

# **Backing Up, Not Backing Down: Mitigating Audience Costs Through Policy Substitution**

(9,695 words excluding title page)

5 June 2018

Erik Lin-Greenberg<sup>1</sup>  
Department of Political Science  
Columbia University  
el2723@columbia.edu

## **ABSTRACT**

Can a leader reduce the audience costs imposed for backing down completely on a threat by opting instead to “back up” to a less hawkish policy? Current research examines the political repercussions of making a threat and then taking no action at all. Real world leaders, however, often “back up” and implement policies that involve some action – for instance, imposing sanctions after threatening military force, rather than backing down entirely. This article proposes a theory of audience cost mitigation through policy substitution: backing up to less hawkish policies – that reduce inconsistency between a leader’s words and deeds – may reduce audience costs. A series of original survey experiment finds support for the theory and demonstrates the population treats inconsistency as a continuum. The findings have implications for domestic politics and crisis bargaining. A leader who backs up is seen as more competent than one who backs down, and foreign audiences view the leader’s future threats and promises are seen as more credible. Backing up can subsequently degrade the credibility of crisis signaling by making it difficult for rivals to discern between credible threats and those that will be backed up on.

---

<sup>1</sup> This project was approved by the Columbia University Institutional Review Board for Human Subjects Research (Protocol IRB-AAAQ5010). I am grateful to Richard Betts, Brian Blakenship, Kolby Hanson, Robert Jervis, Roseanne McManus, Sara Moller, Theo Milonopoulos, Constantino Pischedda, Tonya Putnam, Tara Slough, Jordan Sorensen, Jack Snyder, and participants at the 2017 International Studies Association Conference and the Columbia Security Studies Working Group for their valuable comments and assistance. All errors are my own.

In 2012, President Obama declared a red line, threatening military action if the Assad regime used chemical weapons. When Syrian forces launched a sarin attack in 2013, the Obama administration instead invoked a peaceful United Nations-mediated solution. Critics claimed Obama had sullied America's reputation, while supporters believed he had achieved his policy objectives. Can leaders reduce audience costs incurred for backing down entirely from threats by implementing less hawkish alternatives, like those taken by Obama in 2013?

Much research suggests that governments can credibly signal their intentions during interstate crises by generating domestic audience costs. Since domestic constituents can punish their leaders for being inconsistent, leaders have incentives to avoid bluffing and issue only threats they intend to carry out, making threats credible (Fearon 1994; Schultz 1998; Weeks 2008). Empirical studies on audience costs, however, have unrealistically simplified the inconsistency between threats and actions, typically treating whether a leader follows through as dichotomous (Tomz 2007; Levendusky and Horowitz 2012; Kertzer and Brutger 2016). In other words, leaders either carry out their threat completely or back down by taking no action at all. This simplification is problematic. In practice, there can be significant variation in the degree of inconsistency between a leader's words and deeds, something neglected in existing studies. Some leaders follow through with their threats, while others "back up" to a range of less escalatory actions. If leaders face less punishment for backing up than for backing down, they might renege on threats – albeit incrementally – more than existing theories predict. Backing up can subsequently degrade the credibility of crisis signaling by making it difficult for rivals to discern between credible threats and those that will be backed up on.

This article introduces and tests a theory that suggests leaders face lower audience costs when they minimize the perceived inconsistency between their words and deeds. They do this by

“backing up,” which I define as making a threat then subsequently taking a less escalatory action – for either pre-meditated or post-hoc reasons – instead of doing nothing at all. Does backing up instead of backing down reduce audience costs? If so, why does the public deal less harshly with leaders who back up instead of back down? Most importantly, what does this mean for the credibility of signaling during crises?

To explore these questions, I synthesize the disparate literatures on audience costs, consistency, and policy substitution, and develop a theory on the political repercussions of partially reneging on threats. I test this theory using original survey experiments that expose respondents to a hypothetical invasion of an ally. I manipulate whether the President threatens to deploy ground forces to defend the ally and whether the President follows through on his threat, backs up to airstrikes or sanctions, or backs down altogether. I find that by backing up rather than backing down, leaders suffer lower audience costs. Analysis of qualitative data from the survey suggests that much of this reduction can be attributed to respondents’ beliefs that some form of action is more consistent and desirable than reneging entirely. The experiments also reveal that the future threats of a leader who backs up are seen as more credible than those of a leader who backs down.

This article makes three contributions to the study of audience costs. First, introducing the concept of backing up adds nuance to the inconsistency mechanism that underlies audience costs. Existing studies that treat threat fulfillment as binary overlook instances where leaders lessen audience costs by taking some form of action. Second, it confronts critiques that the public cares little about inconsistency (Snyder and Borghard 2011; Trachtenberg 2012). The experiments demonstrate the public recognizes variation in the *degree* of consistency between a president’s threats and actions, which can affect audience costs and perceptions of a leader’s

reputation and competence. Third, it empirically demonstrates that backing up along a continuum of consistency degrades the credibility of a leader's future threats among a target population, potentially affecting crisis bargaining.

## **MITIGATING AUDIENCE COSTS THROUGH POLICY SUBSTITUTION**

A large body of literature suggests domestic audiences punish their leaders for making a foreign policy threat and then reneging on it (Fearon 1994; Schultz 1998; Tomz 2007; Weeks 2008). The political costs for bluffing make leaders reluctant to issue idle threats (Fearon 1994), helping to differentiate costly (and credible) signals from less credible cheap talk (Jervis 1970; Fearon 1994). Recent scholarship has identified several factors that influence the consequences for reneging on threats. Audience costs can be shaped by the salience of the issue of stake (Clare 2007), the existence of an independent media or whistleblowers (Potter and Baum 2014), and by the leader revealing new information about the situation (Levendusky and Horowitz 2012). Existing research, however, has not examined whether the degree to which a leader follows through with a threat affects audience costs.

### ***Policy Substitution***

Before implementing a foreign policy, leaders generally consider a range of possible options, rather than developing just one course of action. Research on policy substitution suggests that several strategies can be used to address the same geopolitical issue (Most and Starr 1984; Palmer and Bhandari 2000; Clark and Reed 2005). For instance, when deciding how to counter a rival, a leader might choose between war (Gilpin 1981), strengthening alliances (Walt 1990), levying sanctions, or bolstering military power (Waltz 1979). The ultimate decision is informed in large part by the leader's assessment of the costs and benefits for each possible

strategy, taking into account a host of domestic and international factors (Clark, Nordstrom, and Reed 2008).

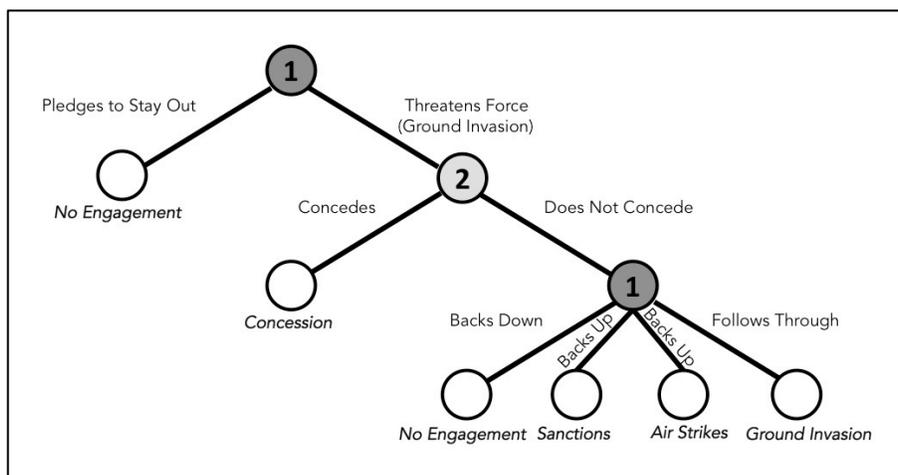
Just as a leader considers multiple strategies before enacting a foreign policy, a leader can also back up to an alternate policy after threatening another strategy. Existing audience cost research has touched upon policy substitution, but does so by varying the initial threat or demonstration of force rather than analyzing different degrees of “backing up” once a threat has been made (Tomz 2007). Critics might argue that backing up is unlikely because leaders avoid making unambiguous threats that commit them to specific actions (G. H. Snyder and Diesing 1978; J. Snyder and Borghard 2011). Even ambiguous threats about “red lines” and “grave consequences,” however, can imply certain types of actions, from which leaders can back up.

Backing up can be either premeditated or post-hoc. When premeditated, a leader may issue a threat to obtain concessions from a rival or to satisfy domestic constituents, knowing that she will later back up to another policy. In post-hoc cases, a leader may make a threat then back up after receiving new information about the situation (Levendusky and Horowitz 2012) or if an alternate policy is discovered to entail lower risk or cost. To be sure, a leader who backs up still faces political costs. But the costs of backing up may be lower than those of backing down since the former reduces the inconsistency between a leader’s words and deeds. In other cases, leaders may prefer backing up to following through on a threat, especially if carrying out the threat generates high political or financial costs. For instance, limited airstrikes might be cheaper than a ground invasion in both blood and fortune, making backing up an attractive alternative.

