to depict situations in which active insurgencies vie with governments for political control with violence already underway. Our model departs importantly from this tradition by depicting conditions prior to the outbreak of war, where insurgency remains only a latent rather than overtly political or violent actor. Thus, whereas models like Berman, Shapiro, and Felter (2011) take the open manifestation of the insurgency as exogenously given, our model clarifies that insurgent entry is a

50Modeling fragmentation within the incumbent regime may also be productive.
strategic choice that internalizes political conditions on the ground. Indeed, our model suggests that insurgents should only exist when citizens extend support to latent competitors following an incumbent’s failure to commit to good governance. Furthermore, following insurgent entry we should only observe violence if the government responds intransigently to insurgent demands. By the time we observe active violence between a government and insurgents—in other words, the start of the Berman, Shapiro, and Felter (2011) model—we are likely dealing with a government that has been unable to credibly and believably convince citizens and insurgents that it will undergo reform.

Given that the government is perceived by citizens as relatively negligent and intransigent thus far, it seems remarkably optimistic for hearts-and-minds strategists to suppose that the citizens could be readily convinced to take the government’s side against the insurgency. As our model clarifies, citizens have no reason to suppose that the government will share anything more than a fraction $\nu_g$ of their post-war spoils, whereas citizens should believe insurgents will share $\nu_i > \nu_g$. The government cannot raise $\nu_g$ by making perfunctory gestures of good governance with such transparently instrumentalist motives. Indeed, if the citizens were so gullible, then the insurgents would not have risked entry in the first place. How, then, can the hearts-and-minds strategy work?

One solution is for the government to recruit informants using privately-distributed incentives. In other words, a strategic government may attempt to exploit the citizens’ collective action problem. Even though every citizen may prefer the insurgents to win, each individual can be bribed or blackmailed to collaborate with the government because they each assume that their private action will render only a negligible impact on the overall success or failure of the insurgency in its competition with the government. If a critical mass of citizens can be recruited in this fashion, then the insurgents’ probability of victory $p$ is meaningfully reduced and the government may perhaps prevail. The calculation becomes even easier if we assume that citizens do not all prefer to be governed by the insurgents. For example, in self-determination
or secessionist struggles, there are often ethnic or sectarian minorities among the citizens who
do not actually share the majority’s nationalist fervor and can be convinced relatively easily to
support the counterinsurgents. The hearts-and-minds COIN strategy can therefore make sense if
we view citizens as heterogeneous.

Our baseline model can be modified to account for such heterogeneity, and researchers
should explore ways to implement these modifications in the future. To sketch one method,
recall that in the baseline model, citizens support insurgent entry whenever an accountability
deficit exists such that \( v_i > v_g \). With a nonzero risk of civil strife, the calculation amends
to \( v_i - \text{risk of violence} > v_g \). If citizens are homogeneous, then as soon as this inequality is
satisfied, all citizens support the insurgency. If, however, we assume that each citizen \( k \) has
an idiosyncratic preference \( \varepsilon_k \) for being ruled by the incumbent, drawn from a continuous and
infinitely supported distribution \( F \) with mean 0, then as \( v_i \) rises or \( \text{risk of violence} \) falls, a larger
mass of citizens—but not all—support insurgent entry. Indeed, a citizen \( k \) supports insurgent entry
only if \( v_i - \text{risk of violence} > v_g + \varepsilon_k \). Thus, whereas the baseline model effectively discretizes
the question into “do the citizens support the insurgents or not,” the introduction of idiosyncratic
preferences alters the entry question to “are the insurgents supported by a critical mass of
citizens?”

With this setup, we can now allow both the government and the insurgency to bribe
individual citizens. Indeed, if citizen \( k \) initially supports the insurgency (\( v_i - \text{risk of violence} > v_g + \varepsilon_k \)), the government must offer her a bribe of at least \( v_i - \text{risk of violence} - v_g - \varepsilon_k \) to
switch sides. Researchers pursuing this line of inquiry could also introduce some probability
that collaborators are caught and executed, which disincentivizes snitching, although this is
counteracted by the risk that even non-collaborators may fall under suspicion of such behavior.\(^51\)
The insurgents’ capacity to keep citizens onside, therefore, is a function of the mean position of
citizen support (\( v_i - v_g - \text{risk of violence} \)), the government’s budget for bribery, the insurgency’s

\(^{51}\)Merrell and Abrahams 2019.
budget for bribery, and the quality of counterintelligence on both sides to catch and execute informants to disincentivize illicit cooperation with the opposing side.

