

Copyright
by
Xuanxuan Wu
2017

The Dissertation Committee for Xuanxuan Wu
certifies that this is the approved version of the following dissertation:

**Bargaining in Public: Resolve and Publicity in
International Crises**

Committee:

Michael Scott Wolford, Supervisor

Terrence L. Chapman

Songying Fang

Patrick J. McDonald

**Bargaining in Public: Resolve and Publicity in
International Crises**

by

Xuanxuan Wu, B.LAW.; M.PHIL.

DISSERTATION

Presented to the Faculty of the Graduate School of

The University of Texas at Austin

in Partial Fulfillment

of the Requirements

for the Degree of

DOCTOR OF PHILOSOPHY

THE UNIVERSITY OF TEXAS AT AUSTIN

May 2017

Dedicated to my grandfather, Jiabi Xie.

Acknowledgments

A journey towards a Ph.D. would be impossible without the support of many people.

First and foremost, I am deeply grateful to my advisor, Scott Wolford. Beyond the constant support for my dissertation and research, Scott Wolford introduces me to the world of game theory, encourages me to explore ideas, and believes in me more than I do. Following his move to the University of Texas was the best decision I have made in my graduate studies.

Other committee members also deserve special acknowledgment. Terry Chapman and Pat McDonald have provided series of comments and feedback on my dissertation drafts and practice talks. Songying Fang have shared invaluable suggestions on how to survive and succeed in the U.S. academia as a foreign scholar.

In addition to my dissertation committee, I have benefited greatly from the generous help of faculty at the University of Texas, including Bethany Albertson, Josh Eisenman, Mike Findley, Wendy Hunter, Nate Jensen, Tse-Min Lin, Amy Liu, Xiaobo Lü, Scott Moser, and Rachel Wellhausen.

My graduate colleagues have provided strong intellectual and mental support along this journey. Dan McCormack and Henry Pascoe offered thoughtful comments on several drafts of the theory chapter. The Women in

IR Workshop initiated by Julie Phillips and Megan Thomas has been very helpful in the development of my dissertation project. I also thank Riitta Koivumaeki, Peter Harris, Andy Stravers, Calla Hummel, Giorleny Altamirano, Caitlin Andrews, Jay Kao, Kyosuke Kikuta, Hans-Inge Langø, Carolina Moehlecke, Iasmin Goes, and Barea Sinno.

Others have had input that deserves acknowledgment, including David Bearce, Bill Clark, Jim Morrow, Shawn Ramirez, Branislav Slantchev, and participants of the EITM Institutes at the University of Michigan and at the University of Houston. I am grateful for the comments from Sam Bell, William Spaniel, Ahmer Tarar, and Leah Windsor at the Online Peace Science Colloquium, with special thanks to Emily Ritter for organizing it.

Another special thanks must go to friends who have made this journey from Colorado to Texas more enjoyable and memorable. These include Pinar Alakoc, Jia Chen, Dasha Djukic, Cody Eldredge, Shan Jiang, Yoo-Sun Jung, Ou Kai, Chao Liu, Sa Liu, Uly Putri, Kim-Lee Tuxhorn, Mindy Zhang, among others.

I also owe a great deal to my parents, Zhengzhen Wu and Lingbo Xie, for their unconditional support. Finally, my deep thanks go to my boyfriend Alex Zhou for his candor, insight, and love for the past twelve years.

Bargaining in Public: Resolve and Publicity in International Crises

Publication No. _____

Xuanxuan Wu, Ph.D.

The University of Texas at Austin, 2017

Supervisor: Michael Scott Wolford

How actors credibly signal resolve has been an enduring question for both scholars and policy makers. The existing literature disagrees on the effects of public threats on signaling resolve and the effects of political constraints on crisis outcomes. This dissertation examines how leaders use public threats to signal resolve from two new perspectives. First, citizens are concerned both about national prestige and about crisis outcomes, the latter of which are shaped by their resolve. Second, leaders adjust their vulnerability to political punishment by controlling the publicity of their threats.

By locating resolve in the public and allowing leaders to choose the level of publicity during a crisis, the dissertation offers an integrated framework to understand the apparently disparate strategies of leaders under different levels of political accountability. Moderately accountable leaders increase the level

of publicity to signal the public's resolve, whereas highly accountable leaders lower publicity to signal the public's resolve. I also identify the conditional effects of political accountability on the risks of war with statistical support. Finally, assuming that leaders know their citizens' resolve no better than foreign targets, I find that dovish citizens may pretend to be supportive of war if leaders tend to increase publicity in any crisis.

This project reconciles the conventional audience costs theory and its critics. It specifies the linkage between a leader's choice of publicity and signaling resolve. It also suggests that political accountability may produce perverse outcomes costly to the public. Finally, I discuss the implication of this project to U.S. foreign policy and mass mobilization in international crisis.

Table of Contents

Acknowledgments	v
Abstract	vii
List of Tables	xi
List of Figures	xii
Chapter 1. Introduction	1
1.1 Rethinking Audience Costs Theory	8
1.2 The Argument	15
Chapter 2. Political Costs, Publicity, and Crisis Bargaining	21
2.1 Why is Backing Down Undesirable?	24
2.2 Audience Costs vs. Policy Costs	27
2.3 The Baseline Model	30
2.4 Endogenous Publicity and Political Costs	34
2.5 Summary	40
Chapter 3. Seeking and Avoiding Accountability: Signaling Resolve to Foreign Targets	43
3.1 The Model	45
3.2 Equilibrium under Complete Information	48
3.3 Equilibrium under Incomplete Information	52
3.4 Discussion	59
3.4.1 Endogenous Publicity and Signaling Resolve	60
3.4.2 Political Accountability and Risks of War	62
3.4.3 Autocratic Advantage and Sources of Uncertainty	65
3.5 Summary	66

Chapter 4. Empirics	71
4.1 Empirical Implications of the Model	72
4.2 Testing the Effectiveness of Threat	76
4.2.1 Threat Failure	77
4.2.2 Political Punishment: Capability vs. Incentive	79
4.3 Research Design	82
4.4 Results	91
4.5 Summary	97
Chapter 5. Dragging Themselves to War: Signaling Resolve to National Leaders	99
5.1 The Modified Model	103
5.2 Equilibrium under Complete Information	106
5.3 Equilibrium under Incomplete Information	110
5.4 Summary	115
Chapter 6. Conclusion	120
6.1 Experiments and Theory Building	126
6.2 Publicity of Threat and Beyond	128
6.3 Policy Implications	132
Appendices	135
Appendix A. Chapter 2 Appendix	136
Appendix B. Chapter 3 Appendix	138
Appendix C. Chapter 4 Appendix	152
Appendix D. Chapter 5 Appendix	156
Bibliography	167

List of Tables

4.1	Components of ACC	85
4.2	Political Accountability Index	87
4.3	Estimating Compellent Threat Failure (1918–2001)	92
C.1	Descriptive Statistics	153
C.2	Estimating Compellent Threat Failure (1918–2001)	154
C.3	Estimating Compellent Threat Failure (1918–2001)	155

List of Figures

2.1	The Baseline Model	31
2.2	F 's Offer	32
3.1	Signaling Resolve to Foreign Target	47
3.2	The Equilibrium Spaces	57
4.1	Equilibrium probability of threat failure	73
4.2	Predicted probability of compellent threat failure	94
5.1	Signaling Resolve to National Leader	104

Chapter 1

Introduction

Featured with extreme aggression, destruction, and mortality, the outbreak of war is of enduring interest to students of international relations. Rationalist explanations for war begin with the observation that war is costly to the parties involved and hence there exists a range of negotiated settlements that the parties involved prefer to war (Fearon 1995). In other words, states fight against each other as a result of the failure to achieve an agreement. One of the main explanations to the failure of negotiated settlements is the existence of uncertainty among states and their incentives to misrepresent it.¹ When negotiating over some disputed goods (e.g. territory, policy change, leadership, etc), states are uncertain about some unobservable characteristics of their opponents, such as their military capabilities or valuations of the disputed goods. One of these unobservable characteristics is resolve, or one's willingness to fight. If an actor is considered more resolute, it indicates that this actor is more willing to fight over the disputed goods and with all else being equal, more likely to achieve a settlement to its favor, than an irresolute actor. Since these unobservable characteristics may shape the decisions over

¹The other approach is to consider war as a commitment problem. Fearon (1995) and Powell (1999) are foundational works on commitment problems in international relations.

bargaining postures and outcomes, each state has incentives to bluff about its own characteristics in order to boost its bargaining leverage. The incentives of misrepresent thus lead to the incentives to disbelieve, which reduces the effectiveness of diplomatic negotiations and increases the risks of war. Following this logic, states can learn their opponents' resolve only by involving costly actions such as war.

Hence, a serious question posed to scholars and policy makers is how to make interstate communications more informative without resorting to war. One general solution is costly signaling: to increase the credibility of communication, a truly tough state must be willing to pay extra costs that all other not-so-tough states are unwilling to pay. It is the extra costs that improve the effectiveness of communication, making it more informative. In addition to war, scholars examine other mechanisms that enable states to incur extra costs and thus reveal their resolve. Examples of costly signaling mechanisms include generating a threat that “leaves something to chance” (Schelling 1960), military mobilization (Fearon 1997, Slantchev 2005), and diplomatic reputation (Guisinger and Smith 2002, Sartori 2002). Each mechanism differs in who imposes extra costs and how. Among the existing mechanisms, audience costs focus on the domestic interactions between leaders and the public, with extra costs imposed by the public.²

²Some studies also include international audience (i.e. foreign states or other international actors) when discussing audience costs theory. I put the works on international audience in the category of “diplomatic reputation.” Unless specified, an “audience” in this dissertation refers to a domestic public.

According to audience costs theory, a domestic public may improve the effectiveness of crisis communication among states if the public can tie their leader's hands. In the context of crisis diplomacy, tying hands refers to extra costs for being caught bluffing—an actor threatens to use force if its demand is not satisfied yet fails to carry out its threat later. If leaders can show their hands tied, they are less likely to bluff and therefore the information they communicate is more credible. The existing literature argues that domestically, leaders can make their hands tied by the public. In his foundational work on audience costs, Fearon (1994) argues that citizens can witness how leaders communicate as crisis is a public event, and that citizens can punish their leader for being caught bluffing, which they believe damages national prestige. Hence, a domestic public is considered the “audience” and the costs for leaders being caught bluffing is labeled “audience costs.” The more audience costs leaders are subject to, the more credible their threats of war are with foreign opponents. The literature suggests that the amount of audience costs vary with political regimes and the level of public attention involved in crises (Baum 2004, Fearon 1994, Smith 1998). Following this logic, we should expect the audience cost mechanism to be more effective in revealing true resolve when the public pays more attention and/or when threats are issued by democratic leaders, as people in democracies are more capable of punishing leaders. However, we do observe cases inconsistent with the above logic: either democratic leaders were caught bluffing without much punishment, or they were able to achieve settlements by deliberately reducing public attention

to their communications with foreign opponents.

In 2012, U.S. President Barack Obama made a widely-noticed threat to Bashar al-Assad of Syria, warning that “a red line for us is we start seeing a whole bunch of chemical weapons moving around or being utilized” (Kessler 2013). Despite the evidence of Assad targeting civilians with chemical weapons, a year later, Obama called off a military attack on Syria and decided to seek congressional approval. Obama’s new move was largely considered as inconsistent with his initial threat, since he had known that neither American people nor the Congress was interested in another military intervention in Middle East and that it mattered to “have the public with him” (Goldberg 2016). A similar scenario emerged when Chen Shui-bian, the then-President of Taiwan, challenged Beijing by proposing a referendum on joining the United Nations under the name “Taiwan” and yet backed down. Chinese leaders viewed Chen’s proposal of the UN referendum as a challenge to the status quo, raising tension across the Taiwan Strait (Lauge 2008). To draw wide attention, Chen conducted campaigns and scheduled the referendum on the same day of the Presidential election. Concerned about the prospect of war against mainland China, the majority of Taiwanese citizens chose not to vote in the referendum. Chen backed down after the referendum was declared invalid due to a low turnout.³

On the other hand, leaders sometimes choose to reduce public attention

³While the presidential election had a turnout rate at 76%, the turnout rate of referendum was 35%, less than the 50% threshold to validate the result.

to crises. In September 1970, when the Nixon Administration detected that the Soviet Union was building a support base for submarines in the Cuban bay of Cienfuegos, President Richard Nixon decided not to “force a public confrontation” with the Soviet Union (Siniver 2008). Instead, the Nixon Administration maintained a low-profile, but not fully secret, strategy regarding the developments in Cuba. Initially, the knowledge of the submarine base was leaked to the press, with few details and a non-threatening tone (Crall and Martin 2013).⁴ But soon U.S. officials were instructed neither to confirm press inquiries nor to raise this issue in external meetings (Crall and Martin 2013). Kissinger was constantly playing down the severity of the crisis with the press and avoiding the development of a “crisis atmosphere.” The Nixon administration played a low-profile strategy and finally reached an agreement with the Soviet Union.

According to audience costs theory, President Obama and President Chen were both supposed to pay large audience costs due to the public concerns of national prestige, yet their hands seemed “untied” by their citizens with little support for war. For Obama’s case, despite some criticisms about his handling of the Syria crisis, the Congress had no interest in authorizing a military strike. Similarly, Taiwanese people forced Chen to back down from the threat of formal independence, with the majority of them not casting

⁴(Siniver 2008) suggested that the initial knowledge of the base was leaked to the press because a State Department spokesman mistakenly mentioned the situation in Cuba during a press briefing. But it is hard to verify whether the leak was unintentional and it did not involve any details as well.

votes. If public threats tie a leader's hands and increase the credibility of threats (Fearon 1997), why did Obama and Chen make high-profile threats when their citizens seemed *unwilling* to fight? More generally, how do public threats affect crisis bargaining? Why do leaders make high-profile threats sometimes and low-profile ones at other times?

The existing literature cannot address these questions for two reasons. First, most works rely on the assumption that people, as domestic audiences, are biased against backing down, which they believe hurts national prestige with negative consequences in the future (Fearon 1994, Guisinger and Smith 2002). However, it is unclear why people should always prefer escalating a crisis to war if they are the direct victims of a costly war (Chiozza and Goemans 2004). What information would public threats reveal or disguise if the public cares both about their leader's consistency and about the substance of crisis outcomes? Considering the public as the direct victims of war leads to further questions regarding the locus of resolve: Is resolve located in leaders or the public? Extant works on audience costs either assume a unitary-state framework or examine the resolve of leaders without further justification. Nonetheless, if the public suffers most from war, it seems more reasonable to consider a state's resolve located in the public and examine how leaders signal the public's resolve. Second, while extant works consider the levels of publicity as well as of audience costs fixed largely by regime types, leaders are able to manipulate public attention with careful plans and thus manage their exposure to political costs. Previous examples show that policy makers are able to manoeuvre

political resources to increase or decrease public attention to foreign crisis, by repeated statements, the support of political party, the control of bureaucracy, etc. If we take into account the strategic control leaders have over public attention, the effectiveness of public threats in revealing information may no longer be an advantage exclusive for democratic leaders.

This dissertation studies how leaders use public threats in crisis bargaining using a combination of theoretical and empirical models. First, after carefully reexamining the key assumptions of audience costs theory, it establishes a new crisis-bargaining model to show how leaders strategically use public threats to address foreign opponents' uncertainty about the public's resolve. The new model enables us to identify when the uncertainty about the public's resolve matters in crisis bargaining, as opposed to the uncertainty about leaders, and to outline the conditions under which increasing public attention signals the public's resolve. Second, it shows how the interactions between leaders and the public can reveal to foreign opponents private information about the public's resolve. By separating the public's preferences about policy substance from its concerns about national prestige, it highlights two factors that explain the effectiveness of public threats in reducing uncertainty: political accountability to the public's policy preferences, and the distribution of military power between states.⁵ I construct a new index of political accountability and test the key hypotheses derived from the theoretical model. Finally, it explores the role of public threats in crisis bargaining if leaders are

⁵Throughout, I use "political accountability" or "accountability" for short.

also uncertain about their citizens' resolve. Whether the public reveals its true resolve, or lack of it, depends on how the public anticipates its leader to manipulate public attention. In particular, it shows how the tendency of mobilizing public attention under all circumstances may prevent leaders from learning the public's true preferences.

1.1 Rethinking Audience Costs Theory

This dissertation examines the domestic mechanism of signaling resolve in international crisis. While there are various domestic actors that affect leaders' international behaviors, such as constituents (Slantchev 2005, Smith 1998), political elites (Saunders 2015, Weeks 2008) or opposition parties (Arena 2015, Ramsay 2004, Schultz 1998), I focus on the domestic interaction between leaders and the general public, excluding the effects of other political incentives inherent in opposition parties and political elites. Consistent with audience costs theory, the dissertation begins with the notion that domestic political punishment can serve as a costly signal in international bargaining, but it further discusses challenges to audience costs theory by revisiting three questions: Whose resolve do leaders signal? Why does the public punish leaders? How do leaders signal resolve?

What is resolve in international relations? Kertzer (2016) summarizes three camps studying the concept of resolve. The first camp, most about deterrence theory, highlights the importance for states to build or maintain "reputations for resolve." Since it is related to emotional belief, scholars in

the first camp also suggest the difficulty of changing others' perception of resolve.⁶ The second camp is the crisis bargaining literature in which resolve is considered as an actor's private information, indicating the extent to which the actor is willing to take risk in exchange for advantages in negotiations.⁷ In the third camp, scholars of international security define resolve as "willingness to suffer" and use this concept to explain military outcome. Scholars in the third camp also examine resolve at a "lower level of analysis" such as military units or "the public at large."

While these camps study different aspects of resolve, they are not necessarily contradictory with each other. Resolve can be private information that helps an actor advance its bargaining power in negotiations. In the context of security crisis with risks of war, the willingness to "take risk" is equivalent to the willingness to "suffer from war."⁸ From the perspective of uninformed actors, learning one's resolve can be a lengthy process, depending on their prior beliefs and the informed actors' incentives. Finally, unless a state is assumed as a unitary actor, students of international relations should specify whose resolve they examine. My definition of resolve integrates insights from all the three camps. With the focus on crisis bargaining, I define a state's resolve as its willingness to suffer from war if it cannot achieve a peaceful agreement with

⁶See Mercer 1996; Press 2005; Hopf 1994.

⁷See Jervis 1970; Schelling 1966; Fearon 1997; Snyder and Diesing 1977; Fearon 1994; Fearon 1995.

⁸As Kertzer (2016) recognizes as well, rationalist scholars of conflict have long integrated crisis bargaining models and models of military outcomes: negotiations during crises are affected by the anticipation of war outcome, whereas war is also a continuation of the bargaining process. For example, see Wagner 2000; Filson and Werner 2004.

other states. Conceptually, resolve refers to morale, determination, “nerves,” or sensitivity to the loss in war, which is separate from capabilities or material resources such as military power, population, and wealth. Considering resolve as a cost of potential outcomes, I assume that resolve is situational, varying with specific crises and disputed goods: actors are more resolute if the stakes are high. In this sense, I do not distinguish between an actor’s resolve and the stakes (or valuation) of goods in the following theories.⁹

Assuming that states are unitary actors, models of crisis bargaining examine why private information about resolve may lead to war and how states signal their resolve (Fearon 1995, 1997). If we investigate a domestic mechanism of signaling resolve, it is essential to unpack unitary states and specify whose resolve is signaled to foreign opponents. Resolve refers to the willingness to bear costs of war. If breaking unitary-state assumptions, we should specify which actor—the leader or the public—bears costs in the event of war. The majority of works on audience costs either assumes unitary states or imposes costs of war on leaders without further reasoning (Debs and Weiss 2014, Ramsay 2004, Slantchev 2006, Smith 1998). In the conflict literature, the term—“the cost of war”—often indicates military fatalities and economic destruction. It refers to the societal aspects of costs “imposed on the population” (Fearon 1995, pp. 379). While leaders may have to bear political costs for fighting, it is the public who directly suffers from the societal costs of war

⁹Alternatively, resolve can be considered as an enduring trait. Most works in this category examine reputations for resolve.

associated with casualties and economic destruction (Chiozza and Goemans 2004). In this dissertation, I investigate how leaders signal the public’s resolve to foreign opponents during security crises. Focusing on the public’s resolve has two advantages. First, it is consistent with a state’s resolve—the societal cost of war—in the unitary-actor framework. Second, it facilitates our understanding of how the public weighs fighting against a negotiated settlement, which, as I argue next, affects its evaluation of the leader’s performance.

A costly signaling mechanism requires an understanding about the source of extra costs. The conventional audience costs theory suggests that the public imposes extra costs on leaders for bluffing due to its concerns about “international credibility, face or honor” (Fearon 1994, pp. 581). Therefore, the conventional source of extra costs originates from the public’s concerns about the consistency of leaders. If a leader issues a threat and then backs down, the public can punish the leader for the latter’s inconsistency between words and deeds in order to restore national prestige. According to this logic, backing down is rarely a desirable option. Recent survey experiments have confirmed that the concerns of consistency or prestige affect individual evaluations of how their leader handles foreign crises (Levy et al. 2015, Tomz 2007). In addition to prestige concerns, experimental works show that individuals’ policy preferences also affect their approval of leadership: dovish people tend to disapprove leaders of using force or threatening to use force (Kertzer and Brutger 2016).

Recent experimental works imply that there exist two competing log-

ics of domestic punishment. If the public is concerned about consistency, the leader is always subject to extra costs for backing down from an initial threat: this is the conventional logic of audience costs which increase the credibility of threat. If the public is concerned about the substance of crisis outcomes—a peaceful settlement or an escalation to war, backing down can be a desirable outcome for those unwilling to bear the costs of war. Consider the earlier example about the UN referendum in Taiwan. According to several surveys before the referendum, the majority of Taiwanese respondents agreed to increase the “international space” of Taiwan, that is, more recognition and participation in intergovernmental institutions, but they became more cautious when asked about potential conflicts with mainland China. In the question about whether respondents would be willing to join war if mainland China attacked after the referendum, only 37% of responses provided a positive answer. Only 30% of respondents agreed to conduct the referendum after the open opposition from the U.S.¹⁰ Therefore, it is essential to disentangle these competing logics in the new mechanism. The public is considered not only as a distant audience judging its leader’s performance, but also as a direct participant in the event of war. For the second logic, the public’s resolve is the key to understand how the public weighs militarized conflicts against backing down with a peaceful settlement.

Finally, the literature on audience costs seems unclear about how the

¹⁰Mainland Affairs Council, Republic of China (Taiwan), 2008. 2007 Opinion Polls Analysis on Cross-Taiwan Strait Relations (in Chinese). <http://www.mac.gov.tw/ct.asp?xItem=56143&ctNode=6333&mp=1>

amount of extra costs varies, which affects the strength (or costliness) of signals delivered to foreign opponents. I argue that the problem lies in the ambiguity between a public event and a public threat. In his original piece on audience costs, Fearon (1994, pp. 577) recognizes the public aspects of crises: “measures such as troop deployments and public threats make crises public events in which domestic audiences observe and assess the performance of the leadership.” Nonetheless, Fearon and other scholars view the level of publicity constant across crises and attribute the variation of extra costs only to political or institutional differences, namely, the audience’s capability to punish their leader. Their assumptions of publicity create two problems. First, despite the ability to impose political punishment, citizens in democracies usually pay low attention to political affairs (Baum 2004, Zaller 1992). Hence, democratic citizens may be unable to generate sufficiently large political costs to make costly signaling work. Second, private diplomacy can be effective in settling a peaceful agreement or achieving cooperation (Kurizaki 2007, Yarhi-Milo 2013). In fact, most challenges to audience costs theory are concentrated on whether democratic citizens are (more) capable of punishing their leaders. What if there are other factors that affect the variation of political costs?

In addition to political differences, the level of publicity—how much attention the public pays to a crisis—determines the amount of extra costs imposed on leaders. While the effects of regime types on extra political costs are largely exogenous, national leaders have much stronger control over the level of publicity. For instance, authoritarian leaders increase public atten-

tion and involvement in foreign affairs by granting unusual permission to anti-foreign protests (Weiss 2013). Nonetheless, I argue that the strategic control of publicity is *not* a unique advantage for authoritarian leaders. Although democratic leaders are much more limited in controlling mass media or repressing the freedom of speech, they are able to influence the level of publicity through their interactions with mass media and the resources they can employ from administrations or political parties. During the Cienfuegos crisis, the Nixon administration was able to prevent a “crisis mood” among the public even after news media were aware of Soviet Union’s nuclear facility. Kissinger and the NSC staffs were cautious about the wording and tones in their statements and constantly declined to disclose further details to news media. Hence, the Cienfuegos incident was not exactly “private,” but it managed to receive low public attention throughout the negotiation process.

I argue that the distinction between public and private diplomacy is misleading in studying the domestic signaling mechanism of crisis bargaining. On one hand, leaders from neither side can keep a crisis completely private, as the underlying issue of a crisis (e.g. the development of military facilities in Cuba, the promotion of formal independence of Taiwan, the humanitarian crisis in Syria) is often a heated topic to the public and mass media.¹¹ On the other, despite the public aspects of crisis, actual exchanges between leaders are rarely available to the public immediately. During the Cuban Missile Crisis,

¹¹If a negotiation can be maintained in complete secrecy, the issue involved can hardly generate a crisis, such as Nixon’s secret visit to the People’s Republic of China. See Yarhi-Milo 2013.

what the American public observed was merely an open warning by President John Kennedy, rather than the exact negotiation process between the U.S. and the Soviet Union. What affects extra political costs imposed by the public, I argue, is not a leader's choice between public and private diplomacy, but a leader's choice of publicity given the occurrence of a crisis. If both democratic and authoritarian leaders are able to manipulate the publicity of threats, the amount of extra costs conveyed in signals is determined both by regime types and by the endogenous publicity. Modeling the control of publicity may offer an integrated explanation about why democratic leaders do not have a unique advantage in signaling resolve.

1.2 The Argument

To further motivate research questions, Chapter 2 lays out more details about the challenges to audience costs theory and the new perspectives presented in the previous section. I propose two types of political costs leaders are subject to: audience costs and policy costs. Audience costs—matching the public's concerns of national prestige—are imposed when a leader's action is inconsistent with his initial statement. Policy costs are imposed when a leader's decision between war and peace does not match the public's preference about crisis outcomes. I further suggest that policy costs indicate a leader's political accountability to the public's policy preference. After clarifying the public's concerns and the associated political costs, I begin with a stylized model of crisis bargaining, in which the public expresses its policy preference during the

bargaining process. Introducing the public's resolve and policy costs provides new insights to audience cost theory. A leader's bargaining power may be undermined by his political accountability to the public's policy preferences. If the public is considered unwilling to fight and allowed to punish its leader for implementing threats, the leader's bargaining power shrinks as her policy costs increase.

Chapter 3 presents the core theory on how leaders use the publicity of threats to signal the public's resolve in crisis bargaining. Building upon the baseline model in Chapter 2, the full model not only allows a public to be concerned both about national prestige and about policy substance, but also enables a leader to choose the level of public attention. Once the model takes into account the variation of the public's policy preferences, increasing public attention does *not* always tie a leader's hands. Analyzing the model under complete information yields an important finding different from the conventional wisdom: going public ties a leader's hands only when the leader is accountable to a hawkish public rather than a dovish public. This deviation from the conventional wisdom thus changes our understanding about the signaling behaviors of leaders.

When the public's resolve is unknown to the foreign opponent, the foreign opponent can reduce its uncertainty by observing the leader's strategy of publicity, which depends not only on the leader's political accountability to the public's policy preferences but also on the distribution of power. First, greater publicity does not always signal the public's resolve. When leaders

are highly accountable (e.g. democratic leaders), surprisingly, decreasing the level of publicity signals the public's resolve. Second, democratic leaders do not always have an advantage in signaling the public's resolve. The effects of political accountability on credible signaling are conditioned on the distribution of power between a challenging state and a foreign target. When the foreign target is relatively strong, democratic leaders—highly accountable to the public's policy preferences—credibly communicate the public's resolve and therefore avoid the escalation of crisis. When the foreign target is relatively weak, democratic leaders are no better at signaling the public's resolve than less democratic leaders. With a resolute public, democratic leaders may choose to hide the true information, provoking the resolute public to support war. Without directly suffering from war, leaders are better off by fighting with public support if the foreign target is relatively weak.

Taking into consideration different sources of public preferences, the theory presented in the previous chapter yields two new findings. First, without the pressure from the domestic public, unaccountable leaders achieve peaceful agreements with no need to signal. Second, the credibility of highly accountable leaders in communicating resolve is conditioned on the distribution of power: in particular, these leaders have *no* advantage in delivering credible signals when their countries have strong military capabilities relative to those of targets. Chapter 4 provides the empirical support for the second finding: the effects of political accountability on bargaining outcomes are conditioned on the distribution of power between adversaries. When a chal-

lenging state is relatively weak, the political accountability of its leader has a non-linear relationship with crisis escalation: the risk of war reaches the highest when the leader is moderately accountable, but declines when the leader becomes highly accountable or not accountable at all. When a challenging state is relatively strong, increasing the leader's political accountability has no effect in delivering credible threats: the pacifying effects of high accountability disappears. By modifying one of the most recent measures of Audience Costs Capacity, I construct a new index for political accountability. Using the Militarized Compellent Threat dataset, I examine how political accountability affects the likelihood that a challenger's threat receives compliance from its target. The empirical result is consistent with the propositions above: a challenger's political accountability affects the outcome of threats, yet such effect is conditioned on the relative strength between a challenger and its target.

Chapter 5 extends the previous model by (a) further assuming that the public's resolve is unknown both to leaders and to foreign opponents, and (b) making the domestic public express their preferences about crisis outcomes before the leader's and the foreign target's moves. The modified model shows that war is possible under complete information, as the public support for war may make peace costly to *leaders* and thus eliminate the bargaining range with foreign opponents. This chapter further explores the conditions under which the public's expression reveals or masks its true preference about war and peace. I narrow down the scope such that the public's support for war can eliminate the bargaining range between two states, as I am particu-

larly interested in whether war may occur due to a leader's overestimation of the public's resolve. The modified model shows that the public will express their true preferences only when the leader makes distinct choices of publicity based on their expressions: maximizing public attention given their support for war, yet minimizing public attention given their opposition against war. Without such distinction, the public has few incentives to reveal its true preference. Most importantly, the public will always appear supportive of war if the leader is not expected to reduce the level of publicity after anti-war expressions. Amplified by a high level of publicity, anti-war expressions would make the leader "look bad" in front of foreign opponents and potentially lead to an offer worse than fighting. The leader's tendency to constantly engage public attention may prevent the irresolute public from revealing its dovish preference. It yields some new implications to the research on nationalism or public mobilization. The undifferentiated tendency of mobilizing the public in crisis can be ineffective or even counterproductive, not only because it appears incredible to foreign observers but also because it prevents leaders from learning the true preferences of their people.

