Regime Type, Balance of Capabilities, and Level of Escalation in International Crisis Bargaining

Xiaoyan Qiu
Jingzhou, Hubei Province, China

B. A., Fudan University, 2013

A Thesis presented to the Graduate Faculty
of the University of Virginia in Candidacy for the Degree of
Master of Arts

Department of Politics

University of Virginia
May 2015

Acknowledgements

I would like to thank my advisor Professor Todd Sechser and reader Professor Jonathan Kropko for their guidance and help throughout the whole project. During our meetings, Professor Sechser' advices were inspiring and extremely helpful in guiding me through decisions about topic choice, research question and data processing. His suggestion helped me grasp the core point of my thesis, how to revise the draft and how to explore the project further. I also would like to thank the reader of my thesis, Professor Kropko, who gave me a lot of methodological help before the project, carefully went through my draft, and provided many useful methodological as well as editorial advice that is of great value for later development of the project.

Regime Type, Balance of Capabilities, and Level of Escalation in International Crisis Bargaining

Abstract: The existence and meaningful influence of audience costs in international crisis bargaining have been hotly contested. Instead of crisis outcome, this article focuses on actor behavior in crisis bargaining, and explores the interactive effects of regime type and observable balance of capabilities on target's escalatory behavior. This article finds out that democratic states are more cautious in target selection, and do select themselves into crises they are more likely to prevail on the basis of observable balance of capabilities. Also, the direction and degree of regime type's influence on target's escalation as well as crisis outcome are context dependent and vary across different degrees of balance of capabilities. Audience costs do exist and significantly influence international crisis bargaining, but its role is far more complex than previously theorized and tested. My findings suggest that audience costs generation is more of a strategic choice than a pure invariable attribute.

Introduction

Do democracies make more credible coercive threats than their autocratic counterparts? An increasing literature in international relations argues that democratic states are better at signaling their resolves and true preferences in international crises.[1] Also, constrained by domestic opposition parties, democratic states are less likely to bluff and misrepresent information in crisis bargaining, meaning that the threats they make are more likely to be genuine.[2] Therefore, the target of a threat made by a democratic state should be less likely to resist or take further steps of escalation.[3] Following this logic, democratic states make more credible coercive threats compared with their

[1] See Fearon 1994a, Smith 1998; Guisinger and Smith 2002; Ramsay 2004; and Slantchev 2006.
[2] The most prominent work about influence of domestic whistle-blowing opposition parties on international crisis bargaining is done by Schultz 2001a. Also, see Ramsay 2004 for the role of opposition in international crisis bargaining. Potter and Baum 2013 bridges Schultz 2001a's theory with audience costs theory through media access, and explores institutional heterogeneity within democracies, further confirming the existence and influential role of audience costs.
[3] See Partell and Palmer 1999 for research about escalation. See Schultz 1999 and 2001a for research about reciprocation, which is defined as escalation beyond "no militarized actions" by the target state.

nondemocratic counterparts.

One reason that *ex post* inefficient wars happen is that states have incentives to misrepresent information in crisis bargaining so as to gain more leverage in division of the disputed goods (Fearon 1995). Information asymmetries are thus further exacerbated by states' rational behaviors of concealment, deception, and bluff. There are several attempts to bridge informational asymmetries into theories about regimes type and war.

One that has drawn the widest attention is audience costs theory. According to Fearon (1994a), "audience costs" refer to the costs leaders would bear if they back down from previous commitment to use force given their demands are not met by the target. It is generally agreed that democratic states are on average better at generating audience costs than their non-democratic counterparts, mainly because democratic leaders are domestically more accountable.[4] Fearon (1994a) argues that higher "audience costs" imposed on democratic leaders because of domestic accountability help them better signal resolve in crisis bargaining. By sending signals that carry audience costs too high for unresolved challengers, resolute challengers distinguish themselves from bluffers, and thus make their threats more credible. Knowing they are going to pay high audience costs if their bluffs are called, irresolute democracies automatically refrain from bluffing, thus "select themselves" into crises they are more likely to prevail. By contrast, domestically less accountable non-democracies states can make limited probes and then back down if resisted without suffering too much audience costs, rendering their unselective threats less credible.

Another theory proposed by Schultz (2001a) argues that open political competition together with the existence of legal opposition parties which are generally associated with democratic regimes can further increase a government's ability to reveal its true preference in foreign decision making.[5] In a political system where the government's foreign policy decision making can be publicly debated and legally challenged by opposition parties, the decision makers' ability to conceal or misrepresent

[4] Weeks 2008 argues that there is institutional heterogeneity within non-democratic states, and some autocracies are also quite accountable domestically and thus also able to credibly signal their resolve through audience costs mechanism. But on the whole, democracies are more accountable than non-democracies. An exception is Brown and Marcum 2011, who argue that non-democracies are more accountable domestically than democracies because they have smaller winning coalitions that can more effectively monitor and sanction leaders' diplomatic mistakes.

[5] Also see Schultz 1998.

their genuine intentions, that is, whether they are bluffing or not, is constrained. A government with only little domestic support for war cannot easily bluff without being questioned and challenged by the opposition parties, and thus has less incentive to lie. On the contrary, the threat from a government with strong support for war from within can be seen as more credible since it is somewhat "confirmed" by the opposition parties. Therefore, democratic states in the face of stronger domestic constraints should be more selective in sending threats, and thus increase the credibility of the threats they do make. Ramsay (2004) addresses a similar logic, pointing out that the opponent's rhetoric about foreign policy can have important effects in crisis bargaining, and compared with states without elections, democracies generally characterized by institutionalized electoral competition can get better offers from foreign rivals in crisis bargaining.

A third theory of democratic credibility involves the media freedom and information transparency that are also generally related to democratic states. Siegel (1997) proposes that free media that allow public discussion of state policy increase the transparency of a state, making it easier for decision makers of the adversary to obtain information about the government's preferences and constraints. The existence of free press can better inform the adversary about the military capability, the extent of public support for war and other constraints on waging war, and thus make it easier to "guess" whether the leader is bluffing or not.

Among them, the audience costs theory has been heatedly debated for a long time. As direct observation of the existence and magnitude of audience costs encounters severe problems because of partial observability and strategic selection[6], empirical tests of it focus mainly on indirect implications the theory predicts. Empirical researches on coercion generally support democratic credibility arguments. Partell (1997), Rioux (1998), Partell and Palmer (1999), Gelpi and Griesdorf (2001), and Schultz (2001a) all find evidence that democracies are more likely to prevail in international crisis than their non-democratic counterparts. Most importantly, by showing that challenges issued by democratic states are less likely to be reciprocated, Schultz (1999) demonstrated that the

[6] First, audience costs are less often observed than they actually exist, since they are only paid if the leader backs down after escalation. Second, states with ability to generate higher audience costs also have structural reasons to strategically avoid invoking audience costs too high for them to bear, implying that the probability of incurring audience costs also depends on their value. See Schultz 2001b for more details.

advantages democratic states enjoy in coercive diplomacy are brought about by more credible signal due to domestic accountability.

However, in recent years, democratic credibility arguments have been seriously challenged both theoretically and empirically.

Finel and Lord (1999) explore historical crises-1898 Fashoda crisis and 1967 The June War-and find clear evidence against transparency argument. Rather than revealing useful information and facilitating crisis resolution, transparency instead hinders conflict resolution by creating confusion and ambiguity with excess amount of non-authoritative information and undermining behind-the-scenes efforts at conflict resolution. Potter and Baum (2010) also express concern from the perspective of political communication, pointing out that media in democratic states are far from "neutral". Both elite messages and media filter can sometimes distort the information that reaches the domestic as well as international audiences.

Trachtenberg (2012) scrutinizes several historical cases and finds out little evidence suggesting audience costs playing even a limited role in influencing the crisis outcomes, neither does he find leaders reneging on their promises suffered serious political punishment *ex post*. Snyder and Borghard (2011) provide a more systematic critique of audience costs argument by pointing out that (1) leaders prefer policies with considerable degree of flexibility and seldom lock themselves in by clear commitment, (2) domestic audiences care more about substance of foreign policy than consistency between leaders' words and deeds, (3) domestic audiences care about state reputation and national honor independent of explicit threat, and (4) authoritarian targets of democratic challengers usually do not perceive audience costs as the theory predicts.[7] More importantly, Downes and Sechser (2012) convincingly challenge the validity of most statistical tests on democratic credibility theories by pointing out that the MID and ICB data sets widely used in these tests are not appropriate for testing democratic credibility of coercive threats, because they contain lots of minor skirmishes and alerts that are not coercive practices, and code the results of coercion incorrectly[8]. After correcting for the errors, they present another empirical analysis showing that democratic threats are not

[7] Also see Mercer 2012.

[8] See Downes and Sechser 2012. Conflation of victory achieved by pure military force and coercive success is the most serious defect.

more likely to be successful using the Militarized Compellent Threats (MCT) data set they compiled. McManus reported in Chapter 6 of her dissertation (2014, p. 175) that her empirical finding does no support audience costs theory because audience costs theory would predict that the statements of resolve would be more effective when a conflict is more salient and when the president is less secure. By contrast, the results indicate that salience has no impact on the effectiveness of statements, and presidential insecurity in office actually makes statements of resolve less effective.

Some of the critiques are quite convincing. Others, however, are worth contesting. And by doing so, they provide new opportunities to look into the underlying mechanisms as well as interactive and curvilinear effects unexplored in early literature.

First, audience costs are more of a choice than an invariable attribute. In order to avoid punishment in unfavorable situations, the leaders of states with ability to generate high audience costs sometimes choose not to deliberately invoke and increase audience costs.[9] Therefore, it is reasonable to see leaders try hard to keep crisis in low profile and avoid full commitment. However, a systematic difference between democracies and non-democracies remains strong that democratic leaders' attempts to avoid public attentiveness, scrutiny and punishment are less effective than those of the non-democratic leaders, because of the existence of whistle-blowing opposition parties and free media. Second, though audience costs are modeled with continuous value in several important studies[10], payment of audience costs is manifested in electoral outcome, a dichotomous variable for each candidate. Consequently, it is problematic to assume that a leader does not suffer from any audience costs *ex post* given he is not removed from office: it is a probability rather than necessity. More importantly, magnitude of audience costs is a strategic choice; states not only select themselves into crises they are more likely to prevail (Schultz 2001b), but also choose magnitude of audience costs through different level of escalation[11]. When there is little hope the other side will concede, rational leaders back down before audience costs are too large for them to bear-sometimes but not

[9] See Baum 2004, U.S. presidents prefer to handle international crisis without public attentiveness or scrutiny unless they are confident of success or important strategic interests are at stake.
[10] See, for instance, Fearon 1994a, and Schultz 1999 and 2001a for direct model of audience costs. Seen Smith 1998, Guisinger and Smith 2002, and Ramsay 2004 for model of electoral advantage, which is also presented as continuous.
[11] Detailed analysis is presented in the next part.

always, before they lose votes enough to cause electoral failure. Audience costs are difficult to observe exactly because of leaders' strategic choice. Thus lack of observation does not equal to nonexistence. Third, research shows that democratic states characterized by legal opposition parties still refrain from bluffing independent of audience costs[12].