I illustrate the crisis bargaining model with policy substitution as a stylized game in Figure 1. The players are the same as in the classic crisis bargaining model, as are the strategies each player can adopt in their first move. However, if State 2 does not concede to State 1’s

threat, the policy substitution variant of the model expands the strategy set for State 1’s second move. Rather than choosing only between “Backs Down” and “Follows Through,” State 1 now has additional “Backs Up” options: airstrikes and sanctions. In an actual crisis, the strategy set might include many more backed up means that vary in escalation. Indeed, one frequently cited dataset identifies 22 distinct levels of militarized action a state can take during crises (Jones, Bremer, and Singer 1996). Further, the initial threat need not be a ground invasion. A leader might instead threaten sanctions or some other action and then follow through, back up, back down, or back in (i.e. escalate).

**FIGURE 1: Crisis Bargaining Model with Policy Substitution**



This model has played out during recent crises. As described earlier, President Obama publicly warned Bashar al-Assad in August 2012 that the use of chemical weapons would be “a red line” that would change Washington’s calculus on “military engagement.”<sup>2</sup> Even if President Obama did not intend his comment to be an explicit coercive threat, many of his closest advisors, the media, and segments of the American public viewed it as such. When Syrian forces launched

<sup>2</sup> President Obama’s “red line” remains a topic of debate. For more on this debate see, Glenn Kessler, “President Obama and the ‘red line’ on Syria’s chemical weapons,” *The Washington Post*, 6 September 2013.

sarin attacks on opposition-held territory a year later, many believed this crossing of the red line would lead Obama to use force. Indeed, ten days after the attack President Obama announced, “After careful deliberation, I have decided the United States should take military action against Syrian regime targets (Obama 2013).” Actors around the world viewed this threat as credible: the Pentagon prepared for military strikes (Chollet 2016, 8–26), news outlets frequently referenced the president’s red line, and Syrian military forces braced for a U.S. attack (Baker and Weisman 2013).

Military action, however, never materialized as President Obama decided to “explore another option” in the face of public opposition to military operations and a lack of support from Congress and close allies (Chollet 2016, 5). Instead of fulfilling his threat, Obama arguably backed up to a multinational agreement under which the Syrian regime turned over its chemical weapons for destruction (Chollet 2016, 8–26). Obama’s actions align with the concept of post-hoc backing up. After issuing a threat of military action, President Obama faced opposition to following through on his threat. Given the limited domestic and international support, carrying out strikes could have generated high political costs at home and abroad. Doing nothing at all, however, could have led to negative political consequences. Rather than backing down entirely or following through with an attack, Obama “backed up” to an intermediate policy.

American presidents are not the only ones who back up from threats. In December 2001, militants attacked the Indian Parliament. The following day, Indian Prime Minister Atal Bihari Vajpayee publicly accused Pakistan of supporting the attacks and called for a “war to the finish” to halt Pakistani-backed terrorism (Sood and Sawhney 2003, 94). India also demanded that Pakistan stop cross-border attacks, extradite suspected terrorist leaders, and block terrorist finances. Days after issuing his threat of war, Vajpayee launched Operation Parakram,

dispatching 800,000 troops along the border in New Delhi's largest mobilization since the 1971 India-Pakistan War. Although Pakistan took limited steps to outlaw certain terrorist groups and signaled a willingness to act against some terrorist leaders, the actions were not enough to stop additional terrorist attacks or India's mobilization (Sood and Sawhney 2003, 95).

In May 2002, terrorists attacked four Indian military outposts, killing dozens of soldiers and their families. India again signaled impending war by recalling soldiers from leave and issuing new warnings. The Army chief announced "the time for action had come" while Prime Minister Vajpayee told frontline soldiers to "prepare for a decisive battle (Sood and Sawhney 2003, 96–97)." Despite these threats, the war did not occur, even though Pakistan never met India's demands. Instead of war, Vajpayee arguably backed up on his initial threat by taking other military moves short of a ground invasion. In addition to the mobilization, India conducted a highly publicized test of a ballistic missile capable of striking Pakistan at the height of the crisis (Dugger 2002).

Although several factors may have driven Vajpayee to back up, scholars suggest he was under heavy pressure from the United States, which feared conflict between two nuclear-armed actors (Sechser and Fuhrmann 2017, 155–61). When Vajpayee finally demobilized his forces in October 2002, he was criticized for backing up on his threat of war. Backing down entirely or following through, however, could have entailed greater financial and political costs. Had Indian forces invaded, Pakistan was postured to launch conventional operations and irregular warfare in Kashmir, which could have mired India in an expensive and bloody counterinsurgency campaign (Sood and Sawhney 2003, 174). By considering a continuum of "backed up" options like those in Syria and India, this project captures real world dynamics that current research neglects.

### ***Inconsistency: A Continuum***

Much research suggests audience costs stem from inconsistency between a leader's threats and actions, but the literature has adopted an overly simplistic view of inconsistency, treating it as a binary variable. Tomz (2007, 835), for instance, defines inconsistency as "saying one thing and doing another." This conceptualization overlooks gradations in "doing another" that potentially have political consequences. To better capture this continuum, I argue that during crises leaders can substitute a range of policies that vary along dimensions like financial cost, operational risk, and degree of escalation. Scholars have used these characteristics to categorize crisis actions along ordinal scales that range from low intensity moves like diplomatic demarches to full-scale war (Jones, Bremer, and Singer 1996; Brecher et al. 2017). The public should perceive actions that are closer together on these continuums as more consistent with each other than actions that are further apart. For instance, airstrikes should be considered more consistent with a threat of invasion than a diplomatic demarche as the costs, risks, and intensity of airstrikes and invasions are more similarly related.

Past research suggests individuals generally favor consistency over inconsistency and do so for both psychological and instrumental reasons. From a psychological perspective, individuals experience cognitive inconsistency and dissonance when a set of cognitive elements – pieces of information, beliefs, and opinions – diverge. Divergence may stem from a variety of factors including a diminishing correspondence between an individual's fundamental beliefs and a policy option (Converse 1964), or a decreasing similarity between one's preference and an alternative option (Brehm and Cohen 1959). This divergence of cognitive elements occurs along a continuous spectrum in which increased dissimilarity between one element and another increases dissonance (Brehm and Cohen 1962, 11–17). To reduce this dissonance, individuals

typically prefer alternative policies that minimize divergence from the policies they initially hoped to obtain. Indeed, experimental studies have unsurprisingly shown that people prefer alternatives that are qualitatively similar to their most favored option (Brehm and Cohen 1959).

Actors may also prefer consistency for more rationalist reasons. Individuals may have a certain ideal point for a leader to fulfill her promises. The less deviation between this ideal point and the actions a leader eventually takes, the more public support the action will have. If the public believes their state suffers reputational consequences when a leader is inconsistent, they may fear a larger deviation between a leader's words and deeds will result in larger adverse consequences. As a result, the public may prefer an action that is qualitatively similar in terms of perceived risks, costs, and military intensity to an initially promised action over a qualitatively dissimilar alternative.

### **How Backing Up Shapes Audience Costs and Crisis Signaling**

Regardless of whether distaste for inconsistency is grounded in psychological or rationalist logics, the preference for consistency suggests leaders should face lower audience costs when there is less divergence between their words and deeds. By backing up to actions that are qualitatively similar in risk, cost, or military intensity to the initially threatened policy, instead of backing down altogether, the leader potentially reduces perceived inconsistency and the associated audience costs. For instance, after a leader pledges an invasion, implementing a less escalatory military action might be viewed as more consistent than taking only diplomatic measures or no action at all. This leads to two testable hypotheses on inconsistency and audience costs:

*H<sub>1</sub>: Domestic audiences will impose lower audience costs on a leader that backs up than on a leader who backs down.*

*H<sub>2</sub>: As the qualitative difference (in risk, cost, or military intensity) between the initial threat and the implemented policy decreases, so too will the perceived inconsistency of the leader's actions.*

Although the subject of debate (Mercer 1996; Huth 1997; Copeland 1997), scholars increasingly find that leaders suffer reputational costs when they make idle threats and promises (Brutger and Kertzer 2018). Because reputations are “belief[s] about a trait or behavioral tendency of an actor, based on that agent’s past behavior (Dafoe, Renshon, and Huth 2014, 375),” the public may hold leaders accountable for actions they believe degrade the state’s reputation. Reputations, however, are multi-dimensional, meaning that multiple intertwined traits contribute to an actor’s reputation in a given situation (Huth 1997; Brutger and Kertzer 2018). For crisis bargaining, two types of reputation are particularly important: a leader’s signaling reputation and her reputation for resolve. A signaling reputation indicates a leader fulfills her explicit threats and promises (Jervis 1970), while reputation for resolve means a leader has a tendency to stand firm in particular classes of disputes (Huth 1997; Dafoe, Renshon, and Huth 2014). Strong reputations in both traits allow a leader to more convincingly issue threats that compel or deter rivals.

While recent studies examine factors underlying reputation costs (Brutger and Kertzer 2018), they do not investigate whether the degree of inconsistency between a leader’s words and deeds matters. If statesmen are judged on how willing they are to carry out explicit threats and their steadfastness during disputes, the amount of inconsistency between a leader’s past threats and actions ought to shape her reputation. Backing down completely on a threat should have a significant negative effect on a leader’s signaling reputation and reputation for resolve as she has demonstrated that her threats cannot be trusted and that she is unwilling to stand firm. In

contrast, backing up might tarnish a leader's signaling reputation because she fails to fulfill explicit threats, but may preserve some reputation for resolve by demonstrating she is willing to do something, even if not the initially threatened action. In the minds of both domestic and international actors, the future threats issued by a leader who backs up might be viewed as more credible than those issued by a leader who previously backed down.