Chapter 6 summarizes the dissertation's contributions to the literature of international conflict and costly signaling, suggesting directions for future research. It presents a more integrated framework which reconciles the conventional audience costs theory and subsequent critics. The main findings reinforce the conventional wisdom that signaling resolve increases a leader's credibility in crisis bargaining, but they challenge the linear linkage between

increasing public attention and signaling resolve. The conventional mechanism of audience costs—increasing public attention signals resolve—exists when leaders are moderately accountable (i.e. to the resolute type of people only). When leaders are highly accountable (i.e. to all types of people), increasing public attention makes backing down easier and thus signals flexibility. Therefore, highly accountable leaders will decrease public attention to enhance credibility and bargaining advantage. In terms of future research, I first discuss how experimental works and theory development can reinforce each other in advancing our understanding on the domestic impact in crisis bargaining. I also suggest more research on how leaders make public threats, including data collection and more experimental works on how individuals evaluate their leaders conditioned on different levels of attention. Finally, I explore the implications in this dissertation to U.S. foreign policies and the growing trend of nationalism and populism.

Chapter 2

Political Costs, Publicity, and Crisis Bargaining

Why do leaders choose to increase public attention during crisis bargaining? How does greater public attention affect the outcomes of crises? The conventional wisdom suggests that increasing public attention, as a costly signal of resolve, advances a state's bargaining advantage in crisis bargaining. The underlying rationale is that bargaining in public may generate so-called audience costs on the challenger who backs down from its initial threat, making it more likely to escalate and thus signaling greater resolve (Fearon 1994, 1997). To further examine audience costs as a domestic mechanism of costly signaling, subsequent works explore mainly two questions: why backing down is considered undesirable and how political costs are imposed on leaders. In this chapter, I review the literature on both questions and discuss their limitations. In particular, I challenge the notion why backing down is considered unpopular *only* and discuss when standing firm may incur other political costs as well. Second, while the literature attributes the variation of the amount of political costs to political institutions, I explore the role of leaders in determining the amount of political costs. These discussions propose alternative perspectives which lay the ground for new assumptions and models in the

following chapters.

I first review the literature on the domestic attitudes towards backing down from initial threats. While the majority of works attributes the unpopularity of backing down to the concerns of national prestige or reputation, recent works point out that the preferences about policy substance also shape the public's attitudes towards backing down. The concerns of national prestige refer to the (in)consistency between a leader's words and deeds, which affects national honor and reputation in future bargaining. The concerns of policy substance emphasize the public's calculation of benefits and costs between escalation and a peaceful agreement in a given crisis. However, the extant literature treats the public's policy preference as fixed or exogenously determined. I argue that it is critical to directly model the variation of the public's preference of policy substance, not only because it affects the public's attitudes towards backing down but also because it constitutes the essence of "resolve"—a major source of uncertainty in crisis bargaining—at the state level. I introduce a baseline model with the following features: First, the attitudes towards backing down are determined both by reputation concerns and by policy preferences; Second, the public's policy preferences are shaped by its payoffs from war and peace.

When examining *how* audience costs are imposed on leaders, scholars on audience costs theory debate about whether democratic leaders are indeed subject to greater audience costs than their authoritarian counterparts. The general notion is that the more democratic the challenger is, the more polit-

ical costs are generated and thus the more effective a signal is. Challenges to this notion suggest either that citizens in democracies are unable to generate high audience costs due to low attention to politics (Baum 2004, Zaller 1992), or that autocratic leaders are also subject to audience costs under certain circumstances (Weeks 2008). Nonetheless, the existing debate focuses on the exogenous effects of regime types on the domestic audience, such as how much the audience is capable of coordinating to impose political punishment or whether the audience pays attention to foreign affairs. Instead of focusing entirely on the audience side, I argue that individual leaders, independent of institutional variations, are able to adjust political costs by altering the level of public attention to crises. The amount of political costs is jointly determined by exogenous political environments as well as the level of public attention chosen by leaders.

Before I move on, it is worth mentioning that I focus on the general public as the domestic audience. Scholars refer domestic audiences to a variety of actors, such as constituents (Slantchev 2005, Smith 1998), political elites (Saunders 2015, Weeks 2008) or opposition parties (Arena 2015, Ramsay 2004, Schultz 1998). In this dissertation, I concentrate on the general public and its preferences, therefore excluding the effects of other political incentives inherent in opposition parties and political elites (e.g. political competition).

2.1 Why is Backing Down Undesirable?

The enduring problems in conflict studies are why war occurs if a peaceful agreement is *ex post* less costly and how states can prevent it. One of the general explanations attributes the outbreak of war to the problem of uncertainty among states. War occurs when one has some unobservable characteristics, unknown to the other, and has an incentive to deliver wrong information to advance its own bargaining leverage (Fearon 1995). The incentive to hide true information leads to ineffective communication and a risk of war. Under this logic, the tragedy of war can be prevented if states solve the problem of ineffective communication. A key unobservable characteristic in the choices between war and peace is one's subjective assessment of the costs of war, or resolve. The smaller the costs a state assesses for using force, the more resolute the state is. How do states address the uncertainty over each other's resolve in international crises? To make communication more effective, in principle, a resolute type must distinguish itself from an irresolute type with costly signals: a signal must involve sufficient costs such that the irresolute type chooses to reveal the true information rather than pretend the resolute one.

Audience costs theory proposes a domestic signaling mechanism in crisis bargaining. In principle, when a state threatens to use force, it has stronger incentives to carry out the threat if it engages additional costs for backing down. With all else being equal, as the amount of additional costs increases, the state will be more likely to escalate when its demand is rejected, implying that its communication is more credible (Fearon 1997). If the amount of ad-

ditional costs is sufficiently large, a resolute state is able to fully distinguish itself from an irresolute state, solving the informational problem. While the additional costs for backing down can be imposed either by domestic or by international audience (hence “audience costs”), the majority of literature focuses on the domestic side. How does the amount of audience costs increase? The conventional wisdom suggests that it increases with the audience’s capability of punishing its leader, determined by external political institutions.

But why is backing down a bad idea for the public? The original discussion highlights the “loss of international credibility, face or honor” (Fearon 1994, pp. 581). Since escalating and then backing down hurt “the state’s reputation for threatening the use of force only when serious” (*ibid*), a leader is judged to have performed poorly for making an empty threat. Citizens have an incentive to remove leaders caught bluffing and thus restore credibility (Guisinger and Smith 2002). In other words, a leader is able to signal his resolve because his audiences care about the values of consistency.

Another reason, followed with more theoretical development, is associated with the competence of leaders. However, scholars vary in terms of how they think the public evaluates “competence.” First, some argue that a leader is considered incompetent when people “find out that his behavior was different from what they would have done if they had the same information” (Downs and Rocke 1995, Slantchev 2006). It implies that in the decisions of war and peace, people may have different payoffs of war from their leader.¹ In the con-

¹In the rationalist framework, the difference may result from the disagreement over the

text of crisis bargaining, while backing down may render a leader incompetent among resolute audiences, carrying out threats may also be labeled incompetent if the public believes that peace is more preferable. Modeling competence in this way would require some prior knowledge of what the public prefers in the first place. In other words, inconsistency may not be politically costly to leaders once we relax the assumption that the public always rejects the idea of backing down.

Recent survey experiments offer new insights to reconcile the contradictory assumptions about domestic preference. Consistency is only one of the dimensions upon which people evaluate their leaders. On the one hand, various studies show that people care about whether the leader is consistent between words and deeds (Davies and Johns 2013, Tomz 2007). On the other, there is evidence that in addition to consistency, preferences over war and peace also affect the public approval of their leader (Chaudoin 2014, Kertzer and Brutger 2016). What do these experimental findings imply to theorists? In addition to the costs of backing down after an empty threat, leaders also suffer from launching escalation with which the audiences disagree. If leaders worry about the costs of backing down alone, the case where Taiwanese forced their president to back down is puzzling. However, it makes sense if we consider the strong domestic opposition against war which the former president could not afford to ignore. To understand how the public influences its leader, we need to theorize what people prefer in the first place. That is, instead of assuming

costs of war or the valuation of goods or both.

“backing down is a bad policy”, we need to specify when the public prefers escalation, which involves militarized conflicts, to backing down. Hence, modeling the public’s payoffs should focus on its role as a participant in potential military conflicts rather than as a distant audience who is aloof from war and only judges its leader after a crisis ends.

2.2 Audience Costs vs. Policy Costs

To examine the domestic mechanism of audience costs, scholars argue that audience costs originate from the concerns about incompetent leaders to whom people delegate the power of foreign policies (Slantchev 2006, Smith 1998). As people are uncertain about the competence of leaders, a competent leader uses public threats to separate himself from an incompetent one. However, if a leader is considered incompetent when people “find out that his behavior was different from what they would have done if they had the same information” (Downs and Roche 1995), it is unclear why backing down is always considered undesirable by the public and hence implies incompetence. Recent experimental studies confirm the conjecture that people worry about the loss of prestige or reputation (Levendusky and Horowitz 2012, Tomz 2007), but some also provide new evidence that individuals who prefer peaceful solutions disapprove of a leader committed to escalation (Kertzer and Brutger 2016).² Independent from the concerns of national honor, people vary in the degrees of tolerance of war, which shapes different preferences for war and peace. If

²See a non-militarized example in Chaudoin (2014).

the public imposes audience costs on leaders for backing down due to prestige concerns, it is reasonable to contend that the public can also punish leaders for the consideration of “policy substance” Snyder and Borghard (2011) or “merits of policy” in terms of national interest (Levy 2012).

To my knowledge, the literature on audience cost theory has not explicitly modeled the variation of the public audience’s preferences for war and peace. Most works on the domestic origin of audience costs assume a “good” policy exogenously given, as they focus on how the public overcomes informational asymmetry to learn the competence of leaders (Debs and Weiss 2014, Slantchev 2006). Instead, I explicitly model the public’s preferences for war and peace by adopting Fearon’s rationalist framework yet locating the costs of war in the domestic public. Fearon (1995) first assumes that states are unitary actors and models how unitary states calculate benefits and costs of war. The assumption—a state bears the costs of war—is appropriate if one seeks to understand war and peace on the level of interstate interaction. If we are interested in how the interactions between leaders and the public affects war and peace, we should reconsider which actor—the leader or the public—bears the costs of war. According to Fearon’s notion on the “second-image” mechanism, “[...] [leaders] will enjoy various benefits of war without suffering costs imposed on the population (Fearon 1995, pp. 379).” In other words, if we unpack a unitary state, it is people rather than their leader who bears the *societal* costs associated with casualties and destruction (Chiozza and Goemans 2004,

Fearon 1995).³ The minimum offer people accept depends on how many costs they are willing to bear in the event of war. Small costs of war indicate the public willingness to fight, or what I call the public's resolve. Modeling the public's resolve allows me to specify the conditions under which people support backing down and would impose punishment if leaders choose the opposite.

Therefore, I assume that the public is concerned both about national prestige and about policy substance in the following models. The concerns of national prestige refer to the (in)consistency between a leader's words and deeds, which affects national honor and reputation in future bargaining. Under this logic, it is undesirable if a leader's subsequent behavior is inconsistent with his statement. This is consistent with the original audience costs theory: after a leader threatens to use force, backing down is undesirable because it hurts national honor and future bargaining.⁴ The concerns of policy substance indicate the public's calculation of benefits and costs between escalation and a peaceful agreement in a given crisis. Similar to the rationalist unitary-state framework (Fearon 1995), whether the public prefers crisis escalation or a peaceful settlement depends on the relative military capabilities and the costs of war the public is willing to bear.

The analytical distinction between prestige concerns and policy sub-

³Unless specified, the costs of war in this article refers to the societal costs as indicated above.

⁴Although this is not the focus here, according to the logic of national prestige, it is also undesirable if a leader launches militarized conflict after being committed to a peaceful negotiation.(Levy et al. 2015)

stance is crucial: while the pressure of prestige concerns always discourages leaders to back down, the preferences of policy substance may or may not work in the same direction, depending on whether people prefer peace to war. Indeed, the UN referendum was invalid because the Taiwanese public were unwilling to fight against mainland China. I assume that leaders are therefore subject to two types of political costs due to different concerns: (a) the concern of national prestige generates audience costs that always encourage leaders to stand firm, and (b) the concern of policy preferences generates policy costs which may or may not encourage leaders to back down, depending on the public's resolve. The greater policy costs are, the more accountable a leader is to the public's preference for policy substance.

2.3 The Baseline Model

To illustrate how two types of political costs differ, below I present a baseline model with complete information. Two states—Home (H) and Foreign (F)—have disputes over some divisible goods, such as territory, policies, rights and privileges, etc. The disputed goods are currently under F 's control, with the value normalized to one. While F is a unitary state, H is composed of a leader (L) and a domestic audience (A). The model begins with the foreign opponent's response to the leader's challenge.⁵ F proposes to H a new share,

⁵In the full model, I am interested in how leaders choose the publicity of threats. Whether leaders make threats is not the focus of this chapter.

$x \in [0, 1]$.⁶ Observing the proposal, A chooses whether to support fighting before L decides whether to stand firm or back down. F receives $1 - x$ if the game ends with peace; otherwise it receives its payoffs of war, $1 - p - c_F$. L either gains her new share after a peaceful agreement, or fights a winner-take-all war, defeating F with probability p . L 's political costs depend on A 's prestige concerns and policy preferences. If A does not support war, L pays audience costs, r , for backing down, or policy costs, d , for standing firm. If A supports fighting, L incurs both r and d for backing down as it makes an empty threat *and* fails to deliver what A prefers; otherwise, L escalates and receives its war payoff: with probability p , L wins the whole share and with probability $1 - p$, it pays d for losing a popular war. A bears the costs of war, c_A . Whether A supports fighting depends on its new share of goods, x , on one hand, and its payoffs of war, $p - c_A$, on the other. Figure 2.1 presents the baseline model.

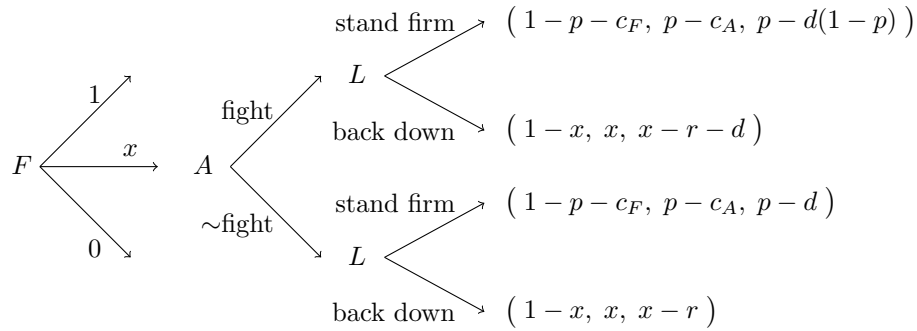


Figure 2.1: The Baseline Model

From F 's perspective, the minimum offer to avoid war makes both A

⁶Similar to Tarar and Leventoglu (2009), I treat the disputed goods divisible in all the models.

and L prefer peace to war, that is, A does not support fighting *and* L backs down. The first condition requires an offer no less than $p - c_A$, which I call A 's reservation value. Similarly, the second condition yields L 's reservation value, $p - d + r$. To avoid war, F picks the *greater* reservation value between two options so that both L and A would prefer a peaceful agreement to a costly war. If A has the greater reservation value, F proposes $p - c_A$, which also prevents L from standing firm. If L has the greater reservation value, offering $p - c_A$ would not be enough to achieve peace because L is able to withhold domestic oppositions against war and still chooses to stand firm; instead, F must offer $p - d + r$.⁷ It is obvious that which offer F proposes depends on the relative values between c_A and $d - r$. Figure 2.2 illustrates a non-linear relationship between F 's offer (x) and L 's policy costs (d).⁸

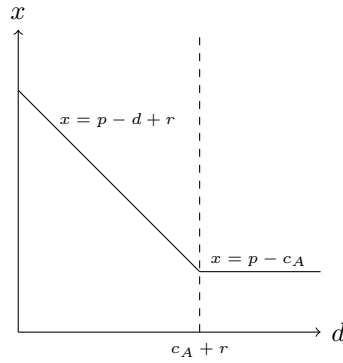


Figure 2.2: F 's Offer

⁷Formally, the following strategies constitute F 's best response functions in the subgame perfect equilibrium of the baseline model: when $d \geq c_A + r$, F proposes $x = p - c_A$; when $d < c_A + r$, F proposes $x = p - d + r$. See a complete proof in Appendix A.

⁸I assume that $p = 0.6$, $c_A = 0.4$, and $r = 0.2$.

Introducing the public's resolve and policy costs provides new insights to audience cost theory. First, a leader's bargaining power may be undermined by her political accountability to the public's policy preferences. Once we allow the public to consider the costs of war and to punish leaders for standing firm when they do not support fighting, the leader's bargaining power shrinks as her policy costs increase. In particular, if policy costs are small ($d < c_A + r$), F proposes L 's reservation value, $p - d + r$, which decreases with the values of policy costs. L 's bargaining power decreases with its policy costs: the more L is accountable to the public's policy preferences, the weaker bargaining power L has. In other words, an authoritarian leader may be able to take a tough position and thus gain a better deal than a democratic leader as the former is less constrained by the domestic inclination against a costly war.

Second, though no war occurs under complete information, the baseline model highlights whose preferences (and the knowledge about them) matter in the bargaining outcome under different situations. As shown above, if $d < c_A + r$, F achieves a peaceful outcome by proposing L 's reservation value, implying that the key to avoid war lies in the information about L 's political costs rather than the information about the public's costs of war. Whether F knows the value of c_A does not affect its bargaining strategy and the overall outcome. However, if $d \geq c_A + r$, A has the greater reservation value and therefore, F must propose $p - c_A$ to secure an agreement, which does not increase with L 's political costs. F 's full knowledge of c_A is necessary to guarantee a peaceful outcome.

The second finding sheds lights on the debates over whose resolve—leaders’ or people’s—matters more in international conflicts (Chiozza and Goemans 2004, Horowitz, Stam and Ellis 2015, McGillivray and Smith 2008, Wolford 2007).⁹ In reality, the opponent’s uncertainty about resolve can exist in both sources and affect the probability of war. Overall, the knowledge about the public’s resolve matters more if the opponent bargains with democracies, the knowledge about individual leaders is more critical if it deals with autocracies.

2.4 Endogenous Publicity and Political Costs

Although the origin of audience cost theory has little to do with the debate of democratic peace (Schultz 2012), the majority of the subsequent works focuses on whether democracies are capable of generating larger audience costs than authoritarian regimes. On one hand, scholars modify or strengthen the original theory by specifying the conditions under which democracies can generate greater audience costs or providing empirical evidence. Democratic leaders are subject to audience costs or domestic punishment with the presence of free media (Slantchev 2006) or opposition parties (Ramsay 2004). Recent empirical works provide evidence that democratic leaders are indeed imposed audience costs in the form of domestic legislative failure (Gelpi and Grieco 2015). On the other hand, other works are skeptical about the evidence of

⁹The original debate puts individual leaders in contrast with unitary states. As discussed above, the resolve of states resides in the public when we examine the domestic mechanism of states.

audience costs in democracies or the unique advantage that democratic leaders have in generating audience costs. Some scholars argue that large audience costs would not be generated in democracies because people tend to pay low attention to political affairs (Baum and Potter 2008, Trachtenberg 2012), whereas others suggest that autocracies may be as capable of generating audience costs as their democratic counterparts (Weeks 2008, Weiss 2013). As a result, there is no agreement regarding whether democratic leaders are more capable of signaling resolve in crisis bargaining (Downes and Sechser 2012, Kurizaki and Whang 2015, Moon and Souva 2016). To reconcile the disagreement, I have proposed another dimension of political costs—policy costs based on the public’s policy preferences—in addition to the conventional audience costs. In this section, I further argue that another problem is the insufficient attention to the influence of leaders in managing the level of public attention.

Leaders are capable of adjusting political costs by controlling the publicity of threats. Extant works usually consider political costs as exogenously determined by regime types or institutional arrangements (Slantchev 2006, Smith 1998). Recently, scholars also argue that more salient issues generate greater audience costs in democracies (Gibler and Hutchison 2013). However, neither political costs nor public attention is entirely predetermined and moreover, they are connected.¹⁰ A recent experiment shows that a stronger statement of threat results in more public disapproval if leaders do not follow

¹⁰While some scholars study the exogenous variation of public attention, for example, the access to mass media (Potter and Baum 2014), I focus on the leader’s endogenous management of public attention.

it through (Trager and Vavreck 2011). Leaders can choose not only whether to make public or private threats (Kurizaki 2007), but also *how* to make public threats. For example, leaders can decide the rank of speaker, the context of statements (McManus 2016), and the tone or rhetoric of threats (McManus 2014); they can create other opportunities to increase the salience of crisis, such as referendums, congressional debates (Debs and Weiss 2014),¹¹ the permission of anti-foreign protests (Weiss 2013).

Given the existence of a crisis, I argue that leaders can choose to increase or decrease the level of publicity to the crisis. First, leaders can decide how to make a public threat by choosing speaker, occasion, rhetoric, frequency, etc. For instance, the higher a speaker's rank is, the more attention the public will pay to the statement. During the Cienfuego crisis, the Nixon administration tried to draw this incident out of public attention, not to creating crisis mood among the public. Apart from Kissinger's ambiguous statements, President Nixon rarely made any public comments on this matter. Instead, Nixon remained impassive and maintained his planned schedule, including a trip to Europe (Crall and Martin 2013, pp. 196). Second, leaders can decide whether to engage additional mass mobilization to manipulate the level of publicity. Extant works suggest that both democratic and authoritarian leaders are able to mobilize public attention, though the specific techniques they adopt may

¹¹While Debs and Weiss (2014) discuss the role of public debates in revealing external circumstances—a component of the costs of war—to audience and foreign adversaries, they consider public debates as an unconditional feature of democracies, not as a leader's strategic choice.

vary. Since mass mobilization is usually forbidden or restrained under authoritarian regimes, authoritarian leaders are able to show their intention to draw public attention by temporarily permitting mass protests in foreign crises. This technique does not apply to democratic leaders: anti-foreign protests may well be spontaneous from the bottom, given an overall high level of political freedom in democracies. The techniques of additional mass mobilization available to democratic leaders include increasing the involvement of legislative body or political parties. Take the Cienfuegos crisis as an example again. One of the major concerns of Kissinger and the National Security Council was that the U.S. Congress would “build up” a crisis. In addition to few comments to the press, Kissinger refused to cooperate with the Congress when asked to share information (Siniver 2008).

Considering both exogenous environments and strategic choices reconcile the conventional audience costs theory and its critics. A leader’s threat to use force is more credible if he suffers from more political costs for backing down, which, according to the conventional wisdom, tends to occur in democracies.¹² In other words, threats issued by a democratic leader are more credible, advancing his bargaining position and reducing the risk of war. This notion receives theoretical challenges in two aspects. First, scholars disagree on the amount of political costs generated in democracies. Some suggest that citizens in democracies may be unable to generate large political costs as they

¹²See Schultz’s (2012) review about the development of audience costs theory, in particular, how the theory has become one explanation of democratic advantage in crisis bargaining or democratic peace.

pay low attention to political affairs (Potter and Baum 2014, Saunders 2015, Zaller 1992). Given a certain level of public attention, nevertheless, democratic leaders are able to further increase or decrease public attention through various measures listed above. Second, studies show that authoritarian leaders are sometimes able to tie their hands and signal resolve. By strategically raising public involvement in international crisis, authoritarian leaders increase the amount of political costs for backing down (Weiss 2013). In the next chapter, I will model a leader's political costs as an interaction between political environments and the leader's choice of publicity.

How does the endogenous control of publicity affect political costs? The magnitudes of policy costs increase when leaders choose high-profile threats. Associated with frequent media coverage, sustained open debates, etc, increasing publicity raises the importance of policy preferences and holds leaders more accountable to public preferences on policy substance. During the Fashoda crisis, when the foreign secretary of Britain, Lord Salisbury, released to the public the details of negotiation with France, he hoped that "people would be able to approach the issue in a more sober way."¹³ However, the magnitudes of audience costs decrease with publicity *if* people prefer peace to war. Recent experiments show that when people are asked to consider both consistency and policy preferences, the dovish respondents do not punish leaders for

¹³See Courcel to Delcassé, 6 October 1898, *Documents Diplomatiques Français (1871-1914)*, series 1, vol. 14, 633-35. Cited in Trachtenberg (2012).

backing down (Kertzer and Brutger 2016). Since great publicity makes policy preferences more important, national prestige becomes a secondary concern to the dovish people. With close attention, if the dovish people prioritize the preference for peace, they are *less* likely to impose audience costs on leaders for backing down.¹⁴ With low attention to crisis, however, people downplay the importance of policy preferences—which makes inconsistency a relatively salient issue—and thus blame leaders for damaging national prestige.¹⁵ Overall, if policy preferences are compatible with prestige concerns, that is, if people prefer war to peace, great publicity further raises the importance of both considerations. If the two concerns are incompatible, great publicity prioritizes the consideration of policy preferences as it makes the public put more weight on *current* policy substance—whether to fight or not—than on prestige concerns that may affect *future* bargaining.

Foreign opponents can therefore infer the public’s resolve by observing how leaders make public threats, for a leader’s political costs are associated both with the choices of publicity and with public resolve. The relationship

¹⁴It also applies to international audiences (Fearon 1994). If international audiences observe that the leader backs down without public support for fighting, they are more likely to believe that the leader is subject to domestic pressure instead of desire for bluffing.

¹⁵Building on Zaller’s (1992) work, some scholars suspect that people have independent preferences as they simply take cues from the elites they support (Berinsky 2009, Saunders 2015). However, Zaller argues that elite disagreement can divide public opinion of policies only among the politically aware, who “pay enough attention to elite discourse” and follow particular ideological implications; otherwise, people are more likely to evaluate leaders based on their own preferences. If leaders choose to increase public attention to crisis, with the incumbency advantage, they are able to reach not only the politically aware ones, but also a much broader public audience.

between leaders and the public is similar to a two-level game in the context of crisis bargaining (Putnam 1988), except that leaders do not directly suffer from war, which creates the possibility that war may be rational for leaders “if they will enjoy various benefits of war without suffering costs imposed on the population” (Fearon 1995, pp. 379). On the international level, a leader, constrained by the domestic public, bargains with a foreign opponent. On the domestic level, the public express preferences about whether to accept the opponent’s offer and decide, upon the leader’s final move, how to impose political costs on leaders. To minimize political costs, leaders have incentives either to signal or to mask the types of their people.¹⁶

2.5 Summary

In this section, I review the literature on the domestic signaling mechanism in crisis bargaining with a particular focus on audience costs theory. Through extra political costs imposed by the public for backing down, audience costs theory proposes a causal mechanism of how domestic politics *can* reduce uncertainty between states and thus affect the outcome of crisis bargaining. While audience costs theory introduces the possibility of solving informational problems of war through domestic politics, scholars in this research program disagree over when and why signaling resolve is more effective. I argue that the existing literature has not addressed two fundamental modeling questions

¹⁶In reality, a leader may not know the resolve of his people with certainty, but it is reasonable to contend that he has better knowledge than his foreign counterpart. See Tarar (2001).

in audience costs theory. First, why do leaders incur political costs only when they back down from an earlier threat? Can the public also impose political costs for carrying out a threat of using force? Second, how do the amount of political costs vary? Is it determined by external political difference only (i.e. regime type, institutional constraint, etc)?

I provide answers to these two questions by synthesizing the current literature on audience costs, especially recent experimental research on the theory's microfoundation. I first challenge the notion that the public always disapproves of backing down and introduce policy costs along with the conventional audience costs. While audience costs discourage leaders from backing down, how policy costs affect escalation decisions is conditioned on the public's policy preferences. I propose to explicitly model policy preferences by switching the locale of resolve from leaders to the public. Whether people support fighting depends on their willingness to fight or what I call the public's resolve, independent of the concerns of national prestige. I then construct a baseline model that enables the public to express its policy preferences and to impose both types of political costs on leaders. The baseline model shows that knowledge about the public's resolve is crucial to the bargaining outcome. Finally, I relax the assumption that political costs are purely determined by political institutions. Rather, leaders are able to adjust the magnitudes of political costs by choosing the level of public attention to their threats. By observing how leaders make public threats, foreign adversaries are therefore able to gain information about the public's resolve during crisis bargaining. In

the next chapter, I construct a crisis-bargaining model with the new features formalized.

Chapter 3

Seeking and Avoiding Accountability: Signaling Resolve to Foreign Targets

This chapter examines how a leader signals the public's resolve by choosing a level of publicity during crisis bargaining. Drawing key assumptions from the previous chapter, I construct a crisis-bargaining model that (a) allows a public to care both about national prestige and about policy substance, and (b) enables a leader to choose a level of public attention. While the public pays the costs for war if negotiation fails, the leader is subject to political costs for her choice between war and peace, related to two concerns of the public. First, the leader pays audience costs for hurting national prestige, that is, not executing the initial threat. Second, the leader pays policy costs for not following the public's preferred policy. If the public prefers war to a peaceful agreement, prestige concerns and policy preferences are compatible, as both concerns discourage leaders from backing down. If the public prefers a peaceful agreement to war, however, it suggests that the public's concerns are incompatible or competing.

The leader chooses a level of public attention to adjust political costs. When the public's concerns are incompatible, the model assumes that in-

creasing public attention makes the public put more weight on *current* policy substance—whether to fight or not—than on prestige concerns that may affect *future* bargaining. In this chapter, I analyze the model with complete information and one-sided information where a foreign opponent does not know the costs of war for the public (or the public’s resolve). In Chapter 5, I further relax the assumptions about the information structure, allowing the public’s resolve unknown both to its leader and to the foreign opponent.

Once the model takes into account the variation of the public’s policy preferences, increasing public attention does not always tie a leader’s hands. I first analyze the model with complete information. In equilibrium, a leader can tie her hands by increasing public attention to foreign crises if the public is dissatisfied with a peaceful agreement. If the public prefers a peaceful agreement to a militarized conflict, increasing public attention to foreign crises unties a leader’s hands instead. The conventional wisdom that going public ties a leader’s hands exists only when she is accountable to a hawkish public rather than a dovish public. This deviation from the conventional wisdom thus changes our understanding about the leader’s signaling behaviors.

