A Preference for War: How Fairness and Rhetoric Influence Leadership Incentives in Crises

MATTHEW S. GOTTFRID
US Department of State

AND

ROBERT F. TRAGER
University of California, Los Angeles

We conduct a survey experiment to examine the effects of international compromise, war, and foreign government rhetoric on presidential approval. We find that, in certain conflicts, popular approval tracks fairness heuristics—leaders seeking to maximize voter approval prefer equitable divisions of disputed goods and are risk acceptant for divisions below this threshold. Moreover, aggressive rhetoric by a foreign leader increases domestic leaders’ expected approval from war, decreases the value of compromise, and provides them with powerful incentives to fight harder. Thus, leaders motivated by popular approval have preferences that are inconsistent with the non-satiated, risk-averse preferences defined over shares of an objective good—that is, with those that much of the rationalist literature on conflict assumes. Fairness heuristics and the rhetorical framing of disputes during the conflict process may be at least as important as material factors in understanding why some disputes result in war.

Analysts often ascribe the meanest ambitions to leaders in international politics. They say that the anarchic system induces expansionist tendencies and leaders therefore face material incentives to seize as much of contested goods as they can (for example, Mearsheimer 2001; Fearon 1995). Yet, individuals often show great concern for appearing fair. This leads to more moderate behavior than their material self-interests otherwise dictate (for example, Güth, Schmittberger, and Schwarze 1982; Kühneman, Knetets, and Thaler 1986; Thaler 1988; Güth 1995; Lind 2001; Kapstein 2006; Albin 2001). Further, an individual’s willingness to accept an outcome—or engage in costly resistance to it—often depends as much on whether that outcome appears fair as it does on the material outcome itself. In this article, we provide strong evidence that these dynamics influence leader incentives in international relations (IR). Fairness heuristics play an important role in explaining political preferences in disputes and thus leader popularity is sometimes highest at a 50-50 division of a disputed good. Moreover, the use of bellicose rhetoric by foreign leaders frames compromise as imposition, which shifts domestic leaders’ political incentives toward war.

Much of the scholarly literature on conflict makes a different set of consequential assumptions about leader preferences. In a seminal article, Fearon (1995) demonstrates that a peacefully negotiated settlement that both sides prefer to war always exists. This preference for peace holds even when leaders desire as much of the disputed good as possible, provided that both actors are not risk loving. This finding is central to the bargaining model of war that dominates much analysis of interstate conflict; most formal models of conflict represent preferences in this way. Consequently, decades of scholarship have focused on the causal mechanisms that arise from this rationalist framework: asymmetrical information, commitment problems, and issue indivisibilities. Although leader preferences are fundamental to these explanations of conflict, scholars have yet to demonstrate the universality of these commonly assumed preferences. The literature fails to systematically analyze the relationship between marginal changes in crisis outcomes and the incentives of leaders.

Many leading scholars of international politics call for a greater focus on the determinants of leader preferences. Keohane (1993, 294) emphasizes the need for theories of interest and preference formation. Moravcsik (1997, 544) argues that we need to “take preferences seriously” by explicitly studying preference formation prior to analyzing strategic interaction. Powell (1994, 318) notes that many prominent theories “lack a theory of preferences over outcomes.” Fearon and Wendt (2002, 54) defend a division of labor in which rationalists do not investigate the origins of preferences. But they correctly warn that “... through a process of forgetting what we are doing, what starts out as merely an analytical convenience can become something more than that, a tacit assumption about what the world is...”

Matthew S. Gottfried is a Foreign Affairs Officer in the Office of Opinion Research at the US Department of State. His research interests include the relationship between political terrorism and public opinion, quantitative methods, elite interviewing, crisis bargaining, and Middle East politics.

Robert F. Trager is an Associate Professor of Political Science at UCLA. His primary field of expertise is in diplomatic communication, applied game theory, and international history. Professor Trager’s book on diplomacy is forthcoming with Cambridge University Press.

Authors’ note: We gratefully acknowledge support from the Institute for Global Cooperation and Conflict, the Lynde and Harry Bradley Foundation, and the UCLA Department of Political Science, as well as insightful comments from Allan Dafoe, Arman Grigorian, Ron Gurantz, Susan Hyde, Galen Jackson, Chaim Kaufmann, Kouke Imai, Jeff Lewis, Dom Levin, Paul MacDonald, Nuno Monteiro, Neil Narang, Kevin Narizny, Chad Nelson, Philip Potter, Steven Spiegel, Mike Tomz, Marc Trachtenberg, seminar participants at Harvard, Lehgh, and Yale Universitives, and the ISQ anonymous reviewers. Remaining errors and omissions are our own. The views expressed in this article are those of the authors and not necessarily those of the US government.