*H<sub>3</sub>: The public will view a leader who backs up as having a less adverse impact on the state's ability to issue credible threats and promises in the future.*

In addition to affecting a leader's reputation, inconsistency between a leader's words and deeds can lead the public to deem their leader as incompetent. Incompetence can have political consequences: incompetent leaders are voted out of office (Smith 1998) or face difficulty promoting domestic legislation (Gelpi and Grieco 2015). Competence, however, is not directly observable, so the public makes inferences about leader competence, in part, by examining past actions. How previous behavior affects judgments on competence is both theoretically and empirically controversial. According to one logic, leaders that fulfill their threats are seen as competent because they expect to fare well in conflicts, while incompetent leaders who renege on threats signal weakness (Smith 1998). More recent scholarship suggests leaders are judged for the costs and benefits of their actions, with more competent leaders producing better policy outcomes (J. Snyder and Borghard 2011; Debs and Weiss 2016).

The matter is far from settled and studies suggest the public views leader competence as multi-dimensional. Voters likely consider several factors, including inconsistency, when making inferences about leader competence. A leader who backs up and takes some action may be seen as signaling greater capability than a leader who reneges entirely on a threat. In some cases, a leader

may even be able to claim her backed up action is superior to the initially threatened option or backing down, particularly if she reveals new information that altered strategic conditions (Levendusky and Horowitz 2012). Indeed, in describing his response to Syrian chemical weapon use, President Obama said he was “very proud of this moment...ultimately it was the right decision to make (Goldberg 2016).”

*H<sub>4</sub>: The public will view a leader who backs up as more competent than a leader who backs down.*

## **METHODOLOGY**

I test these hypotheses using a series of original survey experiments. Because leaders seek to avoid political punishment, they strategically select out of issuing empty threats, making audience costs difficult to study with observational approaches (Tomz 2007). In recent years, experimental methods have become increasingly common in international relations research, particularly for studying subjects like audience costs where observational data are scarce (Hyde 2015). Although survey experiments allow researchers to precisely identify causal effects by randomizing and controlling the application of treatments (Gerber and Green 2012, 1–17), they do not overcome all research challenges. Most significantly, there is considerable debate surrounding external validity and whether findings gathered under tightly controlled experimental conditions mean anything in the real world (Hyde 2015). Some critics claim audience cost experiments overstate the backlash that leaders face for backing down (Clare 2007; Kurizaki and Whang 2015). Others suggest artificially constructed experimental scenarios unrealistically feature explicit “committing threats” that leaders rarely make (J. Snyder and Borghard 2011).

These critiques may be overstated. First, international relations experiments are grounded in the plausible assumption that experimental subjects apply conceptual and cognitive processes in a manner similar to non-experimental subjects in the real world (Schelling 1961, 55; Levendusky and Horowitz 2012, 328). Second, constituents may perceive even veiled threats as binding. Indeed, many Americans viewed Obama’s off-handed “red line” remark as a threat of force, and subsequently criticized him for backing up. Survey experiments therefore remain useful tools for studying international relations, a field where manipulating actual crises is unrealistic for obvious practical and ethical reasons.

***Experimental Design***

My experiment draws from the setup introduced in Tomz’s (2007) audience cost study, and exposes respondents to a scenario involving the U.S. president’s decision to employ military force. All respondents are told that, “A country sent its military to take over a neighboring country in order to get more power and resources. The leader of the state being invaded asked the United States for assistance.” I then randomly assign respondents to one of five experimental conditions that vary whether the president threatened to use force and whether he followed through on his initial threat (experimental design in Table 1).<sup>3</sup>

**TABLE 1: Experimental Design**

	<i>Follows Through</i>	<i>Backs-Up 1</i>	<i>Backs-Up 2</i>	<i>Backs Down</i>
<i>No Threat</i>	Stays Out <i>n</i> =208	n/a	n/a	n/a
<i>Threatens Invasion</i>	Invades <i>n</i> =198	Airstrikes <i>n</i> =208	Economic Sanctions <i>n</i> =195	No Military Action <i>n</i> =197

*Note:* The vertical axis represents whether the president initially threatens invasion. The horizontal axis captures the implemented action. *n* is the number of respondents in each condition.

<sup>3</sup> See Appendix C, Section 1 for balance tables and randomization checks.

In the control group the president pledges to stay out of the conflict and takes no action. In the other four experimental conditions, all respondents were told “the U.S. President said that if the attack continued, the U.S. military would use ground forces to push out the invaders.” I then vary whether the president follows through with this threat, backs up by substituting an intermediate policy of sanctions or airstrikes, or backs down entirely by taking no action.<sup>4</sup> Sanctions and airstrikes represent backed up options as both are commonly used tools of coercive diplomacy that are less escalatory than an invasion, but more hawkish than doing nothing at all.

To prevent priming, I avoid potentially loaded phrases such as the president “followed through with” (“backed down from”) his threat. The treatments also include no specific countries other than the United States to avoid influencing respondents’ opinions. To control for policy outcomes, the attacking country “continued to invade” and “took over 20-percent of its neighbor’s territory,” regardless of the president’s statements or actions. To control for opinions based on U.S. casualties, “No Americans were killed” in any of the five conditions.

Respondents are then asked for their opinion of the U.S. president’s handling of the situation, which they rate on a five-point scale. To better identify respondents’ decision-making logic and to assess the mechanisms underpinning the experimental findings, the survey instrument also asks respondents several questions about their perceptions of consistency, reputation, and presidential competence in the context of the scenario. I also collect data on several covariates including demographics and respondent preferences toward international

---

<sup>4</sup> Appendix A includes the full survey instrument.

intervention. I also include a manipulation check to ensure subjects received and internalized the intended treatment.<sup>6</sup>

As with any study, there are limits on the generalizability of findings. First, I limit manipulations to the key variable of interest. Like other recent studies (Trager and Vavreck 2011; Levy et al. 2015), I do not vary factors like casualties, and the adversary's regime type, strength and motive for invasion. By manipulating only whether the president threatens to invade and his subsequent action, I maximize statistical power and focus on the effect of gradations of consistency between a leader's words and deeds. To be sure, varying the factors above would increase generalizability. For instance, the American public might levy lower audience costs on a president that reneges on a threat to invade a democratic or militarily powerful state. Second, I only field the experiments in the United States. Past studies find that audience costs exist in other democratic and non-democratic states (Weeks 2008; Davies and Johns 2013), meaning there may be cross-national variation in the intensity and causes of audience costs. Respondents in other states may hold different foreign policy preferences, views on national honor, and expectations of presidential conduct than Americans.<sup>7</sup> For example, citizens of countries with extensive international relationships, like the United States, may put greater weight on reputational consequences of inconsistency than respondents from less interconnected states. I leave these tests to future research.

---

<sup>6</sup> See Appendix C, Section 2. I do not drop respondents who fail the manipulation check. Doing so may produce biased estimates of the treatment effect (Gerber and Green 2015, 211-252).

<sup>7</sup> Even individuals from different regions of the same country may hold different perceptions of reputation. See Dafoe and Caughey (2016).

### ***Sampling and Implementation Strategy***

The experiment was fielded on 1,006 adults across the United States in December 2016. Respondents were recruited using Amazon Mechanical Turk (MTurk) and randomly assigned to treatment groups using Qualtrics, an online survey platform. Social scientists have increasingly turned to MTurk as a means of gathering data more quickly and inexpensively than traditional survey methods (Buhrmester, Kwang, and Gosling 2011; Berinsky, Huber, and Lenz 2012; Goodman, Cryder, and Cheema 2013). Although crowd-sourced labor markets like MTurk allow researchers to efficiently build research samples, they lack the representativeness of national probability samples (Berinsky, Huber, and Lenz 2012). The MTurk sample used in this experiment underrepresents blacks, Hispanics, and households with higher incomes, and overrepresents younger people and individuals with at least a high school degree. The gender balance and balance of military veterans and non-veterans are similar to a national sample.<sup>8</sup>

While MTurk samples are less representative than national probability samples, studies on risk perception conducted using MTurk samples have produced results similar to those obtained from nationally-representative samples (Berinsky, Huber, and Lenz 2012). These findings suggest the cognitive processes of MTurk respondents are similar to those of the broader American population. Because of this, MTurk samples have been used in recent audience cost experiments (Chaudoin 2014; Levy et al. 2015).

## **FINDINGS**

### ***Average Treatment Effects: Audience Costs***

To assess the effects of policy substitution on audience costs, I examine respondents' approval ratings of the president's handling of the hypothetical scenario. Like earlier studies, I

---

<sup>8</sup> Appendix C, Section 4 compares demographic characteristics of the experimental sample with a national sample.