When the public’s resolve is unknown to the foreign opponent, the foreign opponent can reduce its uncertainty by observing the leader’s strategy of publicity, which depends not only on the leader’s political accountability to the public’s policy preferences but also on the distribution of power. First, greater publicity does not always signal the public’s resolve. When leaders are highly accountable (e.g. democratic leaders), a high-profile threat may suggest

the lack of resolve among the public as the irresolute type of people would impose smaller audience costs on leaders for backing down. Second, democratic leaders do not always have an advantage in credibly communicating the public's resolve. The effects of political accountability on credible signaling are conditioned on the distribution of power between the challenging state and the foreign opponent. When the foreign opponent is relatively strong, democratic leaders—highly accountable to the public's policy preferences—credibly communicate the public's resolve and therefore avoid costly war. When the foreign opponent is relatively weak, a democratic leader with an resolute public chooses not to reveal the information about her people and therefore risks fighting a costly war with domestic support. In the latter scenario, fighting with public support brings leaders greater payoffs than a peaceful agreement does because it is easier to defeat a weak opponent.

3.1 The Model

I develop a Bayesian game in which (a) the opponent is uncertain about the public's resolve and (b) the leader chooses the publicity of threats before the opponent's offer. Both A and F receive the same set of payoffs as in the baseline model. The sequence of the model goes as follows. First, Nature (N) chooses A 's costs of war, generating an resolute public ($c_A = \underline{c}_A$) with probability ϕ and an irresolute public ($c_A = \bar{c}_A$) with the complementary probability, where $0 < \underline{c}_A < \bar{c}_A$. Next, L specifies a level of publicity, $0 \leq \lambda \leq 1$, when making public threats. F then proposes an offer based on which

A chooses whether to fight or not. L makes the ultimate decision of whether to stand firm (a costly war) or back down (a peaceful agreement). Nature reveals A 's type to L and A only, while the distribution of A 's type is common knowledge. Everything else is complete information for all players.

As discussed above, I assume that great publicity increases L 's policy costs for going against A 's choice, but decreases L 's audience costs if A does not support fighting and L backs down. If L does not follow A 's policy preference, greater publicity increases the amount of policy costs (d) by λ times. In particular, if L backs down after A 's opposition against fighting, greater publicity decreases the amount of audience costs (r) by λ times. Without loss of generality, I assume that L chooses between a low-profile threat ($\lambda = 0$) and a high-profile threat ($\lambda = 1$). When $\lambda = 0$, L receives the same political costs and benefits as in the baseline model. Below I discuss the changes of L 's political costs when $\lambda = 1$.

First, great publicity increases L 's policy costs when it goes against A 's choice, formally, by λ times. If A opposes fighting but L stands firm, a high-profile threat increases L 's policy costs by λ times and yields $d(1 + \lambda)$. More precisely, if L stands firm despite A 's opposition against fighting, a high-profile threat ($\lambda = 1$) yields $p - 2d$ for standing firm. If A supports fighting but L backs down, a high-profile threat doubles L 's costs—both policy and audience costs—and L receives $x - 2(d + r)$. Second, great publicity decreases L 's audience costs if A does not support fighting and L backs down. As discussed above, great publicity allows people to prioritize policy preferences and thus,

they are willing to reduce the leader's audience costs, by λ times. Thus, if A does not support fighting, L receives x for backing down after a high-profile threat ($\lambda = 1$). In other words, great publicity leads to *no* audience costs if L backs down with A 's support for peace, implying that F will *not* have to compensate L 's audience costs to persuade L to back down. Finally, if A chooses to fight and L stands firm, L receives $p - d(1 - p)$ regardless of its choice of publicity. Figure 3.1 presents the full model.

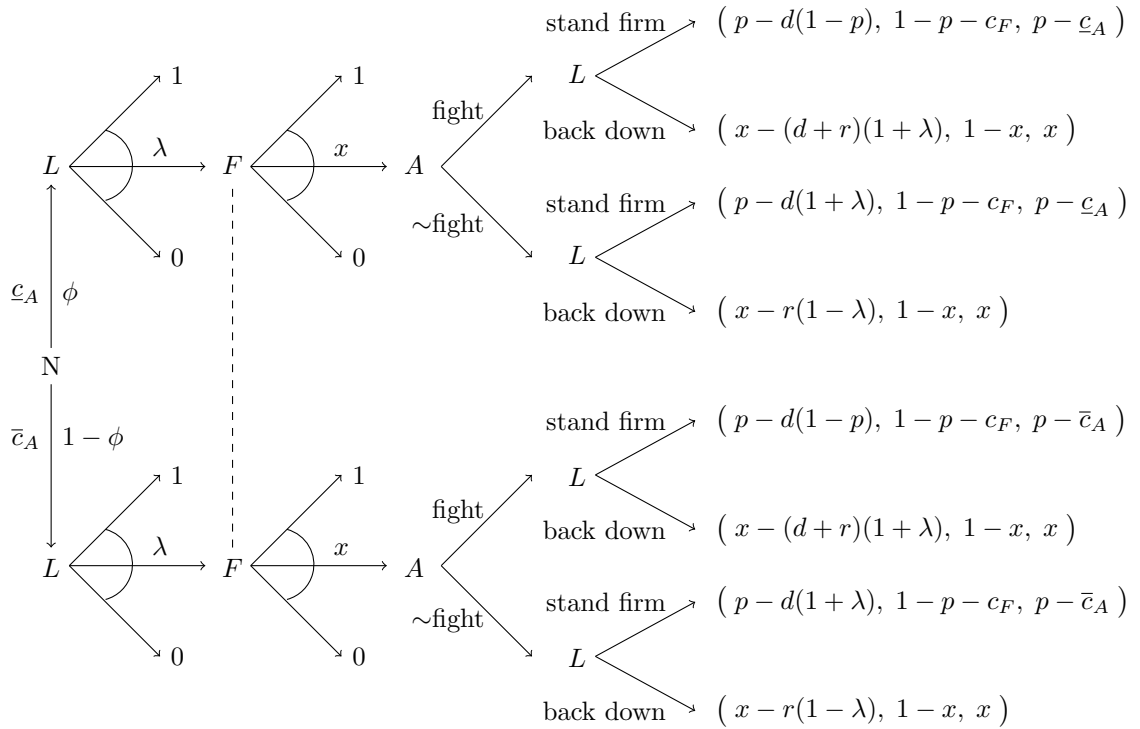


Figure 3.1: Signaling Resolve to Foreign Target

3.2 Equilibrium under Complete Information

I analyze the full model under complete information to highlight changes from the literature and implications for the incomplete-information game. First, I identify the condition under which F 's offer does not depend on the values of c_A . If the information of A 's type is irrelevant to F 's strategies, L would have no need to signal under incomplete information. Then I discuss the links between L 's choice of publicity and the type of its people, which sheds light on how the choice of publicity shifts F 's beliefs of the public's resolve or formally, A 's type. Proposition 1 summarizes L 's and F 's strategies as follows.

Proposition 1. *The following sets of strategies are each part of a Subgame Perfect Equilibrium:*

- (1) *When $d < \underline{c}_A$, L chooses a low-profile threat and F offers $p - d + r$, regardless of A 's type.*
- (2) *When $\underline{c}_A \leq d < \bar{c}_A$, L with the irresolute A chooses a low-profile threat and F offers $p - d + r$; L with the resolute A chooses a high-profile threat and F offers $p - \underline{c}_A$.*
- (3) *When $d \geq \bar{c}_A$, L always chooses a high-profile threat; F offers $p - \bar{c}_A$ to the irresolute A and $p - \underline{c}_A$ to the resolute A .*

Similar to the baseline model, F proposes an offer which both L and A prefer to escalation in order to avert war. There are four options given L 's choice of publicity and A 's type: to prevent A from fighting, F should offer

$p - \underline{c}_A$ to the resolute A or $p - \bar{c}_A$ to the irresolute type; to render L indifferent between standing firm and backing down, F proposes $p - d + r$ to L with a low-profile threat ($\lambda = 0$) and $p - 2d$ to L with a high-profile threat ($\lambda = 1$). By assumption, we have

$$p - \underline{c}_A > p - \bar{c}_A \Leftrightarrow \underline{c}_A < \bar{c}_A, \quad (3.1)$$

$$p - d + r > p - 2d \Leftrightarrow -r < d. \quad (3.2)$$

The first inequality shows that F proposes a greater offer to the resolute type of A than the irresolute type. In the second inequality, F 's offer is greater given a low-profile threat than given a high-profile threat. It implies that if L goes against A 's support for peace, a low-profile threat increases L 's bargaining leverage because low public attention reduces L 's difficulty of standing firm. War never occurs under complete information. F proposes an offer which meets the greater value between L 's and A 's reservation values, given A 's type and L 's choice of publicity.

L 's choice of publicity depends directly on its anticipation about what F would offer. First of all, if F 's offer is expected equal to A 's reservation value, $p - c_A$, L always chooses a high-profile threat.¹ A is satisfied with $p - c_A$ and prefers backing down to standing firm. If L decides to follow A 's policy preference, a high-profile strategy makes A more aware of L 's implementation

¹Here c_A refers to both \underline{c}_A and \bar{c}_A .

of the popular policy and thus reduces L 's audience costs for backing down. Thus, L chooses $\lambda = 1$ and receives $p - c_A$. Second, if F 's offer is equal to L 's reservation value, we know from Inequality 3.2 that L would prefer $\lambda = 0$ and receives $p - d$. L 's choice of publicity is determined by the relative values between its own policy costs and A 's costs of war. Given A 's type, there are three possible situations: $d < \underline{c}_A$, $\underline{c}_A \leq d < \bar{c}_A$, and $\bar{c}_A \leq d$.

When $d < \underline{c}_A$, L always prefers a low-profile threat and F proposes $p - d + r$, regardless of A 's type. The underlying logic is as follows. If d is sufficiently small, L can turn down an offer accepted by A . As L dictates the decision of escalation, F 's priority is to propose an offer which L prefers to escalation. To maximize F 's offer, L opts for a low-profile threat which minimizes the influence of A 's anti-war preference. Since F 's offer, $p - d + r$, is not a function of either type of c_A , the bargaining between F and L does not require the knowledge of A 's type. Since the public's resolve does not affect the outcome of crisis bargaining given $d < \underline{c}_A$, L has no need to address the informational problem when the private information of c_A is introduced later.

When $\underline{c}_A \leq d < \bar{c}_A$, L 's choice of publicity varies with the type of people it has. On one hand, L with the irresolute A still prefers a low-profile threat and F offers $p - d + r$ to render L indifferent between standing firm and backing down. On the other, L with the resolute A deviates to a high-profile threat and F proposes $p - \underline{c}_A$. With a moderate value of d , L backs down as long as the resolute A chooses not to fight. Given the resolute A , therefore, F 's goal is to render the resolute A rather than L indifferent between war

and peace. F proposes $p - \underline{c}_A$ which explains why L deviates to a high-profile threat. When A does not support war, a high-profile threat helps L relieve from audience costs.

When $d \geq \bar{c}_A$, L always prefers a high-profile threat to a low-profile one, regardless of A 's type. L 's escalation decisions are highly constrained by A 's policy preferences, that is, L always backs down as long as A prefers a peaceful agreement to war. Thus, F proposes an offer equal to A 's reservation value based on the latter's type. Formally, F offers $p - \underline{c}_A$ to the resolute A and $p - \bar{c}_A$ to the irresolute A . Again, L considers a high-profile threat preferable, as F 's offer never compensates L 's audience costs.

Solving the model with complete information sheds light on whether and how the opponent's uncertainty over the public's resolve matters to crisis bargaining. Whether the public's resolve shapes the opponent's offer is conditioned on the leader's policy costs. When policy costs are extremely small, an offer that the public accepts may be rejected by the leader. To avoid the bargaining failure, the opponent thus proposes an offer that renders the leader indifferent between standing firm and backing down. The opponent's strategies are determined by the leader's political costs rather than by the costs of war. Hence, the knowledge of the public's resolve does not matter to the bargaining outcome and leaders always keep low-profile threats to maximize bargaining advantage.

With the increase of policy costs, since it is more difficult for leaders to stand firm against the possibility of public support for peace, the opponent

may be able to compensate the public's reservation value to avoid war. When policy costs are moderate, on one hand, the opponent still proposes $p - d + r$ to the irresolute people and their leader; on the other, as the leader with the resolute people follows public opinion, the opponent offers $p - \underline{c}_A$ and the leader adopts high-profile threats to minimize audience costs. What do the changes of offers imply to the opponent if it is uncertain about the public's resolve? Under moderate policy costs, the real uncertainty is about whether the opponent faces the resolute people or the leader unaccountable to the irresolute public, rather than about which type of people it faces. When policy costs are sufficiently large, leaders always follow public opinion. It is the public who dominates the decision of escalation. The opponent proposes $p - \underline{c}_A$ to the resolute people and $p - \bar{c}_A$ to the irresolute one. Since its offers are tailored to different values of c_A , the opponent must figure out which type of people it deals with in the incomplete-information game.

3.3 Equilibrium under Incomplete Information

Now I turn to the incomplete-information game in which L and A , as a grouped entity (i.e. the Home country), keep the private information of c_A from F . F infers A 's type by observing L 's choice of publicity. From the previous discussion, we know that A 's resolve may or may not shape F 's offers and the links between publicity and A 's resolve are conditioned on policy costs. In this section, I analyze L 's signaling strategies under three conditions of policy costs identified in Proposition 1. Each condition differs in (a) how

F 's uncertainty affects its strategies, if any, and (b) what message L 's choice of publicity conveys about A 's type.

Proposition 2. *When $d < \underline{c}_A$, no signaling is necessary; L chooses a low-profile threat regardless of A 's type.*

When $d < \underline{c}_A$, under complete information, F proposes $p - d + r$ to both types of A . Since the offer does not require the knowledge of A 's type, A has no need to update beliefs and thus no signaling is required. L always chooses a low-profile threat and ends bargaining with a peaceful agreement. When $d \geq \underline{c}_A$, however, F 's offers are shaped at least partially by A 's costs of war, which gives L incentives to either reveal or mask A 's type. The key difference between the remaining two conditions is how the choice of publicity updates F 's belief. To summarize, when $\underline{c}_A \leq d < \bar{c}_A$, a high-profile threat signals the public's resolve in the sense that it increases F 's belief of facing the resolute A ; when $d \geq \bar{c}_A$, in contrast, a low-profile threat signals the public's resolve or increases F 's belief of facing the resolute A . Proposition 3 describes L 's signaling strategies of public threats when $\underline{c}_A \leq d < \bar{c}_A$.

Proposition 3. *When $\underline{c}_A \leq d < \bar{c}_A$, the following sets of strategies are each part of a Perfect Bayesian Equilibrium: L with the resolute A chooses a high-profile threat; L with the irresolute A randomizes between a high-profile and a low-profile threat.*

With moderate policy costs, the leader's choice of publicity varies with the type of people under complete information (Proposition 1): a high-profile

threat given the resolute people, and a low-profile threat given the irresolute people. Under incomplete information, L with the resolute A stays with a high-profile threat but L with the irresolute A bluffs probabilistically, randomizing between a high-profile threat and a low-profile one. Neither a separating nor a pooling equilibrium exists. On one hand, L with the irresolute A sometimes chooses a high-profile threat as bluffing is not too costly. On the other hand, L with the irresolute A has no incentives to entirely deviate from the type of its people. Pooling on high-profile threats would not update F 's beliefs, further reducing F 's offer to the amount smaller than $p - d$, an offer L receives by revealing the irresolute people.² In other words, L with the irresolute A would be worse off if completely deviating from the type of its people. When observing a high-profile threat, F randomizes its offer between a greater offer always accepted, and $p - \bar{c}_A$ which the resolute A rejects and leads to escalation.³ Therefore, while a high-profile threat partially updates F 's belief of the public's resolve, it is unable to eliminate the risks of war.

When $d \geq \bar{c}_A$, F 's goal is to render people, rather than the leader, indifferent between war and peace. Since F 's offer does not compensate L 's audience costs, under complete information, L would always choose a high-profile threat to minimize audience costs regardless of the types of people.

² Formally, if L with the irresolute A always bluffed, F would offer $p - 2d$ or $p - \bar{c}_A$, depending on the relative values between $2d$ and \bar{c}_A . Given $\underline{c}_A \leq d < \bar{c}_A$, neither of them is greater than $p - d$.

³The larger offer is either $p - d + r$ if $\underline{c}_A > d - r$ or otherwise $p - \underline{c}_A$. In either of the cases, L with the resolute A chooses a high-profile threat and receives an offer greater than or equal to $p - d$.

Given F 's uncertainty over A 's type, L with the resolute A must be willing to pay additional costs that the other type would not want to, *if* it prefers a peaceful agreement to war. Hence, a low-profile threat serves as a costly signal to demonstrate the public's resolve as people would impose greater costs on leaders for backing down. The principle of costly signaling applies here not only because leaders with the resolute people have to pay additional costs in order to be distinguished from the irresolute type, but also because those leaders would rather *not* pay additional costs with the absence of informational problems. The second factor tends to be overlooked in the discussions of costly signaling. Proposition 4 specifies L 's signaling strategies in detail.

Proposition 4. *When $d \geq \bar{c}_A$, the following sets of strategies are each part of a Perfect Bayesian Equilibrium:*

- (1) *If $d(1 - p) \geq r + \underline{c}_A$ and $r \geq \bar{c}_A - \underline{c}_A$, L with the resolute A chooses a low-profile threat and L with the irresolute A chooses a high-profile threat;*
- (2) *If $d(1 - p) \geq r + \underline{c}_A$ and $r < \bar{c}_A - \underline{c}_A$, L with the resolute A chooses a low-profile threat; L with the irresolute A randomizes between a high-profile and a low-profile threat;*
- (3) *If $d(1 - p) < r + \underline{c}_A$, L chooses a high-profile threat regardless of A 's type.*

A separating equilibrium requires two conditions simultaneously: (a) L with the resolute A does not hide the true information, and (b) L with the irresolute A does not bluff. The former condition is guaranteed when

a peaceful agreement makes L better off than fighting with the support of the resolute A , which yields $d(1 - p) \geq r + \underline{c}_A$. Substantively, the political costs for losing a popular war, $d(1 - p)$, outweigh the costs for a peaceful agreement after revealing the public's resolve, $r + \underline{c}_A$. The latter condition is satisfied by sufficiently large audience costs or $r \geq \bar{c}_A - \underline{c}_A$. Considering the sufficiently large costs for mimicking the resolute type, L with the irresolute A would rather reveal the lack of resolve than suffer political costs by deviating from the type of its people. If the former condition holds but not the latter, we observe a semi-separating equilibrium where L with the irresolute A bluffs probabilistically: since losing a popular war is still costly, L with the resolute A prefers a high-profile threat to be separated from the irresolute type; however, a smaller amount of audience costs makes bluffing less costly.

When $d(1 - p) < r + \underline{c}_A$, counterintuitively, L chooses a high-profile threat regardless of the type of people, meaning that L with the resolute A masks its resolve and traps its people into war. Abandoning the opportunity of information updates increases the risks of war, but why does L with the resolute A prefer war to a peaceful agreement? Without paying the costs of war, leaders sometimes prefer fighting with public support to a peaceful agreement. But since war is inherently costly to the public and leaders always follow public preferences, why would the public support fighting in the first place? An informational account of war suggests that bargaining fails when the uninformed opponent is unable to update information and proposes an offer to be rejected by the resolute type (Fearon 1995). Put it in the context,

war occurs when F fails to update information and proposes an offer that the resolute A rejects. As the public does not deliver information directly, war occurs when L with the resolute A hides the true information and mimics the irresolute type with a high-profile threat. On the surface, the leader stands firm “because of” the support from the resolute people, but if we examine the underlying logic closely, it is the leader who blinds the opponent to the true information and indirectly provokes the resolute people to support fighting. When L with the resolute A considers a popular war more attractive than a peaceful agreement, L strategically induces a suboptimal demand doomed to be rejected by her own people.

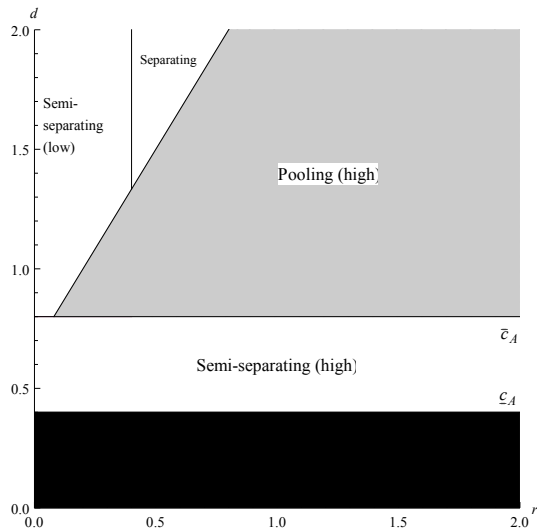


Figure 3.2: The Equilibrium Spaces

Figure 3.2 plots L 's equilibrium strategies in each proposition as a func-

tion of L 's policy costs and audience costs.⁴ Proposition 2 is represented by the black bottom area where F 's offer meets L 's reservation value. The uncertainty over the public's resolve does not matter. L and F bargain as if with complete information and divide the disputed goods peacefully. L always chooses low-profile threats to maximize F 's offer. The middle area in white illustrates Proposition 3 when policy costs are moderate. As L with the resolute A switches to a high-profile threat, L with the irresolute A has incentives to bluff probabilistically for a better deal. A high-profile threat increases F 's belief of facing the resolute A , however, without full certainty. Finally, L 's strategies of publicity in Proposition 4 are divided in the top area. When policy costs are large, L with the resolute A can reveal the information of its resolve by lowering publicity (the top-left area in white), or hide the information, leaving the risks of war open (the top-right area in gray). Which direction L with the resolute A takes, depends on whether fighting with public support is more profitable than a peaceful agreement. If the answer is negative, L with the resolute A sends a costly signal to demonstrate the public's resolve by choosing a low-profile threat. If the answer is positive, however, L with the resolute A masks the true information, induces a suboptimal offer only to be rejected by its own people and then escalates to war with their support.

⁴ The resolute L 's choice of public threats is specified in parentheses. The values of other parameters are as follow: $\underline{c}_A = 0.4$, $\bar{c}_A = 0.8$, $p = 0.6$.

3.4 Discussion

The conventional wisdom emphasizes the public's role as an audience concerned about whether leaders keep their promises (Fearon 1994), but scholars tend to overlook the role of the public as a direct participant in the event of war. In addition to prestige concerns, people are concerned about "policy substance" by weighing the payoffs of war against the payoffs of a peaceful agreement (Snyder and Borghard 2011). When the public prefers war to a peaceful agreement, prestige concerns are compatible with policy preferences and thus the unconditional notion seems reasonable that great publicity signals resolve. However, when people prefer peaceful solutions, which makes the two concerns incompatible, it is unclear why great publicity still signals the public's resolve. The existing literature also ignores the initiatives of leaders in managing the amount of political costs. I argue that leaders are able to choose various strategies to increase or decrease the level of publicity of their threats. Considering the public's policy concerns and the leader's choice of publicity helps us reconcile the conventional audience costs theory and its critics.

In this section, I first identify the strategic linkages between the leader's choice of publicity and signaling the public's resolve under different political environments. Policy costs indicate the extent to which the public holds leaders accountable to their policy preferences, an exogenous feature determined by political regimes. The conventional logic of audience costs theory exists only when leaders are moderately accountable: given moderate policy costs, greater publicity ties a leader's hands, increasing the opponent's belief of facing

a resolute public. When leaders are highly accountable, in contrast, lower publicity increases the opponent's belief of facing a resolute public. The findings also uncover the conditional links between publicity and signaling the public's resolve and highlight the advantages and disadvantages of leaders in preventing war under different political regimes. The probability of crisis escalation is relatively high when leaders are moderately accountable. Highly accountable leaders are able to fully address the informational problem and avoid war only when their opponents have strong military capabilities. While the majority of discussions focus on a comparison between moderate and high accountability, I close this section by discussing the advantage of unaccountable leaders in avoiding war and what it implies to the research on informational problems of war.

3.4.1 Endogenous Publicity and Signaling Resolve

The model explores how leaders choose the publicity of threats to signal the public's resolve in crisis bargaining. The audience costs theory highlights public threats as a mechanism of signaling resolve, solving the informational problem of war (Schultz 2012). Although scholars debate about whether public threats issued by democratic leaders are more credible, the underlying mechanism remains unchallenged: public threats signal resolve of fighting as they make leaders more difficult to back down than private threats. The possibility that people favor peace and can impose policy costs on leaders for escalation casts doubt on the link between publicity and signaling resolve. When peo-

ple's policy preferences are taken into account, the model shows that greater publicity does not always tie a leader's hands.

Increasing the publicity of threats makes leaders more difficult to back down only when leaders are moderately accountable. With moderate policy costs, leaders back down only when the resolute public chooses not to fight. Under complete information, leaders make distinct choices of publicity based on the public's type: if the resolute public chooses not to fight, leaders will back down and thus prefer great publicity to minimize political costs for backing down; if the irresolute public chooses not to fight, leaders will dominate the final decision of escalation and thus prefer lower publicity to minimize the negative impact of anti-war pressures in bargaining. Under incomplete information, leaders with the resolute public still choose great publicity, but leaders with the irresolute public sometimes make high-profile threats in order to mimic the resolute type. Upon observing high-profile threats, the opponent increases its belief of facing the resolute type of people.⁵ By considering the variation of the public's policy preference, Proposition 3 identifies a critical condition for the conventional logic of audience costs theory to work: increasing publicity signals the public's resolve only when the public who holds leaders accountable is relatively willing to fight. This is an implicit condition in audience costs theory which assumes no political costs for standing firm.

When policy costs are sufficiently high, leaders decrease the publicity

⁵There remain some degrees of uncertainty.

of threats to signal the public's resolve. It is easier for leaders to back down not despite but because of great public attention. With high policy costs, leaders are highly accountable and always back down if the public chooses not to fight, regardless of the latter's type. Since the public dominates escalation decisions, under complete information, the opponent is able to propose an offer that guarantees a peaceful settlement. To minimize political costs for backing down, leaders always choose to increase public attention: with close attention, the public prioritizes its policy preference in the current crisis over the concern about future reputation, which makes it easier for leaders to back down. In other words, increasing publicity *unties* a leader's hands. Under incomplete information, there exists conditions under which leaders with the resolute public have incentives to signal the public's resolve and receive a better offer. Based on the logic of costly signaling, leaders with the resolute public must be willing to pay additional costs to distinguish themselves from the irresolute type. Since increasing publicity *unties* a leader's hands, leaders decrease the level of publicity to signal the public's resolve.

Result 1. *When leaders are moderately accountable, the choice of greater publicity signals the public's resolve. When leaders are highly accountable, the choice of lower publicity signals the public's resolve.*

3.4.2 Political Accountability and Risks of War

A peaceful agreement is possible if states are able to solve the problem of uncertainty. With moderate accountability, leaders are able to reduce the

risks of war to some extent but the overall risks remain relatively high. With high accountability, whether leaders reduce the risks of war depends on the opponent's relative military capability. If the opponent has strong military capabilities, leaders with the resolute public do not want to fight a strong target, distinguishing themselves from the irresolute type. If the opponent is sufficiently weak, leaders with the resolute public have few incentives to address uncertainty and secure peace in the first place.

The semi-separating equilibrium in Proposition 3 indicates that moderately accountable leaders cannot entirely remove the opponent's uncertainty and thus the risks of war. It corresponds to recent works on the strategic permission of anti-foreign mass movements in weakly institutionalized regimes. First, the finding that leaders with the resolute people tend to increase public attention during crises explains why we often observe aggressive anti-foreign protests but very few antiwar moves in weakly institutionalized or transitioning regimes. Second, it illuminates why anti-foreign or nationalistic protests is a double-edged sword for leaders in foreign crises. On one hand, leaders may be able to signal the public's resolve and increase bargaining leverage by allowing anti-foreign protests; on the other, the finding that leaders with irresolute people bluff probabilistically confirms the notion that the tolerance of anti-foreign protests may lead to war. Critically, the risks of war emerge not from the hawkish people who are more willing to fight, but from the opponent's inability of fully distinguishing "sincere" from "manufactured" protests (Weiss 2013). Moreover, it is impossible for the opponent to distinguish be-

tween two types of protests simply by looking for external traits of protests, for the problem is the incentive of bluffing inherent in leaders with the irresolute people.

Surprisingly though, peace is not always in the best interest of leaders with high accountability. Without directly suffering from war, leaders with the resolute people may consider fighting with public support more profitable than a peaceful agreement. Given that war is costly to people, the only way to mobilize their support is using a high-profile threat to mimic the irresolute type and induce a suboptimal offer. The pooling equilibrium of high-profile threats in Proposition 4 yields new implications to the literature of public threats and coercive diplomacy. First, Snyder and Borghard (2011) use historical cases to contend that the opponent does not perceive the logic of audience cost theory or costly signaling in general, but the pooling equilibrium provides an alternative explanation consistent with the logic of signaling—the opponent is simply *unable* to update information when the informed leader sometimes forgoes the signaling opportunity deliberately. Second, while the conventional wisdom emphasizes the disadvantage of flexibility in signaling as it unties a leader's hands, the potential advantage of flexibility is less explored (Schultz 2012, Trachtenberg 2012). Flexibility refers to great public attention as it makes backing down easier. The pooling equilibrium implies that flexibility may bring advantages not to the public but to leaders. To induce public support for war, leaders with the resolute people sometimes mimic the irresolute type and therefore prefer flexibility to tying hands. Third, the space of the

pooling equilibrium expands when the opponent is weak, implying that leaders of great powers are more reluctant to update the opponent's beliefs. It provides a domestic explanation of why threats against weak opponents tend to fail.⁶ Empirically, when examining the effects of public threats on crisis escalation, we should be cautious about the cases with power asymmetry.

Result 2. *When the target is strong, highly accountable leaders discourage war, but when the target is weak, highly accountable leaders encourage war.*

3.4.3 Autocratic Advantage and Sources of Uncertainty

The model shows that unaccountable leaders, or autocrats, may have an advantage in achieving peaceful agreements precisely because of small policy costs. This is different from audience costs theory which suggests that the inability to generate audience costs puts autocrats at a disadvantage. If we deviate from the unitary-state framework and assign the costs of war to people, small policy costs allow autocrats to render public preferences irrelevant, removing the risks of war driven by the opponent's uncertainty over the public's resolve. Given small policy costs, an autocrat can afford to ignore the public's preferences and turns bargaining into a "one-level" game. The opponent's goal is not to figure out how much people are willing to fight, but to propose an

⁶Sechser (2010) argues that threats against weak targets fail because the challenger and the target have different expectations about the likelihood of future interactions. However, it is unclear why power asymmetry contributes to the divergence of expectations. In my model, power asymmetry increases the challenger's prospects of victory, which makes fighting with public support more profitable.

offer which the autocrat prefers to escalation. A peaceful agreement can be achieved easily if the opponent knows the autocrat's costs for escalation.