The most persuasive criticism comes from Downes and Sechser (2012), who expose important drawbacks of MID and ICB data sets in testing democratic credibility theories. However, since they conducted the analysis that reports negative results only with data on compellent threats, whether their conclusion could be extended more broadly to all coercive practices including deterrent and compellent threats remains doubtful. More importantly, most of the empirical tests criticized by them only test arguments premised on non-interactive effect of each independent variable and focus mainly on crisis outcome[13], leaving rich empirical implications about actor behavior and crisis process generated by audience costs theory unexplored. For instance, Potter and Baum (2013) break down the unitary attribute of democracy and look for the interactive effects of opposition parties number and degree of media access on crisis outcome. Their results are robust even using the MCT data set compiled by Downes and Sechser (2012). Also, leaders' strategic choice to avoid audience costs is situation-dependent, and public willingness to punish bluffers is not invariant.[14] Focusing mainly on crisis outcome waste the theoretical richness of democratic credibility arguments. And failure to account for the interactive and curvilinear effects under different conditions would obscure the role that audience costs play in international crisis bargaining.

In this article, I discuss the utility of increasing escalatory steps in international crisis bargaining, explore how selection effect works through observable measures of balance of capabilities and strategic interests, and test the effects of regime type on crisis actors' escalatory behaviors.

The remainder of this article proceeds as follows. I begin by explaining the

[12] See Schultz 1998 and 2001a. According to Downes and Sechser 2012, Schultz 2001a actually argues that democratic leaders will pay audience costs only if opposition parties also opposed the threat. However, it is slightly different from audience costs theory that in Schultz's theory, audience costs are constant, not increasing as escalation level increases, and its presence or not does not influence the model. See Schultz 2001a Chapter 3 for more details.
[13] See Partell and Palmer 1999 for test involving level of escalation.
[14] See Baum 2004 and Clare 2007.

relationship between level of escalation and audience costs in detail. I then clarify the self-selection mechanism, why it is important for audience costs theory (and democratic credibility theories at large), and what empirical implications the selection effect generates. Next, I explore the effect of regime type conditional on different degrees of balance of capabilities. Finally, I statistically test the hypotheses generated in theoretical analysis using the MID data set (Palmer et al. 2015).

THEORETICAL ANALYSIS

Audience Costs and Level of Escalation

Domestic audience costs are generated because failure to go though commitment carries political consequence for the leader. It is possible that a leader who makes a commitment and then backs down will be punished in election or loses support as failure indicates the leader's incompetence (Smith, 1998)[15]. For audience costs to effectively influence international crisis bargaining, there are several requirements. First, domestic audiences should be able to observe the leader's political actions in an international crisis. Second, domestic audiences must have the ability to politically punish their leader, whether through election or other ways. Third, domestic audiences care about international crisis and are willing to punish the leader because he does not handle the crisis well. The former two requirements are more likely to be fulfilled in democratic states, which are in general characterized by legal existence of opposition parties able to expose the incumbents' misconducts, higher level of media freedom, stronger constraints

[15] Fearon's model (1994a) explored the important role domestic audience costs play in international crisis bargaining without explaining how audience costs come into being in the first place. Some critics of audience costs theory pointed out this inadequacy, arguing that there is no reason for the public to punish leaders who bluff, because bluffing sometimes is an optimal strategy of handling foreign affairs. Also, public should be unwilling to punish leaders making compromises when the public is against war. Smith (1998) provided a micro-foundation for the generation of audience costs by showing that domestic audiences punish leaders called on bluffing because it is a signal of leaders' incompetence, and the audiences want to retain competent leaders while remove incompetent ones. Guisinger and Smith (2002) offers another explanation of why audience costs arise by pointing out that the voters punish leaders who issue empty threats and get called by the opponent because they ruin the state's clean diplomatic reputation and jeopardize the state's credibility and long-time interests in future crises. These explanations point to a similar logic that the public do not solely criticize leaders for bluffing-a strategy commonly adopted in business, politics, and even gambles; on the contrary, they criticize and punish leaders for being called on bluffing-a signal of incompetence and an outcome that jeopardizes national interests in the long run.

on administrative powers, and more effective election system through which the voters could politically punish incompetent leaders without too much costs. Therefore, democratic states are better at generating audience costs.

In literature, higher level of escalation is assumed to generate larger audience costs[16], but no theory has provided a satisfactory explanation of why this is the case. In this article, I argue that higher level of escalation is associated with higher salience of the issue in dispute and therefore wider knowledge of the commitment the leader makes. The public can impose political punishment on leaders who back down from their previous commitment only when they are aware of the misconduct and disprove it. If few people notice the crisis, whether the leader goes through his commitment does not matter as losing of few votes is bearable. On the contrary, if a crisis enters public forum and is heatedly discussed by the citizens, failure of handling the dispute exposes the leader's incompetence to more voters, and thus significantly increases the leader's likelihood of being punished in future election.

More importantly, level of escalation carries information about severity of the international crisis. There is considerable evidence in literature suggesting that the mass public care more about domestic political affairs than foreign policies (Holsti, 2004). Most of the public either lack professional knowledge about international relations or remain aloof over foreign policies that bear less consequence for their own lives. This results in a domestic political atmosphere where only a relative small part of so-called

[16] In Fearon's (1994a) model, audience costs increase with time for three main reasons. First, delay in crisis could be a signal of the leader's higher value for war. Because under the opposite conditions, a leader with higher value for peaceful agreement should be more impatient to get a deal and "enjoy whatever benefits a resolution would allow" (p. 580). Second, delay means the leader is willing to risk the probability of inadvertent war. Third, delay runs a genuine risk of being preemptively attacked, and hence might credibly signal the leader's higher value for war than backing down. However, none of the explanations firmly holds. Delay in response is usually accompanied by domestic divergence on how to response. Existence of opposition, especially those come from political competitors, could be a signal of the leader's lack of domestic support, and therefore reinforce the other side's belief that it is facing an unresolved adversary. Also, inadvertent war is rarely a serious problem in normal international crisis without involvement of nuclear weapon. Except for extremely unstable regimes, supreme leaders of most states retain control over the state power, able to prevent skirmishes and unauthorized military clashes from getting out of hand. Surprise preemptive attack is less of a concern, at least in the early stages of international crisis. More importantly, lots of leaders delay in response only to wait public attention on the crisis to fade away. With less people focusing on the crisis, leaders are less likely to be held accountable for later concessions. More persuasive evidence is that though Fearon (1994a) modeled audience costs to be increasing as time passes, he used level of escalation in further discussion.

"attentive public" pays a considerable and stable attention to foreign affairs, let alone influencing foreign decision-making (Potter and Baum, 2010; Baum, 2004).

Situation changes when international crisis happens. The public is less likely to be attracted by insipid routine interaction with other states than by intensive conflicts over a disputed issue that might be directly related to their own welfare. The later is exactly the case of international crisis.

However, not all crises are similarly "attractive". Study shows that the public pays more attention to strategically important issues (Clare, 2007). Another factor, as I argue here, is whether the outcome of international crisis is directly related to domestic public's own welfare. Higher level of escalation, especially actual use of force and full-scale war, in international crisis leads to wider public attention domestically because higher level of escalation indicates that the crisis is more serious and carries higher risks, thus gains larger media coverage and public attentiveness. Media spare no efforts in reporting consequential and dramatic events while spend very little space on issues of low hostility. As the main resource of political information for the mass public, media coverage is closely associated with public attentiveness. Even if the mass public notice reports on crisis with low escalation level, for instance, a crisis at its early stage when the leader only threats to use force, they are more likely to think it is not a big deal and could be easily resolved soon, and thus discard the ongoing crisis. However, display of troop on boarder or taking some similar actions is not something trivial, and would attract more media coverage and public attention. More importantly, nowadays in most countries, actual use of force and real war would trigger heated discussion. This is true at least in democratic states with free media and strong opposition parties that prevent the leader from hiding their political actions. Use of force and (even limited) war are sensitive topics among mass public because these policies have huge influence on their lives and the national state as a whole. In the first place, use of force and actual war consume huge amount of material as well as financial resources, the costs of which eventually will be burdened by the public. Second, death is unavoidable in war. Loss of young, promising lives is never a pleasant topic in modern states. Democratic states are especially sensitive

to fatality in war. In conclusion, higher level of escalation is associated with higher issue salience and wider public attention.[17]

The association might be doubted in autocratic states. In the first place, it is much easier for government's political actions in an autocratic state to escape from public attention or discussion by government media control so that the mass public only knows what the government wants them to know. A public threat can be widely dissipated among the public while a troop can be deployed secretly without public attention. Salience of the crisis is more in the leader's control. However, if the purpose of displaying military force in a crisis is to convey resolve to the target, what is the meaning of doing it in secret? Unless in a completely isolated country, such as North Korea, no leader can display military force to its target state while holding back such information from its own citizens completely, as long as there is access to outside information, either through foreign media, worldwide social media or immigrant connections. Deployment of troops secretly in preparation for war so that the state can gain preemptive advantage is not part of the international crisis bargaining process, and therefore is not our focus here. Even in countries where media freedom is to some extent restricted and people are not allowed to publicly criticize the government's policies, private discussion is still possible. Also, people will be allowed to publicly state their support for the government's decision to protest, threat, display troops, and use force against the target state. Public support is also a kind of attention, though in a non-democratic state it does not incur audience costs as much as in a democratic state if the leader fails to carry out the threat. In conclusion, level of escalation is still associated with salience of the crisis in autocratic states.

Apart from issue salience and public attentiveness, level of escalation also indicates issue severity. Domestic audience does not punish every empty threat equally. They are more likely to punish leaders backing down on issues with high strategic interests, issues that carry huge material loss or sever political consequence once the leaders back down (Clare, 2007). Level of escalation is an important indicator for domestic audiences, who generally do not acquire professional knowledge about foreign

[17] A quick empirical assessment is done with McManus (2014)'s replication data, which could be downloaded from http://www.prio.org/JPR/Datasets/. In the test, higher hostility level of the initiator (the United States) is associated with increase of salience of the crisis in U.S. media (measured by the average number of New York Times articles per day mentioning the dyadic MID adversary during the dyadic MID time frame) at 95% significance level (p=0.012).

affairs, to judge the severity of the international crisis. The judgment on the basis of rough connection has a logical reason. The higher level of escalation, the more committed a leader is. Also, a leader is more likely to deeply commit if high strategic interests are at stake. Therefore, when the public observes higher level of escalation in a crisis, they are more likely to conclude that the strategic interests at stake are high, and hence more likely to punish the leader if he backs down.