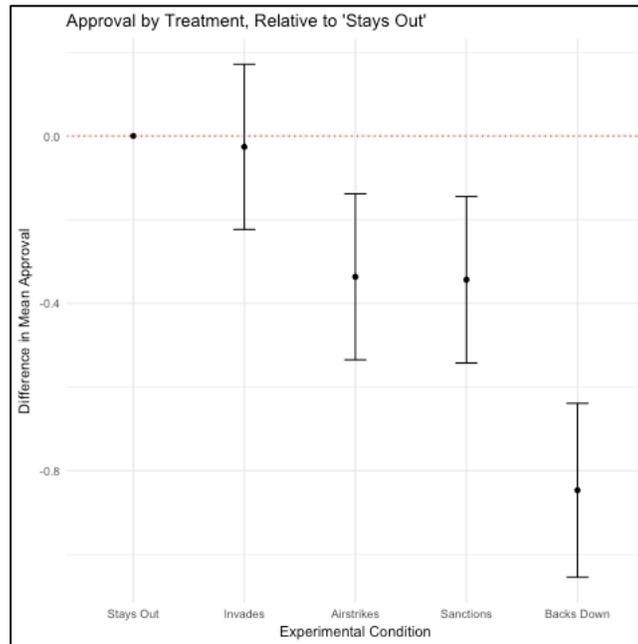
interpret decreased mean approval as the audience cost imposed on a leader. Because the experimental design holds constant policy outcomes, differences in approval can be attributed to variation in the president's handling of the situation. Figure 2 illustrates the difference in mean approval across experimental conditions, with the condition where a leader stays out after pledging to stay out as a baseline.

The findings offer strong support for hypothesis H<sub>1</sub>. The implementation of alternate policies – that lie between carrying out the full threat and doing nothing – appear to mitigate some of the audience costs associated with backing down. When the president follows through with his threat, his mean approval is 3.359 on a five-point scale. When he backs down entirely, the mean approval drops by 0.82-points to 2.538.<sup>9</sup> When the president backs up instead, his approval decreases less than it does when he backs down (Airstrikes:  $\mu=3.048$ ; Sanctions:  $\mu=3.041$ ). Although a leader faces higher audience costs for sanctions than for airstrikes, the difference between the two is not statistically discernable, a surprising finding I explore later.

---

<sup>9</sup> As an alternate measure of audience costs, I measure the percentage of respondents in each experimental condition who either disapprove or strongly disapprove of the president's handling of the situation (Appendix C, Section 3). When the president backs down (stays out after pledging to invade), 58.9% of respondents disapprove or strongly disapprove. In contrast, when the president stays out after to pledging to stay out, only 22.1% disapprove or strongly disapprove. This 36.8-percentage point treatment effect is considerably larger than the 16-percentage point treatment effect in Tomz (2007). Several factors may have contributed to this difference. First, I fielded my survey experiment in December 2016, just after a contentious presidential campaign that highlighted issues of reputation and credibility. This may have led some respondents to dole out more punishment on leaders seen as inconsistent or untrustworthy. Second, a smaller portion of my sample disapproves of staying out (22.1% vs. 33% in Tomz). Overseas interventions in places like Afghanistan, Libya, and Syria that occurred in the decade between Tomz's and my experiment may have left the population more war-weary and more supportive of avoiding interventions.

FIGURE 2:



To more systematically examine these findings, I employ ordinary least squares models to test their statistical significance. I regress approval rating (measured on a five-point scale) on experimental condition (coded as a binary variable). The findings are statistically significant ( $p < .01$ ), indicating that backing up allows leaders to avoid some of the audience costs associated with backing down altogether.<sup>10</sup> Several additional models interact experimental condition with various covariates to assess whether respondent characteristics contribute to heterogeneous treatment effects. With the exception of preferences toward intervention, which I discuss later, covariate interactions generally have no statistically significant effect on approval ratings.<sup>11</sup>

### *Average Treatment Effects: Consistency, Reputation, and Competence*

To examine whether respondents judge a leader's consistency along a continuum, the survey instrument asks respondents "how consistent was the President's handling of the situation

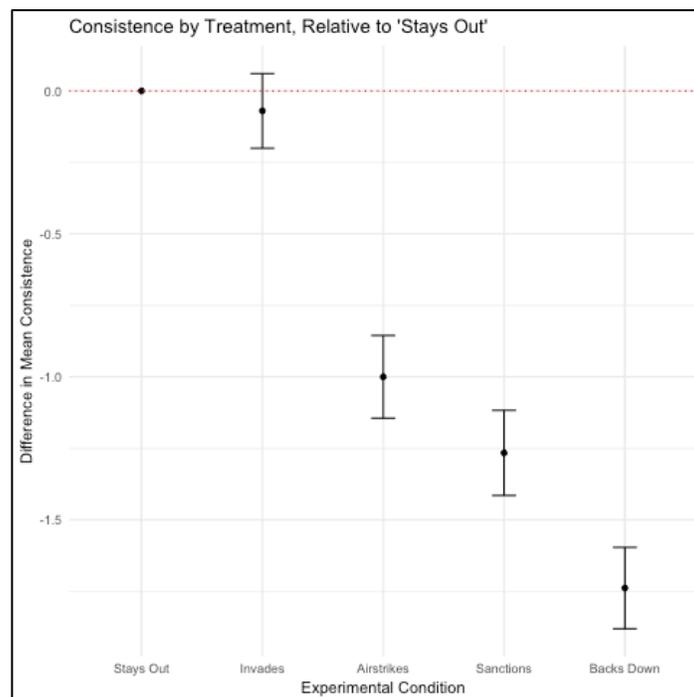
---

<sup>10</sup> See Appendix C, Section 3 for regression analysis for each of the dependent variables (approval, consistency, reputation, and competence).

<sup>11</sup> See Appendix C, Section 3 for discussion of heterogeneous treatment effects.

with his initial statement?” In line with expectations (H<sub>2</sub>), respondents perceive policies as more consistent when they diverge little from the initially threatened action in terms of risk or degree of military escalation. As Figure 3 illustrates, a president who follows through on his threats or promises – either by honoring a pledge to stay out or a threat to invade – receives relatively high consistency ratings. Unsurprisingly, perceived consistency decreases as a leader backs up or backs down from a threat. The public perceives backing up to another military option (airstrikes) as more consistent with the initial threat than backing up to a non-military alternative (sanctions). When examined in conjunction with audience cost measures, behavior viewed as more consistent is generally correlated with lower audience costs. However, while a leader who implements sanctions is considered less consistent than one who launches airstrikes, the audience costs levied in the two backed-up experimental conditions are not statistically discernable. This suggests the public may consider factors other than inconsistency when imposing audience costs, something I examine in more detail below.

**FIGURE 3:**



If distaste for inconsistency underlies audience costs, leaders should suffer punishment for any inconsistency between their words and deeds, regardless of whether they back out of a threat or back into conflict (Levy et al. 2015). To assess whether variation in consistency affects leaders who “back in,” I field a follow-up experiment where a leader launches sanctions, airstrikes, or invades after pledging to stay out of the conflict. The findings suggest the public views inconsistency along a continuum when leaders back in, and that gradation in inconsistency contributes, in part, to the extent of audience costs.<sup>12</sup>

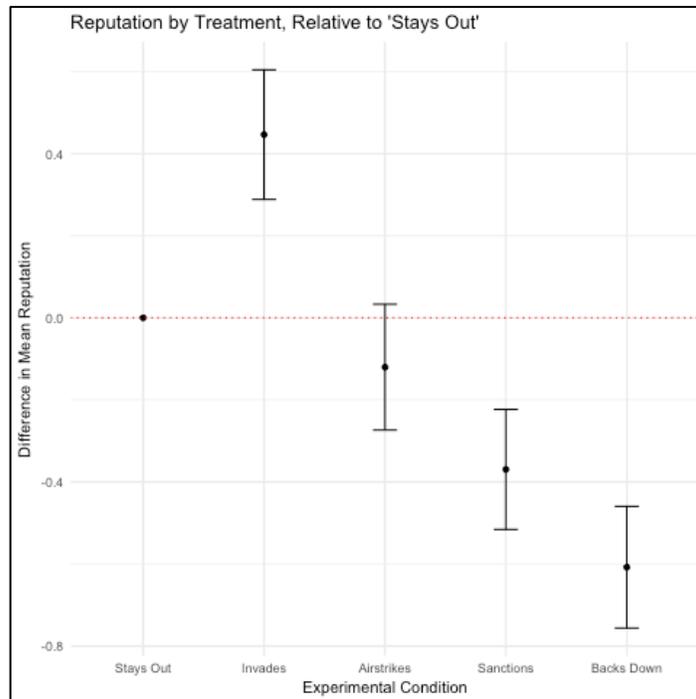
According to one logic of audience costs, the public punishes leaders that renege on threats because inconsistency damages a state’s reputation in future bargaining situations and because bluffing leaders are seen as weak or incompetent. Since the public ascribes different levels of inconsistency to different “backed-up” or “backed-down” policies, I examine whether the degree of inconsistency between a leader’s words and deeds leads to variation in perceived reputational consequences and presidential competence. To solicit feedback on reputational effects, the survey instrument asks respondents to rate “How much did the President’s handling of the situation hurt or improve the reputation of the United States in the world?” on a five-point scale ranging from “Significantly Hurt” to “Significantly Improve.” Because reputation can mean different things to different people, I draw from Levy et al. (2015, 997) and ask a question about specific reputational consequences: “Based solely on the U.S. President’s handling of the situation, how likely is it that other countries will believe threats and promises made by the U.S. President in the future?” Respondents answer on a four-point scale ranging from “Very Unlikely” to “Very Likely.”