The result raises the importance of the source of uncertainty in the studies of crisis bargaining. If the opponent is uncertain about the resolve of leaders or unitary states as assumed in audience cost theory, small policy costs are ineffective in distinguishing the resolute type from the irresolute one. However, if it is uncertain about the resolve of people who directly bear the costs of war, small policy costs make the leader immune from public pressures and thus remove the source of uncertainty. The advantage of autocrats depends on the assumption that the opponent knows the leader's costs for escalation. This is not an implausible assumption as it maybe easier for the opponent to know a leader's costs than the public's, for instance, the resolve of new leaders can be learned through frequent interactions over time (Wolford 2007).

3.5 Summary

The analysis begins with the puzzle why democratic leaders sometimes raise public attention during crises when the public seems unwilling to fight, which is inconsistent with the conventional wisdom that increasing publicity signals resolve. Drawing from recent developments of audience cost theory, I build a crisis-bargaining model that highlights two elements of domestic interactions between the public and leaders during crisis bargaining: the public's policy preferences based on their resolve, and the leader's control of publicity of threats to adjust political costs. In addition to audience costs driven by

prestige concerns, the public can impose policy costs on the leader whose policy does not match their preferences, that is, the public can hold the leader accountable to their policy preferences. Leaders use publicity to affect the relative importance of policy preferences and prestige concerns: when the public supports war, great publicity increases the importance of both policy preferences and prestige concerns; when the public supports peace, great publicity allows the public to weight policy preferences more than prestige concerns. Under complete information, leaders with low accountability to public preferences prefer low public attention, while leaders with high accountability lean towards great public attention. With the opponent uncertain about the public's resolve, the model uncovers whether uncertainty matters to crisis bargaining and if so, how leaders choose the publicity of threats to influence the opponent's beliefs under different levels of political accountability to public preferences:

- Under low accountability, the public's resolve does not affect crisis bargaining; leaders always choose low publicity without changing the opponent's beliefs.
- Under moderate accountability, great publicity increases the opponent's belief of facing the resolute people, but no separating equilibrium exists.
- Under high accountability, low publicity increases the opponent's belief of facing the resolute people; surprisingly, leaders with the resolute

people sometimes choose great publicity to mask their true resolve and escalate with public support.

The model provides an integrated framework to understand the roles of the public's resolve in crisis bargaining under different political environments. It first shows that a negotiation may end peacefully despite the public's resolve unknown to the opponent, if a challenging leader is unaccountable to public preferences.⁷ By separating the public's costs of war from the leader's political costs, the model specifies the conditions under which the information about the public's resolve matters. From the opponent's perspective, the leader's costs are more important when the challenging leader is autocratic, whereas the public's resolve is more important when the challenging leader is at least moderately accountable. It reconciles the debates over the locale of resolve in international relations, particularly whose resolve matters in crises (Saunders 2011, Singer 1961, Wolford 2007). Second, if the opponent's uncertainty increases the risks of war, how leaders signal the public's resolve also depends on their political accountability to public preferences. When leaders are moderately accountable, increasing the publicity of threats raises the opponent's belief of facing the resolute people. Under moderate political accountability, leaders follow the preferences of the resolute people but not the irresolute people. Therefore, great publicity further amplifies the hawkish voices, increasing the difficulty of backing down. When leaders are highly accountable to the

⁷The autocrat's advantage requires the full information about the leader's costs and benefits, which is beyond the scope of this paper.

public regardless of its type, low publicity raises the political costs for backing down, increasing the opponent's belief of facing the resolute people. Taking policy preferences into account yields a different story of how open threats signal the public's resolve.

If war is driven by the opponent's uncertainty over the public's resolve, great political accountability does not necessarily reduce the risks of war. Autocrats may have advantage in promoting peace as low political accountability renders the uncertainty over the public's resolve irrelevant to bargaining. In contrast, it is impossible to completely avoid war when leaders are moderately accountable, often in transitioning states. Empirical analyses show that states tend to experience more conflicts in the period of democratization or unstable regime changes (Gleditsch and Ward 2000, Mansfield and Snyder 1995). Instead of arguing that people in transitioning states are more aggressive or have stronger nationalist appeals, the model attributes bargaining failures to the incentives of bluffing inherent in leaders with the irresolute people. Hence, while increasing publicity alleviates the opponent's uncertainty, it fails to remove the possibility of war. Finally, if democratic leaders are considered highly accountable, the model shows that whether democracy promotes peace depends on the distribution of power. Democratic leaders tend to reveal the true information and achieve peaceful agreements if it is less likely to defeat the opponent. War is more likely to occur, however, when leaders of great powers target weaker states. The strong prospects of winning induce democratic leaders to hide the information of the public's resolve and trap the resolute people into war.

High-profile threats issued by democracies are sometimes uninformative, not because the opponent does not understand the signaling logic (Snyder and Borghard 2011), but because the opponent is unable to update information under the signaling logic.

Chapter 4

Empirics

How do leaders use public threats to communicate resolve in crisis bargaining, and how does political accountability affect the probability of crisis escalation? Scholars have not reached the consensus on the relationship between political accountability and the effectiveness of public threats in communicating resolve. Yet we know that leaders are often careful in choosing the publicity of their threats. While the conventional audience cost theory contends that the effectiveness of public threats increases with political accountability and that democratic leaders are more capable of signaling resolve by increasing public attention than their autocratic counterparts, other scholars challenge audience costs theory with the lack of empirical evidence (Downes and Sechser 2012), and the inability to parse out the sources of political costs—whether the public dislikes backing down due to prestige concerns or hawkish preferences (Snyder and Borghard 2011). Taking into consideration different sources of public preferences, the theory presented in the previous chapter yields two new findings. First, without the pressure from the domestic public, unaccountable leaders achieve peaceful agreements with no need to signal. Second, the credibility of highly accountable leaders in communicating resolve is conditioned on the distribution of power: in particular, leaders of great powers have *no* advantage

in credible signals.

This chapter examines the empirical support for the second finding, that is, the conditional relationship between political accountability and threat effectiveness. The effects of political accountability on bargaining outcomes are conditioned on the distribution of power between adversaries. When the challenging state is relatively weak, highly accountable leaders have advantage in solving informational problems and avoiding crisis escalation. When the challenging state is relatively strong, however, a highly accountable leader is no better in delivering credible threats than a moderately accountable one and thus, the pacifying effects of high accountability disappears. Empirically, I argue that after the theoretical model distinguishes policy costs from audience costs, it is more appropriate to use the extant measures of audience costs to operationalize policy costs than audience costs. By constructing a new index for political accountability, the empirical result illustrates a similar conditional relationship between a challenger's political accountability and the likelihood of target's voluntary compliance in the Militarized Compellent Threat dataset.

4.1 Empirical Implications of the Model

In this section, I derive testable hypotheses from the theoretical model. While the theoretical model provides an equilibrium relationship between political accountability, strategies of public threats, and crisis escalation, the equilibrium space shown in the previous chapter only identifies the regions in which war is possible, without a straightforward link between the param-

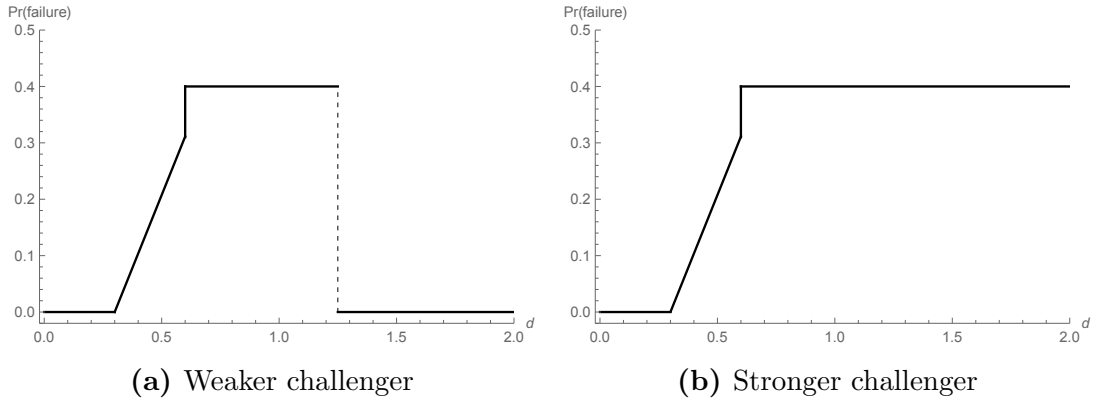


Figure 4.1: Equilibrium probability of threat failure

eters and the probability of threat failure. Therefore, I translate the logic of the theoretical model to the probability of threat failure as a function of the challenger’s political accountability and its national power relative to the target’s.

Figure 4.1 presents the graphs of equilibrium probabilities of threat failure for a weaker challenger and a stronger one, respectively.¹ A comparison of two graphs indicate that the probabilities of threat failure appear the same under low and moderately accountability but, as illustrated in Proposition 4, differ under high probability. Below I will first discuss the predicted patterns of threat failure under low and moderate accountability and then how the patterns under high accountability vary with the distribution of power between two states. This section closes with two testable hypotheses regarding the conditional effects of political accountability on the rate of threat failure.

¹Parameters are fixed at $\underline{c}_A = 0.3$, $\bar{c}_A = 0.6$, $c_F = 0.05$, $r = 0.7$ (ensuring a separating equilibrium), and $\phi = 0.4$ (ensuring that F will risk war given its prior belief).

The model predicts no risk of threat failure under low accountability, that is, leaders are never accountable for the public's support for peace. If a target is uncertain about the public's resolve *only*, Proposition 2 in the previous chapter suggests that unaccountable leaders are always able to settle a peaceful agreement. In reality, it is possible that even though the information about the public does not matter, the risks of war emerge if a foreign target is uncertain about something else. For instance, threats issued by an unaccountable leader *can* fail if the opponent is uncertain about the leader's costs rather than the people's. In sum, when unaccountable leaders make public threats, the probability of war or threat failure is zero in the equilibrium but we should expect a nonzero but low rate of failure in the empirical model, considering the potential uncertainty about the leader's information and the aggregation bias.

When leaders are accountable to the resolute public only, the risks of threat failure increase with the amount of political accountability. The semi-separating equilibrium suggests that while leaders with the resolute public stay true to its type, leaders with the irresolute republic have incentives to mimic the resolute type probabilistically. As the amount of political accountability grows, such incentive becomes stronger, further reducing the effectiveness of high-profile threats as a signal of resolve.

When leaders are always accountable to the public's support for peace, the equilibrium probability of war is determined by a leader's political account-

ability and its relative military power.² According to Proposition 4, leaders play the pooling strategy with high risks of war if $d \leq d^* = \frac{r + \underline{c}_A}{(1-p)}$; otherwise, they play the separating strategy with no risks of war. As the value of d increases, whether leaders can switch from the pooling strategy to the separating one depends on the position of the threshold, d^* , which is determined by the distribution of power between the challenging state and the opponent. If the challenging state is weak, the threshold is relatively low ($p = 0.1$), meaning that high political accountability can easily encourage leaders to choose the separating strategy and lead to a peaceful settlement (Figure 4.1a). If the challenging state is sufficiently strong ($p = 0.6$), the threshold to switch to the separating equilibrium can be very high, such that even the most accountable leaders play the pooling strategy and the risks of war remain high (Figure 4.1b).

Put together, when the challenging state is weak, we expect a concave downward relationship between the leader's political accountability and threat failure. It begins with a nonzero yet low rate of threat failure for unaccountable leaders. The probability of threat failure then increases with the leader's accountability and reaches the peak, before decreasing rapidly to zero again. High political accountability encourages the challenging leaders to choose the separating equilibrium and eliminates the risks of war, because the challenging

²To focus on the conditional effects of political accountability, I assume audience costs at a moderate value ($r \geq \bar{c}_A - \underline{c}_A$) to ensure the comparison between the pooling equilibrium and the separating equilibrium. If $r < \bar{c}_A - \underline{c}_A$, according to Proposition 4, highly accountable leaders simply play a semi-separating strategy, similar to the strategy of moderately accountable leaders.

leaders prefer a peaceful settlement to being defeated by a strong target.

When the challenging state is strong, the concave downward relationship disappears. As discussed earlier, the patterns of threat failure do not change when leaders are moderately accountable or not accountable at all. In other words, the rate of threat failure are expected low under low accountability and then gradually increases. However, as political accountability continues to grow, its effect on threat failure quickly becomes constant. When it is easy to defeat a weak target, high political accountability encourages the challenging leaders to stay with the pooling strategy and fight with public support if necessary. Below I summarize two hypotheses regarding the relationship between the challenging state's power, the challenging leader's accountability and the probability of threat failure.

Hypothesis 1. *When the challenging state is weak, there is a concave downward relationship between the leader's political accountability and the probability of threat failure.*

Hypothesis 2. *When the challenging state is strong, there is no clear relationship between the leader's political accountability and the probability of threat failure.*

4.2 Testing the Effectiveness of Threat

Before moving to the details of research design, this section reviews existing empirical analyses on how political institutions influence the effective-

ness of threats. The debate of appropriate empirical analyses covers nearly all aspects of research design from the scope of samples, the construction of dependent variables and independent variables, to model specifications. Two observations emerge from the review of empirical analyses. First, it is more appropriate to focus on militarized compellent threats and their outcomes to test the hypotheses derived from the model which begins with a compellent threat. Second, since the extant measure of audience costs highlight the capability of audiences to punish leaders rather than the incentive of punishment, these measures do not necessarily capture audience costs in a way distinguishable from policy costs as separated in the model.

4.2.1 Threat Failure

To test the effectiveness of information revelation in conflicts, scholars focus on the target's response but disagree on specific measures and the selection of sample. Earlier works adopt the Militarized Interstate Dispute (MID) dataset and identify whether the target reciprocates with force as the failure of threats and hence the failure of information revelation (Schultz 1999, 2001, Weeks 2008). However, as Downes and Sechser (2012) show, whether a target responds with force is not necessarily associated with acquiescence. Some studies simply consider the outcome of threat as failure if a target uses force by the definition of the hostility of level in the MID dataset. Again, this approach does not directly measure whether a target rejects or concedes to the challenge. Even though a dispute ends with a target's concession, it is

essential to examine why it concedes—by negotiation or violent compulsion— which the extant outcome variable in the MID dataset is unable to parse out (Downes and Sechser 2012). Another problem lies in the sample selection in the MID dataset, in which many of the disputes are incidental or minor skirmishes without an explicit coercive threat from disputants (Sechser 2011). Empirical analyses using the International Crisis Behavior dataset are subject to similar problems.

Alternatively, it is more appropriate to use the Militarized Compellent Threat (MCT) dataset to test the hypotheses derived in my model for the following reasons. First, the MCT dataset focuses on cases in which an explicit compellent threat can be identified. A compellent threat is defined as “an explicit demand by one state (the challenger) that another state (the target) alter the status quo in some material way, backed by a threat of military force if the target does not comply” (Sechser 2011, pp. 380). By excluding irrelevant cases such as incidental skirmishes or military exercises, the disputes selected to the MCT dataset match the setup of my model which begins with a challenger’s threat. Second, as the MCT dataset measures whether and how much a target concedes voluntarily to a challenger’s initial demand. By focusing on *voluntary* concession, it eliminates the possibility of concession by virtue of violent compulsion. Hence, it is appropriate to use this dataset to examine whether a target concedes to a compellent threat short of military confrontation—an indicator of success in coercive demand.

4.2.2 Political Punishment: Capability vs. Incentive

In terms of theoretical variables, empirical analyses of audience cost theory look for domestic institutional features to capture the variation of audience costs. The majority of empirical analyses adopts political regime types, measured by the 21-point Polity score, as an indicator of audience costs (Marshall and Jaggers 2013). This approach of measurement—linking audience costs to regime types or more specifically Polity scores—forms the connection between audience cost theory and democratic peace, which, as Schultz (2012) points out, did not exist originally and leads to inconsistent results and challenges. While Schultz (1999, 2001) demonstrates empirical support for democratic advantage with the MID dataset, Downes and Sechser (2012) find no evidence in the MCT dataset that democracies are more effective than autocracies in achieving success in compellent threats. Challenges come from both sides: on one hand, by breaking down the types of autocracies, Weeks (2008) show that some autocratic leaders are also able to generate audience costs; on the other, others are suspicious about whether audience costs generated by democratic leaders are “small to negligible” due to low public attention or the leader’s ability to talk their way out of being punished (Schultz 2012).

A recent effort to improve empirical analysis of audience cost theory looks for an alternative measure of audience costs. Instead of focusing on the leader’s ability to punish domestic dissidents, Uzonyi, Souva and Golder (2012) redirect the attention to the costs the principal bears for punishing the agent (leader). They conceptualize “audience cost capacity” (ACC) as a function of

the availability of alternative leaders and the level of political participation. The key difference from Polity scores or other measures of regime types is that ACC focuses on *ex post* punishment, highlighting the “institutions that allow for the leader to be punished,” rather than on the arrangement that constrain the choice set of a leader (pp. 765). The ACC score is better than a dichotomous variable of regime types in the sense that (a) its conceptualization directly captures the audience’s capability of punishment, and that (b) it allows for variation both within democratic regimes and within autocracies. However, as I describe in more detail below, because the ACC score captures the audience’s capacity of punishment, it is a more appropriate measure of audience costs *if* a leader is concerned about audience costs only.

If a leader’s political costs contain more than one type, however, neither ACC nor the measures of regime types is able to distinguish audience costs from the other types of costs. The existing measures of audience costs all seek to capture the audience’s *capability* to punish the leader, rather than their *incentive* of punishment. While Fearon (1994) originally attributes the audience’s disapproval of backing down to concerns of national prestige, subsequent works simply identify the disapproval of backing down as that of bad policy: leaders pay audience costs for implementing bad policy. As Uzonyi, Souva and Golder’s (2012) discussion illustrates,

“Audience costs are the mechanism through which the audience attempts to limit moral hazard (Fearon 1994). If the leader is subject to punishment for *bad foreign policy*, she is tied more closely

to the risk associated with her policies and is less likely to engage in foreign policies that do not benefit her audience. Thus, to reduce moral hazard, the audience must be able to punish the leader, exposing her to the cost associated with a *poor foreign policy*” (pp. 768).

Assuming away a variety of incentives leads to a serious challenge to the literature of audience cost theory: even though an audience opposes backing down, it is unclear whether the opposition is driven by the concerns of national prestige or the preferences of hawkish policy (Snyder and Borghard 2011). The distinction between prestige concerns and policy preferences has important implications to the content of signal conveyed in public threats, especially if an audience has dovish preferences and supports backing down instead. As discussed in the previous chapter, the previous literature did not examine the variation of domestic policy preferences and its implications to a leader’s signaling strategies. To fill this gap, I model how an audience decides policy preferences independent from prestige concerns, separating the dimension of policy costs, driven by policy preferences, from the dimension of audience costs driven by prestige concerns. Since the new theoretical model includes multiple dimensions in political costs, it is inappropriate to use the variables of political features—both ACC and regime types—to measure the prestige-driven audience costs.

Instead, I argue that the ACC score is a more appropriate candidate to measure policy costs, or political accountability, after the model specifies

two types of political costs. In the theoretical model, political accountability is represented by the amount of policy costs that the public can impose on the leader for deviating from their policy preferences. The ACC score matches the operationalization of political accountability in two aspects. First, the ACC score, composed of the availability of election competitions and the degrees of political participation, emphasizes the potential bottom-up punishment after a leader has made a decision. Policy costs in the model are also realized after the leader finally decide whether to stand firm or back down. Hence, the ACC score is better than Polity scores whose variation is mostly driven by changes in executive constraints. Second, the ACC score is designed to, as quoted above, measure the costs a leader must pay for implementing “bad foreign policy,” or policies that “do not benefit her audience.” Without modeling the audience’s preference between war and peace, previous studies used to link the ACC score and other political variables to audience costs. Since my theoretical model singles out policy preferences as well as policy costs, it is reasonable to connect the ACC score—measuring the audience’s capability to punish leaders for bad foreign policy—with policy costs or more generally, political accountability.

4.3 Research Design

The review of previous empirical works on public threats yields two implications. First, the MCT variable fits my theoretical model more closely than the other variables—the MID reciprocation and the ICB outcome—in

measuring threat failure. Second, the measures of political features are no longer suitable to capture the variation of audience costs when a leader's political costs consist of both audience costs and policy costs. In this section, I further discuss the choice of the dependent variable, theoretical variables and controls, and the model specification.

I sample on the MCT dataset which includes the militarized compellent threats from 1918 to 2001. As discussed above, a militarized compellent threat is defined as “an explicit demand by one state (the challenger) that another state (the target) alter the status quo in some material way, backed by a threat of military force if the target does not comply” (Sechser 2011). The unit of analysis is the compellent threat, as my theory focuses on the outcome of threat and considers the initiation of threat as given.

The dependent variable, *Threat failure*, is a dichotomous variable where the failure of a compellent threat is coded 1 and the success is equal to 0. In the theoretical model, a successful threat means that both A and F prefer an offer proposed by F to fighting a costly war, whereas the outbreak of war (i.e. L chooses to stand firm) indicates the failure of threat. Following Downes and Sechser (2012), a threat is considered successful if it satisfies both of the conditions: (a) a threat achieves full compliance from the target, and (b) the target decides to fully comply with fewer than 100 military fatalities. Otherwise, a threat fails to achieve its objective. Among 242 compellent threats, 93 of them succeeds with full compliance and small casualties while 149 threats fail.

The first independent variable, *Challenger's accountability*, refers to policy costs (d) in the theoretic model, indicating the extent to which people from the challenging state are able to punish their leader for unfavorable policies. As discussed in the previous section, I argue that most of the existing measures of audience costs capture the *capability* of people to punish the leader, rather than the *incentive* of punishment. This is one of the reasons why, despite the empirical tests of audience costs theory, some scholars disagree with the conclusion drawn from the evidence—that prestige concerns of the public contributes to the tying-hand effect. Instead, they contend that these tests at best show the international consequences of being able to hold leaders accountable (Baum 2004, Snyder and Borghard 2011). Essentially, the public may hold leaders accountable for unfavorable policies, which does not mean that the public is always object to the decision of backing down. To address these criticisms, my theoretical model explicitly separates audience costs from policy costs. In my model, the term of audience costs is restricted to how much the public values prestige and honesty, whereas I use policy costs to represent people's capability to hold leaders accountable for unfavorable policies without predetermining what unfavorable policies are. The hypotheses derived from the theoretical model depict the probability of threat failure as a function of political accountability rather than of audience costs. Nonetheless, since most of the existing measures of audience costs actually capture the public's capability of political punishment, they are considered as candidates to measure a leader's political accountability.

Among the existing measures, I choose the “audience cost capacity” (ACC) score as a candidate to measure *Challenger’s accountability*. In contrast to other measures that focus on *ex-ante* constraints leaders have, ACC highlights the *ex-post* feature of punishment. Instead of using regime types, ACC examines “how institutions affect the specific costs that the principal bears for challenging the incumbent” (Uzonyi, Souva and Golder 2012, pp. 766). In the theoretical model, political accountability, from high to low, indicate whether the entire public, the hawkish portion, or none of them can hold leaders accountable. It matches the conceptualization of ACC that shifts the focus to the principal’s costs of holding leaders accountable. The operationalization of ACC consists of two components: (a) the cost of exit for individual challengers, captured by the Polity measure on openness of executive recruitment (*xropen*), and (b) the cost of mobilizing against the agent, captured by the Polity measure on restrictions on political participation (*parcomp*). Table 4.1 lists the indications of these two components.

Table 4.1: Components of ACC

<i>xropen</i>	Openness of Executive Recruitments	<i>parcomp</i>	Competitiveness of Political Participation
0	unregulated	0	unregulated
1	closed	1	repressed
2	dual executive-designation	2	suppressed
3	dual executive-election	3	factional
4	open	4	transitional
		5	competitive

Nonetheless, I do not use the original ACC score. The original coding rule attaches more importance on *xropen* than on *parcomp*. The ACC score is zero, meaning no costs at all, unless there is some form of election (i.e. $xropen > 2$). In other words, the original coding rule indicates that public participation can have some constraints on leaders only with the presence of election. However, the literature has suggested that leaders are sometimes constrained by the pressure of people even without an election (Weeks 2008, Weiss 2014).³ By lumping these moderate constraints together with those without public participation at all, the original ACC score fails to capture the distinction between low policy costs and moderate ones, which is related to the main findings of the model.

I construct *Political Accountability Index* by modifying the “audience cost capacity” (ACC) score. Political accountability is coded zero if the value of *xropen* is smaller than two. If *xropen* is equal to two, the score of political accountability ranges from zero to two, depending on the competitiveness of political participation (*parcomp*). If some form of election is present (i.e. $xropen > 2$), the score of political accountability ranges from one to three as political participation becomes more competitive. Table 4.2 presents the coding rule in detail.

Therefore, the value of *Challenger’s accountability* ranges from 0 (no accountability) to 3 (high accountability). When the variable *Challenger’s ac-*

³See the coding details in (Uzonyi, Souva and Golder 2012).

Table 4.2: Political Accountability Index

<i>xropen</i>	<i>parcomp</i>		
	0–2	3–4	5
0	0	0	0
1	0	0	0
2	0	1	2
3	1	2	3
4	1	2	3

countability equals zero, the challenging state has extremely limited political participation without an open competition of leadership; as the value increases, it suggests an openness of leadership recruitment and the growing population in political participation.⁴ To test the concave relationship between threat failure and political accountability, I include the second independent variable, *Challenger’s accountability*², by squaring the variable *Challenger’s accountability*.

I include three sets of control variables to tease out the effects of political accountability. The first set of control variables separates other domestic factors that may influence threat outcomes through the challenging leader’s accountability. Though not explicitly modeling the target’s domestic politics, we should expect that the targeted leader’s political accountability may also affect the challenging leader’s accountability and the outcome of threat. Previous studies model crisis bargaining in which both the challenger and the target have domestic audiences, suggesting that the challenging leader may be more

⁴See the coding details in (Uzonyi, Souva and Golder 2012).

cautious with compelling threats considering the target's domestic pressures (Kurizaki 2007, Tarar and Leventoğlu 2009). Without controlling for the targeted leader's political accountability, therefore, we might underestimate the effects of *Challenger's accountability*. I code *Target's accountability* from the *Political Accountability Index*. Despite my focus on the principal's capability of punishing its agent, political regimes have other mechanisms to affect war and peace (Bueno de Mesquita et al. 1999, Debs and Goemans 2010, Russett and Oneal 2001). To control for the correlation between regime types and political accountability, I include *Challenger's Polity score* and *Target's Polity score*, both drawn from version IV of the Polity dataset (Marshall and Jaggers 2013).

The second set of control variables captures similarity of external interests and external sources of conflicts. Following Schultz's (2001) operationalization, I include the following variables: (i) *Contiguity*, to measure whether two states share a land border or are separated by 150 miles of water or less; (ii) *Alliance portfolio similarity*, measuring the similarity of alliance portfolios within a given dyad; and (iii) evaluations of *Status quo* for both the challenger and the target, a proxy of their satisfactions with the international system. As the evaluations of external circumstances may also affect leaders' incentives of signaling resolve in crises, separating these factors make sure that the effects of independent variables illustrated in statistical results focus on the leaders' political payoffs from war and peace.

For both the challenger and the target, some types of goods are typi-

cally more valuable than other types, which may affect the outcome of crisis by influencing public preferences and thus the extent to which leaders are accountable to the public. Therefore, the third set of control variables includes four binary variables regarding the type of issues in dispute—*Territory*, *Leadership*, *Policy*, and *Others*—obtained from the MCT dataset (Sechser 2011). The majority of threats are related to territorial disputes or policy disagreement. Finally, since leaders issued threats during World War II not to signal private intentions but to fulfill other purposes—for example, to coordinate with members in a military coalition (Wolford 2015), I include binary variables for each year during World War II (1939–1945).

While most of empirical works on threat credibility control for the power relationship between states, I do not include it in the main analysis to follow closely the implications of the theoretical model. The hypotheses suggest that the effects of the challenger’s political accountability on threat outcomes may differ, depending on whether the challenger is categorized as weak or strong relative to the target. Within each category, however, the specifics of the distribution of power have little impact on how the challenger’s political accountability affects threat outcomes. Hence, I exclude measures of the challenger’s relative power as control variables from the main result. For a comparison to extant empirical works though, I follow Schultz’s (2001) analysis to include in the appendix (i) *Challenger’s relative power*, measured by the challenger’s share of dyadic military capabilities, (ii) three dichotomous

variables indicating the status of major power in a given dyad.⁵ As shown in the appendix, the effects of independent variables are similar.⁶

The hypotheses suggest that the effects of political accountability on threat failure differ in the challenging state's relative strength. Under high political accountability, as Figure 4.1 shows, the challenging state's relative strength serves as a division between the pooling equilibrium where the risk of war driven by the target's uncertainty remains unsolved, and the separating equilibrium where highly accountable leaders avoid war by solving the target's informational problem. The more powerful the challenging state is, the more the space of separating equilibrium shrinks. To examine the conditional effects of political accountability on threat credibility, therefore, it is critical to set a cutpoint of the challenging state's relative strength. I propose two candidate cutpoints to separate stronger challengers from weaker ones. The first one is its mean value (0.67) of *Challenger's relative power* and the second one is the median (0.78). Challengers are considered weaker if its relative military strength is below the cutpoint, or stronger otherwise.

Since the public's resolve may vary with each crisis, I use binary probit

⁵*Major-Major dyad* is coded one if both the challenger and the target are considered as major powers, and zero otherwise. *Major-Minor dyad* is coded one if the challenger is a major power and the target is a minor power, and zero otherwise. *Minor-Major dyad* is coded one if the challenger is a minor power and the target is a major power, and zero otherwise.

⁶I use the EUGene program to generate the Polity scores and the challenger's relative military capabilities (Singer, Bremer and Stuckey 1972). Unless specified, the rest of control variables are obtained from Downes and Sechser's (2012) data.

regression models with standard errors clustered on the challenger.⁷ The full model specification is as follows:

$$\Pr(\text{Failure}=1) = \Phi(\beta_{0i} + \beta_1 \text{Challenger's accountability}_i + \beta_2 \text{Challenger's accountability}_i^2 + \beta_k \Omega_i + \mu_i)$$

From Hypothesis 1, we expect that given a weaker challenger, the probability of *Threat failure* is initially small at a low level of *Challenger's accountability* and as the value of accountability increases, the probability of failure reaches the highest point before declining finally. Formally, the sign of β_1 is expected positive while the sign of β_2 is expected negative. Given a stronger challenger, nevertheless, Hypothesis 2 implies no statistical significance in either β_1 or β_2 .