© The Author (2016). Published by Oxford University Press on behalf of the International Studies Association. All rights reserved. For permissions, please e-mail: journals.permissions@oup.com
We consider an instance when leader preferences might not take the commonly assumed form. Social psychologists, economists, and behavioral biologists have long explored how fairness motivates behavior, but scholars of IR generally neglect this mechanism. We examine how the prospects of unfair negotiated settlements can produce risk-loving behavior in leaders and how fairness concerns provide leaders with incentive not to pursue ever larger shares of a disputed good. We also expand on the literature on “voice,” in particular, to show that leaders have less incentive to accept settlements that appear imposed upon them. Thus, while aggressive public rhetoric can be used to signal credibility (Fearon 1994; Tomz 2007; Trager and Vavreck 2011) and mobilize support (for example, Baum 2002), such public rhetoric also makes it more difficult for the state’s adversaries to compromise and avoid war.3

We investigate the influence of fairness heuristics on approval and thus on leader incentives. Scholars cannot easily study these dynamics with observational data, however, owing to strategic selection effects (Lewis and Schultz 2003; Signorino 1999). Therefore, we use a randomized survey experiment administered on a large, nation-wide sample of the US adult population. The experiment involves a real-world—and still developing—dispute in which the United States and Russia have competing claims over a resource-rich territory in the Arctic. We independently and randomly vary three factors that influence public evaluations of presidential performance in international crises: (i) the aggressiveness of the foreign leader’s rhetoric, (ii) whether a negotiated settlement was reached or war occurred, and (iii) how much of the resource-rich territory each country ultimately obtained. After respondents read about the dispute and outcome, they were asked whether they approved or disapproved of how the president handled the situation.

The results provide three central contributions to the understanding of international crisis dynamics. First, non-satiated (monotonically increasing) utility functions do not always represent the preferences of leaders who seek to maximize voter approval. Contrary to commonly made assumptions, we find that negotiating for shares of the disputed good above 50 percent actually decreases presidential approval. In international contexts like the one analyzed below, democratic leaders lack a direct political incentive to strive for a highly favorable division of the good in question.

Second, risk-averse (globally concave) utility functions do not always represent the preferences of such leaders. For shares of a disputed good below 50 percent, in particular, we find that the implied leader preferences are risk loving (or convex). This is because many voters draw a much larger distinction between increases in the share of the disputed territory from 30 percent to 50 percent than between increases from 0 percent to 30 percent. Thus, domestic audiences generate incentives for peace by not demanding ever more of disputed goods; they also generate incentives for war, however, by inducing risk-loving preferences in leaders.

Third, foreign rhetoric dramatically affects presidential incentives for war. Aggressive foreign rhetoric increases domestic presidential approval following a successful war by 18 percentage points. Such rhetoric also decreases approval at compromises below a 50-50 split of the disputed territory and provides leaders with an incentive to fight harder to achieve victory. Thus, aggressive rhetoric can powerfully influence foreign leader incentives toward conflict.

We argue that scholars of IR should devote increased attention to the processes that generate non-satiated, risk-averse utility functions, as well as those that produce other sorts of preferences. Non-satiated preferences over disputed goods may not be the norm in international politics. Overall, our findings indicate a fundamental contradiction between two commonly made assumptions: that leaders seek to maximize their prospects of election and that, at the same time, leader preferences can be represented as rationalist theories of war commonly assume. Alongside the now substantial literatures on how the political contexts of conflicts affect public support (for example, Baum and Groeling 2010; Gartner 2008; Berinsky 2009; Gelpi, Feaver, and Reifler 2009; Horowitz and Levendusky 2011; Trager and Vavreck 2011; Jentleson 1992; Gottfried 2014) and leadership preferences (for example, Jervis 2004; Levy 1997; Larson 1985; McDermott, Fowler and Smirnov 2008), our results show that fairness heuristics and the rhetorical framing of disputes during the conflict process may be at least as important as material factors in understanding why some disputes result in war.4

We first discuss how fairness considerations influence preferences over political outcomes. The subsequent section addresses the inherent difficulty in measuring leader preferences and outlines our specific hypotheses. We then describe the survey experiment and research design and present and discuss the experimental results.

Fairness and Political Preferences

Voluminous evidence indicates that fairness concerns drive much of human interaction. In international politics, for example, fairness concerns appear to have dramatically restricted the bargaining space in the Arab-Israeli conflict (Goldbard 2009; Hassner 2009). Both sides have appealed to fairness discourses, often in calls to action or the rejection of compromises viewed as unjust. For instance, Marwan Barghouti, a prominent leader of Fatah’s young guard, proclaimed, “Nothing will help. Only a just agreement…and no one has the right to give up on it.”5

Behavioral economics literature demonstrates that individuals show a concern for fairness that motivates their decisions. Consider ultimatum games, where one party proposes a division of a monetary amount as an ultimatum and the other party can choose to accept the offer or decline it, in which case both get nothing. The standard game-theoretic solution suggests that the proposer should offer nothing to the respondent and the respondent should accept every positive amount. Yet, in laboratory settings, 20 percent to 30 percent of proposers offer a 50-50 split of the money (Camerer 1997; Güth, Schmittberger, 3See Casp (2001).

4All supporting data and Online Appendices are available at www.roberttra-ger.com. For other recent work on the origins of leader preferences, see, for example, Kertzer (2013), Fanis (2011), Chiozza and Goemans (2011), and Horowitz, McDermott, and Stam (2005).

5For work in the liberal tradition related to preference formation, see the literature cited in Moravcsik (1997).

In addition to works cited above, see Leventogu and Tarar (2005), Baum (2004), Baum and Groeling (2010), Trachtenberg (2012); but see also Kurizaki (2007).
and Schwarze 1982; Kahneman, Knetsch, and Thaler 1986; Thaler 1988; Guth 1995). Fairness concerns intensify when choices are made in public (Andreoni and Bernheim 2009). We find similar results across cultures (Oosterbeek, Sloof, and Van De Kuilen 2004), suggesting that they have fundamental biological bases that may even extend to other mammalian species (Brosnan and de Waal 2003).