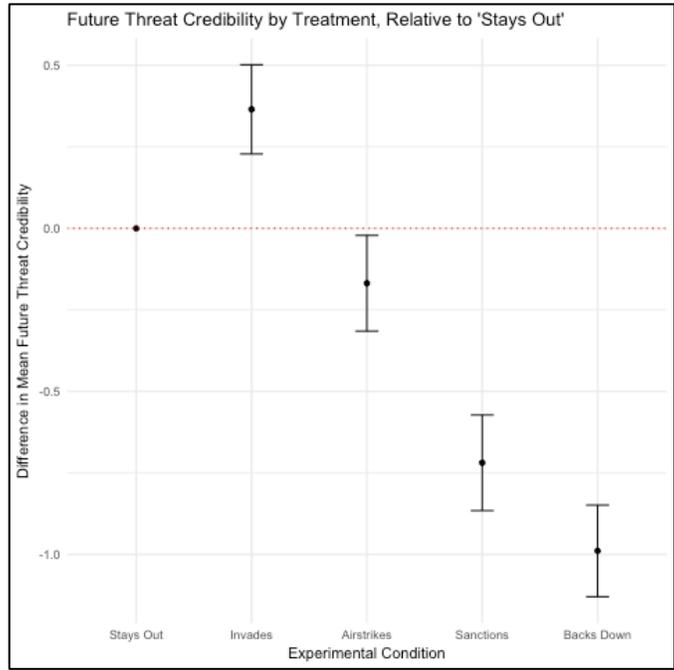
---

<sup>12</sup> See Appendix D, Section 2 for the follow-up experiment.

In line with hypothesis H<sub>3</sub>, the public views leaders who back up as having a less adverse effect on the state’s reputation than a leader who backs down. As Figures 4 and 5 illustrate, the state’s reputation and the credibility of future threats and promises is highest when the president follows through on his threat and lowest when he backs down entirely. Of the backed up policies, airstrikes result in, on average, a higher mean reputation and future threat credibility than sanctions. These findings suggest the public believes gradations in consistency matter in dictating a state’s reputation or a leader’s signaling reputation. Inconsistency, however, is not the only factor shaping reputational consequences. The mean reputation and future threat credibility for following through with a pledge to stay out are approximately 0.4-points lower than those for following through with an invasion, suggesting the public believes that not coming to the aid of an ally in need weakens Washington’s foreign policy reputation.

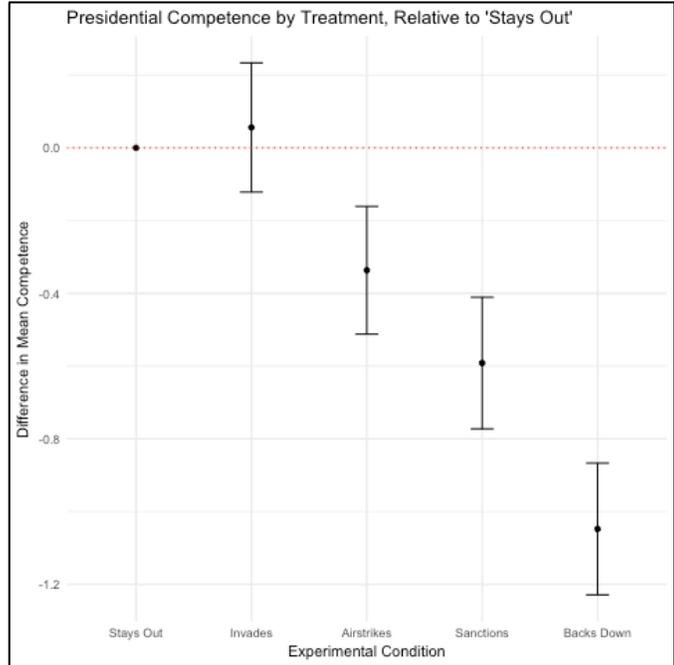
**FIGURES 4 and 5:**





As Figure 6 illustrates, the public considers a leader who backs up as more competent than a leader who backs down, providing support for hypothesis H<sub>4</sub>. Similar to the findings on reputation, the highest mean competence ratings are seen when the president follows through with his threat of ground forces and lowest when the president backs down entirely.

**FIGURE 6:**



### *Heterogeneous Effects: Hawks and Doves*

At first glance the nearly equal audience costs for airstrikes and sanctions is puzzling. Under hypothesis H<sub>1</sub>, audience costs should be higher when inconsistency between a leader's words and deeds is higher. Despite respondents viewing sanctions as more inconsistent than airstrikes, they do not, on average, appear to levy higher audience costs for sanctions. The experimental findings, however, need not be inconsistent with theoretical expectations. Dividing the sample into foreign policy hawks and doves and decomposing audience costs into belligerence and inconsistency cost components reveals that gradations in consistency still affect audience costs.

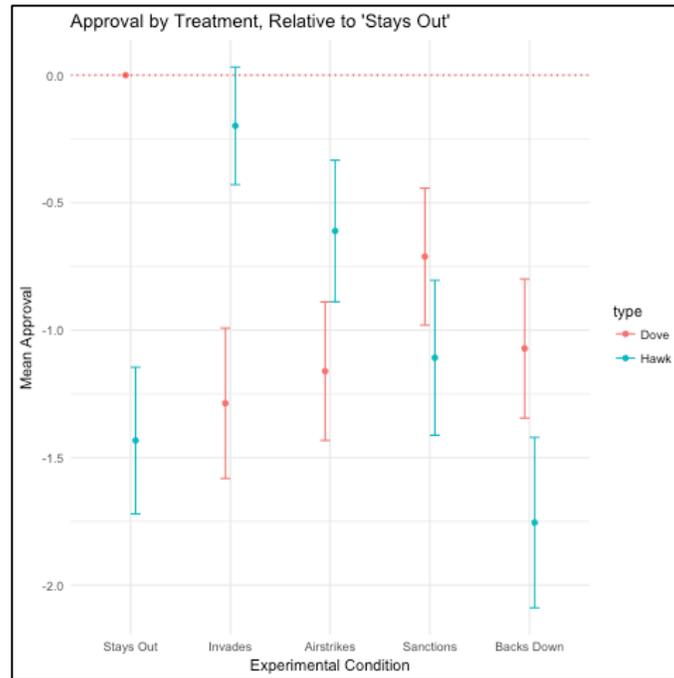
First, I examine approval and consistency ratings of backing up to airstrikes and sanctions among foreign policy hawks and doves.<sup>13</sup> On average, hawks ( $\mu_{\text{consistency}}=2.347$ ) and doves ( $\mu_{\text{consistency}}=2.31$ ) ascribe roughly the same consistency rating to airstrikes, but as Figure 7 illustrates, doves levy significantly higher audience costs ( $\mu_{\text{approval}}=2.798$ ) than hawks ( $\mu_{\text{approval}}=3.347$ ) when a leader backs up to airstrikes. When a president backs up to sanctions, the perceived degree of consistency is again similar among hawks ( $\mu_{\text{consistency}}=2.0$ ) and doves ( $\mu_{\text{consistency}}=2.026$ ), but hawks impose much higher audience costs ( $\mu_{\text{approval}}=2.85$ ) than doves ( $\mu_{\text{approval}}=3.247$ ). These divergent costs average out when audience costs are calculated for the full sample, resulting in audience costs for airstrikes and sanctions that are not statistically discernable. Although hawks and doves appear to judge consistency similarly, the approval

---

<sup>13</sup> Hawks (n=324) either agree or strongly agree that “the United States needs to play an active role in solving conflicts around the world.” Doves (n=412) disagree or strongly disagree with this interventionist sentiment. Although data on foreign policy preference are collected post-treatment, treatment generally appears to have little effect on preferences toward intervention. This suggests self-reported preferences can be used as a proxy of background attitudes toward intervention. See Appendix C, Section 3.

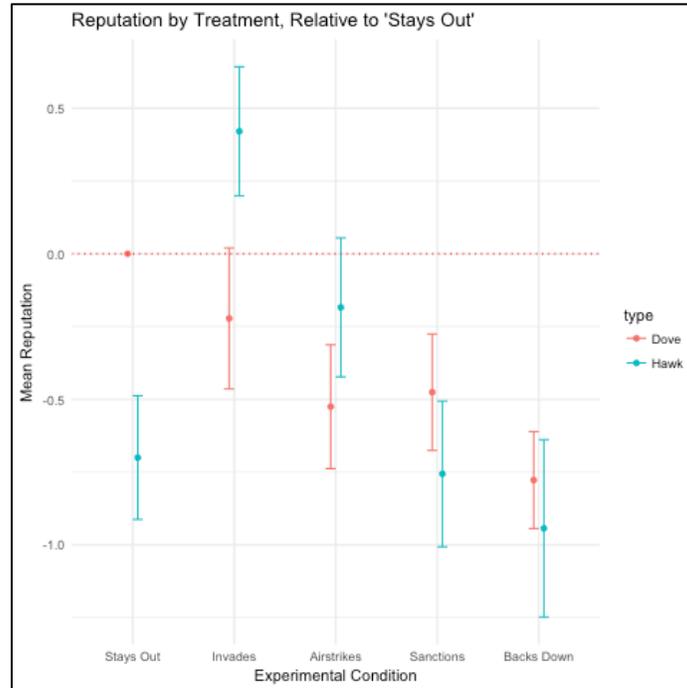
ratings suggests hawks are more likely to punish for inconsistency, while any cost doves might impose on leaders for inconsistency are tempered by their preference for the implementation of a less belligerent policy.