4.4 Results

Table 4.3 reports the coefficients of independent and control variables, with standard errors in parentheses. I separate the sample of weaker challengers (Model 1) from that of stronger challengers (Model 2), using the mean of *Challenger's relative power* (0.67) as a cutpoint. In Model 3 and 4, I estimate the failure of threats issued by weaker challengers (Model 3) and stronger challengers (Model 4), using the median of *Challenger's relative power* (0.78) as a cutpoint. The effects of dummy variables for World War II years are controlled but not reported in the tables.⁸

⁷Results are similar when standard errors are clustered by the MCT threat, as shown in

Table 4.3: Estimating Compellent Threat Failure (1918–2001)

<i>Threat failure</i>	Model 1 ($p < 0.67$)	Model 2 ($p \geq 0.67$)	Model 3 ($p < 0.78$)	Model 4 ($p \geq 0.78$)
Challenger's accountability	3.505*** (0.787)	0.301 (0.733)	2.846*** (0.697)	0.449 (0.915)
Challenger's accountability ²	-1.158*** (0.222)	-0.087 (0.178)	-0.836*** (0.201)	-0.229 (0.218)
Target's accountability	1.074*** (0.353)	-0.199 (0.209)	0.810** (0.252)	-0.311 (0.225)
Challenger's polity score	-0.041 (0.046)	0.001 (0.037)	-0.067 [†] (0.041)	0.045 (0.039)
Target's polity score	-0.178*** (0.047)	0.015 (0.030)	-0.150*** (0.042)	0.020 (0.038)
Contiguity	-0.712 (0.504)	-0.510 (0.343)	-0.880 [†] (0.450)	-0.419 (0.408)
Alliance portfolio similarity	1.629* (0.784)	0.378 (0.385)	1.648** (0.605)	-0.046 (0.524)
Challenger's status quo	1.288 [†] (0.730)	0.078 (0.437)	0.799 (0.603)	0.115 (0.522)
Target's status quo	-3.932*** (0.648)	0.541 (0.523)	-2.821*** (0.388)	0.748 (0.687)
Territory	-0.357 (0.346)	0.210 (0.285)	-0.117 (0.298)	0.526 [†] (0.277)
Government	-2.110* (0.623)	-1.278*** (0.356)	-2.064** (0.589)	-0.767* (0.365)
Policy	-1.178 [†] (0.610)	0.374 (0.330)	-1.252** (0.444)	0.768** (0.280)
Other	-0.725*** (0.268)	0.196 (0.349)	-0.223 (0.380)	0.113 (0.439)
Constant	-2.444* (1.241)	0.470 (0.862)	-2.020* (0.841)	0.627 (1.081)
Observations	94	130	113	111
Model χ^2	112.12	107.07	98.74	80.64

Standard errors in parentheses. Dummy variables for world war years are included but not reported here.

[†] $p < 0.1$, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

The results are consistent with the hypotheses described above. In each of the model, the coefficient of *Challenger's accountability*² has a negative sign, indicating a quadratic relationship between *Challenger's accountability* and the likelihood of compellent threat failure, with the parabola facing downward. In Model 1, both the quadratic coefficient and the linear one are statistically significant at $p < 0.001$, implying that the non-linear effects of *Challenger's accountability* in the group of weaker challengers. The non-linear effects are insignificant in Model 2, however, when a challenging state is relatively stronger than its target, which is consistent with Hypothesis 2. Changing the cutpoint from the mean to the median of *Challenger's relative power* yields similar results. Next, Model 3 and 4 show a similar pattern using the median of *Challenger's relative power* (0.78) as a cutpoint. Model 3 suggests that when a challenging state is weaker, that is, its relative power is below the median, compellent threats are more likely to succeed if the challenging leader is unaccountable or highly accountable, but more likely to fail if the leader is moderately accountable. When a challenging state is stronger (Model 4), the quadratic relationship is again statistically indiscernible.

Since it is difficult to assess the substantive effects of quadratic terms by examining coefficients and standard errors in probit regressions, I present

the appendix.

⁸Some of the dummy variables are omitted by the computing program due to collinearity, which is why the sum of observations in the sub-samples are fewer than the observation numbers of the entire sample.

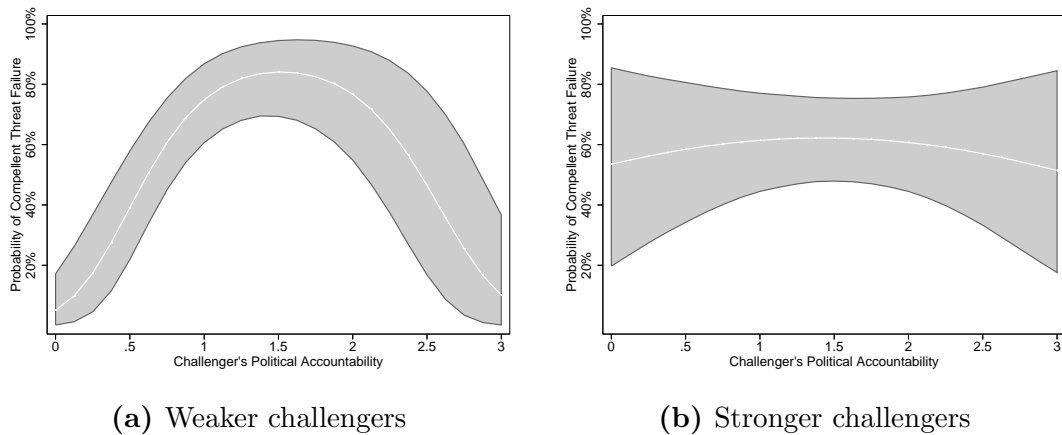


Figure 4.2: Predicted probability of compellent threat failure

the figures of the predicted probability of compellent threat failure as a challenging leader becomes more politically accountable. I generate the predictions based on specifications in Model 1 and 2, which separate weaker challengers from stronger ones with the mean of *Challenger's relative power*. To make the predicted probability more concrete, I present the simulated results of the probability of threat failure for a contiguous dyad, holding other variables at the in-sample mean.⁹ Figure 4.2 illustrates the estimations of threat failure, with a weaker challenger on the left and a stronger one on the right. When a challenger is weaker (Figure 4.2a), the risks of threat failure initially increase in *Challenger's accountability*, after maintaining at a high level (with the failure rate over 85%) in the middling values of accountability, and then decline

⁹Following (Downes and Sechser 2012), I use *Clarify* to generate all the charts with 90% confidence intervals (King, Tomz and Wittenberg 2000). Except issue types, I estimate the probability of threat failure between two contiguous adversaries during the non-World War II period, holding other variables at the in-sample mean.

quickly to less than 60% once the challenger's ACC score exceeds 2. When a challenger is stronger, however, Figure 4.2b shows a much flatter curve with the probability of failure ranging roughly between 50% and 60% across the spectrum of political accountability.

This is consistent with the equilibrium predictions shown in Figure 4.1. On one hand, high political accountability encourages the challenging leader to send clear signals to avoid war when it is unlikely to defeat the target; but when it is likely to achieve military victory, high political accountability discourages the leader with resolute people from signaling credibly as fighting would easily bring the leader political benefits. We observe a similar pattern when leaders demand for changes in both territory and policy alone.

In addition to the effects of independent variables, there are other results worth discussing. The results of the variables of target's domestic politics suggest the importance of differences between measures. In Model 1 and 3, where the theory expects that the credibility of threat increases in the challenging leader's accountability, two measures of the target's domestic politics present opposite effects on the probability of threat failure. First, the outcome variable is positively associated with *Target's accountability*. With all else being equal, the more a *targeted* leader is accountable to her constituents, the more she is likely to reject the threat at the risk of war. This result is consistent with the implication of previous models where domestic audiences are present in both the challenger and the target: a threat is more likely to fail because with the presence of domestic audiences, the targeted leader less likely

to concede (Kurizaki 2007, Tarar and Leventoğlu 2009). Second, the probability of threat failure declines with *Target's polity score*, again, in Model 1 and 3. Cheibub, Gandhi and Vreeland (2010) show that the variation in Polity scores is mostly driven by changes in the dimension of executive constraints, which refers to “the extent of institutionalized constraints on the decision making powers of chief executives” (Marshall and Jaggers 2013, pp. 24). Thus, a negative relationship between *Target's polity score* and threat failure implies that a targeted leader, under greater *institutionalized* constraints on the decision making power, seems more likely to concede, consistent with Filson and Werner's (2004) theory.¹⁰ In sum, Model 1 and 3 demonstrate some evidence on the target's side that democratic regimes can “both inform and constrain,” which suggests that scholars should be more cautious in matching empirical measures of political regimes with theoretical mechanisms.

A comparison across issue types suggests that territorial threats issued by weaker challengers are no more likely to fail than those issued by strong challengers. A potential explanation is that some of the territorial threats are associated not with an informational problem but with a commitment problem, the latter of which, as Moon and Souva (2016) suggest, is irrelevant to the information-based theory presented here. If a territorial threat presents a credible commitment problem, a currently stronger target worries that its concession would lead to a large and rapid shift in the distribution of power

¹⁰The correlation coefficient between *Target's accountability* and *Target's polity score* is 0.542 with statistical significance.

favor to the currently weaker challenger, hence rejecting the challenger's threat and launching a preventive attack.¹¹ Nevertheless, I am refrained from reading too much from the differences between issue types without strong theoretical support.

4.5 Summary

This chapter conducts an empirical analysis on the conditional relationship between political accountability and threat effectiveness derived from the theoretical model. The distribution of power between adversaries conditions the effects of political accountability on bargaining outcomes. When the challenging state is relatively weak, the model identifies a non-linear relationship between political accountability and threat effectiveness, such that crisis bargaining are most likely to fail when the challenging leader is moderately accountable, but more likely to succeed when the leader is highly accountable or not accountable at all. When the challenging state is relatively strong, however, a highly accountable leader is no better in resolving private information problems than a moderately accountable one, implying that high political accountability has little pacifying effect in crisis bargaining. Empirically, I argue that after the theoretical model distinguishes policy costs from audience costs, the extant measures of audience costs, such as the ACC score, approximate to the operationalization of policy costs. Building upon the ACC score, I create

¹¹Given a relatively smaller sample of observations, I keep the territorial disputes associated with a commitment problem according to Moon and Souva (2016), producing a conservative test.

the Political Accountability Index to illustrate both the openness of executive recruitment and the openness of political participation. With the Political Accountability Index, the empirical result illustrates a similar conditional relationship between a challenger's political accountability and the likelihood of target's voluntary compliance in the MCT dataset.

Chapter 5

Dragging Themselves to War: Signaling Resolve to National Leaders

In Chapter 2, I propose two new perspectives to examine the domestic impact on crisis bargaining: (a) the public's concerns both about consistency and about substantive crisis outcomes, the latter of which is mainly determined by the public's resolve; (b) a leader's strategic control of publicity when issuing threats. Building on these observations, in Chapter 3, I use a game-theoretic model to examine how leaders strategically choose the level of publicity of their challenges to signal the public's resolve to foreign targets. The conclusion in Chapter 3 requires a key assumption that leaders have full knowledge, or at least better knowledge than foreign targets, about the amount of resolve their people have—that is, how willing their people are to fight—in a given foreign crisis.

However, there exist some situations under which national leaders are no better informed of the public's preferences than their foreign counterparts. The knowledge about the public's preferences requires the smooth transmission of public opinion to the top as well as the accumulation of governing experiences. While these two requirements seem natural in a well-functioned

and stable government, they are more likely to fail during a regime transition and/or a leadership turnover. A regime transition is often associated with changes in how national governments gather and spread information and how different groups of people are represented at the national level (Mansfield and Snyder 1995). While new leaders may have incentives to demonstrate their individual resolve (Wolford 2007), they are less experienced in gathering and assessing public opinion, especially on foreign affairs. A transition of leadership may also come with changes in the domestic political coalition, shifting the composition of relevant domestic audiences and their policy preferences (Bueno de Mesquita et al. 2003, Croco 2011, Leeds, Mattes and Vogel 2009, Licht 2010, Weeks 2012). Hence, with changes in political regimes, individual leadership, or political coalitions, leaders' prior knowledge about the public's preferences may not apply to new situations. Even without changes in regime or leadership, facing a new enemy may also generate a new set of preferences among the public, unknown to leaders.

Take China's intervention to the Korean War as an example. While scholars have long focused on the uncertainty of the US about China's resolve, they pay less attention to the fact that Chinese leaders were uncertain about whether the Chinese people would tolerate the huge costs of fighting against the US. The historical evidence suggests that even before the success of the Inchon landings in September 1950, Mao Zedong was determined to support North Korea, including military intervention if necessary (Yang 1999, pp. 373). What Mao and other Chinese leaders worried about was the public's resolve—

whether the Chinese public would be willing to fight another war against the most powerful country immediately after a century-long domestic turmoil and foreign invasion. If the public lacked strong resolve, the leaders feared that such a costly intervention would stimulate “reactionaries” who remained hostile to the new Communist regime and create social instability (Chen 1994, pp. 128).

When does the public reveal or hide its preferences about war and peace if such information is unknown to leaders? How does the additional uncertainty of leaders affect crisis bargaining? I modify the model in Chapter 3 by making the domestic public express their preferences about crisis outcomes before the leader’s and the foreign target’s moves. This is more likely if the public has some existing knowledge about the underlying issue of crisis or the underlying issue has been covered before the emergence of crisis, such as territorial disputes. For example, the tensions on the South China Sea have been accumulating in the past few years, featured with the Hague Tribunal’s ruling and China’s continual construction of military bases; if a security crisis occurs in this area, the public should be able to gather basic information and express their preferences before leaders decide whether to further mobilize public attention or to nip it in the bud.

With complete information, I first identify the condition under which (a) the public’s preferences about war and peace vary with its type, and (b) the leader maximizes the level of publicity given public support for war and minimizes publicity without public support for war. Under this condition, I further analyze the model with incomplete information. In particular, I

examine when and why one type of people mimics the other type when the risk of war is relatively high (i.e. small costs of war for the foreign target).

There exists an equilibrium in which the irresolute people will mask its unwillingness to fight and thus suffer from crisis escalation, even though they would not express support for war under complete information. Why would the irresolute people drag themselves into war? The key lies in the leader's off-the-equilibrium strategy of publicity, that is, what the irresolute people expect their leader to do *if* they do not support war initially. When the irresolute people are unwilling to fight, in fact, it is in their best interest that the leader minimizes public attention to the bargaining process: by reducing the influence of irresolute people as much as possible, the leader can maximize the foreign target's offer, which benefits both the leader and the people. However, if the irresolute people expect their leader to maintain a moderate level of publicity during crisis bargaining, their expression against war would make the leader "look bad" in public, further jeopardizing his bargaining position to the extent that a peaceful offer can be worse than their payoffs from war. Therefore, it is possible that the irresolute people would rather fake their support for war from the very beginning.

I identify another pooling equilibrium in which the resolute people pretend to be the irresolute type. Although the resolute public and the foreign target are both willing to fight if necessary, they also have incentives to prevent a crisis from escalating as well. In this pooling equilibrium, war can be avoided for two reasons. First, since the resolute people pretends to be irreso-

lute by opposing war, the domestic anti-war opinion makes both sides always prefer negotiation to war. In other words, a bargaining range always exists between two states. Second, the foreign target decides not to tolerate the risk of war and hence proposes an offer that both types of people prefer to war. The equilibrium exists when neither the resolute people nor the foreign target is extremely willing to fight.

5.1 The Modified Model

The modified model is presented in Figure 5.1. The first change is the information structure: Nature (N) reveals the public's (A) type, resolute (\underline{c}_A) or irresolute (\bar{c}_A) to A only this time. Both the leader (L) and the foreign target (F) only know the distribution of A 's type such that the prior probability of encountering the resolute A is ϕ .

Consistent with the new information structure, secondly, I change the sequence of the game by making the public move before the leader and the foreign target. As mentioned earlier, this is possible if the public faces an enduring enemy or the underlying issue of crisis has been well known, such as territorial disputes. Since the public is considered a direct participant in a crisis and gains a share of external bargaining with the foreign opponent, the public will have incentives to boost its leader's bargaining position if allowed to. By allowing for the public's input *before* the external bargaining, it may open for novel explanations for war.

Formally, after Nature (N) determines A 's type, resolute or irresolute,

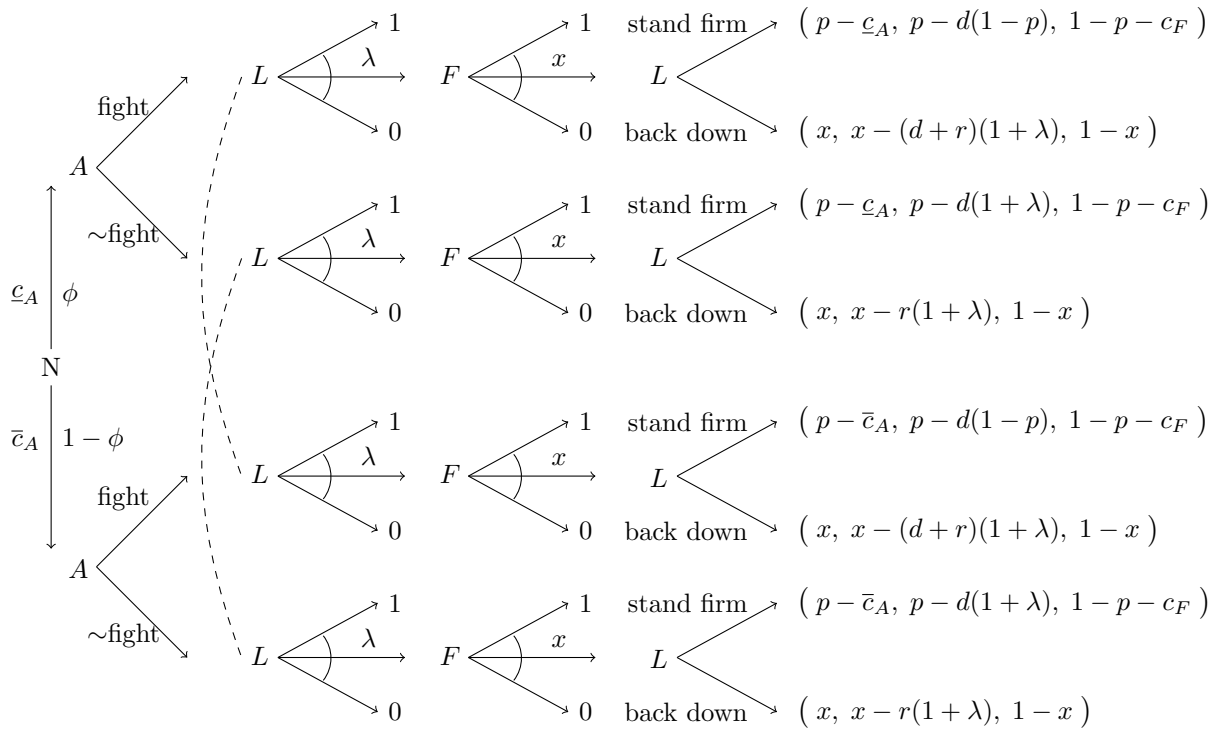


Figure 5.1: Signaling Resolve to National Leader

A first expresses its opinion about war: support or no support. After observing A's expression, L then decides a level of publicity, $\lambda \in [0, 1]$, when issuing an open challenge. The rest of the sequence remains the same: F offers a share of the goods, $x \in [0, 1]$, and finally L decides whether to stand firm or back down.

The third change is related to the leader's payoff for backing down when the public does not support fighting. Recall that in Chapter 3, the public chooses whether to fight or not *after* it pays a certain level of attention and observes the target's offer. If the bargaining process draws great attention

(i.e. $\lambda = 1$) from the public who then prefers a peaceful offer to war, in this case, I argued that the public would relieve their leader from reputation costs if the latter chooses to back down. Formally, if the public chooses not to fight, in Chapter 3, the leader's payoff for backing down is $x - r(1 - \lambda)$.

In the modified model, the public expresses its opinion about war *before* the leader chooses the publicity of his threats and the target proposes its offer. After the actual offer proposed by the target, the public does not further express its opinion about crisis escalation (or not). Without the immediate expression of public opinion, it is more difficult to justify that the leader should be relieved from reputation costs if he chooses to back down. Therefore, if the leader backs down, reputation costs always increase with the level of publicity, that is, $r(1 + \lambda)$. In the modified model, the public's initial opinion about crisis outcomes may affect the leader's choice of publicity, but it does not matter to the amount of reputation costs.

In summary, there are three major changes compared to the original one in Chapter 3 (Figure 3.1). First, the public's resolve (\underline{c}_A or \bar{c}_A) is unknown both to the leader and to the foreign target. Second, when a crisis occurs, the public expresses its preferences about war and peace first; after that, the leader decides a level of publicity of his threat and then the foreign target proposes an offer. Third, since no direct input of public opinion is involved after the target's offer, I assume that the leader's reputation costs always increase with the level of publicity he chooses earlier.

5.2 Equilibrium under Complete Information

In this section, I analyze the model under complete information. Although the results are straightforward and not particularly interesting, we can see how the new sequence changes some basic dynamics among the three actors, highlighting the strategic role of the public. In particular, if the foreign target is willing to make great concessions, the public takes advantage of that and always shows support for war as doing so involves few costs; if the foreign target is tough enough, the public becomes more cautious and decides whether to support fighting based on their costs of war. Proposition 5 describes part of the Subgame Perfect Equilibrium.

Proposition 5. *The following sets of strategies are each part of the Subgame Perfect Equilibrium.¹*

(1) If $c_F \geq c_F^l$:

- Both types of A support fighting;
- L is indifferent between $\lambda = 0$ and $\lambda = 1$ given A 's support for fighting, but chooses $\lambda = 0$ without A 's support for fighting;
- If $c_F \geq c_F^h$, F offers x^l given $\lambda = 0$ or x^h given $\lambda = 1$; if $c_F^l \leq c_F < c_F^h$, F offers x^l given $\lambda = 0$ or zero given $\lambda = 1$.²

(2) If $c_F < c_F^l$:

¹See all the proofs in the appendix.

² $c_F^l = r + dp$; $c_F^h = 2r + d + dp$; $x^l = p + r + dp$; $x^h = p + d + 2r + dp$.

- *Both types of A support fighting if $\underline{c}_A < \bar{c}_A < d - r$; only the resolute A supports fighting when $\underline{c}_A < d - r \leq \bar{c}_A$; neither types of A supports fighting when $d - r \leq \underline{c}_A < \bar{c}_A$;*
- *L's strategy is the same as above.*
- *F offers zero with A's support for fighting; F offers $p - d + r$ with A's opposition against fighting.*

One of the most obvious changes in the modified model is that F bargains only with L whose position is influenced by A 's preference between war and peace. Recall the model in Chapter 3: Since F proposes an offer before the moves of A and L , F is able to propose an offer that both of them prefer to war. In the modified model, F bargains with L who alone makes the decision of whether to escalate a crisis to war. A 's preferences between war and peace only affects L 's bargaining position. For instance, with A 's support for fighting, L would suffer from greater political costs (policy and reputation costs) for backing down and yet pay very small costs for standing firm.³ In other words, F has to offer a huge share of goods in order to avoid war. Knowing that its support for war would strengthen L 's bargaining leverage, A may have an incentive to express support for fighting, which will make L excessively aggressive.

Regardless of its true willingness to fight, A always supports fighting and strengthens L 's position, if the former never has to pay the real costs

³The amount of political costs grows further with a higher level of publicity.

of war. If $c_F \geq c_F^l$, war is too costly to F such that F is willing to make great concessions to compensate L 's costs for backing down. Since peace is guaranteed in this case, both types of A will take advantage of F 's weakness and act (or pretend) tough without considering their payoffs from war.

A becomes more cautious when the risk of war is real. If $c_F < c_F^l$, A 's support for fighting would strengthen L 's bargaining position to the extent that it eliminates the bargaining range between L and F , even when L minimizes the publicity of its threat. In this case, A must compare its war payoff to an agreement without its support for war ($p - d + r$). The cut point is $d - r$: A shows support for fighting when its costs of war are smaller than $d - r$, but opposes fighting if its costs of war exceed this threshold. Obviously, if the costs of war for both types are too large or too small, A 's move does not vary with its type. There exists the condition, $\underline{c}_A < d - r \leq \bar{c}_A$, under which only the resolute A supports fighting.

Proposition 5 shows that war can occur under complete information. Using the unitary-actor framework, the existing literature suggests that a costly war can be avoided under complete information, because a peaceful transfer of payment is possible between two states.⁴ Even when we break the unitary actor assumptions, as Chapter 3 demonstrates, this insight holds as long as the foreign opponent can propose an offer directly to the public who bears the costs of war. In this chapter, I assume that the public only influences

⁴We do not consider the commitment problem here.

the leader's bargaining position, rather than directly bargain with the foreign opponent. By increasing the costs for backing down, the public support for war can advance the leader's bargaining position: on one hand, the leader will face severe political punishment for backing down given the public support for war; on the other, the leader does not bear the societal costs of war directly. However, it is possible that the public support for war eliminates the bargaining range between the leader and the foreign opponent (i.e. $c_F < c_F^l$). Why does the public still choose to support fighting if such choice guarantees the escalation of crisis? The public will support fighting if its opposition against war damages the leader's bargaining position such that it produces an offer smaller than its payoffs of war, or formally if $c_A < d - r$. With all else being equal, the public is more likely to support and risk war if the difference between policy costs and audience costs increases.

The equilibrium with complete information also sheds light on where to look for the equilibria when the public's resolve is unknown. First of all, it is meaningful to examine how uncertainty matters to crisis bargaining only when the risk of war is real. If F 's costs of war are too high, A always expresses support for war to maximize F 's offer, with any concern about the danger of war. Whether L or F has knowledge about A 's resolve does not affect the bargaining outcome. Therefore, I will analyze the model with incomplete information $c_F < c_F^l$. Second, even though the risk of war is real, the uncertainty about A 's type does not matter if the costs of war for both types are either too small ($\underline{c}_A < \bar{c}_A < d - r$) or too large ($d - r \leq \underline{c}_A < \bar{c}_A$). Again,

this is different from the previous model in which F tailors its offer sometimes based on A 's costs of war. In the modified model, while A 's initial opinion may influence L 's bargaining position, the value of c_A is irrelevant to the exact offer proposed by F . In addition to $c_F < c_F^l$, therefore, I further restrict the condition to $\underline{c}_A < d - r \leq \bar{c}_A$ when introducing incomplete information in the next section.

5.3 Equilibrium under Incomplete Information

When does A reveal or hide its type to the other players? There exist multiple equilibria under incomplete information. When $c_F < r + dp$ and $\underline{c}_A < d - r \leq \bar{c}_A$, Proposition 5 shows that A 's choice of fighting varies with its type. It is not surprising that a separating equilibrium exists such that war occurs with the resolute A only. Given their distinctive strategies under complete information, I am interested in when one type pretends to be the other type, in particular, whether the irresolute A fakes resolve and drags itself to war. This section focuses on the existence and conditions of two pooling equilibria. After briefly introducing the separating equilibrium, I begin with the first one in which both types of A supports fighting, and then discuss the second pooling equilibrium in which both types oppose fighting. As the following discussion shows, a key condition of pooling strategies is that L 's off-the-equilibrium strategy cannot differ too much from its on-the-equilibrium strategy.

Proposition 6. *If $\underline{c}_A < d - r \leq \bar{c}_A$, the following sets of strategies are each part of a Perfect Bayesian Equilibrium:*

- *The resolute A supports fighting; the irresolute A opposes fighting;*
- *L chooses $\lambda = 1$ given A 's support for fighting, and $\lambda = 0$ given A 's opposition against fighting;*
- *F offers zero given A 's support for fighting, and $p - d + r$ given A 's opposition against fighting.*

Proposition 6 presents a separating equilibrium: war occurs with the presence of the resolute A only. F and L are able to update their knowledge about A 's type by observing its strategy. When the irresolute A reveals its type by opposing war, L needs to minimize the negative impact of domestic anti-war opinion on its bargaining position; formally, L chooses $\lambda = 0$ to minimize the amount of policy costs for standing firm. Given $\lambda = 0$, F proposes $p - d + r$ which both L and the irresolute A prefer to war. Since F 's offer is greater than the irresolute A 's war payoff, the irresolute A has no incentive to mimic the resolute type, which would simply lead to war.

If A is the resolute type, it reveals resolve with the support for war. A 's support for war boosts L 's bargaining position by reducing L 's costs for standing firm drastically. Given the existing assumptions, there is no bargaining range between F and L if A expresses support for war. Hence, F would rather offer L nothing and both sides end up fighting. With smaller costs

of war, the resolute A will not deviate from the support for war, as its war payoff is greater than $p - d + r$. In this separating equilibrium, even though leaders fully update their information, it does not always guarantee a peaceful outcome: war still occurs when A is the resolute type. This is similar to the equilibrium under complete information: knowing that its support for war eliminates the bargaining range between F and L (i.e. $c_F < r + dp$), the resolute A is willing to support fighting and escalate the crisis to war when $\underline{c}_A < d - r \leq \bar{c}_A$.

Proposition 7. *If $\underline{c}_A < d - r \leq \bar{c}_A < (1 + \lambda)(d - r)$, the following sets of strategies are each part of a Perfect Bayesian Equilibrium:*

- *Both types of A support fighting, which leads to war;*
- *L chooses $\lambda = 1$ given A 's support for fighting, and $\lambda \in (0, 1)$ given A 's opposition against fighting;*
- *F offers zero given A 's support for fighting, and $p - (1 + \lambda)(d - r)$ given A 's opposition against fighting.*

If either F or L alters its strategy after A 's opposition against war, a different equilibrium may emerge. Both types of A express their support for fighting in Proposition 7. It is not very surprising to see that the irresolute type mimics the resolute type *per se*. In most of the signaling games, the irresolute type usually faces no risk of war: it has the option of backing down if mimicking

the resolute type fails to achieve a better offer. In fact, the irresolute type's threat is incredible precisely because of no risk involved. However, one of the key features in this pooling equilibrium is that war occurs regardless of A 's type. A 's support for war will eliminate the bargaining range between L and F . Therefore, once it decides to mimic the resolute type, the irresolute A knows for sure that the support for fighting will lead to crisis escalation, a path it is unwilling to take under complete information.