Outside the laboratory, in public settings with shared bargaining power and repeated interactions, we likely see even stronger preferences for appearing fair. Indeed, many psychological studies describe “fairness heuristics” that regulate many aspects of social interaction. Rather than calculating an optimal action, individuals use much simpler considerations of fairness to guide action (Lind 2001). As a result, in the absence of another marker of fair division, the 50-50 norm proves common (Bloom 1986, Veugelers and Kesteloot 1996, Dasgupta and Tao 1998, Agrawal 2002, Deweaver and Roumasset 2002). Violations lead to dissatisfaction and protest. Evidence indicates, for instance, that equitable agreements help to sustain a durable peace following civil wars (Druckman and Albin 2011).

However, scholars of IR pay almost no attention to fairness as a determinant of actor behavior. The few exceptions largely focus on economic relations (Kapstein 2006, 2008) and international law (Franck 1995). Some recent scholarship also investigates how norms of justice impact international negotiations and the content of political agreements (Albin 2001; Zartman and Faure 2005; Albin and Druckman 2012).

We further apply two of the largest literatures that investigate the sources and effects of fairness judgments to the origins of preferences over international outcomes. In “equity theory,” perceptions of just or fair claims to goods relate to the amount of effort that individuals or groups put into the development of those goods. If neither side of a dispute puts in effort to develop a good or resource, both sides enjoy equal claim on the good in question; fairness corresponds to a 50-50 division. 6 While many critique equity theory, substantial evidence supports its core prediction that contribution influences fairness judgments—even if other factors also matter (Folger 1988). The raft of studies cited above also underscoring that the 50-50 outcome is perceived as the most fair in the absence of another fairness prime.

Other literature shows that having a “voice” in determining an outcome leads to a much more positive response to that outcome (for example, Thibaut and Walker 1975; Lind, Kanfer, and Earley 1990; Tyler 1989). Aggressive actions or rhetoric that denies voice lead to negative responses. Some consider this the single most replicated finding in all of the psychological literature on fairness. It also tracks with studies of the effects of honor replicated finding in all of the psychological literature on negative responses. Some consider this the single most

**Popular Approval and Leader Preferences**

We seek to understand how fairness heuristics influence leader preferences and incentives over different international crisis outcomes. Biases inherent in analyses of observational data render this a difficult task. Strategic selection effects provide one source of estimation bias. Leaders choose one policy from many choices, and a variety of factors influence their preferences over outcomes. Signorino (1999) demonstrates that most large-N statistical studies do not account for the statistical expectations implied by strategic interdependence. Some studies incorporate strategic interaction into statistical models by deriving estimators directly from extensive-form games (for example, Signorino 1999; Signorino and Yilmaz 2003; Lewis and Schultz 2003), but this approach has its limitations. Since players are never afforded the opportunity to play all moves because of the sequential nature of the formal model, the utilities associated with different outcomes cannot be placed on comparable scales (Lewis and Schultz 2003). Some strategies help fix these scales, but they rely on untestable restrictions. These estimators also require researchers to make precise choices about model specification—with dramatic effects on estimation results—with little theoretical or empirical foundation.

How then can we measure leader preferences in international crises? Public opinion tells us much about the political incentives of democratic actors because they are motivated by elections (Downs 1957; Aldrich, Sullivan, and Borgida 1989; Sobel 2001; Bueno de Mesquita, Smith, Siverson, and Morrow 2003). The public tends to be uninformed about general foreign policy issues, but in the United States (Holsti 1996) the public expresses its passions in times of crisis and war (Mueller 1973; Brody 1991; Zaller 1992; Berinsky 2009; Baum and Groeling 2010). Many studies indicate that public approval strongly influences democratic decision making (Edwards 1976; Ragsdale 1984; James and Oneal 1991; Kernell 1997; Ostrom and Job 1986), especially because approval correlates highly with the probability of winning elections (Campbell and Lewis-Beck 2008).

Political leaders do not blindly follow opinion polls. Expectations of voter approval weigh alongside other factors in their decision making. Leaders could ignore changes in approval on salient issues, such as national security, when in crises. But doing so risks serious costs that could jeopardize their political survival and the fulfillment of their policy agenda. Despite debate about whether democratic leaders can avoid public opinion constraints on their foreign policy choices through deceptive strategies, we know that public opinion strongly influences leader preferences and that deceptive strategies involve political risks.

Thus, we use presidential approval as a measure of leader preference. We take levels of approval at different crisis outcomes to represent the incentives of leaders over those outcomes and serve as a proxy for their preferences, at least for leaders who prefer to maximize voter approval (cf. Tomz 2007; Trager and Vavreck 2011; Horowitz and Levendusky 2012; Gottfried 2014; Schultz 2001). Thus, the empirically identifiable leader preference function that we discuss below is the mapping from crisis outcomes to presidential approval. We measure presidential approval in a randomized survey experiment. This allows us

---

6Canonical citations from the vast literature on equity theory include Adams (1965), Akerlof and Yellen (1990), and Berscheid, Boye, and Walster (1968).

7See Lind, Kanfer, and Earley (1990). According to Tyler (1989), this finding relates particularly to disputes as opposed to collective action environments.