**FIGURE 7:**



Audience costs, however, are not the only outcome where hawks and doves have divergent views. As illustrated in Figure 8, hawks believe following through with an invasion has a significant positive effect on reputation, while staying out, backing up, or backing down generates adverse reputational consequences. Doves, in contrast, believe that either following through or renegeing on a threat leaves a state's reputation worse off to pledging to stay out. This suggests that that hawks support following through in large part for reputational concerns in addition to distaste for inconsistency.

**FIGURE 8:**



To examine the extent to which gradations in consistency shape audience costs, I draw from Kertzer and Brutger (2016) and decompose audience costs associated with sanctions and airstrikes into belligerence and inconsistency components. Kertzer and Brutger suggest that leaders are punished both for saying one thing and doing another and for threatening to use force. Doves punish leaders for threatening to use force, while hawks – who are less likely to disapprove of belligerence – punish leaders primarily for inconsistency. This “double barreled” logic of audience costs predicts doves will levy higher belligerence costs for threatening bellicose airstrikes than more pacific sanctions. Hawks should impose higher inconsistency costs for renegeing on a threat of airstrikes than a threat of sanctions, since doing nothing is more inconsistent with airstrikes than sanctions.

I test these predictions using a follow-up experiment that is nearly identical to the main experiment, but slightly alters the president’s initial threat in order to generate the data needed to

calculate belligerence and inconsistency costs.<sup>14</sup> The leader in the follow-up experiment first pledges to stay out, or threatens sanctions or airstrikes. She then follows through with her initial pledge or threat, or reneges by taking no action. Using the approach introduced by Kertzer and Brutger (2016), I calculate belligerence costs by subtracting the mean approval of following through with a pledge to stay out of the conflict from the mean approval of following through with a threat of sanctions or airstrikes. I then calculate inconsistency costs by subtracting the mean approval of following through with a threat from the mean approval of reneging on a threat. The sum of these two components yields the total audience cost.

**TABLE 2:**

<i>Backs down from</i>	<b>Audience Cost</b> $\mu_{22} - \mu_{11}$	<b>Belligerence Cost</b> $\mu_{21} - \mu_{11}$	<b>Inconsistency Cost</b> $\mu_{22} - \mu_{21}$	<b>Belligerence Cost Fraction</b> $\frac{\mu_{21} - \mu_{11}}{\mu_{22} - \mu_{11}}$	<b>Inconsistency Cost Fraction</b> $\frac{\mu_{22} - \mu_{21}}{\mu_{22} - \mu_{11}}$
Sanctions (Full Sample)	-1.038	0.047	-1.085	-0.045	1.045
Airstrikes (Full Sample)	-0.953	0.040	-0.993	-0.042	1.042
Sanctions (Hawks)	-0.599	0.686	-1.285	-1.15	2.15
Airstrikes (Hawks)	-0.49	0.766	-1.256	-1.56	2.56
Sanctions (Doves)	-1.396	-0.676	-0.72	0.484	0.516
Airstrikes (Doves)	-1.422	-0.996	-0.426	0.70	0.30

As Table 2 illustrates, gradations in both consistency and belligerence matter. Among hawks, inconsistency costs comprise a larger fraction of audience costs when a leader reneges on a threat of airstrikes than on a threat of sanctions. This aligns with the theoretical expectations that inconsistency shapes punishment since reneging and doing nothing is more dissimilar and inconsistent with airstrikes than sanctions, yielding greater inconsistency costs. Doves, on the other hand, levy higher belligerence costs for threatening airstrikes than for threatening sanctions, suggesting that belligerence is also viewed along a continuum. The different punishments levied by hawks and doves helps explain why overall audience costs for sanctions

<sup>14</sup> See Appendix D, Section 3.

and airstrikes are similar even if sanctions are considered less consistent of an alternate policy than airstrikes.

### ***Micro-Mechanisms: Explaining Experimental Outcomes***

To further unpack whether individual respondents' distaste for inconsistency underpins approval for the president's handling of the hypothetical crisis, I analyze qualitative data in a manner similar to Tomz (2007). I ask respondents in the main experiment to "write a sentence or two telling us why you approve/disapprove of the way the President handled the situation." I then manually coded the 1,006 responses into one of five categories to assess *why* respondents gave lower approval ratings to leaders who failed to carry out their initial threat.<sup>15</sup> Analysis of these qualitative data supports the theory that inconsistency is measured along a continuum: respondents are more likely to cite inconsistency as justification for decreased approval as the actions become less qualitatively similar to the initial threat.

One group of respondents believed the president did the right thing. Some of these respondents said the president's decision helped save American lives, protected an ally, or simply "seemed to get the job done." A second group of respondents disapproved of the actions because they were opposed to international involvement or military escalation. Respondents in this category justified their positions with statements like, "I think we need to stick more to protecting our own. We don't need to interfere in everyone else's conflicts[.]" or, "I don't approve of war. We need to stop killing each other." The third group of respondents held their positions because of the president's inconsistency. Some subjects believed reneging on a threat

---

<sup>15</sup> I base my classifications on Tomz 2007. Tomz introduces three broad classifications: right thing to do, dislike of escalation/overseas involvement, and inconsistency (Tomz 2007, 834-835). I add two additional categories – Presidential weakness and lack of information/expertise – to capture responses with greater nuance. A second individual coded all responses, producing an 85.2 % inter-coder agreement rate. Coding guide in Appendix B.

could harm America’s reputation by “mak[ing] it look like we do not honor our agreements” and because “foreign enemies will not take threats by the United States seriously.” Others suggested that inconsistency highlighted the president’s weakness or lack of honor. A fourth group of respondents cited presidential incompetence or poor planning without mentioning inconsistency. A fifth group thought they lacked sufficient information or expertise to make a valid judgment.

**TABLE 3: Micro-Mechanisms by Experimental Condition**

<i>Justification for Approval Rating</i>	<b>Stays Out</b> <i>Pledges to Stay Out, Stays Out</i>	<b>Follows Through</b> <i>Threatens Invasion, Invades</i>	<b>Backs Up</b> <i>Threatens Invasion, Implements Air Strikes</i>	<b>Backs Up</b> <i>Threatens Invasion, Implements Sanctions</i>	<b>Backs Down</b> <i>Threatens Invasion, Takes No Action</i>
Right Thing to Do	18.3 (13.0-23.5) n=38	58.1 (51.2-65.0) n=115	40.4 (33.7-47.1) n=84	30.8 (24.3-37.2) n=60	10.7 (6.4-15.0) n=21
Opposed to International Interference	42.3 (35.6-49.0) n=88	25.8 (19.7-31.8) n=51	19.2 (13.9-24.6) n=40	12.8 (8.1-17.5) n=25	14.2 (9.4-19.1) n=28
Inconsistency	0.5 (-0.46-1.4) n=1	0.0 (0.0-0.0) n=0	28.4 (22.2-34.5) n=59	39.5 (32.6-46.3) n=77	61.9 (55.1-68.7) n=122
Ineffective Plan/Incompetence	21.2 (15.6-26.7) n=44	7.6 (3.9-11.3) n=15	4.8 (1.9-7.7) n=10	7.2 (3.6-10.8) n=14	6.6 n=13 (3.1-10.1)
Insufficient Information to decide	17.8 (12.6-23.0) n=37	8.6 (4.7-12.5) n=17	7.2 (3.7-10.7) n=15	9.7 (5.6-13.9) n=19	6.6 (3.1-10.1) n=13

*Note:* Entries are percentages. Percentages within each column may not sum to 100 due to rounding. Range in parentheses is a 95% confidence interval.

The results support the commonly theorized foundations of audience costs and offer greater insight into how domestic publics perceive inconsistency. As predicted by hypothesis H<sub>2</sub>, respondents in the “backs down” condition are, on average, the most likely to justify their disapproval by citing concerns about inconsistency between a president’s threats and actions, while respondents in the “backs up” conditions are, on average, less likely to reference inconsistency. The proportion of subjects citing inconsistency as justification for approval ratings increases as the president’s actions become more qualitatively different from the initially threatened invasion. Similarly, the percentage of respondents who believe the president’s actions

were the “right thing to do” decreases as the implemented action becomes more qualitatively different from an invasion. These findings suggest the public views acts of policy substitution along a continuum of inconsistency. Indeed, some respondents explicitly described the divergence between the president’s words and deeds. One subject explained that sanctions were “too different from the action promised[,]” while another suggested sanctions were “more of an economic slap on the wrist than anything...it wasn’t really intervention as promised.” In contrast, several respondents in the “follows through” condition explained that the president “did the right thing” in part because his actions were consistent with his initial threat.