In a situation where actors have incentive to misrepresent information, signals that do not carry costs are not able to credibly reveal the actor's true preference and hence convey no useful information. International crisis bargaining is exactly such a situation, where states involved in the crisis have incentive to exaggerate their military capabilities or willingness to fight so as to get a better deal in negotiation (Fearon, 1995). Under such circumstances, threats unaccompanied by costs are empty, costless signals that do not tell much about the crisis actor's true preference between war and backing down. On the contrary, escalations carry costs, and therefore whether a state is willing to escalate, and to what a degree it is willing to escalate can be useful in differentiating resolved states from irresolute ones.

Higher escalation in crisis is associated with higher costs.[18] Apart from audience costs, escalation carries other physical as well as reputational costs. The former includes the financial and organizational costs connected with mobilization and deployment of the troops (Fearon, 1994a; Fearon, 1997). Except for direct financial expenditure, mobilization also entails huge opportunity costs. During mobilization, productive forces, such as human resources and raw materials that could have been devoted to profitable economic activities are reallocated to military production, which expects no refund even if the other side backs down. The higher the level of escalation, the higher the sunk costs. Deployment of troops is more costly than oral threat of force, while partial or full mobilization in preparation for war is more costly than deployment of troops.

[18] In Fearon (1994a)'s model, escalation is costly for states because it carries audience costs. However, here I am trying to explain why audience costs increase as states escalate to higher level. It is necessary to differentiate audience costs and other costs incurred by escalation. The later is more appropriate in explaining why higher escalation level indicates deeper commitment of the leader. Though in equilibrium, states with ability to generate higher audience costs are more likely to escalate high, and thus in turn states willing to escalate to higher level are more resolved. However, using such logic to explain why higher audience costs arise from higher level of escalation amounts to circular reasoning. Nevertheless, we need to notice that in realty the circular effect does reinforce each other.

The later mainly refers to extra costs a state has to pay in future crises if it develops a reputation for certain attributes in historical crisis management. It could be a reputation for lack of resolve for war. Potential adversaries are more likely to challenge less resolved states. With a reputation for weakness, a state is more likely to be challenged and pay costs it could have avoided. It could also be a reputation for a propensity to bluff and make limited probes. A state that challenges in the first place and then backs down in many crises is likely to be regarded as habitual bluffer, and therefore more likely to be resisted in future crises. There could be situation where the state is actually making a genuine threat; nevertheless its threats are discarded as bluffs. Under such circumstances, a crisis that could have been resolved by the target's concession now has to end up with a costly war. The difference between the state's payoff for war and successful coercion is the reputational costs the state suffers. It could also be a reputation for diplomatic inconsistency that renders the state's diplomatic statement as well as military threat in the future less credible in general (Guisinger and Smith, 2002). No clear evidence shows that escalating to higher level and then backing down bears higher reputational costs than backing down at relatively lower escalation level. However, people are more likely to learn from events of higher salience and greater influence (Jervis, 1976, p. 239). This means that other states are more likely to recall and learn lessons from severer crises in the history. Higher level of escalation, as argued previously, is associated with higher issue salience and severity, which might leave deeper expression on international audiences. If a state backs down after escalating to higher level, its weakness, bluff, or diplomatic inconsistency is more likely to be remembered. Following this logic, higher level of escalation carries higher reputational costs.

Since higher escalation level carries higher costs, only leaders who are highly committed are likely to escalate high. In contrast, repeated statement of resolve unaccompanied by "real" political actions implies lack of resolve. Sometimes even domestic audiences understand that oral threat of force issued by the leader might be just bluffing and do not take it as deep commitment. For instance, Chinese government was mocked at by Chinese netizens because it repeatedly issued "strong condemnation" of Japan's behaviors regarding Diaoyu Islands yet never took any further actions beyond

patrol ships. Even Chinese citizens know that the government was just speaking diplomatic cliché and would not attack Japan at least recently.

Leaders are more committed when strategic interests at stake are high. States have different evaluation of the value in dispute. The higher a state's evaluation of the disputed good, the higher the state's value for war, and thus the higher the likelihood it is resolved. Therefore, the mass public is more likely to conclude that important strategic interests are at stake after observing high level of escalation, and thus more likely to punish the leader who backs down after escalating high. In conclusion, higher level of escalation is associated with higher audience costs.

Selection Effect and Balance of Capabilities

Audience costs cast influence on international crisis bargaining mainly through selection effect. Selection bias happens when observations are non-randomly selected into the sample, so that the result is biased systematically and cannot accurately estimate the true relationship between two variables (King, Keohane, and Verba, 1994). When it happens, selection bias in political research usually distorts the real effect of variables under study, and generates a negative effect that researchers strive to minimize as much as possible (Heckman, 1979). The situation is exactly the opposite in audience costs theory: selection effect is one of the main mechanisms that accounts for the reason why states with ability to generate higher audience costs enjoy advantages in international crisis bargaining. To test hypotheses generated by audience costs theory, we do not expect the sample of crises we collect to be an unbiased representative of all possible situations in which a crisis might happen; rather, states make strategic decisions as whether to challenge in the first place and whether to resist once challenged. As pointed out by Schultz (1999, p.256), conditional on the initiator's having chosen to challenge the target, audience costs theory predicts certain phenomena and relationship in the sample. It is precisely because of selection effect-certain states select themselves into the crises-that we observe the empirical outcome.

Selection effect is the result of audience costs carried along with escalation in the crisis bargaining. Previous model of crisis bargaining reveals that states with ability to generate higher audience costs are significantly less likely to initiate "limited probes" in

foreign policy (Fearon, 1994a, p. 585). The logic lies in the fact that escalation is more costly for such states. Knowing that they are going to pay huge amount of audience costs if they are caught bluffing, or if they escalate to a point when the audience costs are too high to bear so that they have to fight a costly war they are initially unwilling to fight, states with higher audience costs would refrain from the pulse to make limited probes. For them, the costs associated with high risks of being caught issuing empty threats or being forced into an unavoidable costly war outweigh the possible gains they might achieve from bluffing. As audience costs increase, unresolved states are increasingly reluctant to challenge in the first place if they are not sure the type of its opponent or remain doubtful about whether they can prevail in the crisis. On the contrary, states with low audience costs are relatively more willing to initial limited probes if the possible gains from bluff are high enough to offset the audience costs they would suffer once their threats get called. Therefore, states with low audience costs are more emboldened to bluff. Consider the extreme case in which the leader is completely not held accountable domestically for backing down after any level of escalation. The leader suffering from no audience costs will have an incentive to challenge in any crisis regardless of its own resolve. By doing so, he can be no worse off. If there is a slight chance that the target will concede, he gains by getting the disputed good without fighting; if unfortunately his threat is resisted, he can immediately back down without suffering from any costs. Either way, he is at least as good as not disputing the status quo. In conclusion, states with ability to generate higher audience costs, being aware of severer consequence they will bear once their bluffs get called, are more cautious in choosing their targets. In general, they select themselves into crises they are more likely to prevail.

How do states make initial evaluation about whether they are likely to prevail in the international crisis bargaining or not? Consistent with much of the literature on international crisis (Fearon, 1994a; Baum, 2004; Schultz, 2001a), I argue that (1) observable balance of capabilities and (2) strategic interests at stake are the two most important indicators the states use to *ex ante* gauge their chance of prevailing in the crisis. To clarify possible confusion, this argument is distinctively different from the claims that militarily superior states are more likely to prevail and that states with important strategic interests at stake are less likely to back down. The later two claims' line of logic has been

criticized by Fearon (1994a, p. 586; 1994b). Balance of capabilities and strategic interests at stake do not directly determine the outcome of crisis bargaining; rather, these two indicators provide relatively credible information about the opponent's possible value for war and thus shape the state's initial belief about its opponent's resolve. On the one hand, increasing balance of capabilities in favor of the opponent increases the probability that the opponent will win on the battlefield and decreases the opponent's costs of war; on the other hand, more strategic interests at stake indicates that the opponent has higher value for issue in dispute: both are associated with opponent's higher value for war and hence higher resolve. Observing military superiority of the opponent and high strategic interests for the opponent, the potential challenger should be more likely to believe that it is facing a resolved opponent and that it has little chance of coercing the opponent into concession. A state with ability to generate higher audience costs thus filters these resolved opponents out from its target list; while a state with low or no audience costs might still challenge because its costs for making limited probes are bearable.

Consequently, if the initiator is a high-audience-costs state, in this case a democratic state, after self-selection into the crises on the basis of above criteria, observable balance of capabilities and strategic interests at stake should have no significant influence on the target's decision to escalate in later crisis. Empirical evidence regarding the later supports the prediction: democracies rarely contend over territorial issues, which are consistently salient and involve important strategic interests for national states; but beyond issues of extreme salience, democratic challengers are able to choose crises they are easy to prevail.[19] The analysis above leads to my first hypothesis.

H1 *Ceteris paribus*, when the initiator is a democratic state, observable balance of capabilities should have no significant influence on the target's probability of escalation.

Diverging from Fearon's conclusion that all "rational states will 'select themselves' into crises on the basis of observable measures of relative capabilities and interests and will do so in a way that neutralizes any subsequent impact of these measures" (Fearon, 1994a, p. 586), my theory predicts that the selection effect does not happen for all types of initiators. On the contrary, I argue that the effect depends on the initiator as well as the target's regime type. Non-democratic states with low audience costs on

[19] See Gibler and Hutchison 2013.

average are less cautious than their democratic counterparts in choosing promising crises. Consequently, in a crisis initiated by a non-democratic state, observable balance of capabilities and strategic interests at stake should remain as a feasible indicator to help one side guess the other side's chance of prevailing in the crisis bargaining.

How ensuing crisis unfolds also depends on the target's regime type. Facing non-democratic initiators, a democratic target also selects itself into crises it has more confidence in prevailing based on observable clues about the initiators' true resolve, such as balance of capabilities and strategic importance of the disputed issue for the initiators.