8Chiozza and Goemans (2004), for example, infer leader preferences by analyzing political survival following crises and war but cannot rule out the possibility that leaders enter certain wars precisely because they are believed to be less likely to result in the leader’s ouster (regardless of their actual outcome).

9For both sides of the debate, see Schuessler (2010), Trachtenberg (2006), and Reiter (2012).
to avoid biases associated with strategic selection effects because we observe approval at all crisis outcomes.

**Political Incentives for Compromise over Divisible Goods**

We now turn to a discussion of the specific hypotheses we will test. The first two do not correspond to our expectations, but cohere with assumptions made in the IR literature on crisis bargaining. Many formal models of crisis bargaining assume that leader preferences in disputes can be represented by utility functions that are non-satiated and weakly risk averse, which is equivalent to weakly decreasing marginal returns in the disputed good. So, for example, prior to the outbreak of World War II, Hitler should have derived greater satisfaction from acquiring the first half of Czechoslovakia than from the second.\(^{10}\)

We use the term “risk aversion” in a specific technical sense: if an actor may choose a gamble over a smaller and larger amount of some good, risk aversion implies that the actor would prefer some certain middle amount to the gamble.

These assumptions imply that a mutually preferred division of a disputed good always exists. Thus, it matters a great deal if these assumptions turn out to be unsubstantiated in certain contexts. If, for instance, leader preferences are not weakly risk averse and the costs of war are small enough relative to potential gains, both sides may see no negotiated solution that they prefer to conflict. Rationalist explanations for leadership decisions to engage in conflict would not strictly require incomplete information, commitment problems, or issue indivisibilities (see Fearon 1995).\(^{11}\)

Bargaining equilibria would then have somewhat different properties; in some models, the amount of information conveyed by cheap-talk diplomacy would increase (Trager 2011).

Despite the fundamental importance of these assumptions, only a few studies rigorously examine their veracity. Chiozza and Goemans (2004) estimate a series of hazard models to ascertain the effects of international conflict on leader tenure. They find that leaders do not face a higher risk of being removed from office following wars than following crises. Thus, they conclude that “for contestant leaders, war is not negative-sum and not ex post inefficient” (2004, 616). Using an experimental approach, Trager and Vavreck (2011) investigate presidential approval for outcomes of a canonical international crisis bargaining model. They show that US presidents experience high mass approval when they win concessions from foreign states. These, however, constitute the entirety of the disputed good. Thus, these studies fail to reveal the preferences of leaders over the range of possible peaceful divisions of a disputed good.

Our first two hypotheses, therefore, follow from these commonly made, but untested, assumptions: that leader preference functions, which we operationalize as the mapping from crisis outcome to presidential approval, are globally, weakly concave (risk averse) and non-satiated.

**Hypothesis 1a (Non-Satiation):** Using leader popularity as a measure of leader preference, the implied leader utility function over shares of a disputed good is non-satiated.

**Hypothesis 2a (Risk Aversion):** Using leader popularity as a measure of leader preference, the implied leader utility function over shares of a disputed good is globally, weakly concave.

We expect, however, that the universal concern for fairness will be evident in public reactions to their leaders’ behavior in certain international disputes. This implies predictions that differ from the Non-Satiation and Risk Aversion hypotheses. Consider the prediction of equity theory in a conflict over a newfound resource in which neither side enjoys a prior claim. Equity theory predicts that a 50-50 division of the resource will be considered most fair and therefore most highly approved. Highly favorable bargaining outcomes would therefore prove less popular than more moderate bargaining outcomes—violating non-satiation. Thus, we expect presidential approval to be highest at a 50-50 split of a disputed good in such contexts.\(^{12}\) This is stated as Hypothesis 1b.

**Hypothesis 1b (Fairness Heuristic):** Using leader popularity as a measure of leader preference, the implied leader utility function will be highest at a 50-50 division of the disputed good (in the absence of a countervailing fairness prime).

Further, because a 50-50 split will be perceived as fair, it will garner much higher approval than surrounding outcomes. If this difference in approval is high enough, it implies convexity of the presidential utility function in some neighborhood below the 50-50 division. Thus, the implied leader utility function will not be globally concave. This is stated as Hypothesis 2b.\(^{13}\)

**Hypothesis 2b (Risk Acceptance):** Using leader popularity as a measure of leader preference, the implied leader utility function will be convex in some neighborhood below the 50-50 division of the disputed good (in the absence of a countervailing fairness prime).

**Foreign Aggression and Voice**

Many think that aggressive rhetoric produces benefits in international negotiations. It enhances the position of the threatening state by generating domestic “audience costs” if a state later backs down, allowing states to negotiate for greater concessions from adversaries (Fearon 1994; Leventoglu and Tarar 2005; Tomz 2007; Goddard 2006; Trager and Vavreck 2011). A reputation for aggressive foreign policy even makes it easier for a leader to initiate

\(^{10}\)O’Neill (2001) argues that risk preferences cannot be applied to territory, in part because different but equally sized portions are unlikely to have the same inherent value.

\(^{11}\)Powell (2006) points out that in infinite horizon models, costs of conflict can be arbitrarily small relative to the benefits of winning a war. Moreover, even when actors are risk acceptant, a negotiated settlement that randomly allocates the good with probability determined by the military balance is still preferable to war. As Powell notes, however, commitment problems make such solutions impracticable.