### 3.5.4 Insurgent Fragmentation

Our baseline model assumes the insurgency is a unitary actor, but in reality they are often quite fragmented. Insurgent fragmentation should in general have positive and negative consequences for citizens. On the one hand, fragmentation may create positive competitive pressure among insurgent groups to win citizen support. Indeed, the insurgency’s commitment to the citizens $v_i$ may itself be underwritten by competitive pressures among insurgents. On the other hand, insurgent groups may engage in counterproductive internecine violence that inhibits their capacity to present a coherent front against the government.

To sketch one possible extension along these lines, let us adopt the argument of Krause (2017), where the insurgency must first resolve its internecine power disputes, perhaps through an extended period of violence, until a hegemon emerges. Only then, Krause argues, can the insurgency focus on its ‘external’ adversary (the government, in our model). This argument lends itself easily to our setup, since we can simply suppose that the insurgents pay a ‘hegemonization’ cost $c_{\text{hegemony}}$ after entering the political arena but before confronting the government. In some sense, then, we can think of the insurgency’s entry cost $c_f$ from the baseline model as including a cost of hegemonization along with other costs $c_f^*$, so that $c_f = c_f^* + c_{\text{hegemony}}$. If costs of hegemonization are high, then overall costs of insurgent entry are high, and this disincentivizes insurgent entry overall. Imaginably, internecine violence also takes a toll on citizen support.

---

52 See, among others, Christia 2012, Krause 2017, and Lake 2019. Note that Lake also addresses the possibility that insurgent groups may be “too powerful” relative to the state, and in these cases will facilitate state failure. Our model clarifies this logic by demonstrating that when citizens can influence the relative combat success of insurgents and incumbents, they can leverage each actor against the other. Although we emphasize throughout the paper that this allows citizens to wield the threat of latent insurgency to hold governments accountable, influential citizens can also side with the government to prevent the emergence of overly powerful or otherwise undesirable insurgent groups.

53 We might also interpret this as the cost of engaging in costly signaling so as to attract supporters, as in the terrorist “outbidding” strategy (Kydd and Walter 2006).
for the insurgency, so the anticipated support level $\sigma$ is really the difference between support prior to internecine violence and the disillusionment loss associated with internecine violence: $\sigma = \sigma^* - \text{disillusionment}$. The probability of victory in the ‘external’ fight against the government, $p(\sigma)$, therefore internalizes these shifts in citizen support.

Likewise, spoiler violence may occur when insurgent movements fail to consolidate prior to their negotiations with the government. Suppose, for example, that one insurgent group within the broader insurgent movement accepts the government’s offer for a peace deal. If another insurgent group, however, is dissatisfied with the terms of the deal, it may spoil the deal with violence. Accordingly, to achieve peace, the government would have to sweeten the peace deal with a premium large enough that the insurgent movement can buy off the splinter group. If the splinter group is very demanding, however, then the government may be unwilling to pay the premium, and civil war will ensue. From a modeling perspective, we can simply add a premium $c_{\text{spoiler}}$ to the government’s minimum acceptable offer $g_2$, so $g_2 = p_{\text{max}}(1 - c) - \frac{c_i}{1 - \nu_i} + c_{\text{spoiler}}$. For a large premium $c_{\text{spoiler}}$, the government prefers to eschew peace and take its chances in war. Thus, fragmentation narrows the continuum of feasible bargains, making insurgent entry a less attractive option for citizens overall.

### 3.5.5 Outside Players

The baseline model analyzes civil conflict within an environment in which civilians, insurgents, and governments operate independent from international influence. More realistically, countries are embedded in the international system and domestic strife invariably attracts foreign interference. Indeed, Jones writes that “of the 181 insurgencies since 1946, 148 cases (82 percent) involved some form of outside support” (2017). In this section, we explore the implications of our theory when we introduce the presence of interested foreign actors. As we see in the autarkic setup, citizens can achieve domestic accountability by carefully calibrating an insurgent threat to government rule. This balance, however, may be upset by foreign interference when these
external actors are naive to the dynamics we outline in this paper.

In addition, even well-informed foreign powers have regional or global agendas to which the interests of citizens in a client country play decidedly second fiddle.\(^{54}\) From the perspective of some foreign powers, the prospect of civil strife in a client country is an inconvenient nuisance that should be suppressed.\(^{55}\) If the U.S. seeks to deter Iranian aggression by docking a naval fleet at Bahrain’s deep-water harbor, agitation by Bahraini citizens for a more accountable government may annoy American officials in two ways. First, unrest may impede the United States’ capacity to deter Iran and, second, domestic unrest may distract the Bahraini government from their cooperation with the United States. In this context, the U.S. may provide support to the Bahraini regime, whether tacitly or explicitly, in order to raise the odds of government victory against a would-be Bahraini insurgency, thus deterring insurgent entry and relieving the Bahraini government of the need to institutionalize.