Accordingly, in the face of a non-democratic initiator, a democratic target only uses observable measures of military capabilities as a reference in escalation decision at early stage of the crisis, that is, decision on whether to reciprocate. Given a democratic target "selects itself" to be involved in the crisis by reciprocation, balance of capabilities should have no significant influence on the target's decision to escalate thereafter. Conversely, a non-democratic target is incautious in choosing what kind of crisis to be involved. As a result, observable measures of balance of capabilities should influence the target's decision to escalate, and the direction of the influence is expected to be in accordance with the common informal claim previously discussed. This generates the following two hypotheses.

H2a *Ceteris paribus*, when the initiator is a non-democratic state, increasing observable balance of capabilities in favor of the initiator will decrease a non-democratic target's probability of escalation to higher level.

H2b *Ceteris paribus*, when the initiator is a non-democratic state, increasing observable balance of capabilities in favor of the initiator will decrease a democratic target's probability of reciprocating the crisis. However, observable balance of capabilities should have no significant influence on the democratic target's probability of further escalation.

Regime Type and Level of Escalation

Fearon (1994a)'s model of "war of attrition" generates several important empirical implications regarding the influence of regime type on state behavior and crisis

outcome. Here, I mainly discuss two of them that are most closely related to states' escalatory behaviors.

First, escalation levels needed to credibly signal one's preference differ for states of different regime types (Fearon, 1994a, p. 585). On average, democratic states can generate higher audience costs than their non-democratic counterparts with each escalatory step. If a certain amount of audience costs is necessary to costly communicate with the opponent, a democratic state can reach the critical amount using fewer escalatory steps. Thus a democratic state is less likely to escalate to higher level than a non-democratic state. Second, probabilities of escalation and concession vary according to difference of audience costs between crisis actors. States with high audience costs are more likely to escalate, while states with low audience costs are more likely to back down (Fearon, 1994a, p. 586).

While Fearon (1994a) models the independent effect of regime type on escalation, I take a step further by arguing that the effect of regime type are not the same across various degree of balance of capabilities. The interactive effect requires more detailed theorization and explanation.

When the initiator is a non-democratic state, which is less cautious in selecting itself into "proper" crises, observable measures of balance of capabilities is a more important indicator of the initiator's value for war, and thus a main source of the target's belief about the initiator's type. Consequently, balance of capabilities, instead of regime type, cast more influence on the target's decision to escalate in the crisis. The logic above leads to my third hypothesis.

H3 *Ceteris paribus*, when the initiator is a non-democratic state, the target's regime type should have no significant influence on its probability of escalation across all degrees of balance of capabilities.

The situation is different when the initiator is a democratic state. Democratic states, as previously argued, are more reluctant to make limited probes and carefully select themselves into crises they are more likely to succeed. After democratic states' self-selection process, balance of capabilities as an observable measure loses most of its value in determining crisis outcome. Though it no longer directly influences the target's probability of escalation, it still forms the underlying conditions under which the target is

more or less likely to behave under the influence of audience costs. Formally, balance of capabilities does not influence the probability of escalation directly, but it defines the target's *ex ante* belief about the distribution of the initiator's type, and thus creates a range in which the effect of audience costs is more salient. I argue that regime type has the most salient effect when other observable indicators are no longer useful or too elusive in helping guess the possible type of the opponent. Therefore, the effect of regime type on probability of escalation is expected to be most salient when the crisis actors are comparable in observable military capabilities. Combined with predictions illustrated by Fearon's model, the fourth hypothesis is given as follows.

H4 *Ceteris paribus*, when the initiator is a democratic state, a democratic target is less likely to escalate to higher level compared with a non-democratic target. The effect should be most salient when the crisis actors are comparable in observable balance of capabilities, and should be indistinguishable when one of the crisis actors has observable military advantages.

Alternative Measures

My hypotheses address the relationship between regime type, balance of capabilities and level of escalation. Escalation, however, has been criticized for having substantial drawbacks in reflecting state's advantage in crisis bargaining. Most models of international crisis offer clear-cut strategies to crisis actors, who are allowed to choose between concession, escalation and fight. However, in real crisis bargaining, it is far more likely for the disputants to reach a peaceful resolution, turn into stalemate, or leave it alone until public attentiveness fades away.[20] A crisis could also end up with a peaceful agreement in favor of the side that "backs down" first on the table. Citing U.S.-Afghanistan 1998 crisis, in which Afghanistan's Taliban regime simply ignored U.S. demand to hand over Osama bin Laden, Downes and Sechser (2012, p. 466) clearly pointed out that low escalation level does no necessarily imply target's concession and initiator's victory.

[20] Most crises (1,708 out of 2,209 in my sample of analysis) in the MID data set ended up with compromise, stalemate and unclear outcomes.

Additionally, I will test my hypotheses with MID data set, the coding deficiencies of which have been well documented (Downes and Sechser 2012). However, MID data set remains to be the most inclusive data set about interstate military conflicts before any data set specific about coercion is available. To check the robustness of tests using MID data set, it is important to replicate the tests using MCT data set, which is free of the deficiencies mentioned above.

Though less escalatory steps do not equal to advantage or favorable outcome in crisis bargaining, it is known by definition that the target's escalation to or beyond actual use of force with considerable amount of fatality is failure for the initiator. This leads us to the following hypotheses specific for tests using MCT data set.

H5 *Ceteris paribus*, observable balance of capabilities should have no significant influence on a democratic initiator's probability of failure.

H6 *Ceteris paribus*, increasing observable balance of capabilities in favor of the initiator will decrease a non-democratic initiator's probability of failure.

H7 *Ceteris paribus*, the target's regime type should have no significant influence on the probability of failure for a non-democratic initiator across all degrees of balance of capabilities.

H8 *Ceteris paribus*, compared with a non-democratic target, a democratic initiator is less likely to fail facing a democratic target. The effect should be most salient when the crisis actors are comparable in observable balance of capabilities, and should be indistinguishable when one of the crisis actors has observable military advantages.

EMPIRICAL TEST

Data and Dependent Variables

I begin to test my hypotheses with MID data set. I use most part of Downes and Sechser's replication of Schultz's data set in his book *Democracy and Coercive Diplomacy* (2001), since Schultz's own replication data is not available. Following Schultz's description in his book about how he constructed his data set based on several other sources, Downes and Sechser (2012) reconstructed a data set that generates the same empirical result with high precision, and thus can be seen as a reliable source. It

contains all the variables I need in this analysis. Instead of MID 2.10 data set and Polity III scores Schultz used when he completed the work in 2001, I update the datasets with the latest versions available, including MID 4.01[21] dataset and Polity IV[22] data set. I also update "S-scores" that evaluate alliance portfolio similarity and the Tau-b scores that measure the similarity of a given state's alliance portfolio to that of the most powerful state in the system (Schultz 2001, p. 266) using the EUGene program (v3.204)[23]. The updated dataset contains a total number of 2,209 observations of militarized interstate crises from 1816 to 2010. Since S-scores and Tau-b scores of states from 2001 to 2010 are not available, after deletion of cases after 2001, I end up with a data set of 1,935 observations. The dependent variable is level of escalation of the target state, an ordinal variable that has 5 values, including no militarized action, threat to use force, display of force, use of force, and war, from low to high. Since the dependent variable is an ordinal variable, I will use ordered logit model to test my hypotheses.

I also use the Militarized Compellent Threats (MCT) data set compiled by Downes and Sechser (2012) to check the robustness of findings with MID data set. MCT data set is comprised of 242 crisis dyads and provides detailed information about 210 compellent threats issued between 1918 and 2001.[24] To keep in accordance with tests using MID data set, I drop all multilateral crises, which leaves me a total number of 187 observations. MCT data set is specifically suited for testing democratic credibility theory, because it is comprised of cases in which political demands accompanied with coercive threats are publicly issued, and it differentiates clearly between victory by force and victory by coercion.[25] The dependent variable for tests using MCT data set is a binary variable that indicates whether the crisis outcome is a coercive failure for the initiator. Compellent threats are successful (*failure_100_ks*=0) if the initiator achieved full compliance from the target without actually using force or the target complied with less than 100 military fatalities on its side. This coding standard reflects the logic that military engagement accompanied by small fatalities has more coercive meaning than tactic value.

[21] Palmer, Glenn, Vito D'Orazio, Michael Kenwick, and Matthew Lane. Forthcoming. "The MID4 Data Set, 2002-2010: Procedures, Coding Rules and Description." Conflict Management and Peace Science.
[22] Marshall, Monty G., Ted Robert Gurr, and Keith Jaggers. 2013. "Polity IV Project: Political regime characteristics and transitions, 1800-2013".
[23] Available at http://eugenesoftware.org/.
[24] Sechser 2011, and Downes and Sechser 2012.
[25] Downes and Sechser 2012, pp. 475-476.

A target conceding in the face of small fatalities should not be considered as failure in military confrontation. Rather, the target yields because more military force is on the way, and small-scale use of force by its adversary successfully conveys the resolve. Since the dependent variable here is a binary variable, I will use logit models to test the final set of hypotheses.

A more detailed discussion of variables included in this analysis is presented below. And descriptive statistics are shown in Table 1and Table 2.

Table 1. Summary of regime types

	Nondemocratic target	Democratic target	Total
Nondemocratic initiator	498	258	756
	65.87%	34.13%	100%
Democratic initiator	147	76	223
	65.92%	34.08%	100%
Total	645	334	979
	65.88%	34.12%	100%

Table 2. Regime type and level of escalation

	Target escalation					
	No militarized action	Threat of force	Display of force	Use of force	War	Total
Nondemocratic Initiator	415	22	102	203	14	756
	54.89%	2.91%	13.49%	26.85%	1.85%	100.00%
Democratic Initiator	146	4	20	47	6	223
	65.47%	1.79%	8.97%	21.08%	2.69%	100.00%
Total	561	26	122	250	20	979
	57.30%	2.66%	12.46%	25.54%	2.04%	100.00%

Independent Variables

The independent variables of interest are the initiator's regime type, the target's regime type and balance of capabilities between crisis actors. In order to test the interactive effects of regime type and balance of capabilities in my hypotheses, I include two-way interaction between each pair of independent variables (*democ_a*democ_b*,

*democ_a*capshare_a*, and *democ_b*capshare_a*), and a three-way interaction term of all three independent variables of interest (*democ_a*democ_b*capshare_a*).

Measures of regime type come from Polity IV data set, which contains information on regime characteristics for every state in the international system from 1800-1994. Following Schultz's method, states are coded as democracies (*democ_a*=1 or *democ_b*=1) if the executive of the state is elected through competitive election system, and the political competition within the system is stable and enduring.[26] Though Schultz's coding has been criticized as somewhat idiosyncratic[27], it could find support in the literature. Studies show that domestic political constraints on executive power are more reliable indicators of international crisis outcome[28], and domestic political stability is associated with less escalatory behaviors and more credible signals in international crisis bargaining[29].