The qualitative data also shed light on why the public punishes leaders for being inconsistent, supporting the reputational mechanisms thought to underlie audience cost theory. Many respondents feared their leader’s idle threats harm the state’s reputation among friends and foes, or paint the leader as untrustworthy or incompetent. Some respondents expressed concerns that not following through on promises would “make it look like we do not honor our agreements with other nations” and “deteriorate our relations with foreign allies.” Other respondents believed that reneging on threats meant that “foreign enemies will not take threats by the United States seriously” and that “other nations will prey on that weakness.” In addition to concerns about future reputation, respondents also suggested that a “President who doesn’t follow through on his promises is a weak leader” and is not “reliable” or “honorable.”

These findings support existing audience cost research, but also reveal that the punishment imposed on leaders is shaped by the degree of deviation between the leader’s initial words and her subsequent deeds. This offers a more nuanced understanding of public perceptions of inconsistency and its reputational consequences.

## IMPLICATIONS FOR CRISIS BARGAINING

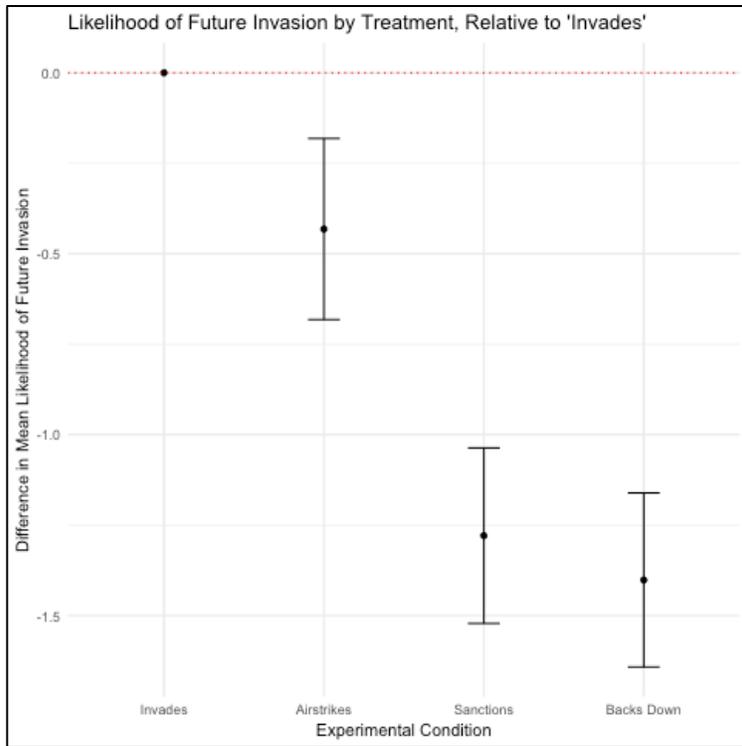
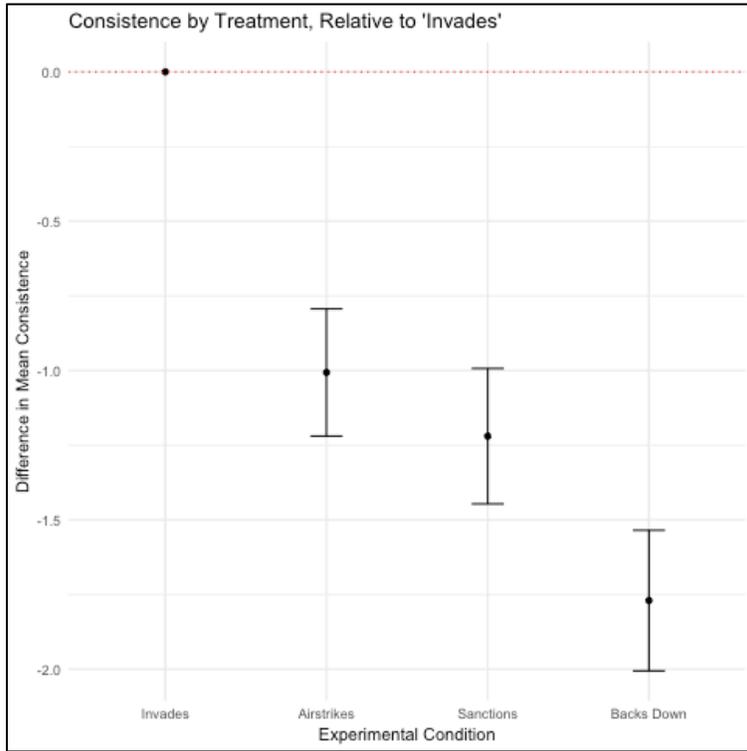
According to the traditional audience cost logic, a leader only makes threats she intends to carry out. However, if a leader faces less punishment for backing up than for reneging entirely, she may have incentives to issue a strong threat before implementing a “backed up” policy. Even if the sending state’s population can levy audience costs on its leader for reneging on threats, threat recipients may have trouble identifying whether an initial threat is credible or cheap talk that will be followed by a less hawkish policy or no action at all. As a result, a state on the receiving end of threats issued by a state that has previously backed up may be less likely to concede to the aggressor’s threat.

To test the effect of backing up on the credibility of crisis signaling, I field a follow-up experiment that places respondents on the receiving end of a threat.<sup>16</sup> Respondents are told that the president of a rival country threatens to invade a U.S. ally if the U.S. does not close its military bases in the allied country. The U.S. does not close its bases and respondents are randomly assigned to one of four experimental conditions that vary the rival’s subsequent action: 1) follows through with invasion, 2) backs up to airstrikes, 3) backs up to sanctions, or 4) backs down and takes no action. To control for outcomes, all respondents are told that no Americans are harmed by the rival’s actions. Respondents are asked to rate how consistent the rival leader’s handling of the situation was with the initial threat. Respondents are then informed “several months later, the rival leader threatens to invade another U.S. ally.” Respondents are asked to rate the likelihood that the rival will actually invade using a four-point scale.

---

<sup>16</sup> See Appendix E.

FIGURES 9 AND 10



As figure 10 illustrates, respondents are, on average, less likely to believe a rival's future warnings if she has reneged on threats in the past. The greater the inconsistency between the rival's past words and deeds, the less seriously their threats are taken in the future. Future threats are most believable when the rival leader has previously followed through on an initial threat to invade ( $\mu=3.38$ ) and least credible when the leader previously backed down entirely ( $\mu=1.983$ ). When a rival backs up to airstrikes ( $\mu=2.95$ ) and sanctions ( $\mu=2.10$ ), her future threats are less likely to be believed than those of a leader who followed through with an earlier threat, but are more credible than those of a leader who previously backed down. Although respondents in the experiment may think differently than statesmen, their responses still provide insights into how leaders judge the credibility of rivals' threats. The degraded believability of backed up threats suggests that backing up damages a leader's signaling reputation, but is less detrimental than backing down altogether.

Some states – namely those with the ability to leverage the full range of diplomatic, military, and economic power – may be more able to back up than others.<sup>17</sup> A state with only ground forces is unable to back up to airstrikes, and a state with weak economic influence cannot credibly enact sanctions. The ability of powerful states to more easily back up appears to counter the assertions of both audience cost theorists and leaders like former U.S. Vice President Joe Biden who once claimed, “big nations can't bluff.”<sup>18</sup>

---

<sup>17</sup> I appreciate Brian Blankenship highlighting this point.

<sup>18</sup> Vice-President Biden made this comment at the 2013 Annual Conference of the American Israel Public Action Committee (AIPAC). See, <http://www.aipac.org/act/attend-events/policy-conference/videos/2013/speeches/biden>.

## CONCLUSION

This article extends research on audience cost theory and crisis bargaining by introducing the concept of “backing up” as an alternative to “backing down.” Rather than assuming a leader has a binary choice of following through with an initial threat or backing down completely, I argue that real world leaders often practice policy substitution. By backing up to an alternate policy rather than completely backing down from an initial threat, leaders can reduce the inconsistency between their words and deeds and subsequently face lower audience costs.

A series of original survey experiments finds strong support for this theory. Respondents view actions like sanctions and airstrikes as more consistent with an initial threat of ground invasion, and therefore dole out less punishment on a leader who backs up than on a leader who backs down. While the experimental findings support the mechanisms underlying audience cost theory, they raise questions about the common claim that audience costs lead to more credible crisis signaling. Since backing up allows a leader to lessen the punishment for deviating from threats, leaders might make a threat, knowing that they can step back without facing the full audience costs of backing down. As the experiment demonstrates, actors on the receiving end of backed up threats are less likely to believe a rival’s future threats. This degraded threat credibility may weaken the deterrent power of threats and decrease the likelihood that a rival will concede during crises.

The finding that consistency is measured along a continuum suggests several pathways for future research. First, future projects might explore whether the findings are generalizable beyond the United States. Most experiments examining audience costs have been fielded in the United States, but studies have suggested that audience costs plague leaders in other states (Weeks 2008; Davies and Johns 2013). Replicating the experiment in other national contexts

may shed light on how variation in inconsistency shapes audience costs and public perceptions of reputation and leader competence. Second, further studies might examine the role of media and politicians in shaping perceptions of inconsistency and the subsequent political effects. Recent research has found that political rhetoric and media framing can affect audience costs (Levendusky and Horowitz 2012; Davies and Johns 2013; Potter and Baum 2014). Can messaging by elected officials or media outlets influence the public's perceptions of inconsistency between a leader's words and deeds and the resulting political consequences? Third, future studies might more deeply explore how backing up affects behavior by states on the receiving end of threats. Are leaders more likely to escalate vis-à-vis a leader who backs up on a threat? Additional studies that assess the continuum of inconsistency will expand our understanding of audience costs and crisis bargaining.