Alternatively, a foreign power may support the insurgency if it dislikes the political leanings of the incumbent government. In the 1980s, the United States armed and funded Afghani and Nicaraguan insurgents because the governments of the states in which those movements operated were Soviet-aligned. As our baseline model shows, an insurgent threat has positive consequences for governance and institutionalization, but only if insurgent entry is conditional on citizen support. That is to say, the insurgents’ expected payoff for entry should be positive with citizen support \((1 - \nu_i)[p_{\text{max}}(1 - c) - \frac{ci}{1 - \nu_i}] - cf > 0\), and negative without citizen support, \((1 - \nu_i)[p_{\text{min}}(1 - c) - \frac{ci}{1 - \nu_i}] - cf < 0\). If a foreign power provides insurgents with arms and funding such that \(ci\) and \(cf\) are small, then the insurgency may feel emboldened to enter the fray even without the blessing of citizens: \((1 - \nu_i)[p_{\text{min}}(1 - c) - \frac{ci}{1 - \nu_i}] - cf > 0\). If this happens, the insurgency is now effectively decoupled from citizens’ interests and operates outside of citizens’ strategic control. The watchdog is ‘off the lead,’ so to speak, and no longer incentivizes accountability reforms on behalf of the incumbent.

\(^{54}\)See Jamal 2012, Lake 2016, Berman and Lake 2019, or Lee 2018, for example. 
\(^{55}\)Berman and Lake 2019
Our argument does not imply that decisive intervention is always harmful to citizens, nor do we insist that conflicts from which neither side emerges victorious are universally beneficial. At the most immediate level, foreign intervention in an ongoing conflict may benefit civilians if it hastens the conclusion of war or reduces the intensity of violence to which residents are directly exposed. In a broader sense, however, the ultimate consequences of foreign intervention hinge on whether it facilitates or disrupts domestic competition in the overall governance market. Even forms of intervention that initially appear successful at suppressing violence may nevertheless prove harmful in the long run if they reduce overall competition or restrict citizens’ capacity to influence the outcome of a violent conflict between an incumbent and insurgent challenger.

In addition to the problems we discuss above, intervention may also produce a chaotic and anarchic post-war environment in which no party emerges as the focal point for civilian support. To pose a meaningful threat to a consolidated government, insurgencies typically must undergo a maturation process in which they ultimately resolve their own internal feuds.\footnote{See especially Krause 2017 and Christia 2012.} Foreign assistance may deliver victory to a group of insurgents ‘too soon,’ without requiring them first to pass through the eye of such organizational maturation. Instead, insurgent infighting ensues in the context of a power vacuum, after the government’s defeat. Likewise, if civilians lack sufficient information about the relative qualities of each combatant they may divide their allegiance in a manner that prevents any individual competitor from growing large enough to benefit from economies of scale in governance provision.

Given these complexities, foreign powers who seek to maximize the benefits of intervention should not necessarily pursue the quickest route to a peace settlement, nor should they always set a goal of wholly dismantling one of the belligerent parties. Instead, the goal of benevolent foreign intervention should be twofold. First, interveners should seek to erect or preserve political conditions in which opposing sides are held mutually accountable through latent competition and their reliance on civilian support. Second, intervention should attempt to resolve and prevent
the types of market inefficiencies that allowed conflict to break out in the first place. Citizens’ interests are best served when intervention succeeds in cultivating a durable political environment that nurtures viable competitors and addresses the incipient causes of conflict.

3.6 Conclusion

This paper should help researchers better understand the political economy of domestic governance and the inherent complexities of foreign intervention. First, the theory suggests that researchers should move beyond traditional debates regarding the conceptual distinctions between “autocracy” and “democracy.” Instead, we should study the mechanisms through which citizens can best manage the tradeoff between government power and accountability. Citizens face a fundamental tradeoff between empowering their government to enforce stability and constraining government power so as to keep their leaders accountable. With a monopoly over the use of force, a sovereign can ensure domestic stability but cannot promise to eschew predatory or extractive behaviors.

In contrast to one popular perspective, we demonstrate that democratization is neither the only nor necessarily the most practical method of resolving this dilemma. Rather, insurgents in many cases represent a potent means by which citizens can secure accountable governance even in the absence of institutionalized constraints. Indeed, we demonstrate that the existence of viable insurgents can benefit citizens even when the insurgents are neither supportive of nor sympathetic with the broader civilian population. In a departure from polarized interpretations of insurgent movements as either violent and opportunistic “antagonists” at one extreme or the well-intended “allies” of civil society at the other, we argue that insurgent groups deserve neither such censure nor such praise. Just as consumers benefit from the existence of multiple suppliers even if all such suppliers are inherently self-interested, citizens can benefit from the existence of latent insurgent competitors even if such insurgents would themselves engage in violent predation.
if given an opportunity.