Why does the irresolute A hide its type, knowing that it would certainly lead to war? The answer lies in the off-the-equilibrium strategies of L and F . Suppose that the irresolute A opposes fighting, we know from the previous discussion that it is optimal that L minimizes the value of λ : by minimizing the level of public attention, L is able to reduce the negative impact of anti-war opinion on its bargaining position as much as possible. If A is the irresolute type, choosing $\lambda = 0$ would maximize F 's offer and thus benefit the home country as a whole. In Proposition 7, however, I specify that L does not reduce λ to zero in the off-the-equilibrium path (i.e. A 's opposition against fighting); as a result, F would be able to avoid war by proposing an offer smaller than $p - d + r$. The off-the-equilibrium offer, $p - (1 + \lambda)(d - r)$, can be smaller than the irresolute A 's payoff from war, if $\bar{c}_A < (1 + \lambda)(d - r)$. Therefore, if L is expected to keep λ at a moderate value, it is possible that the irresolute A would rather fake their support for fighting. As long as the value of \bar{c}_A is not too high, the irresolute A would rather support fighting than make its leader look bad openly during crisis bargaining.

Proposition 8. *If $d - r - \phi d(1 - p) < \underline{c}_A < d - r \leq \bar{c}_A$, the following sets of strategies are each part of a Perfect Bayesian Equilibrium:*

- *Both types of A oppose fighting, and there is no risk of war;*
- *L chooses $\lambda \in (0, 1]$ given A 's support for fighting, and $\lambda = 0$ given A 's opposition against fighting;*
- *F offers zero given A 's support for fighting, and $p - \underline{c}_A$ if $c_F \geq c_F^*$ given A 's opposition against fighting.⁵*

Proposition 8 illustrates the second pooling equilibrium under which neither types of A supports fighting. The bargaining ends with a peaceful agreement not because the problem of uncertainty is solved, but because the resolute A is willing to mimic the irresolute type. The resolute A is more willing to fight than the irresolute type, but it also prefers peace as long as F proposes an offer greater than or equal to $p - \underline{c}_A$.

Without updating its belief about A 's type, F faces a typical risk-return tradeoff: whether to propose a smaller offer which risks war with the resolute A , or to propose a greater offer which both types of A prefer to fighting. The analysis shows that if $c_F \geq c_F^*$, F prefers greater concessions to the risk of war and therefore proposes $p - \underline{c}_A$ given A 's opposition against war. Combined

⁵ $c_F^* = \frac{(1-\phi)(d-r)-\underline{c}_A}{\phi}$. For $c_F^* \leq c_F < r + dp$ to exist, it requires $d - r - \phi d(1 - p) < \underline{c}_A$.

with $c_F < r + dp$, formally, we should have $c_F^* \leq c_F < r + dp$. This range exists as long as the value of \underline{c}_A is not too small, or $d - r - \phi d(1 - p) < \underline{c}_A$. Substantially, given A 's support, it is impossible for F to bargain with L ; given A opposition against fighting, there always exists some bargaining range unless both F and the resolute A are extremely tough.

5.4 Summary

Despite the constant evaluations of public opinion, it is still possible that leaders understand the public's preferences no better than foreign targets. I present a modified model by allowing the public to express its opinion about crisis outcomes before crisis bargaining between its leader and a foreign target. In the modified model, while the public does not directly bargain with the foreign target, public opinion affects the target's offer and the bargaining outcome by changing the leader's bargaining position. With complete information, the model shows that (a) the leader tends to maximize the level of publicity given public support for war and minimize publicity without public support for war, and (b) public support for war may eliminate the bargaining range between the leader and the foreign target. I further analyze the model with incomplete information. In particular, I examine when and why the public reveals or hides its preferences about war and peace when the risk of war is relatively high (i.e. small costs of war for the foreign target).

I identify an equilibrium in which the irresolute people will pretend its willingness to fight and drag themselves to war, even though they would

oppose fighting under complete information. The answer depends on what the irresolute people expect their leader to do and how it affects the target's offer *if* they do not support war initially. Strategically, the public wants to enjoy the share of goods without necessarily suffering from war. If the irresolute people anticipate their leader to minimize their impact in international bargaining, they are more willing to express their true preference without worrying that their dovish opinion might jeopardize the leader's bargaining power. However, if the leader is expected to maintain a moderate level of public influence, for example, by allowing for open debate or protest, the irresolute people would refrain from completely revealing their unwillingness to fight. Even a moderate level of publicity amplifies the public's dovish voice and puts their leader in a weaker position, which may yield a share of goods worse than their payoffs from war.

I identify another pooling equilibrium in which the resolute people pretend to be the irresolute type. Although the resolute public and the foreign target are both willing to fight if necessary, they also have incentives to prevent a crisis from escalating as well. In this pooling equilibrium, war can be avoided for two reasons. First, since the resolute people pretends to be irresolute by opposing war, the domestic anti-war opinion makes both sides always prefer negotiation to war. In other words, a bargaining range always exists between two states. Second, the foreign target decides not to tolerate the risk of war and hence proposes an offer that both types of people prefer to war. The equilibrium exists when neither the resolute people nor the foreign target

is extremely tough.

The modified model yields new findings about how different choices of publicity affect the problem of uncertainty and crisis outcomes. First, war may occur in equilibrium under complete information even without repeated interactions. The previous literature shows that inefficient fighting may occur under complete information when unitary state are able to use war to impose costs on their opponents (Slantchev 2003). This chapter provides a domestic explanation on the occurrence of war under complete information when we break the unitary actor assumptions. In general, war is possible when (a) leaders do not bear the societal costs of war, and (b) foreign opponents cannot directly compensate the public's costs of war. When the public does not directly participate in external bargaining, its preference between war and peace affects crisis outcomes only by changing the leader's bargaining position. Under certain conditions, when the public expresses its support for war to demonstrate resolve, it advances the leader's bargaining position to the extent that any peaceful settlement cannot compensate the leader's political costs for backing down. The public's resolve does not lead to a greater offer, but makes it impossible for both sides to look for a mutually preferable agreement.

Second, the modified model provides a domestic origin of aggressive (hawkish) leaders. The term "aggressive" or "hawkish" indicates that the cost for fighting is small relative to the cost for peace. The majority of literature considers aggressiveness as an exogenous yet private trait of leader and thus examines the conditions under which leaders reveal or hide their true prefer-

ences (Cukierman and Tommasi 1998, Schultz 2005). Instead of focusing on individual traits, this chapter shows how and why a seemingly aggressive leader on the negotiation table is created by an aggressive public. In this sense, a leader serves as an agent of the public. Whether a leader appears aggressive in external bargaining depends largely on the public's resolve. It is complementary to the model of political bias and war established by Jackson and Morelli (2007), who offer an alternative principal-agent model of war: under complete information, war may occur if citizens select a politically-biased leader who sees more gains from war than the country he represents. Jackson and Morelli (2007) examine when citizens select a politically-biased individual to be their leader, while this chapter studies when citizens choose to support fighting, encouraging their leader to act more aggressively in external bargaining.

Finally, by reversing the informational asymmetry in principal-agent games, the modified model focuses on uncertainty about the principal and discusses when the principal may or may not address the informational asymmetry. Chapter 3 shows that increasing publicity may not signal the public's resolve because the leader with the irresolute people sometimes mimics the one with the resolute people—the foreign target is unable to fully distinguish between two types of the public. Many analyses on nationalism or public mobilization in general fall in this category. This chapter discusses how the tendency of publicizing crisis may exacerbate the problem of uncertainty, increasing the risk of war. If people sense that their leader tends to mobilize their attention or involvement in any crisis, they always act supportive of fighting,

which prevents their leader and the foreign target from learning new information. This is a problem of nationalism or public mobilization that existing studies ignore. The undifferentiated tendency of involving and mobilizing the public in crisis can be ineffective or counterproductive, not only because it appears incredible to foreign observers but also because it prevents leaders from learning the true preferences of their people.

Chapter 6

Conclusion

This dissertation begins with the question about the role of public threats in signaling resolve and the risk of escalation in international conflict. According to the conventional wisdom, one of the major reasons to the failure of a peaceful settlement is the existence of uncertainty among bargaining states and their incentives to misrepresent (Fearon 1995). One of the sources of uncertainty is resolve, which I define the costs of war an actor is willing to bear in the event of war. To address the problem of uncertainty, a resolute state sends a costly signal which invokes extra costs on concession and thus makes crisis escalation more attractive, separating itself from an irresolute state (Fearon 1997). What do such extra costs originate from? The audience costs theory proposes an answer from the perspective of domestic politics. After publicly threatening to use force against a foreign opponent, a leader is subject to extra costs—imposed by the domestic public—for making concessions to the opponent; a leader’s threat to use force is more credible if the public is capable of imposing greater costs (Fearon 1994).

The audience costs theory contributes to our understanding of costly signaling by highlighting the public as a critical source of extra costs for con-

cession, but it raises more questions about how the logic of costly signaling operates in a domestic context. For instance, scholars explore other sources of extra costs for concession (Arena 2015, Ramsay 2004, Schultz 1998, Weeks 2008), the rationale of extra costs for concession (Slantchev 2006, Smith 1998, Snyder and Borghard 2011, Tomz 2007), and the initiatives or reactions of leaders in this logic (Debs and Weiss 2014, Kurizaki 2007, McManus 2014, 2016, Weiss 2013). Despite numerous works on domestic audience costs, we need a new domestic mechanism of costly signaling integrating theoretical, experimental, and empirical insights scholars have generated in the past two decades. It is no longer satisfying to argue whether democracies or autocracies have advantages in signaling their resolve credibly. Instead, scholars have shown that both democracies and autocracies *can* have advantages and limitations in credible signaling. A more important question becomes when or under what conditions states with certain political features can credibly signal their resolve.

To establish an integrated framework, Chapter 2 revisits three assumptions related to audience costs theory. First, since resolve refers to the costs of war associated with human costs and economic destruction, I argue that it is the public rather than their leader who bears those costs in the event of war. This is a major deviation from the majority of existing literature in which the public was usually considered as a judge evaluating their leader's performance after the end of crisis. Instead, I focus more on the public's role as a direct participant in militarized conflict if negotiation fails. Related to the first

change, I further assume that the public evaluates their leader's performance based not only on the leader's consistency between words and deeds, but also on whether the leader's final decision represents their own preferences between war and a peaceful settlement. While one of the most frequent challenges to audience costs theory is that the public cares both about national prestige and policy substance, scholars have not suggested how to model the public's policy preference or explicitly theorized how these concerns might affect their evaluations of leadership and ultimately our understanding of costly signaling. With resolve located in the public, it is straightforward to model how the public weighs war against a peaceful settlement, similar to the way scholars model a unitary state's preference in crisis bargaining. Finally, I incorporate the recent evidence that leaders have leeway in adjusting how they issue a public threat, arguing that the amount of political costs leaders are subject to depends both on exogenous institutional difference and on their endogenous control of public attention to crisis.

Building from the new assumptions, Chapter 3 presents a crisis-bargaining model examining how leaders signal the public's resolve by choosing a level of publicity in their threats. With the opponent uncertain about the public's resolve, the model uncovers whether uncertainty matters to crisis bargaining and if so, how leaders choose the publicity of threats to influence the opponent's beliefs under different levels of political accountability to public preferences. The model shows that a leader's signaling pattern—how a level of publicity signals the public's resolve—varies with a leader's political accountability to

the public's policy preferences. If leaders are rarely accountable to the public's policy preferences, the information about the public's resolve does *not* affect crisis bargaining; leaders always choose low publicity with nothing to do with signaling the public's resolve. If leaders are moderately accountable, a high-profile threat will increase an opponent's belief of facing the resolute people, but leaders with the irresolute people may have incentives to exaggerate the public's resolve. The logic is similar to that of audience costs theory, but it applies to leaders with moderate accountability only. If leaders are highly accountable, the signaling logic is opposite. To avoid war, leaders with the resolute people will choose a low-profile threat while leaders with the irresolute people will choose a high-profile threat. Surprisingly though, leaders with the resolute people sometimes prefer a popular war to a peaceful settlement and thus, those with the resolute people will make high-profile threats to mask the public's resolve and engage into war. In brief, greater publicity is not always a signal of resolve.

Chapter 3 also establishes a conditional relationship between political accountability and the risk of war. Whether higher political accountability reduces war depends on the target's capabilities relative to the challenger's. A higher level of political accountability reduces the risk of war only when a target is strong, as leaders with the resolute people have incentives to address uncertainty and thus avoid war. If a target is weak, a high level of political accountability does not reduce the risk of war, because leaders with the resolute people may gain more benefits by defeating a weak target and thus choose not

to reveal true information. The empirical implication of the model is supported by the statistical results in Chapter 4.

Finally, Chapter 5 explores how leaders use public threats to learn the public's resolve in crisis bargaining. It expands the original model by assuming that the public's resolve is unknown to both leaders and a foreign target. The modified model shows that war is possible under complete information, as the public support for war may make peace costly to *leaders* and thus eliminate the bargaining range with foreign opponents. Under incomplete information, war may also occur if leaders are unable to learn the public's preferences. More specifically, there exist conditions under which the irresolute people exaggerate their willingness to fight even when they anticipate that a peaceful negotiation will fail. If the irresolute people anticipate that their leader tends to raise public attention or involvement into a crisis, regardless of their true preferences, revealing their lack of resolve would put the leader at a disadvantage in crisis bargaining.

By locating the costs of war in the public and allowing leaders to the level of publicity during a crisis, the dissertation offers an integrated framework to understand the apparently disparate strategies of leaders under different levels of political accountability. It explains why moderately accountable leaders can signal public resolve with high-profile threats, and why highly accountable (i.e. democratic) leaders may back down easily after high-profile threats or even give up signaling sometimes. Second, it contributes to the domestic explanations of war and peace by integrating the variation of political

institutions with the variation of domestic preferences. The audience costs theory emphasizes the capability of domestic punishment, yet it is puzzling why citizens in democracies should always support crisis escalation if war is costly to citizens and they are able to punish leaders. The earlier democratic peace literature simply assumed that the public held preferences for peace. By explicitly modeling the interaction between the capability and the incentive of public punishment, I show new insights about how institutions and interests influence the intervening variable of publicity, which influences information revelation and negotiation outcomes. Finally, it also identifies the conditions under which a high level of political accountability produces perverse outcomes costly to the public. By tracing the divergent interests between leaders and the public, the dissertation shows that high political accountability can either increase or decrease the risk of war, depending on the target's relative strength. Democratic leaders are highly accountable to the public's policy preferences, yet they influence crisis outcomes—a peaceful settlement or war—by choosing whether to reveal true information to foreign opponents. When bargaining with a weak target, democratic leaders from powerful states are tempted to provoke the resolute people and trap them into war.

In the rest of this chapter, I discuss its contributions and possible extensions in three aspects. First, I emphasize the connections between experimental works and the theory developed in Chapter 3. Existing experimental works not only confirm some assumptions of audience costs theory, but also provide new micro-foundations to improve the domestic mechanism of signal-

ing. New assumptions and findings developed in this dissertation will also rely on further experimental analysis. Second, by focusing on the endogenous control of publicity, this dissertation examines when increasing publicity enables a leader to signal the public's resolve. Future research in this direction includes (a) constructing measurement on the variation of publicity, (b) collecting data about how leaders make public threats in foreign crises, and (c) conducting experiments to see whether new assumptions in this dissertation have a microfoundation. This chapter closes with discussions on how theoretical insights help us understand the international behaviors of great powers and the impact of anti-foreign mass protests.

6.1 Experiments and Theory Building

Building upon recent developments on audience costs theory, this dissertation reexamines the roles of leaders and the public in crisis bargaining. The emergence of experimental studies in international relations enables scholars to identify the microfoundations of audience costs theory: whether and why the public punishes their leader in a foreign crisis. On one hand, experiments have shown the evidence that individuals are more likely to punish their leader for not carrying out threats as they are concerned about the loss of national prestige (Davies and Johns 2013, Tomz 2007). On the other, scholars gain new insights from survey experiments that individuals also evaluate their leader based on their dispositional preferences between a peaceful settlement and fighting (Kertzer and Brutger 2016). Experimental works advance our

understanding of audience costs theory, by confirming the original assumption of the theory's microfoundation and generating new insights about the public's preferences. They have provided important insights about "how to move forward our thinking about domestic politics, regime type, and strategic interaction in international crises (Schultz 2012, pp. 373)."

To my knowledge, this dissertation is the first attempt to incorporate the new insights from experimental studies to improve the model of audience costs theory or more broadly, a domestic theory of crisis bargaining. Experiments show that the consistency between words and deeds is only one of the dimensions when individuals evaluate their leader's performances. Throughout this dissertation, I emphasize that the public should be considered not only as a distant audience judging its leader's performance, but also as a direct participant in the event of war. The public's new role suggests (a) the importance of modeling the public's preference between militarized conflicts and a peaceful settlement, and (b) a new type of political costs—policy costs—that leaders would pay for not following the public's preferred policy.

With more solid microfoundations, this dissertation presents an integrated framework which reconciles the conventional audience costs theory and subsequent critics. The main findings reinforce the conventional wisdom that signaling resolve increases a leader's credibility in crisis bargaining, but they challenge the linear linkage between increasing public attention and signaling resolve. The original mechanism of audience costs—increasing public attention signals resolve—exists when leaders are accountable to the hawkish (resolute)

type of people only. When leaders are accountable to all types of people, hawkish (resolute) or dovish (irresolute), increasing public attention makes backing down easier and thus signals flexibility. To enhance credibility and bargaining advantage, highly accountable leaders will decrease public attention.

This dissertation also provides new areas for future experimental research on the microfoundation of public threats in costly signaling. Previous experiments on audience costs theory focus more on the features of individual respondents without considering the strategic impacts of leaders. Recently, there is growing interest in experimental studies regarding the role of leaders in public opinion but most of them focus on elite as the source of information.¹ In this dissertation, I contend that the amount of political costs imposed on leaders vary with the level of public attention. For instance, I assume that the public puts more weights on current crises than on future reputation; therefore, greater attention to current crises makes the public weigh their policy preference over national prestige related to future reputation. Future experimental research can examine whether and how individuals evaluate their leader's performance in the context of international crisis given different levels of attention.

6.2 Publicity of Threat and Beyond

Different from audience costs theory, this dissertation is the first attempt modeling a leader's choice of publicity in international crisis. I argue

¹See Bullock (2011), Guisinger and Saunders (forthcoming).

that the distinction between public and private action is misleading, as scholars on audience costs are conflating different concepts such as threat, crisis, negotiation, and diplomacy. As a result, scholars sometimes disagree with each other regarding whether an event or action is considered public, which has direct impact in the validity of audience costs theory. For instance, some scholars consider China's threat to the United States during the Korean War a private threat based on the private mediation by the Indian Ambassador (Fearon 1994); the failure of a private threat is thus consistent with audience costs theory. Other scholars believe that Chinese leaders made public threats from the evidence of several public warnings and propose an alternative theory to explain the failure of those threats (Sartori 2002). One of the most frequently cited examples of private crisis is the Cienfuegos crisis (Kurizaki 2007). However, the U.S. media were fully aware of the military facilities developed by the Soviet Union. What Nixon and Kissinger were manipulating was how to prevent a "crisis mood" given that the public's awareness of this crisis or tension (Crall and Martin 2013). Instead of comparing public to private negotiations, I focus on how a leader further increases or decreases the level of publicity *given* that a crisis is known to the public.

This new approach not only avoids the confusions about public threats as mentioned above, but also separates the effects of individual leaders from those of exogenous political environments, opening areas for further research. In addition to new experiments on the microfoundation of publicity, future research can explore the effects of publicity on crisis outcomes. Due to the

data limitation, I only test the empirical predications about the conditional effects of political accountability on crisis outcomes. In fact, the model in Chapter 3 also yields empirical predictions about the effects of publicity on crisis outcomes. For instance, whether increasing publicity advances a leader's bargaining position depends on a leader's political accountability. If a leader is moderately accountable, increasing publicity partially signals resolve, increasing the leader's bargaining position. If a leader is highly accountable, increasing publicity hurts a leader's bargaining position; the leader either accepts a smaller offer which his irresolute people prefer to war, or fights with the resolute people's support. Further research can build new datasets about the publicity of threats and test these empirical implications.

Scholars can begin with the existing dataset on international crisis or militarized conflict and examine how leaders involve public attention in each case. Despite the differences of political environments, I contend that both democratic and authoritarian leaders are able to manipulate the level of publicity, sometimes with different approaches. I propose two aspects of leaders' behaviors. The first collects the information about public statements, such as the rank of speaker, location, and frequency, etc. If a cross-national comparison is possible, scholars can further investigate the wording or tone of public statements (McManus 2014). The second aspect examines whether leaders take additional actions to increase or decrease public involvement. In democracies, leaders can increase public involvement by holding congressional debates, hearings, conducting referendums, etc (Debs and Weiss 2014). In

authoritarian regimes, leaders can manipulate public involvement through its control of state-owned media and political participation, e.g. editorials in the state-owned media, permission of anti-foreign protests, etc.

This dissertation highlights the initiatives of leaders across different political environments. By allowing leaders to choose the level of publicity, the theory explains why both democratic and authoritarian leaders are able to signal the public's resolve under certain conditions. In addition to the choice of publicity, I propose two directions for future research. The first one is related to the content of threat. The majority of literature examines whether a leader fails to implement threats to use force. Will leaders be punished if they fail to honor promises of staying out of a conflict? On April 6, 2017, U.S. President Donald Trump authorized a military strike against a Syrian airbase after Trump's enduring position about staying out of Syria and Secretary of State Rex Tillerson's statement hinting no intervention of the U.S. On March 30, Secretary Tillerson commented that Assad's long-term status "will be decided by the Syrian people." A recent experiment suggests that individuals also punish leaders for not honoring promises of staying out of a conflict (Levy et al. 2015). Future research can examine why leaders send a benign signal in international crisis and when they break their promises.

Second, while this dissertation takes the public's policy preference as given, future research can explore how the formation of public opinion may affect crisis bargaining. Existing studies, mainly drawn from the literature of American politics, focus more on the role of political elites in shaping public

opinion (Berinsky 2007, 2009, Bullock 2011). The current development in the U.S. politics shows the importance of leaders in shaping public opinion. Future research can explore (a) when leaders can influence public opinion about foreign policies (even with fake news), and (b) how such capability influences bargaining outcomes. In the context of security crisis, an analysis on how leaders shape public opinion on disputed issues may extend the scope of research to the pre-crisis stage. Despite numerous works on the escalation from crisis to war, we do not know much about the pre-crisis stage. One may consider the pre-crisis stage as a process in which leaders can cultivate the public's preferences about disputed goods and its tolerance of war. These preferences may not be clear until leaders bring these issues to public attention.

6.3 Policy Implications

This dissertation has implications for the U.S. foreign policy. It shows that democratic leaders have an advantage in revealing information and avoiding war only when their targets have strong military capabilities. When their targets have weak military capabilities, democratic leaders are either prone to war (given the hawkish public) or put at a disadvantage in bargaining (given the dovish public). Therefore, a democratic great power, such as the U.S., does not reduce the risk of war. Based on this logic, the U.S. should be cautious in involving a crisis with a smaller or weaker state: either the U.S. leaders are likely to risk war as an easy defeat leads to political benefits, or they may receive an unfavorable deal by revealing the lack of domestic resolve. The first

scenario occurs when the American public is considered “resolute” or more tolerate with the costs of war. Two more questions remain to be answered. First, Trump’s “America First” idea is gaining popularity, which may indicate a new wave of isolationism among the American public. Will the rise of isolationism make the American public more or less resolute when dealing with potential international crises? Will their attitudes switch with the feature of crisis—less resolute in a third-party intervention and more resolute in a direct confrontation? Second, how do we consider smaller states with close allies or partners? In particular, do we consider Syria and North Korea as a weak target or a strong target given their external support from Russian and China, respectively? The answers to these questions will be critical in understanding the U.S. foreign policy in the near future.

Also, this dissertation offers new perspectives in understanding mass movement in international crises or tensions. I define mass movement loosely, with various forms in different regime types, such as anti-foreign protests, mass mobilization through party-controlled organizations, referendums, public debates, etc. How does mass movement matter in international crisis? Recent studies suggest that authoritarian leaders can signal resolve to foreign opponents by giving a green light to anti-foreign protest (Weiss 2014), which is confirmed in Chapter 3. I argue that moderately accountable leaders can signal resolve by raising the level of publicity, but the risk of war cannot be fully eliminated as the irresolute type of people has incentives to mimic the resolute one. In other words, mobilizing anti-foreign protests, or mass movement in

general, may signal resolve and increase bargaining leverage to some extent, but it is still a risky move as foreign opponents are unable to fully distinguish between genuine protests and manipulated ones. In Chapter 5, mass movement is viewed as a process through which the public expresses opinion and leaders learn its preference. The key argument is that the tendency of mobilizing mass movement under all circumstances will prevent leaders from learning the public's true preference, leading to a costly war which would not occur if leaders are more selective in mass mobilization.

Appendices

Appendix A

Chapter 2 Appendix

I solve the baseline model under complete information, using backward induction. The solution concept is subgame perfect equilibrium (SPE). Since the focus is how much F should offer in order to prevent war, I further assume that $c_F > r - d$, assuring that F has sufficient incentives to achieve an agreement.

I begin with L 's strategy. Given that A chooses not to fight, L will back down if $x \geq p - d + r$. Given that A chooses to fight, L will back down if $x \geq p + dp + r$.

A chooses whether to fight or not based on its expectation about L 's strategy. If $x \geq p + dp + r$, since L always backs down, A is indifferent between fighting and no fighting. If $x < p - d + r$, since L always stands firm, A is also indifferent between fighting and no fighting. If $p - d + r \leq x < p + dp + r$, L stands firm only when A chooses to fight, meaning that A 's preference over fighting will directly affect L 's decision. In the last situation, A chooses not to fight as long as $x \geq p - c_A$.

F 's goal is to achieve a peaceful agreement with a minimum offer. Since $p + dp + r$ is greater than the other two options, F 's best strategy is to propose

an offer such that L will back down after A chooses not to fight. Formally, F must propose the greater value between $p - d + r$ and $p - c_A$. Therefore, when $d \geq c_A + r$, F proposes $x = p - c_A$; when $d < c_A + r$, F proposes $x = p - d + r$.

Appendix B

Chapter 3 Appendix

Proof of Proposition 1

Define L 's reservation value as the offer that makes L indifferent backing down and standing firm after A does not support war: $p-d+r$ if $\lambda = 0$, or $p-2d$ if $\lambda = 1$. Define A 's reservation value as the offer that makes A indifferent between fight and no fight: $p - \underline{c}_A$ given the resolute type, or $p - \bar{c}_A$ given the irresolute type. Since war is costly to F , under complete information, F can avoid war by proposing an offer that both A and L prefer to war. The key is to identify whose reservation value is larger given L 's choice of publicity and A 's type. To simplify the proof, I first do not specify A 's type, that is, A 's reservation value is $p - c_A$.

When $p - c_A < p - 2d$ or $2d < c_A$, A 's reservation value is always smaller than L 's, regardless of L 's choices of publicity. F should make an offer that renders L indifferent. The offer is $p - d + r$ if $\lambda = 0$ or $p - 2d$ if $\lambda = 1$. A always chooses not to fight. L would receive $p - d$ if $\lambda = 0$ or $p - 2d$ if $\lambda = 1$. Since $p - d > p - 2d$, L prefers $\lambda = 0$.

When $p - d + r \leq p - c_A$ or $c_A \geq d - r$, A 's reservation value is always

greater than L 's, regardless of L 's choices of publicity. F proposes $p - c_A$, making A indifferent between fight and no fight. L would receive $p - c_A - r$ if $\lambda = 0$ or $p - c_A$ if $\lambda = 1$. Thus, L prefers $\lambda = 1$.

When $p - 2d \leq p - c_A < p - d + r$, to make both A and L indifferent, F must propose $p - c_A$ if $\lambda = 1$ but propose $p - d + r$ if $\lambda = 0$. L would receive $p - c_A$ if $\lambda = 1$ or $p - d$ if $\lambda = 0$. Thus, L prefers $\lambda = 1$ if $d \geq c_A$ or $\lambda = 0$ if $d < c_A$.

Put together, if $d < c_A$, L chooses $\lambda = 0$ and F proposes $p - d + r$; if $d \geq c_A$, L chooses $\lambda = 1$ and F proposes $p - c_A$. A and L always accept the offer.

Suppose A has two types. If $d < \underline{c}_A < \bar{c}_A$, A 's type does not affect the equilibrium. All actors play the same strategies as in $d < c_A$. If $\underline{c}_A < \bar{c}_A \leq d$, while L always chooses $\lambda = 1$, A 's type affects F 's offer. F proposes $p - \underline{c}_A$ to the resolute type, and $p - \bar{c}_A$ to the irresolute type.

If $\underline{c}_A \leq d < \bar{c}_A$, A 's type affects both F 's offer and L 's choice of λ . On one hand, since $d \geq \underline{c}_A$, we know from above that given the resolute type of A , L chooses $\lambda = 1$ and F proposes $p - \underline{c}_A$. On the other, we have $d < \bar{c}_A$, meaning that given the irresolute type of A , L chooses $\lambda = 0$ and F proposes $p - d + r$.

The four reservation values identified above are the potential options for F . In any given situation, L 's choice and A 's type yield a pair of reservation

values and F picks the larger value to propose an offer. What F would offer under complete information has strong impacts on L 's choice of publicity under incomplete information. By assumption, we have known that

$$p - d + r > p - 2d \Leftrightarrow -r < d,$$

$$p - \underline{c}_A > p - \bar{c}_A \Leftrightarrow \underline{c}_A < \bar{c}_A.$$

More importantly, we need to compare the reservation values between L and A under different circumstances. Formally, we need compare the values of $d - r$, $2d$, \underline{c}_A and \bar{c}_A .

To facilitate the proofs, I use \underline{A} to represent the resolute A and \bar{A} to represent the irresolute A . $L_{\underline{A}}$ refers to L with \underline{A} and $L_{\bar{A}}$ refers to L with \bar{A} .

Proof of Proposition 2

When $d < \underline{c}_A < \bar{c}_A$, we know from Proposition 1 that L always chooses $\lambda = 0$ and F always proposes $p - d + r$. F 's uncertainty about A 's type does not affect their strategies.

Proof of Proposition 3

When $\underline{c}_A \leq d < \bar{c}_A$, we have the following possible circumstances: ① $d - r < \underline{c}_A < 2d < \bar{c}_A$, ② $d - r < \underline{c}_A < \bar{c}_A \leq 2d$, ③ $\underline{c}_A \leq d - r < 2d < \bar{c}_A$, and ④ $\underline{c}_A \leq d - r < \bar{c}_A \leq 2d$.