For robustness check, I also repeat the analyses with a more widely used method known as the standard Polity method. By subtracting a state's 10-point autocracy score from its 10-point democracy score, we get a polity 21-point polity score ranging from -10 to +10. States are coded as democracies if they score 6 or higher on this scale; otherwise, they are identified as non-democracies. The polity scores generate two dummy variables: *democ_a_st* indicates whether the initiating state is democratic or not, *democ_b_st* indicates whether the target state is democratic or not.

The other key independent variable of interest is balance of capabilities between two crisis actors. It is operationalized by measuring the share of dyadic military capabilities possessed by the initiator. According to Schultz, this measure was constructed using data from COW's National Material Capability data set (v4.0)[30], which reports information on demographic, economic, and military characteristics of every state in the national system. A composite score that contains information about a state's relative capability in several aspects is calculated following the standard practice developed by Singer, Bremer and Stukey (1972). This variable (*capshare_a*) is

[26] Schultz 2001, pp. 128-130. The coding rule is that democracies are coded 1 if xrcomp = 2 or 3 and parreg = 2 or 5 in Polity IV.
[27] See Downes and Sechser 2012, p. 477, footnote 70.
[28] Partell, 1997.
[29] Prins, 2003.
[30] Singer, J. David, and Melvin Small 1993. Available at http://www.correlatesofwar.org/data-sets/folder_listing.

constructed through dividing the score of the initiator by the sum of the capabilities of both the initiator and the target state in the dyad. It ranges from 0 to 1, and higher value indicates increasing balance of capabilities in favor of the initiator.

Control Variables

Building on previous studies, I control for several other variables that also influence the dynamics of international crisis bargaining, including crisis actors' power status, alliance portfolio similarity, evaluation of the status quo, territorial contiguity, and the nature of issue in dispute.

Power status measures the power status of the initiator and the target in the international system. Power itself is a backup for resolve. Threats from states with greater power are naturally more credible because they have enough capability to carry on. Based on the COW classification system, states were coded as being either major or minor powers, and four categories are created: MAJMAJ, indicating that both states were major powers; MAJMIN, indicating a major power initiator and a minor power target; MINMAJ, indicating a minor power initiator and a major power target; and MINMIN, indicating that both states are minor powers. In the model, MAJMAJ is treated as the base category, and the effects of other three categories are evaluated relative to the base category.

Alliance portfolio similarity The existence of a military alliance at the time when the crisis happened might influence the evaluation of balance of capability of both states, and thus influence the effectiveness of information revealing. The dummy variable ally measures whether or not the states in the dyad had a defense pact, neutrality agreement or entente in the year prior to the onset of the dispute. The lag of one year attempts to reduce the possibility that an alliance that ended because of the dispute is mistakenly counted.

Status quo evaluations Two variables are used to measure the initiator and the target's evaluation of the status quo. A state's evaluation of the status quo is determined by the similarity of its alliance portfolio to that of the leading state in the system at the

time of crisis.[31] The leading state was Great Britain before 1945 and the United States after 1945 in the time period covered in this research. The variable ranges from -1 to 1, with a higher value indicating more overlap between the strategic interests of the crisis actor and those of the leading state. These variables are included because a state less satisfied with the system is more likely to initiate a crisis or war, and likely to escalate in the crisis.

Territorial contiguity According to Schultz, this variable is constructed based on data from Correlates of War project. It is a dummy variable measuring whether the states involved in the crisis shared a land border or were separated by less than 150 miles of water. Territorial contiguity is a strong predictor of crises and wars in the literature of security study. Territorial contiguity decreases the cost of transportation of force into the disputed area or the border of the adversary state, and increases the probability of escalation. Also, when the states involved in the crisis share a land border, and both display its troops close to the border, the probability of accidental fire and inadvertent escalation is significantly higher.

Nature of issue in dispute The unordered categorical variable *issuestake* measures the nature of the issue in dispute in each crisis, including territorial dispute, policy change, regime or government change, and other. We might predict that issues of high stakes, such as territorial dispute, need higher level of escalation than those minor policy issues to show the initiator's resolve. In the model, crises over territorial disputes are selected as the base category.

World war cases are considered by several studies to be based on different mechanism compared with those in normal period, and hence are dropped in their models.[32] It is true that during World War, states choose to wage war or not regarding more of alliance relationship and other international political considerations rather than domestic audience costs calculation. However, if crisis actors did negotiate and went through escalatory steps, they preferred to acquire what they wanted without fight. To control for influence of world wars, I follow the method of Downes and Sechser (2012) by presenting two models for each dependent variable, with one excluding world war

[31] The measures used here are called Tau-b scores. Only Tau-b scores before 1993 are available in Downes and Sechser 2012's replication data. I update to include data until 2000 using EUGene program.
[32] See Gelpi and Griesdorf 2001.

cases and another including world war cases yet creating a dummy variable for each year during world war periods.

Results

Core empirical results are presented in Table 3. Model 1 and Model 2 use the MID data set and explore the effect with and without world war cases separately. Model 3 and Model 4 use the MCT data set and both analyses are limited to bilateral cases. Model 4 tests the effect on all bilateral compellent threats, while Model 3 dropped disputes initiated during world war periods.

Since it is difficult to obtain information about the substantive effects based on conditional coefficients and corresponding log likelihoods, I report the marginal effects of key independent variables of interest and predicted probabilities conditional as the casual variables vary after controlling for other variables. Table 4 reports the marginal effect of the initiator's military capability share on probability of escalation given different regime types of crisis actors. Table 5 reports the marginal changes in Model 2 with world war dummies. Table 6 presents the marginal effect of initiator's capability share on the probability of compellent threat failure based on coefficients in Model 3 and Model4.

In accordance with the prediction of H1, it is clearly seen in Table 4 that when the initiator is democratic, we cannot reject the null hypothesis that the initiator's share of military capability has no influence on the target's probability of escalation, after controlling for other variables. The results also provide strong support for H2a, and the effect is largest and most salient on no militarized actions and use of force. According to Table 4, when the initiator is a non-democratic state, 1-point-increase in the initiator's capability share increases the probability that the target will take no military actions by 20.68% ($p=0.004$), and decreases the probability that the target escalates to use of force by 15.82% ($p=0.004$), after controlling for other variables. Table 5 presents similar results, with the marginal changes in probability slightly smaller than those reported in Table 4. However, they are both meaningful and at least significant at $p < 0.05$, and the marginal effects at no militarized actions and use of force are even significant at $p < 0.01$.

Table 3. Ordered logit analysis of MID escalation and logit analysis of MCT failure

	MID Target's escalation		MCT Compellent threat failure	
	1 World wars excluded	2 World wars included	3 World wars excluded	4 World wars included
Dem initiator	-0.20 (0.50)	-0.09 (0.49)	0.12 (1.39)	0.68 (1.44)
Dem target	-0.79* (0.36)	-0.81* (0.35)	-1.94 (1.29)	-2.08+ (1.21)
Initiator capability share	-0.97** (0.34)	-0.98** (0.33)	-0.41 (0.97)	-0.31 (0.83)
Dem initiator # Dem target	-0.30 (0.78)	-0.09 (0.75)	-0.21 (2.65)	-1.81 (2.82)
Dem target # Initiator capability share	1.53** (0.58)	1.59** (0.54)	0.27 (2.05)	-0.53 (2.00)
Dem initiator # Initiator capability share	F0.46 (0.68)	0.39 (0.66)	3.95* (1.89)	4.65* (1.85)
Dem initiator # Dem target # Initiator capability share	-1.20 (1.16)	-1.60 (1.11)	-2.53 (3.67)	-0.91 (3.84)
Contiguity	0.83*** (0.18)	0.67*** (0.17)	-0.12 (0.52)	-0.18 (0.50)
Alliance portfolio similarity	0.16 (0.28)	0.21 (0.27)	0.25 (1.07)	0.23 (0.93)
Status quo Initiator	0.17 (0.25)	0.11 (0.24)	0.41 (0.78)	-0.16 (0.68)
Status quo Target	-0.21 (0.28)	-0.19 (0.27)	-2.20* (1.09)	-0.43 (0.80)
Issue type: Territory as base				
Policy	-1.02*** (0.15)	-0.99*** (0.15)	-1.55 (1.06)	-1.77 (1.10)
Government	0.48 (0.35)	0.47 (0.34)	-0.76 (0.52)	-0.22 (0.47)
Other	-1.37* (0.55)	-1.32* (0.54)	-0.21 (0.60)	-0.13 (0.49)
Power status: MAJMAJ as base				
MAJMIN	0.40 (0.46)	-0.04 (0.42)	1.67+ (0.96)	1.54+ (0.83)
MINMAJ	0.42 (0.49)	0.05 (0.44)	-0.12 (1.69)	-0.53 (1.31)

MINMIN	0.37	-0.03	1.76	1.37
	(0.44)	(0.40)	(1.20)	(1.09)
Cut1 Constant	0.06	-0.37		
	(0.49)	(0.45)		
Cut2 Constant	0.18	-0.24		
	(0.49)	(0.45)		
Cut3 Constant	0.89+	0.42		
	(0.49)	(0.45)		
Cut4 Constant	4.19***	3.50***		
	(0.55)	(0.50)		
Constant			-0.38	-0.28
			(1.05)	(0.90)
N	888	979	149	177
Pseudo R^2	0.082	0.083	0.188	0.195

Notes: Standard errors in parentheses: + p<0.10, * p<0.05, ** p<0.01, *** p<0.001. Dummy variables for world war years are included in Models 2 and 4 but not reported here.

H2b is only partly supported. It is clearly seen in Table 4 and Table 5 that no marginal effects are significant when the initiator is a non-democracy while the target is a democracy, meaning we cannot reject the null hypothesis that when the initiator is a non-democratic state, initiator's capability share has no influence on the democratic target's probability of escalation at all stages, after controlling for other variables. The results do not support H2b's prediction that *ceteris paribus*, when the initiator is a non-democratic state, increasing observable balance of capabilities in favor of the initiator will decrease a democratic target's probability of reciprocating the crisis[33]. In addition, the direction of change stands in opposite to "common wisdom" consistently, though the marginal effects are not significant. Instead of decreasing the target's probability of escalation, increasing balance of capabilities in favor of the non-democratic initiator actually decreases the democratic target's probability of escalating to higher levels, as indicated by negative marginal effect at no militarized actions and positive marginal effects at all higher levels of escalation.

[33] Reciprocation means the target taking some sort of militarized action, i.e. the target escalates beyond "no militarized actions".