\(^{12}\)Fairness is not always considered to imply a 50:50 division. Disputes that relate to deep-rooted historical or cultural claims, for instance, are likely to construct fairness in particular ways through narrative. This implies directions for future research.

\(^{13}\)Note that when the procedure for determining an allocation remains unclear, as it is when agreements are arrived at through closed-door elite negotiation, the outcome will be a particularly strong signal of fairness, and small shares will be judged less fair, a dynamic known as the “Fair Outcome Effect” (for example, Van den Bos et al. 1997, Van den Bos 1999). Outcomes that participants judge as unfair for this reason are much less likely to be accepted in experimental settings (Camerer and Thaler 1995, Holland-Blumoff and Tyler 2008). We hypothesize that such bargains will be much less popular when leaders do accept them. This also drives convexity.
peace negotiations and conclude sustainable settlements. For instance, it may have taken a Nixon to go to China (Schultz 2005).

Yet, aggressive rhetoric may make adversaries less willing to compromise and may cause leader incentives to more closely align with those commonly assumed in rationalist studies. We hypothesize that when a side achieves a favorable outcome following their use of aggressive rhetoric, the other side will view that outcome as imposed. As noted earlier, the literature on voice implies that members of the latter group will view it as less fair and therefore prove more willing to take costly actions to oppose it. Thus, aggressive foreign rhetoric that appeals to threats of force rather than a fair allocation procedure will often produce a highly negative response to low shares in a negotiated outcome.

Relationally, aggressive rhetoric may engage the honor of an adversary with similar effects (O’Neill 1999, Nisbett and Cohen 1996). When a concession follows an aggressive threat, publics may perceive it as “backing down” and react negatively compared to a concession made through moderate negotiations. As a result, aggressive rhetoric makes concessions more politically costly. Thus, several related mechanisms imply the expectation that aggressive foreign rhetoric will diminish the popularity of presidents who make substantial concessions. This is stated as Hypothesis 3.

Hypothesis 3 (Rhetoric and Compromise): Aggressive rhetoric by foreign leaders will diminish presidential approval following less favorable compromise outcomes.

This hypothesis is consistent with studies that show that hostile rhetoric is a significant cause of conflict. For example, Zinnes (1966) observes that the receipt of hostile communications from a foreign leader will cause a decision maker to perceive a hostile environment, inducing her to reciprocate in kind. Lindskold, Betz, and Walters (1986) find that defection in a prisoner’s dilemma increases in the presence of verbal threats, insults, and challenges. Similarly, McDermott, Cowden, and Koopman (2002) conclude that the tone of an opponent’s message has a significant effect on arms races and the level of military procurement.

We also find suggestive evidence from the historical record. In the Cuban Missile Crisis, for instance, US President John F. Kennedy was willing to remove the Jupiter missiles from Turkey in return for the removal of Soviet missiles from Cuba. The President was not willing to do so if the trade would be perceived as a concession in response to aggressive Soviet diplomacy. “If you breathe a word of it in public,” the Soviets were told, “the deal’s off.” Similarly, during the 1973 Yom Kippur War, US Secretary of State Henry Kissinger warned the Soviets that “none [of the discussed compromises] we will do if [Soviet] pressures continue.” Kissinger added, “[US policy will be] extremely tough and to teach the facts of life to people who like to make great speeches and we will see what is more important - a speech or reality. I will be very brutal. That will be our strategy.”

Following a similar logic, we hypothesize that the factors that make compromise less attractive will also make a successfully prosecuted war more popular. Because the targets of aggressive rhetoric are less willing to compromise, they will be more accepting of conflict, especially one that achieves its objectives. As in studies of workplace behavior (Lind 2001), arms races, and crisis dynamics (McDermott, Cowden, and Koopman 2002; Zinnes 1966), they will be more willing to incur costs to resist the imposition of a settlement through force. Hypothesis 4 captures the effect of aggressive foreign rhetoric on approval following successful wars. We do not make a prediction about the effect of aggressive rhetoric on the popularity of unsuccessful wars. While aggressive foreign rhetoric makes the public more accepting of the necessity of war, it may lead the public to demand the war’s successful prosecution.

Hypothesis 4 (Rhetoric and War): Aggressive rhetoric by foreign leaders will increase presidential approval following successful wars.

If aggressive public demands make it more difficult for foreign leaders to back down, there are significant implications for the theory and practice of international negotiations. While such rhetoric demonstrates resolve, it may also generate conflict. Kurizaki (2007) shows that these factors can make costly private statements of commitment meaningful. Leaders may understand that decisions to keep discussions out of the public eye are not a sign of weakness, but sensible attempts to reach a peaceful resolution.

Research Design

Consistent with several past studies (for example, Tomz 2007; Trager and Vavreck 2011; Gottfried 2014), the survey begins by telling respondents that they will read about a situation “our country has faced many times in the past and will probably face again. Different leaders have handled the situation in different ways. We will describe one approach American leaders have taken and ask whether you approve or disapprove of that approach.” All respondents are then presented with a real-world source of tension in the international system: “The United States and Russia had a longstanding dispute over a resource-rich area in the Arctic.” Participants then learn about several

---


15In contrast, when leaders use conciliatory language that appeals to shared fairness norms, the literature on “reciprocal altruism” confirms that fairness judgments will require the need for treatment in kind (for example, Shalir and Taversky 1992, Gintis 2000).