Although our analysis throughout the paper is positive, not normative, it nevertheless also yields a variety of important implications for the conduct of policy. By highlighting the productive role of insurgent threats, our argument challenges approaches to statebuilding that depict the cultivation of internal stability as a prerequisite for generous and accountable governance. Blind faith that a regime will produce better governance and construct binding institutions once violent challengers are suppressed is theoretically naïve. Instead, reforms and concessions require motivation. Put simply, where Tilly argued that the threat of international conflict “made the state” by encouraging governments to pursue efficiency-enhancements, we demonstrate the threat of insurgency is what make states accountable. The risk of instability encourages rulers to offer generous concessions to citizens and also provides a mechanism that makes such promises credible. Although information lapses, commitment problems, and heterogeneity of either citizens or insurgents can provoke violence, there remain conditions in which citizens obtain better outcomes by nurturing an insurgent threat than by suffering the rule of a despotic regime.

This is not to say that policymakers should always tolerate insurgencies. When an insurgency’s decision to enter or avoid the political arena hinges on citizen support, we may interpret their participation as representative of the will of the people. However, there also exist circumstances when insurgents enjoy sufficient funding and arms that their decision to challenge the government is decoupled from citizens’ interests. Under these circumstances, insurgent entry should indeed be deterred, as their war with the government would prove destructive and their victory, if realized, promises no benefit to citizens. Likewise, when assisting citizen-supported insurgencies, international actors should take care not to provide assistance that enables insurgents to either operate independent of citizen support or, alternatively, to wholly replace the regime as new monopolists.

Our model suggests important changes in our understanding of political order and state-building. While we agree with Hobbes and Weber that observable violence should be rare in
well-governed states, we challenge the assertion that such conditions arise from monopolistic government. Instead, incumbent regimes face significant competition from latent actors who could enter the political marketplace if the government misbehaved. Is latent coercion an underlying feature of accountable governance? Do consolidated democracies nevertheless face competition from latent insurgencies? In our model, the threat of insurgent entry motivates the government to undertake institutionalization that obviates insurgency. From that view, only under-institutionalized countries require insurgent threats facilitate accountability and endow the government with the legitimacy Weber emphasizes. Many scholars of social movements and civil society activists maintain, however, that there are no norms or institutions that determined governments cannot erode or reverse. Indeed, in recent years many Americans have expressed surprise at the brittleness of norms that constrain misbehavior and enforce rule of law on those who hold high office. Perhaps the most solicitous way to interpret our model is that insurgent threats motivate governments to hold themselves accountable ‘for the foreseeable future,’ by erecting credible checks and balances that would be ‘hard’ (but perhaps not impossible) to deconstruct. Institutions can fall into decay, and even in the U.S. there may come a time when insurgent threats once again safeguard democracy.

Finally, by presenting a theory in which insurgencies emerge endogenously within a parsimonious civil governance framework, this project also establishes a new path forward for both researchers and policymakers. Our model suggests that insurgencies are not inherently destructive; instead, violence should occur only when governments fail to concede to the terms that insurgencies can credibly demand. As such, rather than pursue better methods of suppressing insurgents or, alternatively, of helping rebels depose and supplant predatory regimes, researchers should attempt to identify and address the various frictions within the governance market that can produce violence. Finally, we should recognize that in situations where these frictions cannot

57 Fukuyama 2014
58 Although this may evoke sentiments akin to those espoused by supporters of the second amendment, our “insurgents” must be better organized and more capable of using violence against the state than are individual citizens.
be entirely resolved, repression and rebellion are inversely related: reductions in the former are achievable only via the introduction or promotion of domestic competitors whose presence fosters instability and the attendant risks of either war or repression. The fundamental tension between reducing structural repression at risk of violent rebellion (or vice-versa) is an ethical challenge that researchers and practitioners have a responsibility to confront.

All chapters of this dissertation are being prepared for submission for publication of the material. The dissertation author was the primary investigator and author of this material.

Chapter 3 is coauthored with Abrahams, Alexei S. The dissertation author was the primary author of this chapter and has permission to use the material contained herein.
Bibliography


134


Fang, Songying and Xiaojun Li (2016). Historical ownership and territorial indivisibility. Tech. rep. working paper, Rice University.


— (2014). Political order and political decay: From the industrial revolution to the globalization of democracy. Farrar, Straus and Giroux.


Sater, William F (1986). *Chile and the War of the Pacific*. University of Nebraska Press.


Umoja, Akinyele Omowale (2013). *We will shoot back: Armed resistance in the Mississippi freedom movement*. NYU Press.


Yarhi-Milo, Keren, Alexander Lanoszka, and Zack Cooper (2016). “To arm or to ally? The patron’s dilemma and the strategic logic of arms transfers and alliances”. *International Security* 41.2, pp. 90–139.