First, I show that the separating equilibrium does not exist in which $L_{\underline{A}}$ chooses $\lambda = 1$ and $L_{\bar{A}}$ chooses $\lambda = 0$. Then, I prove the nonexistence of the pooling equilibrium in which L always chooses $\lambda = 1$. Finally, I present a semi-separating equilibrium in which $L_{\underline{A}}$ chooses $\lambda = 0$ and $L_{\bar{A}}$ randomizes between $\lambda = 0$ and $\lambda = 1$.

I propose a separating equilibrium when $\underline{c}_A \leq d < \bar{c}_A$. Strategies and beliefs are as follows.

- $L_{\underline{A}}$: choose $\lambda = 1$ and back down iff $x \geq p - \underline{c}_A$. $L_{\bar{A}}$: choose $\lambda = 0$ and back down iff $x \geq p - d + r$.
- F : believe $\phi_0 = P(\underline{c}_A | \lambda = 0) = 0$ if $\lambda = 0$ and propose $x^* = p - d + r$; believe $\phi_1 = P(\underline{c}_A | \lambda = 1) = 1$ if $\lambda = 1$ and propose $x^* = p - \underline{c}_A$.
- \bar{A} : accept iff $x \geq p - \bar{c}_A$. \underline{A} : accept iff $x \geq p - \underline{c}_A$.

With $\lambda = 0$, $L_{\bar{A}}$ is offered $p - d + r$. After the audience cost is deducted from $p - d + r$, its actual payoff remains $p - d$. With $\lambda = 1$, it would be mistaken for $L_{\underline{A}}$ and gains $p - \underline{c}_A$. For $L_{\bar{A}}$ to choose $\lambda = 0$, it requires

$$p - d > p - \underline{c}_A \Leftrightarrow d < \underline{c}_A, \quad (\text{B.1})$$

which is contradictory with $\underline{c}_A \leq d$. Therefore, the separating equilibrium does not exist.¹

Next, I prove that $L_{\bar{A}}$ has no incentive to mimic $L_{\underline{A}}$ by choosing $\lambda = 1$. If the pooling equilibrium existed, F would make an offer that only $L_{\bar{A}}$ accepts, but which offer $L_{\bar{A}}$ accepts would depend on the relative values between $p - 2d$ and $p - \bar{c}_A$. Strategies and beliefs are as follows.

- $L_{\underline{A}}$: choose $\lambda = 1$ and back down iff $x \geq p - \underline{c}_A$. $L_{\bar{A}}$: choose $\lambda = 1$ and back down iff $x \geq p - 2d$ under ① or ③, and back down iff $x \geq p - \bar{c}_A$ under ② or ④.
- F : believe $\phi_1 = \phi$ if $\lambda = 1$ and propose $x^* = p - 2d$ under ① or ③, and propose $x^* = p - \bar{c}_A$ under ② or ④; believe (out of equilibrium) $\phi_0 = 0$ if $\lambda = 0$ and propose $x^* = p - d + r$.
- \bar{A} : accept iff $x \geq p - \bar{c}_A$. \underline{A} : accept iff $x \geq p - \underline{c}_A$.

It is easy to show that $L_{\bar{A}}$ has an incentive to deviate from $\lambda = 1$. Under ① or ③,

$$p - 2d \geq p - d \Leftrightarrow d \geq 2d \tag{B.2}$$

¹Another separating equilibrium in which $L_{\underline{A}}$ chooses $\lambda = 0$ and $L_{\bar{A}}$ chooses $\lambda = 1$ exists only when $d \geq \bar{c}_A$, which I will show later.

is false; under ② or ④,

$$p - \bar{c}_A \geq p - d \Leftrightarrow d \geq \bar{c}_A \quad (\text{B.3})$$

does not hold, either. The pooling equilibrium in which L always chooses $\lambda = 1$ does not exist.

I prove that the semi-separating equilibrium exists. Since F 's offers are different given the reservation values of L and A , I present the proof under the condition ① or ③ first, then under the condition ② or ④. Strategies and beliefs are as follows.

- $L_{\underline{A}}$: choose $\lambda = 1$ and back down iff $x \geq p - \underline{c}_A$. $L_{\bar{A}}$: with probability μ , choose $\lambda = 0$ and back down iff $x \geq p - d + r$; with probability $1 - \mu$, choose $\lambda = 1$ and back down iff $x \geq p - 2d$ under ① or ③, and back down iff $x \geq p - \bar{c}_A$ under ② or ④.
- F : believe $\phi_0 = 0$ if $\lambda = 0$ and propose $x^* = p - d + r$; believe $\phi_1 = \phi(\mu)$ if $\lambda = 1$, then with probability β , propose $x^* = p - 2d$ under ① or ③ and $x^* = p - \bar{c}_A$ under ② or ④, and with probability $1 - \beta$, propose $x^* = p - \underline{c}_A$.
- \bar{A} : accept iff $x \geq p - \bar{c}_A$. \underline{A} : accept iff $x \geq p - \underline{c}_A$.

Under the condition ① or ③:

First, I verify that $L_{\underline{A}}$ chooses $\lambda = 1$. If $L_{\underline{A}}$ chooses $\lambda = 1$, with probability $1 - \beta$, it may gain the offer $p - \underline{c}_A$ consistent with its people's type; with probability β , $L_{\underline{A}}$ faces the risks of war, which it would fight with public support. If $L_{\underline{A}}$ chooses $\lambda = 0$, it would be mistaken for $L_{\overline{A}}$ and gain $p - d$. Given ①, $L_{\underline{A}}$ prefers $\lambda = 1$ to $\lambda = 0$ when

$$\beta(p - d(1 - p)) + (1 - \beta)(p - \underline{c}_A) \geq p - d \Leftrightarrow d \geq \underline{c}_A, \quad (\text{B.4})$$

which exists. Since \underline{A} accepts an offer $x \geq p - \underline{c}_A$, it would reject the offer $p - d + r$ given $\underline{c}_A \leq d - r$, that is, part of the condition ③. Therefore, war would still occur if $L_{\underline{A}}$ chooses $\lambda = 0$ and F offers $p - d + r$. $L_{\underline{A}}$ would stand firm with the resolute A 's support and receive $p - d(1 - p)$. Under the condition ③, $L_{\underline{A}}$ prefers $\lambda = 1$ to $\lambda = 0$ when

$$\beta(p - d(1 - p)) + (1 - \beta)(p - \underline{c}_A) \geq p - d(1 - p) \Leftrightarrow d(1 - p) \geq \underline{c}_A. \quad (\text{B.5})$$

Second, I verify that $L_{\overline{A}}$ randomizes between $\lambda = 0$ and $\lambda = 1$. To render $L_{\overline{A}}$ indifferent between $\lambda = 0$ and $\lambda = 1$, it requires

$$p - d = \beta(p - 2d) + (1 - \beta)(p - \underline{c}_A) \Leftrightarrow \beta = \frac{d - \underline{c}_A}{2d - \underline{c}_A}, \quad (\text{B.6})$$

which holds as long as $d > \underline{c}_A$.

Third, to see F randomizes between $p - 2d$ and $p - \underline{c}_A$ upon observing $\lambda = 1$, we have

$$1 - p + \underline{c}_A = \phi_1(1 - p - c_F) + (1 - \phi_1)(1 - p + 2d) \Leftrightarrow \phi_1 = \frac{2d - \underline{c}_A}{2d + c_F} \quad (\text{B.7})$$

$$\phi_1 = \frac{\phi}{(1 - \phi)(1 - \mu) + \phi} = \frac{2d - \underline{c}_A}{2d + c_F} \Leftrightarrow \mu = 1 - \frac{\phi(c_F + \underline{c}_A)}{(1 - \phi)(2d - \underline{c}_A)}. \quad (\text{B.8})$$

For $0 < \mu < 1$ to exist, it requires $\phi < \frac{2d - \underline{c}_A}{c_F + 2d}$.

Under the condition ② or ④:

First, the steps to prove that $L_{\underline{A}}$ prefers $\lambda = 1$ are the same as in (4) and (5). Second, to render $L_{\bar{A}}$ indifferent between $\lambda = 0$ and $\lambda = 1$, F chooses β such that

$$p - d = \beta(p - \bar{c}_A) + (1 - \beta)(p - \underline{c}_A) \Leftrightarrow \beta = \frac{d - \underline{c}_A}{\bar{c}_A - \underline{c}_A}. \quad (\text{B.9})$$

Since $0 < \mu < 1$, $L_{\bar{A}}$ is rendered indifferent given $\underline{c}_A < d < \bar{c}_A$.

Finally, to see F randomizes between $p - \bar{c}_A$ and $p - \underline{c}_A$ upon observing $\lambda = 1$, we have

$$1 - p + \underline{c}_A = \phi_1(1 - p - c_F) + (1 - \phi_1)(1 - p + \bar{c}_A) \Leftrightarrow \phi_1 = \frac{\bar{c}_A - \underline{c}_A}{\bar{c}_A + c_F} \quad (\text{B.10})$$

$$\phi_1 = \frac{\phi}{(1 - \phi)(1 - \mu) + \phi} = \frac{\bar{c}_A - \underline{c}_A}{\bar{c}_A + c_F} \Leftrightarrow \mu = 1 - \frac{\phi(c_F + \underline{c}_A)}{(1 - \phi)(\bar{c}_A - \underline{c}_A)}, \quad (\text{B.11})$$

which further requires $\phi < \frac{\bar{c}_A - \underline{c}_A}{c_F + \bar{c}_A}$.

Proof of Proposition 4

When $\underline{c}_A < \bar{c}_A \leq d$, the differences between L 's and A 's reservation values have three possibilities: ① $d - r < \underline{c}_A < \bar{c}_A < 2d$, ② $\underline{c}_A < d - r < \bar{c}_A < 2d$, and ③ $\underline{c}_A < \bar{c}_A < d - r < 2d$.

First, I prove the existence of a separating equilibrium when $\underline{c}_A \leq d(1 - p) - r$ and $r \geq \bar{c}_A - \underline{c}_A$. Strategies and beliefs are as follows.

- $L_{\underline{A}}$: choose $\lambda = 0$ and back down iff $x \geq \underline{c}_A$. $L_{\overline{A}}$: choose $\lambda = 1$ and back down iff $x \geq \bar{c}_A$.
- F : believe $\phi_0 = P(\underline{c}_A|\lambda = 0) = 1$ and propose $x^* = p - \underline{c}_A$; believe $\phi_1 = P(\underline{c}_A|\lambda = 1) = 0$ if $\lambda = 1$ and propose $x^* = p - \bar{c}_A$.
- \overline{A} : accept iff $x \geq p - \bar{c}_A$. \underline{A} : accept iff $x \geq p - \underline{c}_A$.

Observing $\lambda = 1$, F would propose $p - \bar{c}_A$, which makes \underline{A} support war; with \underline{A} 's support, $L_{\underline{A}}$ would stand firm and gain $p - d(1 - p)$, greater than the payoff $(p - d)$ with $\lambda = 0$.

In ② and ③, below I show the conditions under which $L_{\underline{A}}$ prefers $\lambda = 0$ to $\lambda = 1$ whereas $L_{\overline{A}}$ prefers $\lambda = 1$

$$p - \underline{c}_A - r \geq p - d(1 - p) \Leftrightarrow \underline{c}_A \leq d(1 - p) - r \quad (\text{B.12})$$

$$p - \bar{c}_A \geq p - \underline{c}_A - r \Leftrightarrow r \geq \bar{c}_A - \underline{c}_A \quad (\text{B.13})$$

Given $\underline{c}_A \leq d(1 - p) - r$, the separating equilibrium may exist under ② and ③ only. It is easy to show that the separating equilibrium does not exist under ① as $L_{\underline{A}}$ would deviate to $\lambda = 1$.²

²Under the condition ①, $L_{\underline{A}}$ is offered $p - d + r$ after choosing $\lambda = 0$ and its actual payoff is $p - d$. However, $L_{\underline{A}}$ would be better off if it deviates to $\lambda = 1$. With $\lambda = 1$, as \underline{A} rejects F 's offer $p - \bar{c}_A$, $L_{\underline{A}}$ would gain $p - d(1 - p)$ by standing firm with public support.

Second, a semi-separating equilibrium exists when $\underline{c}_A \leq d(1-p) - r$ and $r < \bar{c}_A - \underline{c}_A$. Strategies and beliefs are as follows.

- $L_{\underline{A}}$: choose $\lambda = 0$ and back down iff $x \geq p - \underline{c}_A$. $L_{\bar{A}}$: with probability μ , choose $\lambda = 0$ and back down iff $x \geq p - d + r$ under ② and $x \geq p - \bar{c}_A$ under ③; with probability $1 - \mu$, choose $\lambda = 1$ and back down iff $x \geq p - 2d$.
- F : believe $\phi_1 = 0$ if $\lambda = 1$ and propose $x^* = p - \bar{c}_A$; believe $\phi_0 = \phi(\mu)$ if $\lambda = 0$, then with probability β , propose $x^* = p - d + r$ under ② and $x^* = p - \bar{c}_A$ under ③, and with probability $1 - \beta$, propose $x^* = p - \underline{c}_A$.
- \bar{A} : accept iff $x \geq p - \bar{c}_A$. \underline{A} : accept iff $x \geq p - \underline{c}_A$.

I verify that $L_{\underline{A}}$ chooses $\lambda = 0$. If $L_{\underline{A}}$ chooses $\lambda = 0$, with probability $1 - \beta$, $L_{\underline{A}}$ is offered $p - \underline{c}_A$ and its actual payoff is $p - \underline{c}_A - r$ with the audience costs deducted; with probability β , $L_{\underline{A}}$ faces the risks of war, which it would fight with public support. If $L_{\underline{A}}$ chooses $\lambda = 1$, it would be mistaken for $L_{\bar{A}}$ and offered $p - \bar{c}_A$. Since this offer is rejected by \underline{A} for sure, $L_{\underline{A}}$ stands firm with public support and gains $p - d(1-p)$. Therefore, $L_{\underline{A}}$ prefers $\lambda = 0$ to $\lambda = 1$ when

$$\beta(p - d(1-p)) + (1 - \beta)(p - \underline{c}_A - r) \geq p - d(1-p) \Leftrightarrow \underline{c}_A \leq d(1-p) - r \quad (\text{B.14})$$

Next, I verify that $L_{\bar{A}}$ randomizes between $\lambda = 0$ and $\lambda = 1$. If $L_{\bar{A}}$ chooses $\lambda = 1$, it reveals its people's type and gains $p - \bar{c}_A$. If $L_{\bar{A}}$ chooses

$\lambda = 0$, with probability $1 - \beta$, it is mistaken for L_A and receives $p - \underline{c}_A$ minus the audience costs r ; with probability β , it gains $p - d$ under the condition ② or $p - \bar{c}_A - r$ under ③. To render $L_{\bar{A}}$ indifferent under ②, β is calculated as follows:

$$\beta(p - d) + (1 - \beta)(p - \underline{c}_A - r) = p - \bar{c}_A \Leftrightarrow \beta = \frac{\bar{c}_A - \underline{c}_A - r}{d - r - \underline{c}_A}; \quad (\text{B.15})$$

Similarly, under the condition ③,

$$\beta(p - \bar{c}_A - r) + (1 - \beta)(p - \underline{c}_A - r) = p - \bar{c}_A \Leftrightarrow \beta = \frac{\bar{c}_A - \underline{c}_A - r}{\bar{c}_A - \underline{c}_A}. \quad (\text{B.16})$$

For both (15) and (16) to hold, it further requires $r < \bar{c}_A - \underline{c}_A$.

To check whether F randomizes between a smaller yet risky offer that only $L_{\bar{A}}$ accepts and a greater offer that guarantees peace, given ②, we have

$$1 - p + \underline{c}_A = \phi_0(1 - p - c_F) + (1 - \phi_0)(1 - p + d - r) \Leftrightarrow \phi_0 = \frac{d - r - \underline{c}_A}{c_F + d - r} \quad (\text{B.17})$$

$$\phi_0 = \frac{\phi}{\mu(1 - \phi) + \phi} = \frac{d - r - \underline{c}_A}{c_F + d - r} \Leftrightarrow \mu = \frac{\phi(c_F + \underline{c}_A)}{(1 - \phi)(d - r - \underline{c}_A)}. \quad (\text{B.18})$$

For μ to exist, it requires $\phi < \frac{d - r - \underline{c}_A}{c_F + d - r}$.

Note that F replaces $p - d + r$ with $p - \bar{c}_A$ when $\underline{c}_A < \bar{c}_A < d - r$. Thus, under the condition ③, Equation (17) is replaced with the following:

$$1 - p + \underline{c}_A = \phi_0(1 - p - c_F) + (1 - \phi_0)(1 - p + \bar{c}_A) \Leftrightarrow \phi_0 = \frac{\bar{c}_A - \underline{c}_A}{c_F + \bar{c}_A} \quad (\text{B.19})$$

$$\phi_0 = \frac{\phi}{\mu(1 - \phi) + \phi} = \frac{\bar{c}_A - \underline{c}_A}{c_F + \bar{c}_A} \Leftrightarrow \mu = \frac{\phi(c_F + \underline{c}_A)}{(1 - \phi)(\bar{c}_A - \underline{c}_A)}, \quad (\text{B.20})$$

which further requires $\phi < \frac{\bar{c}_A - \underline{c}_A}{c_F + \bar{c}_A}$.

Again, there is no such semi-separating equilibrium under ①. Suppose that F observes $\lambda = 0$, its uncertainty about A 's type does not matter as it always proposes $p - d + r$ to avoid war given $d - r < \underline{c}_A < \bar{c}_A$. On the other hand, F proposes $p - \bar{c}_A$ when observing $\lambda = 1$. Since $\underline{c}_A < \bar{c}_A \leq d$, $L_{\underline{A}}$ is better off by switching to $\lambda = 1$.

Third, when $\underline{c}_A > d(1 - p) - r$, I show that $L_{\underline{A}}$ mimics $L_{\bar{A}}$'s strategy by choosing $\lambda = 1$. Strategies and beliefs are as follows.

- $L_{\underline{A}}$: choose $\lambda = 1$ and back down iff $x \geq p - \underline{c}_A$. $L_{\bar{A}}$: choose $\lambda = 1$ and back down iff $x \geq p - \bar{c}_A$.
- F : believe $\phi_1 = \phi$ if $\lambda = 1$ and propose $x^* = p - \bar{c}_A$; believe (out of equilibrium) $\phi_0 = 0$ if $\lambda = 0$ and propose $x^* = p - d + r$ under ① or ②, and $x^* = p - \bar{c}_A$ under ③.
- \bar{A} : accept iff $x \geq p - \bar{c}_A$. \underline{A} : accept iff $x \geq p - \underline{c}_A$.

In the event of $\lambda = 0$, I assign F 's out-of-equilibrium $\phi_0 = 0$, i.e. that F believes that only $L_{\bar{A}}$ would deviate to $\lambda = 0$. The assigned belief satisfies the Intuitive Criterion as $L_{\bar{A}}$ would have an incentive to bluff. When $\underline{c}_A > d(1 - p) - r$, on one hand, $L_{\underline{A}}$ would gain the best offer, $p - d(1 - p)$, by choosing $\lambda = 1$, meaning that it never does better by switching to $\lambda = 0$ regardless of F 's strategy. On the other hand, $L_{\bar{A}}$ could do better than its

equilibrium payoff if F proposes $p - \underline{c}_A$ and $r < \bar{c}_A - \underline{c}_A$, even though this offer is not played off the equilibrium path. In sum, if F ever observes $\lambda = 0$, it believes that only $L_{\bar{A}}$ would have made this choice and proposes $x^* = p - d + r$ under the condition ① or ②, or $x^* = p - \bar{c}_A$ under ③.

I first prove that $L_{\underline{A}}$ has an incentive to choose $\lambda = 1$. When both L s choose $\lambda = 1$, F makes an offer that \underline{A} considers unacceptable. Since \underline{A} supports fighting, $L_{\underline{A}}$ stands firm with the payoff $p - d(1 - p)$. If $L_{\underline{A}}$ chooses $\lambda = 0$, F would believe that it faces $L_{\bar{A}}$ and offer either $p - d + r$ or $p - \underline{c}_A - r$ as specified above. Thus, under the condition ① or ②, $L_{\underline{A}}$ prefers $\lambda = 1$ to $\lambda = 0$ if

$$p - d(1 - p) > p - d \Leftrightarrow dp \geq 0, \quad (\text{B.21})$$

which is always true. Under ③, \underline{A} would reject the out-of-equilibrium offer, meaning that $L_{\underline{A}}$ would also stand firm with public support. $L_{\underline{A}}$ does no better by deviating to $\lambda = 0$. Formally,

$$p - d(1 - p) \geq p - d(1 - p). \quad (\text{B.22})$$

$L_{\bar{A}}$ has no incentive to switch to $\lambda = 0$ either. By choosing $\lambda = 1$, $L_{\bar{A}}$ backs down and gains $p - \bar{c}_A$. Similarly, I show that under the condition ① or ②, $L_{\bar{A}}$ prefers $\lambda = 1$ to $\lambda = 0$ when

$$p - \bar{c}_A > p - d \Leftrightarrow d > \bar{c}_A, \quad (\text{B.23})$$

consistent with the assumption. Given ③, F would offer $p - \bar{c}_A$ if $L_{\bar{A}}$ chooses $\lambda = 0$; however, $L_{\bar{A}}$ would be worse off as the audience costs, r , must be

deducted from the offer, that is,

$$p - \bar{c}_A > p - \underline{c}_A - r \Leftrightarrow r > \bar{c}_A - \underline{c}_A. \quad (\text{B.24})$$

Finally, I check the conditions under which F prefers an offer that only $L_{\bar{A}}$ accepts to an offer that L always accepts regardless of A 's type. Formally, we have

$$\phi(1 - p - c_F) + (1 - \phi)(1 - p + \bar{c}_A) \geq 1 - p + \underline{c}_A \Leftrightarrow c_F \leq \frac{(1 - \phi)\bar{c}_A - \underline{c}_A}{\phi} \quad (\text{B.25})$$

As $(1 - \phi)\bar{c}_A - \underline{c}_A > 0$, it requires $\phi < \frac{\bar{c}_A - \underline{c}_A}{\bar{c}_A}$.

Appendix C

Chapter 4 Appendix

Table C.1: Descriptive Statistics

Variable	Mean	SD	Min	Max	Obs
Threat failure	0.616	0.487	0	1	242
Challenger's accountability	1.4	1.001	0	3	240
Target's accountability	1.072	0.901	0	3	237
Challenger's polity score	0.7	7.399	-10	10	240
Target's polity score	-2.029	6.687	-10	10	239
Contiguity	0.636	0.482	0	1	242
Alliance similarity portfolio	0.576	0.357	-0.545	1	237
Challenger's status quo	0.197	0.416	-0.363	1	242
Target's status quo	0.055	0.303	-0.342	1	237
Territorial issue	0.591	0.493	0	1	242
Leadership issue	0.116	0.321	0	1	242
Policy issue	0.455	0.499	0	1	242
Other issues	0.211	0.409	0	1	242
World war dummy	0.165	0.372	0	1	242

Table C.2: Estimating Compellent Threat Failure (1918–2001)

<i>Threat failure</i>	Model 1 ($p < 0.67$)	Model 2 ($p \geq 0.67$)	Model 3 ($p < 0.78$)	Model 4 ($p \geq 0.78$)
Challenger's accountability	3.563*** (0.864)	0.301 (0.594)	2.867*** (0.746)	0.449 (0.741)
Challenger's accountability ²	-1.183*** (0.254)	-0.087 (0.175)	-0.846*** (0.216)	-0.229 (0.219)
Target's accountability	1.069*** (0.319)	-0.199 (0.194)	0.804** (0.262)	-0.311 (0.212)
Challenger's polity score	-0.043 (0.048)	0.001 (0.033)	-0.067 [†] (0.039)	0.045 (0.045)
Target's polity score	-0.179*** (0.046)	0.015 (0.028)	-0.150*** (0.039)	0.020 (0.031)
Contiguity	-0.722 (0.446)	-0.510 (0.346)	-0.887* (0.391)	-0.419 (0.408)
Alliance portfolio similarity	1.649* (0.668)	0.378 (0.505)	1.651** (0.571)	-0.046 (0.608)
Challenger's status quo	1.396* (0.713)	0.078 (0.389)	0.843 (0.567)	0.115 (0.446)
Target's status quo	-3.976*** (0.792)	0.541 (0.546)	-2.843*** (0.578)	0.748 (0.737)
Territory	-0.343 (0.475)	0.210 (0.344)	0.108 (0.393)	0.501 (0.353)
Government	-2.124* (0.832)	-1.289* (0.494)	-2.070** (0.728)	-0.772 (0.533)
Policy	-1.195* (0.569)	0.374 (0.330)	-1.258** (0.402)	0.725* (0.336)
Other	-0.712 (0.473)	0.196 (0.342)	-0.211 (0.411)	0.157 (0.403)
Constant	-2.465* (1.092)	0.470 (0.805)	-2.020* (0.900)	0.143 (0.825)
Observations	95	130	114	111

Standard errors in parentheses are clustered by the individual case. Dummy variables for world war years are included but not reported here.

[†] $p < 0.1$, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table C.3: Estimating Compellent Threat Failure (1918–2001)

<i>Threat failure</i>	Model 1 (all)	Model 2 ($p < 0.67$)	Model 3 ($p \geq 0.67$)	Model 4 ($p < 0.78$)	Model 5 ($p \geq 0.78$)
Challenger's accountability	1.363** (0.511)	2.119** (0.766)	0.451 (0.705)	2.297** (0.741)	0.457 (0.833)
Challenger's accountability ²	-0.429** (0.160)	-0.965** (0.305)	-0.193 (0.208)	-0.842** (0.221)	-0.212 (0.229)
Target's accountability	0.253 (0.195)	1.626** (0.486)	0.395 (0.313)	0.764* (0.355)	0.124 (0.339)
Challenger's polity score	-0.037 (0.027)	0.028 (0.053)	-0.006 (0.048)	-0.042 (0.058)	0.011 (0.064)
Target's polity score	-0.038 (0.025)	-0.211** (0.059)	-0.007 (0.033)	-0.151** (0.045)	0.011 (0.036)
Major–Major dyad	0.024 (0.526)	2.100 (1.391)	-3.017† (1.545)	0.875 (0.953)	-2.406 (1.553)
Major–Minor dyad	0.699* (0.311)	1.814* (0.747)	0.052 (0.471)	2.072** (0.566)	-0.231 (0.600)
Minor–Major dyad	-1.383* (0.548)	-1.679* (0.753)	<i>(omitted)</i>	-0.339 (0.726)	<i>(omitted)</i>
Initiator's relative power	-0.262 (0.433)	1.141 (1.542)	-0.993 (1.760)	0.885 (1.046)	-1.592 (2.572)
Contiguity	-0.366 (0.266)	-0.709 (0.566)	-0.692† (0.403)	-0.612 (0.485)	-0.691 (0.440)
Alliance portfolio similarity	0.258 (0.436)	2.197† (1.213)	-0.162 (0.652)	1.219 (0.874)	-0.249 (0.726)
Initiator's status quo	-0.009 (0.345)	1.337† (0.783)	-0.048 (0.483)	0.678 (0.612)	0.049 (0.524)
Target's status quo	-0.756† (0.409)	-4.418** (1.164)	1.104 (0.735)	-2.302** (0.724)	1.125 (0.907)
Territory	0.002 (0.245)	-0.010 (0.476)	0.144 (0.343)	-0.166 (0.428)	0.533 (0.373)
Government	-1.141** (0.381)	-1.811* (0.881)	-1.233* (0.493)	-1.903** (0.721)	-0.756 (0.576)
Policy	-0.075 (0.228)	-1.600* (0.626)	0.560 (0.343)	-1.336** (0.428)	0.872* (0.351)
Other	-0.022 (0.267)	-1.103† (0.587)	0.123 (0.377)	-0.453 (0.474)	0.150 (0.400)
Constant	0.116 (0.591)	-1.984 (1.605)	1.611 (1.721)	-1.094 (1.099)	2.201 (2.541)
Observations	229	95	129	114	111
Model χ^2	57.785	56.591	27.391	70.612	26.938

Standard errors in parentheses. Dummy variables for world war years are included but not reported here.

† $p < 0.1$, * $p < 0.05$, ** $p < 0.01$

Appendix D

Chapter 5 Appendix

Proof of Proposition 5

I use backward induction to solve the model under complete information. The solution concept is subgame perfect equilibrium (SPE). To simplify the calculations without losing generality, I assume for now that L makes a dichotomous choice, either $\lambda = 1$ or $\lambda = 0$. Also, since the resolute A 's payoffs are the same as the irresolute A 's payoffs except the value of c_A , I simplify the solution by considering only one type of A with the costs of war c_A , rather than differentiating between \underline{c}_A and \bar{c}_A .

I start with L 's final move and F 's proposed offer. Suppose that A supports fighting: if L chooses $\lambda = 1$, L will back down when $x \geq p + d + dp + 2r$ and F will offer $x = p + d + dp + 2r$ when $c_F \geq d + dp + 2r$; if L chooses $\lambda = 0$, L will back down when $x \geq p + dp + r$ and F will offer $x = p + dp + r$ when $c_F \geq dp + r$. Suppose that A does not support fighting: if L chooses $\lambda = 1$, L will back down when $x \geq p - 2(d - r)$ and F will offer $x = p - 2(d - r)$ when $c_F \geq 2(r - d)$; if L chooses $\lambda = 0$, L will back down when $x \geq p - d + r$ and F will offer $x = p - d + r$ when $c_F \geq r - d$.

Then I analyze L 's choice of λ given different values of c_F and A 's initial preferences. Below I demonstrate that whether $d > r$ or $d \leq r$ does not affect L 's choice.

Assume $d > r$. Given $c_F > 0$, we have $c_F > r - d > 2(r - d)$. Therefore, if A does not support fighting, F will always propose an offer that L prefers to standing firm. In this case, L chooses $\lambda = 0$ since

$$U_L(\lambda = 0 | \sim \text{fight}) > U_L(\lambda = 1 | \sim \text{fight}) \Leftrightarrow p - d > p - 2d. \quad (\text{D.1})$$

If A supports fighting, the choice of λ depends on the value of c_F in one of the three ranges as follows. Suppose $c_F \geq d + dp + 2r$: L is indifferent between $\lambda = 1$ and $\lambda = 0$; no war occurs because F always proposes an offer that compensates L 's war payoff. Suppose $r + dp \leq c_F < d + dp + 2r$: L is also indifferent between $\lambda = 1$ and $\lambda = 0$; but since war occurs only when $\lambda = 1$, I assume that L chooses $\lambda = 0$ instead. Suppose $c_F < r + dp$: L is again indifferent about the choice of λ , however, because F offers $x = 0$ and war occurs in either way.