16See Yamamori et al. (2008). This reaction is especially likely if personal impressions of world leaders indicate that such rhetoric is sincere (Hall and Yarhi-Milo 2012). These findings are consistent with the view that “states” can be thought of as emotional actors and that emotional dynamics are a principle lens through which to understand the international system (Sasley 2011). On longer-term effects of advancing justificatory reasons, see for instance, Risse 2006 and Mitten (2005, 41).

17See also Barnhart (2016). For a different view of the importance of rhetoric, see Krebs and Jackson (2007) and Morin and Gold (2010).


19Quoted in Trachtenberg (1999, 355). Also see Kennedy (1969, 106–109), Schlesinger (1978, 529-525), Bundy (1988, 452-454). In fact, if Khrushchev had turned down the US offer to remove the missiles from Turkey in return for the Soviet’s removing the missiles from Cuba, the US administration was willing to make a public trade, but, critically, it would have had to be proposed by the Secretary General of the United Nations, rather than through an aggressive Soviet demand. See Fursenko and Naftali (1999).

important elements of the crisis to establish the material potential of the territory, the ambiguity of US and Russian claims, and the possibility that this dispute could turn into a militarized conflict. Online Appendix A provides the full text of a sample survey.

The first treatment is the type of foreign rhetoric over the disputed territory. In the first arm, participants read that the Russian leader made an aggressive statement about the disputed territory: “This area is ours by right! If our rights are not respected at the negotiating table, we will see that they are respected on the battlefield!” In the second arm, participants read a moderate Russian statement: “This territorial dispute should be solved by negotiations with the participation of the international community.”

The second treatment is the outcome of the dispute. In one set of outcomes, participants learn that the United States and Russia reached a peaceful agreement that concluded the crisis and we specified the division of territory that went to each side. There were eleven randomly assigned divisions from 0 percent to 100 percent of the territory to the US side by increments of 10 percent. In the other set of outcomes, participants learn that negotiations ended abruptly and a militarized conflict occurred in which approximately 1,500 servicemen and servicewomen died on each side. At the conclusion of the conflict, the United States or Russia took complete control of the territory. Thus, the research design creates a fully crossed 2 x 13 experiment generating 26 distinct treatment groups (see Table 1). After reading one of the randomly assigned vignettes, respondents are asked whether they approve or disapprove of how the president handled the foreign policy situation. To help the respondents digest the information in the vignette, we provided a bullet point summary.

We took steps to help increase the external validity of the experiment. The dispute described in the vignette is based on an ongoing real-world dispute. For more than a decade, tension has been growing between states bordering the Arctic over who has legal rights to the region’s resources. Although the Arctic states have been issuing cooperative public statements thus far, some have been building up their military capabilities to operate in the area. For example, two US nuclear submarines were sent to patrol 150 nautical miles north of Prudhoe Bay, Alaska, in spring 2011, and Russia has increased its missile testing in the region (Macalister 2011; Huebert 2010). This growing military competition could escalate into a conflict such as the one we describe in the survey vignette. In the words of NATO’s top commander, Admiral James G. Stavridis, the situation presents an “icy slope toward a zone of competition, or worse, a zone of conflict” (Berkman 2013).

This scenario shares essential features with other cases. A similar dispute has arisen between Israel and Lebanon, for instance, President Saddam Hussein threatened Ayatollah Ruhollah Khomeini, the Supreme Leader of Iran, “Anyone who tries to put his hand on Iraq will have his hand cut off without hesitation” (Webb 1980, p. A2). During the Kosovo War, President Slobodan Milosevic of Yugoslavia sought to intimidate NATO forces: “We are willing to die to defend our rights as an independent sovereign nation” (Perlez 1999). Following the September 11 terrorist attacks, US President George W. Bush declared that “[US] demands are not open to negotiation or discussion. The Taliban must act, and act immediately. They will hand over the terrorists, or they will share in their fate” (Bush 2001).

We also controlled for elite cue effects that have been found to influence public evaluations of presidential performance (Zaller 1992; Baum and Groeling 2010; Berinsky 2008). We found to influence public evaluations of presidential performance (Zaller 1992; Baum and Groeling 2010; Berinsky 2008).

### Table 1. Randomly assigned treatments

<table>
<thead>
<tr>
<th>Category</th>
<th>Treatment</th>
</tr>
</thead>
<tbody>
<tr>
<td>Foreign Rhetoric</td>
<td>Aggressive: United States receives 100%, Russia receives 0%</td>
</tr>
<tr>
<td></td>
<td>Moderate: United States receives 90%, Russia receives 10%</td>
</tr>
<tr>
<td>Peaceful Outcomes</td>
<td>United States receives 80%, Russia receives 20%</td>
</tr>
<tr>
<td></td>
<td>United States receives 70%, Russia receives 30%</td>
</tr>
<tr>
<td></td>
<td>United States receives 60%, Russia receives 40%</td>
</tr>
<tr>
<td></td>
<td>United States receives 50%, Russia receives 50%</td>
</tr>
<tr>
<td></td>
<td>United States receives 40%, Russia receives 60%</td>
</tr>
<tr>
<td></td>
<td>United States receives 30%, Russia receives 70%</td>
</tr>
<tr>
<td></td>
<td>United States receives 20%, Russia receives 80%</td>
</tr>
<tr>
<td></td>
<td>United States receives 10%, Russia receives 90%</td>
</tr>
<tr>
<td></td>
<td>United States receives 0%, Russia receives 100%</td>
</tr>
<tr>
<td>War Outcomes</td>
<td>United States receives 0%, Russia receives 100%, and both countries lost approximately 1,500 servicemen and servicewomen in the militarized conflict.</td>
</tr>
<tr>
<td></td>
<td>United States receives 100%, Russia receives 0%, and both countries lost approximately 1,500 servicemen and servicewomen in the militarized conflict.</td>
</tr>
</tbody>
</table>