## References

- Baker, Peter, and Jonathan Weisman. 2013. "Obama Seeks Approval by Congress for Strike in Syria." *The New York Times*, August 31, 2013.
- Berinsky, Adam J., Gregory A. Huber, and Gabriel S. Lenz. 2012. "Evaluating Online Labor Markets for Experimental Research: Amazon.com's Mechanical Turk." *Political Analysis* 20 (3): 351–68.
- Brecher, Michael, Jonathan Wilkenfeld, Kyle Beardsley, Patrick James, and David Quinn. 2017. "International Crisis Behavior Data Codebook, Version 12."
- Brehm, Jack W., and Arthur R. Cohen. 1959. "Re-Evaluation of Choice Alternatives as a Function of Their Number and Qualitative Similarity." *The Journal of Abnormal and Social Psychology* 58 (3): 373.
- . 1962. *Explorations in Cognitive Dissonance*. New York: John Wiley & Sons.
- Brutger, Ryan, and Joshua D. Kertzer. 2018. "A Dispositional Theory of Reputation Costs." *International Organization* 3 (72).
- Buhrmester, Michael, Tracy Kwang, and Samuel D. Gosling. 2011. "Amazon's Mechanical Turk: A New Source of Inexpensive, Yet High-Quality, Data?" *Perspectives on Psychological Science* 6 (1): 3–5.
- Chaudoin, Stephen. 2014. "Promises or Policies? An Experimental Analysis of International Agreements and Audience Reactions." *International Organization* 68 (1): 235–56.
- Chollet, Derek. 2016. *The Long Game: How Obama Defied Washington and Redefined America's Role in the World*. New York: PublicAffairs.
- Clare, Joe. 2007. "Domestic Audiences and Strategic Interests." *The Journal of Politics* 69 (3): 732–45.
- Clark, David H., Timothy Nordstrom, and William Reed. 2008. "Substitution Is in the Variance: Resources and Foreign Policy Choice." *American Journal of Political Science* 52 (4): 763–73.
- Clark, David H., and William Reed. 2005. "The Strategic Sources of Foreign Policy Substitution." *American Journal of Political Science* 49 (3): 609–24.
- Converse, Phillip E. 1964. "The Nature of Belief Systems in Mass Publics." In *Ideology and Discontent*, edited by David Apter, 202–61. New York: Wiley.
- Copeland, Dale C. 1997. "Do Reputations Matter?" *Security Studies* 7 (1): 33–71.
- Dafoe, Allan, and Devin Caughey. 2016. "Honor and War: Southern US Presidents and the Effects of Concern for Reputation." *World Politics* 68 (2): 341–81.
- Dafoe, Allan, Jonathan Renshon, and Paul Huth. 2014. "Reputation and Status as Motives for War." *Annual Review of Political Science* 17 (1): 371–93.
- Davies, Graeme A. M., and Robert Johns. 2013. "Audience Costs among the British Public: The Impact of Escalation, Crisis Type, and Prime Ministerial Rhetoric." *International Studies Quarterly* 57 (4): 725–37.
- Debs, Alexandre, and Jessica Chen Weiss. 2016. "Circumstances, Domestic Audiences, and Reputational Incentives in International Crisis Bargaining." *Journal of Conflict Resolution* 60 (3): 403–33.
- Dugger, Celia W. 2002. "India Test-Fires Intermediate-Range Missile." *The New York Times*, January 25, 2002, sec. World. <https://www.nytimes.com/2002/01/25/world/india-test-fires-intermediate-range-missile.html>.

- Fearon, James D. 1994. "Domestic Political Audiences and the Escalation of International Disputes." *The American Political Science Review* 88 (3): 577–92.
- Gelpi, Christopher, and Joseph M. Grieco. 2015. "Competency Costs in Foreign Affairs: Presidential Performance in International Conflicts and Domestic Legislative Success, 1953–2001." *American Journal of Political Science* 59 (2): 440–56.
- Gerber, Alan S, and Donald P Green. 2012. *Field Experiments: Design, Analysis, and Interpretation*. New York: W. W. Norton.
- Gilpin, Robert. 1981. *War and Change in World Politics*. New York: Cambridge University Press.
- Goldberg, Jeffrey. 2016. "The Obama Doctrine." *The Atlantic*, April 2016. <https://www.theatlantic.com/magazine/archive/2016/04/the-obama-doctrine/471525/>.
- Goodman, Joseph K., Cynthia E. Cryder, and Amar Cheema. 2013. "Data Collection in a Flat World: The Strengths and Weaknesses of Mechanical Turk Samples." *Journal of Behavioral Decision Making* 26 (3): 213–24.
- Huth, Paul K. 1997. "Reputations and Deterrence: A Theoretical and Empirical Assessment." *Security Studies* 7 (1): 72–99.
- Hyde, Susan D. 2015. "Experiments in International Relations: Lab, Survey, and Field." *Annual Review of Political Science* 18 (1): 403–24.
- Jervis, Robert. 1970. *The Logic of Images in International Relations*. Princeton, N.J.: Princeton University Press.
- Jones, Daniel M., Stuart A. Bremer, and J. David Singer. 1996. "Militarized Interstate Disputes, 1816–1992: Rationale, Coding Rules, and Empirical Patterns." *Conflict Management and Peace Science* 15 (2): 163–213.
- Kertzer, Joshua D., and Ryan Brutger. 2016. "Decomposing Audience Costs: Bringing the Audience Back into Audience Cost Theory." *American Journal of Political Science* 60 (1): 234–49.
- Kurizaki, Shuhei, and Taehee Whang. 2015. "Detecting Audience Costs in International Disputes." *International Organization* 69 (4): 949–80.
- Levendusky, Matthew S., and Michael C. Horowitz. 2012. "When Backing Down Is the Right Decision: Partisanship, New Information, and Audience Costs." *The Journal of Politics* 74 (2): 323–38.
- Levy, Jack S., Michael K. McKoy, Paul Poast, and Geoffrey P.R. Wallace. 2015. "Backing Out or Backing In? Commitment and Consistency in Audience Costs Theory." *American Journal of Political Science* 59 (4): 988–1001.
- Mercer, Jonathan. 1996. *Reputation And International Politics*. Ithaca, NY: Cornell University Press.
- Most, Benjamin A., and Harvey Starr. 1984. "International Relations Theory, Foreign Policy Substitutability, and 'Nice' Laws." *World Politics* 36 (3): 383–406.
- Obama, Barack. 2013. "Statement by the President on Syria." The White House. August 31, 2013. <https://obamawhitehouse.archives.gov/the-press-office/2013/08/31/statement-president-syria>.
- Palmer, Glenn, and Archana Bhandari. 2000. "The Investigation of Substitutability in Foreign Policy." *The Journal of Conflict Resolution* 44 (1): 3–10.
- Potter, Philip B. K., and Matthew A. Baum. 2014. "Looking for Audience Costs in All the Wrong Places: Electoral Institutions, Media Access, and Democratic Constraint." *The Journal of Politics* 76 (1): 167–81.

- Schelling, Thomas C. 1961. "Experimental Games and Bargaining Theory." *World Politics* 14 (1): 47–68.
- Schultz, Kenneth A. 1998. "Domestic Opposition and Signaling in International Crises." *The American Political Science Review* 92 (4): 829–44.
- Sechser, Todd S., and Matthew Fuhrmann. 2017. *Nuclear Weapons and Coercive Diplomacy*. Cambridge: Cambridge University Press.
- Smith, Alastair. 1998. "International Crises and Domestic Politics." *American Political Science Review* 92 (3): 623–38.
- Snyder, Glenn H., and Paul Diesing. 1978. *Conflict Among Nations: Bargaining, Decision Making, and System Structure in International Crises*. Princeton, N.J.: Princeton University Press.
- Snyder, Jack, and Erica Borghard. 2011. "The Cost of Empty Threats: A Penny, Not a Pound." *The American Political Science Review* 105 (3): 437–56.
- Sood, V. K., and Pravin Sawhney. 2003. *Operation Parakram: The War Unfinished*. New Delhi: SAGE.
- Tomz, Michael. 2007. "Domestic Audience Costs in International Relations: An Experimental Approach." *International Organization* 61 (4): 821–40.
- Trachtenberg, Marc. 2012. "Audience Costs: An Historical Analysis." *Security Studies* 21 (1): 3–42.
- Trager, Robert F., and Lynn Vavreck. 2011. "The Political Costs of Crisis Bargaining: Presidential Rhetoric and the Role of Party." *American Journal of Political Science* 55 (3): 526–45.
- Walt, Stephen. 1990. *The Origins of Alliance*. Ithaca, N.Y.: Cornell University Press.
- Waltz, Kenneth N. 1979. *Theory of International Politics*. Boston, MA: McGraw Hill.
- Weeks, Jessica L. 2008. "Autocratic Audience Costs: Regime Type and Signaling Resolve." *International Organization* 62 (1): 35–64.