Table 4. Marginal effect of initiator capability share on probability of target escalation given different regime types (world wars excluded)

	Marginal effect of initiator capability share on probability of				
	No militarized action	Threat of force	Display of force	Use of force	War
Dem Initiator # Dem Target	3.25%	-0.15%	-0.91%	-2.07%	-0.13%
Dem Initiator# Non-dem Target	11.06%	-0.11%	-1.50%	-8.56%	-0.91%
Non-dem Initiator# Dem Target	-12.06%	0.10%	1.60%	9.37%	0.99%
Non-dem Initiator# Non-dem Target	20.68%**	-0.24%+	-2.95%**	-15.82%**	-1.67%*

Notes: This table is generated on Model 1. + p<0.10, * p<0.05, ** p<0.01, *** p<0.001.

Table 5. Marginal effect of initiator capability share on probability of target escalation given different regime types (world wars included)

	Marginal effect of initiator capability share on probability of				
	No militarized action	Threat of force	Display of force	Use of force	War
Dem Initiator # Dem Target	15.32%	-0.54%	-3.36%	-10.42%	-1.01%
Dem Initiator# Non-dem Target	19.13%+	-0.33%	-2.89%+	-14.13%+	-1.79%
Non-dem Initiator# Dem Target	-16.73%	0.14%	2.02%	12.78%	1.77%
Non-dem Initiator# Non-dem Target	18.19%**	-0.29%*	-2.67%*	-13.50%**	-1.73%*

Notes: This table is generated based on Model 2. + p<0.10, * p<0.05, ** p<0.01, *** p<0.001.

Only H5 gains support from models using MCT data set. When the initiator is democratic, we cannot reject the null hypothesis that the initiator's share of military capability has no influence on the probability of compellent threat failure regardless of target's regime type, after controlling for other variables. However, in contrast to the predictions of H6a and H6b, when both crisis actors are non-democracies, we cannot reject the null hypothesis that initiator's capability share has no influence on probability

Table 6. Marginal effect of initiator capability share on probability of compellent threat failure given different regime types

	World wars excluded	World wars included
Dem Initiator # Dem Target	25.63%	50.90%
Dem Initiator# Non-dem Target	-2.49%	-15.61%
Non-dem Initiator# Dem Target	49.97%*	56.18%**
Non-dem Initiator# Non-dem Target	-7.82%	-6.17%

Notes: + p<0.10, * p<0.05, ** p<0.01, *** p<0.001

of compellent threat failure. What's worse, when the initiator is non-democratic and the target is democratic, increasing balance of capabilities in favor of the initiator increases rather than decreasing the probability that the compellent threats fail. The effects are both very large- *ceteris paribus*, 1-point-increase in the initiator's capability share increases the probability of failure by 49.97% (56.18% with world war cases), and significant (both $p < 0.05$). Also, opposite to the predicted direction, increasing balance of capabilities in favor of the initiator increases instead of decreasing the probability of compellent threat failure.

Now turn to the remaining hypotheses about the effect of target regime type on its level of escalation and compellent threat failure. For more intuitive understanding of the effects, I present the marginal change (with 95% CI) in target's probabilities of taking no militarized actions and escalation to use of force conditional on different initiator regime types and various level of balance of capabilities in Figure 1, 2 and 3[34].

H3 predicts the marginal effects of target's regime type on its probability of escalation to be indistinguishable from 0 across all degrees of balance of capabilities when the initiator is a non-democracy. This is not the case presented in the figures, whether world war cases are included or not. When the non-democratic initiator is extremely weaker in capability than its target, a democratic target is significantly more

[34] To conserve place, I only present effects at escalation level "no militarized actions" and "use of force", which are most salient and constantly meaningful. Others are also significant, but the marginal changes are not enough to cast large influence. For instance, when the initiator is a democracy and the initiator's capability share is 0.55, a democratic target is 3.88% less likely to escalate to display of force compared with a non-democratic target. The effect is significant at $p < 0.05$. However, compared with changes of 18.51% and 13.05% at "no militarized actions" and "use of force", 3.88% is relatively small.

likely to take no militarized actions and less likely to use force compared with its non-democratic counterpart. On the contrary, when the non-democratic initiator enjoys extreme capability advantage, a democratic target is significantly less likely to take no militarized actions and more likely to use force compared with a non-democratic target.

Models using MCT data set generate results consistent with models using MID data set. However, the effects hold firmly when a non-democratic initiator enjoys extreme advantage in observable capability, but become indistinguishable (except for model 4 with world wars, effects are significant at $p < 0.1$ conditional on capshare_a < 0.05) when the non-democratic initiator is extremely weaker. Ceteris paribus, compellent threats issued by extremely advantaged non-democratic initiators (capshare_a > 0.75) are on average 22.15% to 33.42% more likely to fail when the targets are democracies instead of non-democracies, and the effects are significant at $p < 0.05$.

More importantly, we can observe from the figures that marginal effects of target's regime types vary with initiator's capability share. When the non-democratic initiator is militarily inferior, a democratic target is less likely to escalate to higher level than a non-democratic target, and a non-democratic initiator's compellent threat against a democratic target is less likely to fail than against a non-democratic target. Being quite opposite, when the non-democratic initiator is militarily superior, a democratic target is more likely to escalate than a non-democratic target, and the initiator's compellent threat against a democratic target is more likely to end up with failure than against a non-democratic target.

Now turn to the final pair of hypotheses. Consistent with H4, Figure 1 and 2 clearly show that when the initiator is a democratic state, the marginal effects of target regime type are statistically indistinguishable from 0 at low values of initiator's capability share. As predicted, when crisis actors are comparable in observable military capabilities, the effects are most salient. When the initiator is a democracy and the crisis actors are comparable (capshare_a between 0.35 and 0.75), a democratic target is on average16.94% to 20.08% more likely to take no militarized actions than a non-democratic target, and 11.76% to 14.38% less likely to use force, after controlling for other variables, and the effects are significant at $p < 0.05$. However, the effects differ between the model with world war cases and that without when the initiator's capability share is high. When

world wars are excluded, in accordance with H4, the effects are indistinguishable from 0; on the contrary, when world wars are included, the effects are significant at $p < 0.05$.

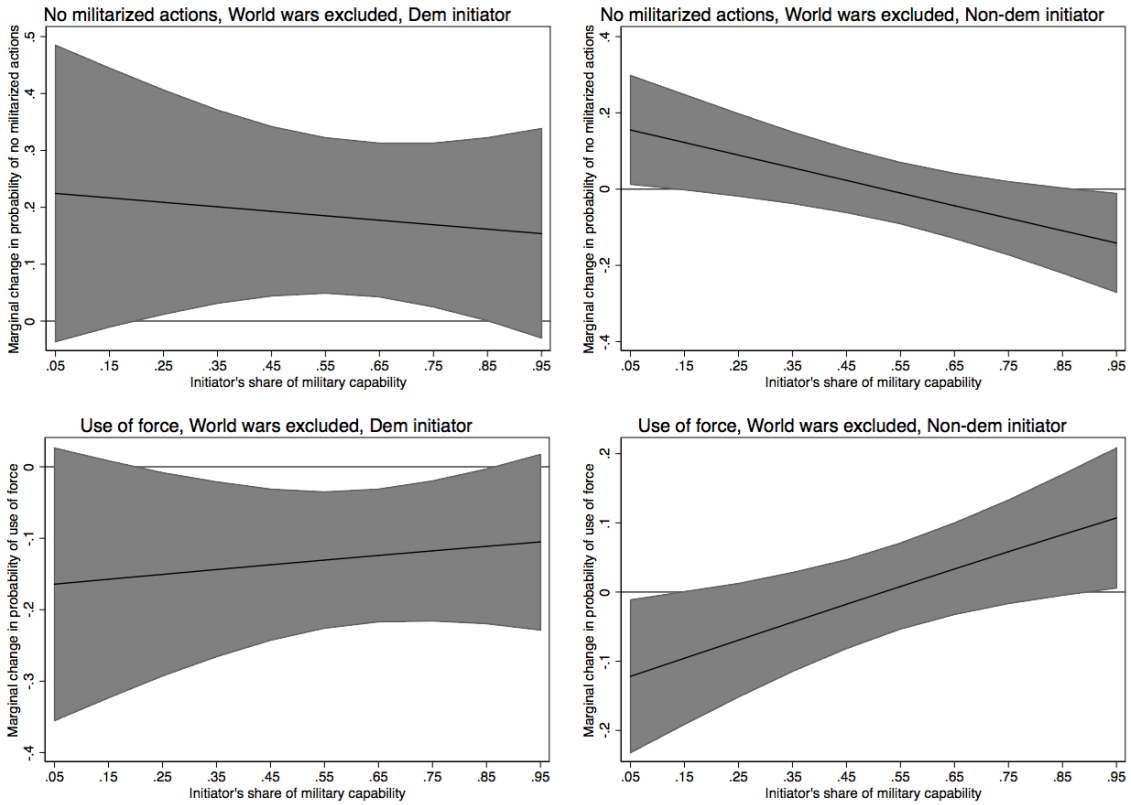

Figure 1. Marginal effect of target regime on probability of escalation, world wars

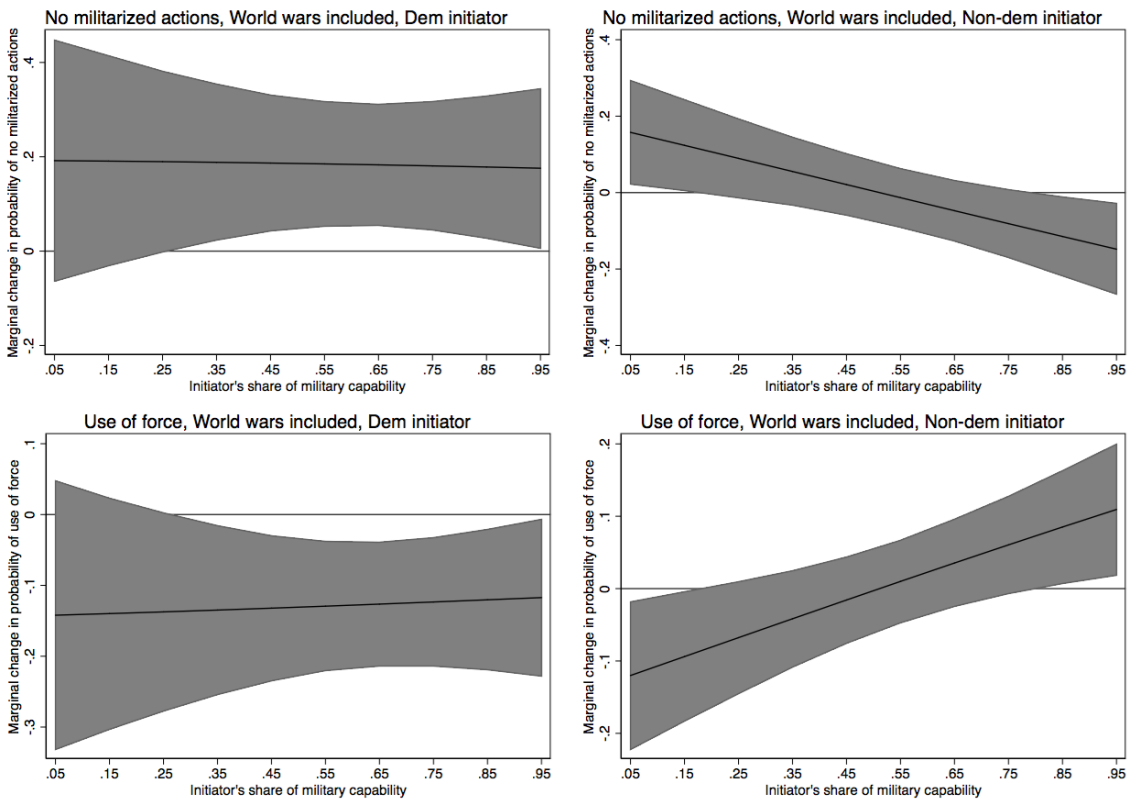

Figure 2. Marginal effect of target regime on probability of escalation, world wars included