Hawkish leaders in international disputes. In August 1980, prior to the outbreak of the Iran-Iraq War, for instance, President Saddam Hussein threatened Ayatollah Ruhollah Khomeini, the Supreme Leader of Iran, “Anyone who tries to put his hand on Iraq will have his hand cut off without hesitation” (Webb 1980, p. A2). During the Kosovo War, President Slobodan Milosevic of Yugoslavia sought to intimidate NATO forces: “We are willing to die to defend our rights as an independent sovereign nation” (Perlez 1999). Following the September 11 terrorist attacks, US President George W. Bush declared that “[US] demands are not open to negotiation or discussion. The Taliban must act, and act immediately. They will hand over the terrorists, or they will share in their fate” (Bush 2001).

We also controlled for elite cue effects that have been found to influence public evaluations of presidential performance (Zaller 1992; Baum and Groeling 2010; Berinsky 2008). Mass opinion tends to approve of policies over which there is an elite consensus and to mirror the partisan divide otherwise. To isolate the effect of foreign rhetoric and crisis outcome on mass opinion, in the survey, opposition elites criticize the president in all treatments.

We administered the surveys to 2,500 US adults from June 21, 2012, to July 3, 2012. This nationwide sample came from SM Audience, an Internet-based panel highly reflective of the US adult population. The sample closely matches the US Census stratification according to gender, age, and geographic region. In fact, the sample matches the US Census more closely than many high-quality, Internet-based panels. The sample is also well-balanced on political ideology: 41 percent self-identified as liberal and 39 percent self-identified as conservative. Stratification by partisanship is comparable to other Internet-based panels. The mean partisanship score on a 5-point scale (Republican = 5) is 2.9 in the Annual National Election 2008-9 Panel and 2.8 in the current study. The subject pool does skew on some measures including a slightly higher Internet usage, income, and education, but no more than other frequently employed Internet-based panels. Online Appendix B provides comparative demographic statistics between our sample, the US Census, and the unweighted Annual National Election 2008-9 Panel.
which provides a benchmark for high-quality, Internet-based samples (Berinsky, Huber, and Lenz 2012).

Results

Concavity and Non-Satiation

The levels of presidential approval at all negotiated divisions of the disputed territory are shown in Figure 1. The bars around each point estimate represent 95 percent confidence intervals. The figure also includes a cubic smoothing spline divided from 0 percent to 50 percent and from 50 percent to 100 percent to display the relationship between the estimates in these zones. The data convincingly reject the Non-Satiation and Risk Aversion hypotheses and confirm the Fairness Heuristic and Risk Acceptance hypotheses. The preferences of leaders who prefer to maximize voter approval are not well-represented by globally concave, non-satiated utility functions, but rather are satiated at equitable divisions.

The negotiated solution that yields the highest estimate of presidential approval is the 50-50 split of the disputed resource-rich territory. The 80 percent approval rating at this outcome is distinguishable from approval at all other outcomes at the 0.05 confidence level (except for approval at a US share of 70 percent, which is significant at the 0.06 level). Approval rises steeply from 29 percent to 80 percent as the US share of the territory increases from 0 percent to 50 percent, but approval actually declines from 80 percent to 66 percent thereafter.

The satiation finding holds when the population is sub-setted along party lines. 22 Democratic approval of the president is 91.3 percent for a 50-50 split and 73 percent on average for shares greater than 50 percent of the territory, a difference of 18.4 percent (p<.0001). Republicans also appear to be more approving of 50-50 divisions (69.5 percent) than larger shares (66.3 percent), though we do not have enough data for the difference to be statistically significant. Thus, there is no evidence that either Republicans or Democrats show higher approval levels as the fraction of the disputed territory increases beyond 50 percent, as the Non-Satiation Hypothesis predicts.23

As a robustness test for satiation, we employ piecewise logistic regression (Muggeo 2003, 2008; Crawley 2007). We first fit a logistic regression model to the peaceful outcome data. The binary measure of presidential approval is modeled as a function of the US share of the disputed territory, a dummy variable for aggressive foreign rhetoric, the interaction between US share and aggressive rhetoric, and several demographic controls. Next, we use a Davies’ test (Davies 1987; Muggeo 2008) to determine if there exists a breakpoint or change in the parametric form. This procedure tests the null hypothesis that there is no broken or segmented relationship between the US share and presidential approval by comparing the slopes at potential breakpoints. If there is no breakpoint, the difference-in-slope of the fitted lines before and after the potential breakpoint is 0. The Davies’ test also estimates the “best” breakpoint for a given set of equally spaced potential breakpoints. 24 The Davies’ test rejects the null hypothesis of no breakpoint (p = 3.782e-16) and indicates that the best estimate for the break is at a 50-50 division of the territory. 25 The regression table is available in Online Appendix E.