Assume $d \leq r$. The three ranges above still exist, meaning that L 's choice remains the same if A supports fighting. We need to check whether L still chooses $\lambda = 0$ if A does not support fighting. The value of c_F may fall into one of these ranges: $c_F \geq 2(r - d)$, $r - d \leq c_F < 2(r - d)$, and $c_F < r - d$.¹ Regardless of L 's choice, F either chooses war with a zero offer or makes L indifferent between standing firm and backing down. Either way, L will still

¹To make sure $2(r - d) < r + dp$, I further assume $r < d(p + 2)$.

receive its war payoff. Therefore, regardless of the range of c_F , L chooses $\lambda = 0$ because (D.1) still holds.

Finally, I analyze the conditions under which A supports fighting or not. I first assume $d > r$. Suppose $c_F \geq r + dp$: obviously, A supports fighting even if A 's support leads to $\lambda = 0$; formally, we have

$$U_A(\text{fight}) > U_A(\sim \text{fight}) \Leftrightarrow p + dp + r > p - d + r. \quad (\text{D.2})$$

Suppose $c_F < r + dp$: A supports fighting when $c_A < d - r$, that is,

$$U_A(\text{fight}) > U_A(\sim \text{fight}) \Leftrightarrow p - c_A > p - d + r. \quad (\text{D.3})$$

Assume $d \leq r$. Given that $c_A > d - r$ always exists, we know from (D.3) that A never supports fighting.

Given $d > r$, below is a summary of equilibrium with two types of A .

- A : if $c_F \geq r + dp$, A always supports fighting; if $c_F < r + dp$,
 - both \underline{A} and \bar{A} supports fighting when $\underline{c}_A < \bar{c}_A < d - r$;
 - only \underline{A} supports fighting when $\underline{c}_A < d - r \leq \bar{c}_A$;
 - neither \underline{A} nor \bar{A} supports fighting when $d - r \leq \underline{c}_A < \bar{c}_A$.²
- L : if A supports fighting, L is indifferent about the choice of λ and stands firm as long as F offers $x < p + d + dp + 2r$ given $\lambda = 1$ or $x < p + dp + r$

²If $d \leq r$, neither types of A supports fighting.

given $\lambda = 0$;³ if A does not support fighting, L chooses $\lambda = 0$ and backs down as long as F offers $x \geq p - d + r$.

- F : if A does not support fighting, F offers $x = p - 2(d - r)$ given $\lambda = 1$ and $x = p - d + r$ given $\lambda = 0$; if A supports fighting,
 - F offers $x = p + d + dp + 2r$ given $\lambda = 1$, and $x = p + dp + r$ given $\lambda = 0$ when $c_F \geq d + dp + 2r$;
 - F offers $x = 0$ given $\lambda = 1$, and $x = p + dp + r$ given $\lambda = 0$ when $r + dp \leq c_F < d + dp + 2r$;
 - F offers $x = 0$ regardless of λ when $c_F < r + dp$.

There exist multiple equilibria under incomplete information. Since I focus on the risks of war and A 's different strategies given its type, to solve the model under incomplete information, I further assume that $c_F < r + dp$ and $\underline{c}_A < d - r \leq \bar{c}_A$. I prove the existence of three equilibria: a separating equilibrium in which only \underline{A} supports fighting; a pooling equilibrium in which both types of A support fighting; another pooling equilibrium in which neither types of A supports fighting.

Proof of Proposition 6

³I assume earlier that L chooses $\lambda = 0$ if $r + dp \leq c_F < d + dp + 2r$ to avoid war.

Strategies and beliefs of the separating equilibrium are as follows.

- A : \underline{A} supports fighting; \overline{A} does not support fighting.
- Updated beliefs: $\Pr(\underline{c}_A|\text{fight}) = 1$; $\Pr(\underline{c}_A|\sim \text{fight}) = 0$.
- L : if A supports fighting, L chooses $\lambda = 1$ and stands firm as long as F offers $x < p + d + dp + 2r$; if A does not support fighting, L chooses $\lambda = 0$ and backs down as long as F offers $x \geq p - d + r$.
- F : if A supports fighting, F offers $x = 0$; if A does not support fighting, F offers $x = p - (1 + \lambda)(d - r)$.

First, I verify L 's choices of λ given its updated beliefs. Given A 's support for fighting, L learns that A is the resolute type and thus would choose $\lambda = 1$. F also updates its belief about A 's type. If the resolute A 's support for war, we know from Proposition 5 that F would propose nothing and fight given $c_F < r + dp$, regardless of the value of λ . Therefore, L is indifferent between $\lambda = 1$ and $\lambda = 0$, both of which brings L its war payoffs with A 's support. Formally, we have

$$U_L(\lambda = 1|\text{fight}) \geq U_L(\lambda = 0|\text{fight}) \Leftrightarrow p - d(1 - p) \geq p - d(1 - p). \quad (\text{D.4})$$

If L and F observe that A does not support fighting, both believe that they face the irresolute A . As F offers $x = p - d + r$, we can verify that L prefers $\lambda = 0$ to $\lambda = 1$ based on the following inequality:

$$U_L(\lambda = 0|\sim \text{fight}) > U_L(\lambda = 1|\sim \text{fight}) \Leftrightarrow p - d > p - d - r. \quad (\text{D.5})$$

Then, I verify F 's strategies given its updated beliefs. A 's support for fighting reveals its resolute type to L and F . Given $\lambda = 1$ and $c_F < r + dp$, there is no bargaining range between L and F , meaning that either $x = 0$ or $x = p - d + r$ would not be sufficient to make L back down. Therefore, F would rather offer nothing. Formally, we have

$$U_F(0|\text{fight}) \geq U_F(p - d + r|\text{fight}) \Leftrightarrow 1 - p - c_F \geq 1 - p - c_F. \quad (\text{D.6})$$

If A does not support fighting, given $\lambda = 0$, L will back down as long as F offers $x \geq p - d + r$. Therefore, F is able to make a peaceful agreement by offering $x = p - d + r$. The remaining portion of the goods is greater than F 's war payoff when

$$U_F(p - d + r|\sim \text{fight}) \geq U_F(0|\sim \text{fight}) \Leftrightarrow 1 - p + d - r \geq 1 - p - c_F. \quad (\text{D.7})$$

This is true given $d > r$.

Finally, I verify A 's strategy consistent with its type. We have already verified that if A chooses to fight initially, the other actors' strategies will lead to war, but no war will occur if A chooses not to fight in the first place. Given the resolute type, \underline{A} chooses to fight when

$$U_{\underline{A}}(\text{fight}) > U_{\underline{A}}(\sim \text{fight}) \Leftrightarrow p - \underline{c}_A > p - d + r. \quad (\text{D.8})$$

This is consistent with one of the initial assumptions, $\underline{c}_A < d - r$. Given the irresolute type, \bar{A} chooses not to fight when

$$U_{\bar{A}}(\sim \text{fight}) > U_{\bar{A}}(\text{fight}) \Leftrightarrow p - d + r > p - \bar{c}_A, \quad (\text{D.9})$$

which also holds given $\bar{c}_A \geq d - r$.

Proof of Proposition 7

Strategies and beliefs of the first pooling equilibrium are as follows.

- A : both \underline{A} and \bar{A} support fighting.
- Updated beliefs: $\Pr(\underline{c}_A|\text{fight}) = \phi$; $\Pr(\underline{c}_A|\sim \text{fight}) = \epsilon$.
- L : if A supports fighting, L chooses $\lambda = 1$ and stands firm as long as F offers $x < p + d + dp + 2r$; if A does not support fighting, L chooses $\lambda \in (0, 1)$ and backs down as long as F offers $x \geq p - (1 + \lambda)(d - r)$.
- F : if A supports fighting, F offers $x = 0$; if A does not support fighting, F offers $x = p - (1 + \lambda)(d - r)$.

Uninformed actors are unable to update beliefs in a pooling equilibrium. First, I verify L 's choices of λ given its beliefs. Similar to (D.4), even if both types of A supports fighting, L makes the ultimate move and has no preferences about the value of λ . Formally, we have

$$U_L(\lambda = 1|\text{fight}) \geq U_L(0 < \lambda < 1|\text{fight}) \Leftrightarrow p - d(1 - p) \geq p - d(1 - p). \quad (\text{D.10})$$

Suppose that A chooses not to fight, I specify that $\lambda \in (0, 1)$ is L 's off-the-equilibrium strategy. As a result, L would back down as long as F offers

$x \geq p - (1 + \lambda)(d - r)$. F 's off-the-equilibrium strategy is $x = p - (1 + \lambda)(d - r)$, which I will verify soon. Below I show that L would not deviate from its off-the-equilibrium strategy, that is,

$$U_L(0 < \lambda < 1 | \sim \text{fight}) > U_L(\lambda = 1 | \sim \text{fight}) \Leftrightarrow p - (1 + \lambda)d > p - 2d. \quad (\text{D.11})$$

This is true as long as $0 < \lambda < 1$.

Next, I verify F 's strategies give its beliefs. Similar to (D.6), if A supports fighting, no bargaining range exists between F and L . Unless F offers $x = p + d + dp + 2r$, which is impossible under this circumstance, any other value of x does not affect the outcome. F would rather make a zero offer and fight. Formally, we have

$$U_F(0 | \text{fight}) \geq U_F(p - (1 + \lambda)(d - r) | \text{fight}) \Leftrightarrow 1 - p - c_F \geq 1 - p - c_F. \quad (\text{D.12})$$

Suppose that A chooses not to fight, I specify that F 's off-the-equilibrium strategy is $x = p - (1 + \lambda)(d - r)$. F would not deviate to $x = 0$ as long as this off-the-equilibrium strategy brings F more than its war payoffs, that is,

$$U_F(p - (1 + \lambda)(d - r) | \sim \text{fight}) \geq U_F(0 | \sim \text{fight}) \Leftrightarrow (1 + \lambda)(d - r) \geq -c_F. \quad (\text{D.13})$$

This is valid given $d > r$.

Finally, I check whether A 's strategy is consistent with its type. Given the resolute type, \underline{A} supports fighting when

$$U_{\underline{A}}(\text{fight}) > U_{\underline{A}}(\sim \text{fight}) \Leftrightarrow p - \underline{c}_A > p - (1 + \lambda)(d - r) \quad (\text{D.14})$$

$$\underline{c}_A < (1 + \lambda)(d - r). \quad (\text{D.15})$$

Given the irresolute type, \bar{A} also supports fighting when

$$U_{\bar{A}}(\text{fight}) > U_{\bar{A}}(\sim \text{fight}) \Leftrightarrow p - \bar{c}_A > p - (1 + \lambda)(d - r) \quad (\text{D.16})$$

$$\bar{c}_A < (1 + \lambda)(d - r). \quad (\text{D.17})$$

Combining (D.15) and (D.17), we have the following condition under which this pooling equilibrium holds:

$$\underline{c}_A < d - r \leq \bar{c}_A < (1 + \lambda)(d - r). \quad (\text{D.18})$$

Additionally, since $0 < p - (1 + \lambda)(d - r) \leq 1$, $(1 + \lambda)(d - r) \leq p$.

Proof of Proposition 8

Strategies and beliefs of the second pooling equilibrium are as follows.

- A : neither \underline{A} nor \bar{A} supports fighting.
- Updated beliefs: $\Pr(\underline{c}_A | \sim \text{fight}) = \phi$; $\Pr(\underline{c}_A | \text{fight}) = \theta$.
- L : if A does not support fighting, L chooses $\lambda = 0$ and backs down as long as F offers $x \geq p - d + r$; if A supports fighting, L chooses $\lambda \in (0, 1)$ and stands firm as long as F offers $x < p + dp + r + \lambda(d + r)$.
- F : if A does not support fighting, F offers $x = p - \underline{c}_A$; if A supports fighting, F offers $x = 0$.

Again, L and F are unable to update their beliefs. First, I verify L 's strategy given its beliefs. If A does not support fighting, L 's payoffs do not vary with A 's type. Given $\lambda = 0$, L will back down as long as F offers $x \geq p - d + r$; therefore, L will back down if F offers $x = p - \lambda(d - r)$. Given $\lambda \in (0, 1)$, L will also back down as long as $x \geq p - (1 + \lambda)(d - r)$ and F 's proposed offer fits in this range. In other words, L will accept F 's proposed offer in both situations. Below is the condition under which L prefers $\lambda = 0$ to $\lambda \in (0, 1)$:

$$U_L(\lambda = 0 | \sim \text{fight}) > U_L(0 < \lambda < 1 | \sim \text{fight}) \Leftrightarrow x - r > x - r(1 + \lambda). \quad (\text{D.19})$$

Note that $x = p - \lambda(d - r)$. Now suppose that A supports fighting. I specify that $\lambda \in (0, 1)$ is L 's off-the-equilibrium strategy. Similar to (D.10), unless F offers $x = p + dp + r + \lambda(d + r)$, which is impossible under this circumstance, any other value of x does not affect the outcome. Formally, we have

$$U_L(0 < \lambda < 1 | \text{fight}) \geq U_L(0 | \text{fight}) \Leftrightarrow p - d(1 - p) \geq p - d(1 - p). \quad (\text{D.20})$$

Next, I verify F 's strategies given its beliefs. If A chooses not to fight, F has two options: either it offers $p - \underline{c}_A$ which guarantees peace with both types, or it offers $p - d + r$ with the risk of fighting with \underline{A} . Formally, F prefers $p - \underline{c}_A$ when

$$U_F(p - \underline{c}_A | \sim \text{fight}) \geq U_F(p - d + r | \sim \text{fight}) \Leftrightarrow \underline{c}_A \geq (1 - \phi)(d - r) - \phi c_F. \quad (\text{D.21})$$

which is possible given $d > r$. Suppose that A chooses to fight, the off-the-equilibrium strategy is $x = 0$. Similar to (D.12), F would rather fight than

make L back down in the off-the-equilibrium scenario. Thus, F would pay its war payoffs, irrelevant to the value of x . Formally, we have

$$U_F(0|\text{fight}) \geq U_F(p - \underline{c}_A|\text{fight}) \Leftrightarrow 1 - p - c_F \geq 1 - p - c_F. \quad (\text{D.22})$$

Finally, I check whether A 's strategy is consistent with its type. Given the resolute type, \underline{A} does not support fighting when

$$U_{\underline{A}}(\sim \text{fight}) \geq U_{\underline{A}}(\text{fight}) \Leftrightarrow p - \underline{c}_A \geq p - \underline{c}_A. \quad (\text{D.23})$$

$$(\text{D.24})$$

Given the irresolute type, \bar{A} does not support fighting when

$$U_{\bar{A}}(\sim \text{fight}) > U_{\bar{A}}(\text{fight}) \Leftrightarrow p - \underline{c}_A > p - \bar{c}_A. \quad (\text{D.25})$$

Combining (D.21) and (D.26), we have the following condition under which this pooling equilibrium holds:

$$(1 - \phi)(d - r) - \phi c_F \leq \underline{c}_A < d - r \leq \bar{c}_A. \quad (\text{D.26})$$

Bibliography

- Arena, Philip. 2015. "Crisis Bargaining, Domestic Opposition, and Tragic Wars." *Journal of Theoretical Politics* 27(1):108–131.
- Baum, Matthew A. 2004. "Going Private: Public Opinion, Presidential Rhetoric, and the Domestic Politics of Audience Costs in US Foreign Policy Crises." *Journal of Conflict Resolution* 48(5):603–631.
- Baum, Matthew A and Philip B.K. Potter. 2008. "The Relationships between Mass Media, Public Opinion, and Foreign Policy: Toward a Theoretical Synthesis." *Annual Review of Political Science* 11:39–65.
- Berinsky, Adam J. 2007. "Assuming the Costs of War: Events, Elites, and American Public Support for Military Conflict." *Journal of Politics* 69(4):975–997.
- Berinsky, Adam J. 2009. *In Time of War: Understanding American Public Opinion from World War II to Iraq*. University of Chicago Press.
- Bueno de Mesquita, Bruce, Alastair Smith, Randolph M. Siverson and James D. Morrow. 2003. *The Logic of Political Survival*. Cambridge: MIT Press.

- Bueno de Mesquita, Bruce, James D Morrow, Randolph M Siverson and Alastair Smith. 1999. "An Institutional Explanation of the Democratic Peace." *American Political Science Review* 93(4):791–807.
- Bullock, John G. 2011. "Elite Influence on Public Opinion in an Informed Electorate." *American Political Science Review* 105(03):496–515.
- Chaudoin, Stephen. 2014. "Promises or Policies? An Experimental Analysis of International Agreements and Audience Reactions." *International Organization* 68(01):235–256.
- Cheibub, José Antonio, Jennifer Gandhi and James Raymond Vreeland. 2010. "Democracy and Dictatorship Revisited." *Public Choice* 143(1):67–101.
- Chen, Jian. 1994. *China's Road to the Korean War: The Making of the Sino-American Confrontation*. Columbia University Press.
- Chiozza, Giacomo and H. E. Goemans. 2004. "International Conflict and the Tenure of Leaders: Is War Still Ex Post Inefficient?" *American Journal of Political Science* 48(3):604–619.
- Crall, Dennis A and Thomas M Martin. 2013. "Cool Hand Nuke: Lessons From the Quiet Diplomacy of the Cienfuegos Non-Crisis." *Foreign Policy Analysis* 9(2):189–201.
- Croco, Sarah E. 2011. "The Decider's Dilemma: Leader Culpability, Domestic Politics, and War Termination." *American Political Science Review* 105(3):457–477.

- Cukierman, Alex and Mariano Tommasi. 1998. "When does it Take a Nixon to Go to China?" *American Economic Review* pp. 180–197.
- Davies, Graeme AM and Robert Johns. 2013. "Audience Costs Among the British Public: the Impact of Escalation, Crisis Type, and Prime Ministerial Rhetoric." *International Studies Quarterly* 57(4):725–737.
- Debs, Alexandre and H.E. Goemans. 2010. "Regime Type, the Fate of Leaders, and War." *American Political Science Review* 104(3):430–445.
- Debs, Alexandre and Jessica Chen Weiss. 2014. "Circumstances, domestic audiences, and reputational incentives in international crisis bargaining." *Journal of Conflict Resolution* 60(3):403–433.
- Downes, Alexander B and Todd S Sechser. 2012. "The Illusion of Democratic Credibility." *International Organization* 66(03):457–489.
- Downs, George W and David M Rocke. 1995. *Optimal imperfection?: Domestic uncertainty and institutions in international relations*. Princeton University Press Princeton, NJ.
- Fearon, James D. 1994. "Domestic Political Audiences and the Escalation of International Disputes." *American Political Science Review* 88(3):577–92.
- Fearon, James D. 1995. "Rationalist Explanations for War." *International Organization* 49(3):379–414.

- Fearon, James D. 1997. "Signaling Foreign Policy Interests: Tying Hands versus Sinking Costs." *Journal of Conflict Resolution* 41(1):68–90.
- Filson, Darren and Suzanne Werner. 2004. "Bargaining and Fighting: The Impact of Regime Type on War Onset, Duration, and Outcomes." *American Journal of Political Science* 48(02):296–313.
- Gelpi, Christopher and Joseph M Grieco. 2015. "Competency Costs in Foreign Affairs: Presidential Performance in International Conflicts and Domestic Legislative Success, 1953–2001." *American Journal of Political Science* 59(2):440–456.
- Gibler, Douglas M and Marc L Hutchison. 2013. "Territorial Issues, Audience Costs, and the Democratic Peace: The Importance of Issue Salience." *The Journal of Politics* 75(04):879–893.
- Gleditsch, Kristian S and Michael D Ward. 2000. "War and Peace in Space and Time: The Role of Democratization." *International Studies Quarterly* 44(1):1–29.
- Goldberg, Jeffrey. 2016. "The Obama Doctrine." *The Atlantic*, April, 2016. <http://www.theatlantic.com/magazine/archive/2016/04/the-obama-doctrine/471525/>. (accessed March 25, 2016).
- Guisinger, Alexandra and Alastair Smith. 2002. "Honest Threats: The Interaction of Reputation and Political Institutions in International Crises." *Journal of Conflict Resolution* 46(2):175–200.

- Guisinger, Alexandra and Elizabeth N Saunders. forthcoming. "Mapping the Boundaries of Elite Cues: How Elites Shape Mass Opinion across International Issues." *International Studies Quarterly* .
- Hopf, Ted. 1994. *Peripheral Visions: Deterrence Theory and American Foreign Policy in the Third World, 1965-1990*. Ann Arbor: University of Michigan Press.
- Horowitz, Michael C., Allan C. Stam and Cali M. Ellis. 2015. *Why Leaders Fight*. New York: Cambridge University Press.
- Jackson, Matthew O and Massimo Morelli. 2007. "Political Bias and War." *The American Economic Review* 97(4):1353–1373.
- Jervis, Robert. 1970. *The Logic of Images in International Relations*. Princeton: Princeton University Press.
- Kertzer, Joshua D. 2016. *Resolve in International Politics*. Princeton University Press.
- Kertzer, Joshua D. and Ryan Brutger. 2016. "Decomposing Audience Costs: Bringing the Audience Back into Audience Cost Theory." *American Journal of Political Science* 60(1):234–249.
- Kessler, Glenn. 2013. "President Obama and the Red Line on Syria's Chemical Weapons." *The Washington Post*, September 6, 2013. <https://www.washingtonpost.com/news/fact-checker/wp/2013/09/06/>

[president-obama-and-the-red-line-on-syrias-chemical-weapons/](#).
(accessed March 25, 2016).

King, Gary, Michael Tomz and Jason Wittenberg. 2000. "Making the Most of Statistical Analyses: Improving Interpretation and Presentation." *American Journal of Political Science* 44(2):347–361.

Kurizaki, Shuhei. 2007. "Efficient Secrecy: Public versus Private Threats in Crisis Diplomacy." *American Political Science Review* 101(3):543–558.

Kurizaki, Shuhei and Taehee Whang. 2015. "Detecting Audience Costs in International Disputes." *International Organization* 69(4):949–980.

Lauge, David. 2008. "Taiwan Prepares for First of 2 Major 2008 Elections." *The New York Times*, January 12, 2008. <http://www.nytimes.com/2008/01/12/world/asia/12taiwan.html>. (accessed March 25, 2016).

Leeds, Brett Ashley, Michaela Mattes and Jeremy S. Vogel. 2009. "Interests, Institutions, and the Reliability of International Commitments." *American Journal of Political Science* 53(2):461–476.

Levendusky, Matthew S and Michael C Horowitz. 2012. "When Backing Down is the Right Decision: Partisanship, New Information, and Audience Costs." *The Journal of Politics* 74(02):323–338.

Levy, Jack S. 2012. "Coercive Threats, Audience Costs, and Case Studies." *Security Studies* 21(3):383–390.

- Levy, Jack S., Michael K. McKoy, Paul Poast and Geoffrey P.R. Wallace. 2015. "Backing Out or Backing In? Commitment and Consistency in Audience Costs Theory." *American Journal of Political Science* 59(4):988–1001.
- Licht, Amanda A. 2010. "Coming Into Money: The Impact of Foreign Aid on Leader Tenure." *Journal of Conflict Resolution* 54(1):58–87.
- Mansfield, Edward D and Jack Snyder. 1995. "Democratization and the Danger of War." *International Security* 20(1):5–38.
- Marshall, Monty G. and Keith Jagers. 2013. "Polity IV Project: Political Regime Characteristics and Transitions, 1800-2012.". Polity IV dataset version 2012. Accessed 5 July 2015 from URL www.systemicpeace.org/polityproject.html.
- McGillivray, Fiona and Alastair Smith. 2008. *Punishing the Prince: A Theory of Interstate Relations, Political Institutions, and Leader Change*. Princeton: Princeton University Press.
- McManus, Roseanne W. 2014. "Fighting Words: the Effectiveness of Statements of Resolve in International Conflict." *Journal of Peace Research* 51(6):726–740.
- McManus, Roseanne W. 2016. "The Impact of Context on the Ability of Leaders to Signal Resolve." *International Interactions* (forthcoming).
- Mercer, Jonathan. 1996. *Reputation and International Politics*. Ithaca, NY: Cornell University Press.

- Moon, Chungshik and Mark Souva. 2016. "Audience Costs, Information, and Credible Commitment Problems." *Journal of Conflict Resolution* 60(3):434–458.
- Potter, Philip B.K. and Matthew A. Baum. 2014. "Looking for Audience Costs in all the Wrong Places: Electoral Institutions, Media Access, and Democratic Constraint." *The Journal of Politics* 76(01):167–181.
- Powell, Robert. 1999. *In the Shadow of Power*. Princeton: Princeton University Press.
- Press, Daryl G. 2005. *Calculating Credibility: How Leaders Assess Military Threats*. Ithaca, NY: Cornell University Press.
- Putnam, Robert D. 1988. "Diplomacy and Domestic Politics: the Logic of Two-level Games." *International organization* 42(03):427–460.
- Ramsay, Kristopher W. 2004. "Politics at the Water's Edge Crisis Bargaining and Electoral Competition." *Journal of Conflict Resolution* 48(4):459–486.
- Russett, Bruce and John Oneal. 2001. *Triangulating Peace*. New York: W.W. Norton and Company.
- Sartori, Anne E. 2002. "The Might of the Pen: A Reputational Theory of Communication in International Disputes." *International Organization* 56(01):121–149.

- Saunders, Elizabeth N. 2011. *Leaders at War: How Presidents Shape Military Interventions*. Ithaca: Cornell University Press.
- Saunders, Elizabeth N. 2015. "War and the Inner Circle: Democratic Elites and the Politics of Using Force." *Security Studies* 24(3):466–501.
- Schelling, Thomas C. 1960. *The Strategy of Conflict*. Cambridge, MA: Harvard University Press.
- Schelling, Thomas C. 1966. *Arms and Influence*. New Haven: Yale University Press.
- Schultz, Kenneth A. 1998. "Domestic Opposition and Signaling in International Crises." *American Political Science Review* 92(4):829–844.
- Schultz, Kenneth A. 1999. "Do Democratic Institutions Constrain or Inform? Contrasting Two Institutional Perspectives on Democracy and War." *International Organization* 53(2):233–266.
- Schultz, Kenneth A. 2001. *Democracy and Coercive Diplomacy*. Cambridge, MA: Cambridge University Press.
- Schultz, Kenneth A. 2005. "The Politics of Risking Peace: Do Hawks or Doves Deliver the Olive Branch?" *International Organization* 59(1):1–38.
- Schultz, Kenneth A. 2012. "Why We Needed Audience Costs and What We Need Now." *Security Studies* 21(3):369–375.

- Sechser, Todd S. 2010. "Goliath's Curse: Coercive Threats and Asymmetric Power." *International Organization* 64(4):627–660.
- Sechser, Todd S. 2011. "Militarized Compellent Threats, 1918–2001." *Conflict Management and Peace Science* 28(4):377–401.
- Singer, J. David. 1961. "The Level-of-Analysis Problem in International Relations." *World Politics* 14(1):77–92.
- Singer, J. David, Stuart Bremer and John Stuckey. 1972. Capability Distribution, Uncertainty, and Major Power War, 1820-1965. In *Peace, War, and Numbers*, ed. Bruce Russett. Beverly Hills: Sage Publishing pp. 19–48. National Material Capabilities Dataset v.4.0, accessed July 20, 2010, available at <http://www.correlatesofwar.org/>.
- Siniver, Asaf. 2008. "The Nixon Administration and the Cienfuegos Crisis of 1970: Crisis-management of a Non-crisis?" *Review of International Studies* 34(01):69–88.
- Slantchev, Branislav. 2005. "Military Coercion in Interstate Crises." *American Political Science Review* 99(4):533–547.
- Slantchev, Branislav L. 2003. "The Power to Hurt: Costly Conflict with Completely Informed States." *American Political Science Review* 97(01):123–133.
- Slantchev, Branislav L. 2006. "Politicians, the Media, and Domestic Audience Costs." *International Studies Quarterly* 50(2):445–477.

- Smith, Alastair. 1998. "International Crises and Domestic Politics." *American Political Science Review* 92(3):623–38.
- Snyder, Glenn H. and Paul Diesing. 1977. *Conflict Among Nations: Bargaining, Decision Making, and System Structure in International Crises*. Princeton: Princeton University Press.
- Snyder, Jack and Erica D Borghard. 2011. "The Cost of Empty Threats: A Penny, not a Pound." *American Political Science Review* 105(03):437–456.
- Tarar, Ahmer. 2001. "International Bargaining with Two-sided Domestic Constraints." *Journal of Conflict Resolution* 45(3):320–340.
- Tarar, Ahmer and Bahar Leventoglu. 2009. "Public Commitment in Crisis Bargaining." *International Studies Quarterly* 53(3):817–839.
- Tomz, Michael. 2007. "Domestic Audience Costs in International Relations: An Experimental Approach." *International Organization* 61(04):821–840.
- Trachtenberg, Marc. 2012. "Audience Costs: An Historical Analysis." *Security Studies* 21(1):3–42.
- Trager, Robert F and Lynn Vavreck. 2011. "The Political Costs of Crisis Bargaining: Presidential Rhetoric and the Role of Party." *American Journal of Political Science* 55(3):526–545.
- Uzonyi, Gary, Mark Souva and Sona N Golder. 2012. "Domestic Institutions and Credible Signals." *International Studies Quarterly* 56(4):765–776.

- Wagner, R Harrison. 2000. "Bargaining and War." *American Journal of Political Science* pp. 469–484.
- Weeks, Jessica L. 2008. "Autocratic Audience Costs: Regime Type and Signaling Resolve." *International Organization* 62(01):35–64.
- Weeks, Jessica L. 2012. "Strongmen and Straw Men: Authoritarian Regimes and the Initiation of International Conflict." *American Political Science Review* 106(2):326–347.
- Weiss, Jessica C. 2013. "Authoritarian Signaling, Mass Audiences, and Nationalist Protest in China." *International Organization* 67(01):1–35.
- Weiss, Jessica C. 2014. *Powerful Patriots: Nationalist Protest in China's Foreign Relations*. Oxford University Press, USA.
- Wolford, Scott. 2007. "The Turnover Trap: New Leaders, Reputation, and International Conflict." *American Journal of Political Science* 51(4):772–788.
- Wolford, Scott. 2015. *The Politics of Military Coalitions*. Cambridge University Press.
- Yang, Kuisong. 1999. *Mao Zedong yu Mosike de En En Yuan Yuan*. Nanchang: Jiangxi Renmin Chubanshe. [Mao Zedong and Moscow: Amity and Enmity].
- Yarhi-Milo, Keren. 2013. "Tying Hands Behind Closed Doors: the Logic and Practice of Secret Reassurance." *Security Studies* 22(3):405–435.

Zaller, John. 1992. *The Nature and Origins of Mass Opinion*. Cambridge university press.