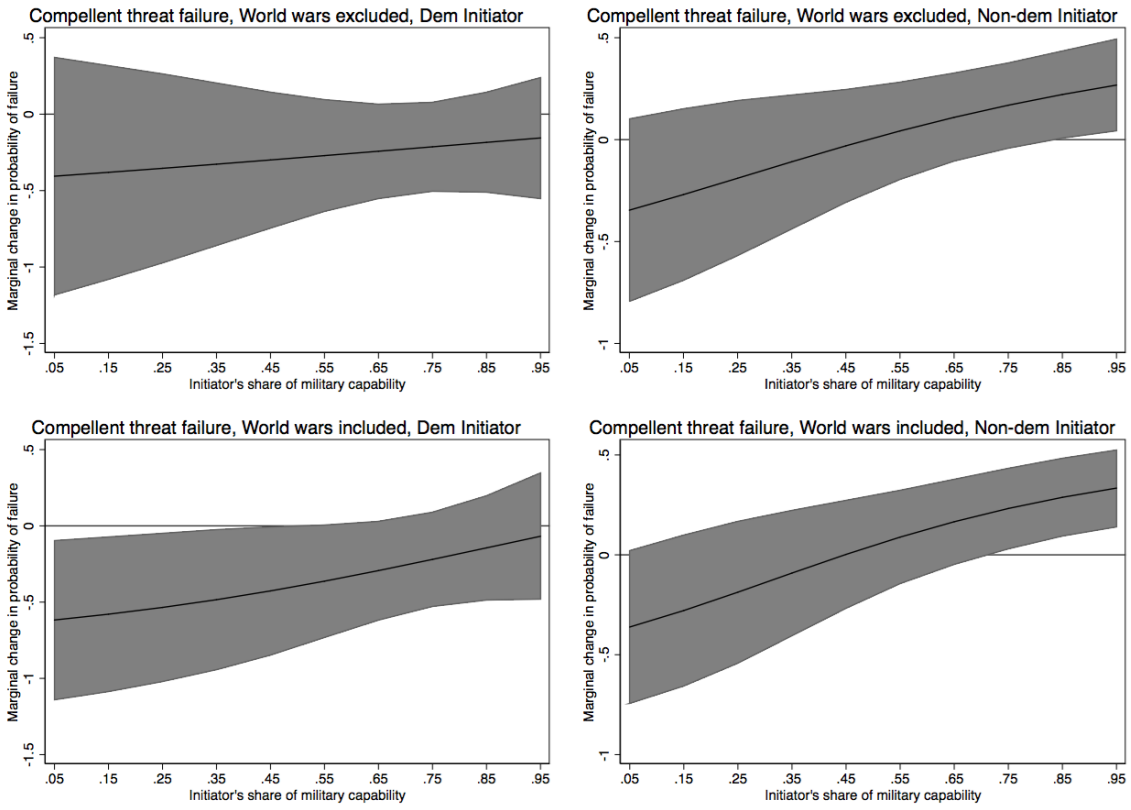

Figure 3. Marginal effect of target regime on probability of compellent threat failure

H8 is only marginally supported. When world war cases are excluded and when the initiator is a democracy, we cannot reject the null hypothesis that regime type of the target has no influence on compellent threat failure across all degrees of balance of capabilities. When world war cases are included, the compellent threat issued by a democratic initiator is significantly less likely to fail when the target is democratic rather than non-democratic, after controlling for other variables. The absolute magnitude of marginal reduction in probability of failure decreases as initiator's capability share increases, ranging from 61.73% (capshare_a=0.05) to 29.33% (capshare_a=0.65), with $p < 0.05$ for militarily disadvantaged initiator (capshare_a < 0.45) and $p < 0.1$ for militarily comparable crisis actors (capshare_a between 0.45 and 0.65). When the initiator enjoys advantages in capability, the effects become smaller and indistinguishable from 0. In addition, the direction of the marginal change, in agreement with models using MID data set, supports my theory. Democratic targets are more likely to take no militarized actions

and less likely to use force than their non-democratic counterparts when the initiators are also democracies. Therefore, compellent threats issued by democratic states are expected to be less likely to fail when the targets are also democracies instead of non-democracies.

Robustness

To check the robustness of analyses above, I replace Schultz's coding of democracy with the standard polity method of democracy coding, which identifies states with 6 or higher on polity score as democracies. To conserve space, the results are briefly summarized here.

All results are similar to those in Figure 1 and 2, except for marginal effects of target regime on probability of escalation when the initiator is a democratic state. In models using standard coding of democracy, the marginal effects are indistinguishable from 0 across all degrees of balance of capabilities. This might suggest that stable political competition and domestic constraints on executive power are better indicators of crisis behaviors than other aspects of democracy.

Discussion

The empirical analyses above confirm the underlying mechanism of audience costs theory. When the initiator is a democracy, we cannot reject the null hypothesis that initiator's capability share has no influence on target's probability of escalation and the probability of compellent threat failure at significance level $p < 0.05$. The effects hold in all models, regardless of what the target's regime type is, which data set is used, whether world war cases are included, and in which way democracy is coded. Democracies do select themselves into crises based on observable measures of balance of capabilities so that the influence of military capability is neutralized or at least mitigated in further escalation.

One side of the selection effect is confirmed; the other half, however, is different from theoretical expectation. My theory argues that given challenged, a democratic target also selects itself into crises it is more likely to prevail according to observable balance of capabilities. Therefore, initiator's capability share should influence a democratic target's choice of whether to reciprocate or not. Given reciprocated, the target's further escalatory

choice is no longer influenced by observable balance of capabilities. Empirical results turn out to be different. In models using MID data set, the effects are not distinguishable from 0 across all degrees of balance of capabilities. This points to the inadequacy of the theoretical argument. The theoretical prediction is based on the assumption that immediate concession given challenged carries no extra costs for the target except for the value in dispute. It is true that the leader does not suffer from "traditional" audience costs, which only have to be paid when the leader backs down from previous commitment, if he has not made any commitment yet. However, this does not mean the leader does not suffer from any costs at all. It is possible that a democratic leader would also be held accountable domestically for doing nothing about the crisis and making downright concessions immediately. Also, since democratic targets are significantly more likely to escalate in face of a extremely powerful non-democratic initiator than non-democratic targets, the effects of balance of capabilities on probability of escalation might be offset.

More severe challenge to the theory comes from models using MCT data set. In contrast to H5, we cannot reject the null hypothesis that balance of capabilities has no influence on the probability of compellent threat failure when both crisis actors are non-democracies. What's worse, exactly the opposite of H6, increasing balance of capabilities in favor of the non-democratic initiator decreases the probability of compellent threat failure in crises against democratic targets. By introducing reputational dynamics into international crisis bargaining process, Sechser argues that in a world of repeated interaction, exceptional military advantage exacerbates the informational asymmetry that prevents the initiator from accurately locating the target's range of acceptable deal and therefore renders the compellent threat more likely to fail.[35] The result is most powerful initiators issue least credible compellent threats, which is consistent with the empirical results here.

This suggests the necessity of differentiating deterrent threats from compellent threats in study of coercion. Two kinds of coercive threats might go through largely different mechanisms and are most influenced by different factors. It is difficult to tell in theoretical models that which mechanism plays a more important role. I only explore the audience costs theory and generate testable hypotheses based on it while discard other

[35] Sechser, 2010.

mechanisms that might also be influential. Only through differential empirical tests can we be able to find out which mechanism applies under different situations.

It is also worth noticing that effects of regime type vary dramatically across different values of initiator's capability share, especially when the initiator is non-democratic. Not only magnitude, but also direction of the effects changes. As presented previously, when the non-democratic initiator is superior in military capability, target being a democracy increases the probability of escalation and decreases the probability of compellent threat failure. However, when the non-democratic initiator is militarily inferior, target being a democracy decreases the probability of escalation. Previous studies grouped different conditions together so that the effects might be neutralized. It further manifests the need to explore interactive effects of key casual variables that influence the international crisis bargaining dynamics.

The marginal effects of target's regime type on probability of escalation are significant when one of the crisis actor enjoys huge military advantage and show opposite directions might be explained by states' strategic choice as to when to evoke audience costs. A democratic leader is more likely to "go public" on a crisis when he expects to achieve a quick and easy victory[36]. Higher audience costs carried by each escalatory steps enable the democratic target to credibly reveal its preference with lower level of escalation[37]. This corresponds to the situation where target enjoys huge military advantages. Also, democratic leaders are driven by domestic constraints to take some actions in response in the face of extremely powerful non-democratic initiators while non-democratic targets are more likely to give in. This corresponds to the situation where the initiator enjoys huge military advantages.[38]

In addition, though not all in support of my theoretical predictions, the effects of target regime type on probability of escalation using MID data set are basically in accordance with those on probability of compellent threat failure using MCT data set. This indicates that the MID data set can still be useful if we choose the right appropriate variables.

[36] See Baum 2004.
[37] See Fearon 1994a.
[38] To be clear, this is a tentative explanation based on extant literature. A more coherent theoretical framework is needed to provide satisfying and inclusive explanations for effects found in empirical tests.

More importantly, the empirical tests using MID data set in this article all reports the largest and most consistent effects at escalation level "no militarized actions" and "use of force". This indicates the choice of reciprocation (escalation beyond no militarized actions) and violent reciprocation (escalation to or beyond use of force) as dependent variables is reasonable and valid.[39] It might be more appropriate to do so instead of using an ordinal variable with multiple values, since it can avoid complicated models and excessive work of result presentation and explanations while at the same time preserve meaningful effects to a large extent.

CONCLUSION

This article empirically tests the important mechanism that contributes to democratic credibility in international crisis bargaining. Democratic leaders, in the face of larger domestic punishment if caught bluffing in crisis bargaining, are more cautious in choosing which crises to contend on, and select themselves into crises they are more likely to prevail. This leads to certain testable observations that find strong support in this article. We cannot reject the null hypothesis that observable balance of capabilities, operationalized as the initiator's capability share, has no influence on the target's probability of escalation in crises initiated by democratic states. On the contrary, in crises initiated by non-democratic states, observable balance of capabilities has significantly effect on the target's probability of escalation, though the direction and degree of effect depend on universe of cases investigated.

This article also explores the interactive effects of regime type and balance of capabilities on probability of escalation. The effects of domestic audience costs in influencing international crisis bargaining are context dependent and contingent on other variables. Specifically, observable balance of capabilities shapes the crisis actors' *ex ante* beliefs about each other's preference, provides the conditions that influence the leaders' strategic choice as whether to evoke audience costs or not, and defines the range in which the effect of audience costs is more salient.

[39] Schultz 1998 first used reciprocation as dependent variable in testing democratic credibility. Many other researchers in literature of audience costs adopted this method. Potter and Baum 2014 added violent reciprocation as the dependent variable in their models, and the results turn to be significant.

My analysis reveals a complex scenario of how regime type influences the crisis actors' behaviors. In crises initiated by democratic states, democratic targets are significantly less likely to escalate to higher level of hostility compared with non-democratic targets when states involved are militarily comparable. In crises initiated by non-democratic states, compared with non-democratic targets, democratic targets are significantly less likely to escalate to higher level of hostility when the initiators are militarily inferior; however, when the initiators are militarily superior, democratic targets are significantly more likely to escalate to higher level of hostility than their non-democratic counterparts.

My findings also address the question as whether audience costs have meaningful effect on coercive diplomacy. As shown in the results, my findings are both consistent and meaningful when they are statistically significant. At escalation level "no militarized actions" and "use of force", difference in regime type amounts to (on average) 10% to 33% change in predicted probability of escalating to certain level. In models using MCT data set, difference in regime type indicates (on average) 30% to 60% difference in predicted probability of compellent threat failure. Compared with the effects of initiator's capability share, the effects of regime type are fairly large and politically consequential.[40]

The complicated non-linear and context dependent effects of audience costs give some initial support for my previous argument that rather than an invariable attribute, audience costs are more of a strategic choice that provides a possible foreign policy tool for those states "able" to generate large audience costs. By exploring how leaders strategically choose to evoke audience costs and how different strategies influence the crisis outcome, we can better understand the dynamics of international crisis bargaining.

Finally, sharing a common inadequacy with literature of democratic credibility theories, my study does no capture an important component of crisis bargaining in the sense that most crises in history end up with compromise, peaceful resolution or stalemate rather than unconditional concession of one side. One reason that *ex post*

[40] As previously discussed, for instance, when the initiator is a non-democratic state, 1-point-increase in the initiator's capability share increases the probability that the target will take no military actions by 20.68% (p=0.004) on average. It should be noted that 1-point-increase in capability share amounts to increase from 0 to 1, which is empirically impossible. To put it more realistically, if the initiator's capability share increase from 0.5 to 0.6, the target's probability of taking no militarized actions is only expected to increase by 2.07% on average, certainly not as large as the effect of regime type.

inefficient wars happen is that states have incentives to misrepresent information in crisis bargaining so as to gain more leverage in division of the disputed goods (Fearon 1995). If escalatory steps in crisis bargaining can credibly signal crisis actor's preference, and therefore to some extent clarify each other's uncertainty about the final result if war ensued, escalation should serve an important role in helping crisis actors find each other's negotiation range, and increase the possibility of reaching a peaceful agreement. Future research in international crisis bargain would benefit in exploring how public contest of resolve by coercive threats influences the contents of negotiated agreements.

REFERENCE

Baum, M. 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.

Bennett, D. S., et al. (2002). Expected Utility Generation and Data Management Program, Version.

Brown, J. N. and A. S. Marcum (2011). "Avoiding Audience Costs: Domestic Political Accountability and Concessions in Crisis Diplomacy." Security Studies **20**(2): 141-170.

Clare, J. (2007). "Domestic audiences and strategic interests." Journal of Politics **69**(3): 732-745.

Downes, A. B. and T. S. Sechser (2012). "The illusion of democratic credibility." International organization **66**(03): 457-489.

Fearon, J. D. (1994a). "Domestic political audiences and the escalation of international disputes." American Political Science Review **88**(03): 577-592.

Fearon, J. D. (1994b). "Signaling Versus the Balance of Power and Interests An Empirical Test of a Crisis Bargaining Model." Journal of Conflict Resolution **38**(2): 236-269.

Fearon, J. D. (1995). "Rationalist explanations for war." International organization **49**(03): 379-414.

Fearon, J. D. (1997). "Signaling foreign policy interests tying hands versus sinking costs." Journal of Conflict Resolution **41**(1): 68-90.

Finel, B. I. and K. M. Lord (1999). "The surprising logic of transparency." International Studies Quarterly **43**(2): 325-339.

Gelpi, C. and J. M. Grieco (2001). "Attracting Trouble: Democracy, Leadership Tenure, and the Targeting of Militarized Challenges, 1918-1992." Journal of Conflict Resolution **45**(6): 794-817.

Gibler, D. M. and M. L. Hutchison (2013). "Territorial Issues, Audience Costs, and the Democratic Peace: The Importance of Issue Salience." The Journal of Politics **75**(04): 879-893.

Guisinger, A. and A. Smith (2002). "Honest Threats The Interaction of Reputation and Political Institutions in International Crises." Journal of Conflict Resolution **46**(2): 175-200.

Haynes, K. (2012). "Lame Ducks and Coercive Diplomacy: Do Executive Term Limits Reduce the Effectiveness of Democratic Threats?" Journal of Conflict Resolution **56**(5): 771-798.

Heckman, J. J. (1979). "Sample selection bias as a specification error." Econometrica: Journal of the econometric society: 153-161.

Holsti, O. R. (2004). "Public Opinion and American Foreign Policy, Revised Edition."

Jervis, R. (1976). Perception and misperception in international politics, Princeton University Press.

King, G., et al. (1994). Designing social inquiry: Scientific inference in qualitative research, Princeton University Press.

Marshall, M. G., et al. (2010). "Polity IV Data Series Version 2010." College Park, MD: University of Maryland. Retrieved from http://www. systemicpeace. org/polity/polity4. htm.

McManus, R. W. (2014a). "Fighting words The effectiveness of statements of resolve in international conflict." Journal of Peace Research **51**(6): 726-740.

McManus, R. W. (2014b). The role of statements of resolve in international conflict, THE UNIVERSITY OF WISCONSIN-MADISON.

Mercer, J. (2012). "Audience Costs Are Toys." Security Studies **21**(3): 398-404.

Palmer, G., et al. (2015). "The MID4 dataset, 2002–2010: Procedures, coding rules and description." Conflict Management and Peace Science: 0738894214559680.

Partell, P. J. (1997). "Executive constraints and success in international crises." Political Research Quarterly **50**(3): 503-528.

Partell, P. J. and G. Palmer (1999). "Audience costs and interstate crises: An empirical assessment of Fearon's model of dispute outcomes." International Studies Quarterly **43**(2): 389-405.

Potter, P. B. K. and M. A. Baum (2010). "Democratic Peace, Domestic Audience Costs, and Political Communication." Political Communication **27**(4): 453-470.

Potter, P. B. K. and M. A. Baum (2013). "Looking for Audience Costs in all the Wrong Places: Electoral Institutions, Media Access, and Democratic Constraint." The Journal of Politics **76**(01): 167-181.

Prins, B. C. (2003). "Institutional instability and the credibility of audience costs: Political participation and interstate crisis bargaining, 1816-1992." Journal of Peace Research **40**(1): 67-84.

Ramsay, K. W. (2004). "Politics at the Water's Edge: Crisis Bargaining and Electoral Competition." The Journal of Conflict Resolution **48**(4): 459-486.

Rioux, J.-S. (1998). "A crisis-based evaluation of the democratic peace proposition." Canadian Journal of Political Science **31**: 263-284.

Schultz, K. A. (1998). "Domestic opposition and signaling in international crises." American Political Science Review **92**(04): 829-844.

Schultz, K. A. (1999). "Do Democratic Institutions Constrain or Inform? Contrasting Two Institutional Perspectives on Democracy and War." International organization **53**(2): 233-266.

Schultz, K. A. (2001a). Democracy and coercive diplomacy, Cambridge University Press.

Schultz, K. A. (2001b). "Looking for audience costs." Journal of Conflict Resolution **45**(1): 32-60.

Sechser, T. S. (2010). "Goliath's Curse: Coercive Threats and Asymmetric Power." International organization **64**(04): 627-660.

Sechser, T. S. (2011). "Militarized compellent threats, 1918–2001." Conflict Management and Peace Science **28**(4): 377-401.

Simonelli, N. M. (2011). "Bargaining over International Multilateral Agreements: The Duration of Negotiations." International Interactions **37**(2): 147-169.

Singer, J. D. and M. Small (1993). National material capabilities dataset, Study.

Slantchev, B. L. (2006). "Politicians, the media, and domestic audience costs." International Studies Quarterly **50**(2): 445-477.

Smith, A. (1998). "International crises and domestic politics." American Political Science Review **92**(03): 623-638.

Snyder, J. and E. D. Borghard (2011). "The Cost of Empty Threats: A Penny, Not a Pound." American Political Science Review **105**(03): 437-456.

Taylor, L. W. (2007). "Nonparametric Estimation of Duration Dependence in Militarized Interstate Disputes." Journal of Applied Statistics **34**(4): 423-441.

Trachtenberg, M. (2012). "Audience Costs: An Historical Analysis." Security Studies **21**(1): 3-42.

Weeks, J. L. (2008). "Autocratic audience costs: Regime type and signaling resolve." International organization **62**(01): 35-64.

Wilkenfeld, J., et al. (2003). "Mediating International Crises: Cross-National and Experimental Perspectives." Journal of Conflict Resolution **47**(3): 279-301.