Figure 1. Presidential approval at peaceful settlement outcomes. (Note. Each estimate corresponds to expected presidential approval for specified negotiated divisions of the disputed territory. Vertical bars provide 95 percent confidence intervals for each estimate. The solid line is a cubic smoothing spline divided from 0 percent to 50 percent and from 50 percent to 100 percent. Values below the dotted line demonstrate risk acceptance.)

22Figures illustrating the effect of partisanship and rhetoric on presidential approval at peaceful settlement outcomes can be found in Online Appendix C.

23As an additional robustness check, parametric regression analysis confirms that respondent demographic features are not driving these results. A figure illustrating the similarity of predicted levels of presidential approval from a fitted regression model to the approval levels shown in Figure 1 can be found in Online Appendix D.

24The test uses k “naïve” Wald statistics for the difference-in-slope, where k is the number of equally spaced potential breakpoints.

25For a sufficiently high number of k potential breakpoints, the estimated break occurs between a US share of 50 percent and 60 percent. Because the experiment specifies a change in share by increments of 10 percent, this implies a change in slope at the 50-50 division of the territory.
We then refit the model taking into account this segmented relationship. A likelihood ratio test confirms an improvement in fit over the linear model ($p = 9.17$ e-17). The coefficient for US share between 0 percent and 50 percent is positive ($\beta = 0.043$, $p < .001$), meaning that presidential approval rises as the US share increases in this range. The coefficient for US share greater than 50 percent, however, is negative ($\beta = -0.013$, $p < .05$). In other words, for favorable divisions of the territory, presidential approval decreases as the US share increases further. Because the analysis demonstrates a global maximum at the 50-50 split of the resource-rich territory, we reject the null hypothesis of non-satiation in favor of the Fairness Heuristic Hypothesis.26

Contrary to the Risk Aversion Hypothesis, the mapping from the US share of the disputed territory to presidential approval is not weakly concave over the entire region of peaceful settlements. If the president’s preferences could be represented by a weakly concave utility function, then approval at the 10 percent through 40 percent shares of the territory would fall at or above the dotted line drawn on Figure 1. Taking the uncertainty of the estimates of approval at 0 percent and 50 percent shares into account, the probability that approval at US shares of the territory strictly between 0 percent and 50 percent is in the range implied by weak concavity is rejected at the 0.05 certainty level for each share between 0 percent and 50 percent. As Figure 1 indicates, this is driven by the high levels of approval elicited by an equal division of the disputed territory. Approval at the 0-100 division and the 10-90 division are nearly equal and are not statistically distinguishable from each other. By contrast, survey respondents drew a distinction between a 40-60 division and a 50-50 division, with approval increasing by a huge 29 percentage points from one to the other ($p < .0001$). This finding holds irrespective of respondent party affiliation.

We perform two additional tests of concavity. First, we show that the data convincingly reject a model constrained to a simple, concave relationship between approval and US share. Model 1 of Table 2 models the binary measure of presidential approval as a function of the aggressive rhetoric treatment and the first four terms of the Taylor series of US share. To improve the model’s fit, we include a dummy variable coded 1 when the US share was 50 percent of the territory and 0 otherwise.27 Closely mirroring the approval levels in Figure 1, the model produces a highly convex relationship between share and approval over portions of the parameter space illustrated by the solid line in Figure 2. Taking estimated approval levels at 0 percent and 50 percent as endpoints, predicted approval levels at interior points are outside the range implied by concavity; this difference is highly significant over most of the interval.

We then tested this model against a constrained model in which the estimated relationship between approval and share is concave. Model 2 of Table 2 constrains the coefficients of the 3rd and 4th terms of the Taylor series on US share as well as the 50 percent dummy variable to be zero.

<table>
<thead>
<tr>
<th>Variable</th>
<th>Model 1</th>
<th>Model 2</th>
<th>Model 3</th>
<th>Model 4</th>
</tr>
</thead>
<tbody>
<tr>
<td>US Share</td>
<td>−2.88</td>
<td>6.12***</td>
<td>4.39***</td>
<td>4.71***</td>
</tr>
<tr>
<td>(US Share)^2</td>
<td>0.27***</td>
<td>−3.96e-02***</td>
<td>−2.65***</td>
<td>−2.84***</td>
</tr>
<tr>
<td>(US Share)^3</td>
<td>−3.76e-05**</td>
<td>(0.11)</td>
<td>(0.55e-02)</td>
<td>(0.83)</td>
</tr>
<tr>
<td>(US Share)^4</td>
<td>1.54e-05*</td>
<td>(1.69e-03)</td>
<td>(0.83)</td>
<td></td>
</tr>
<tr>
<td>50% Dummy</td>
<td>0.92***</td>
<td>1.07***</td>
<td>1.05***</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Variable</th>
<th>Model 1</th>
<th>Model 2</th>
<th>Model 3</th>
<th>Model 4</th>
</tr>
</thead>
<tbody>
<tr>
<td>Aggressive Rhetoric</td>
<td>−0.15</td>
<td>−0.12</td>
<td>−0.54**</td>
<td>−0.51*</td>
</tr>
<tr>
<td>Aggr. Rhet.</td>
<td>(0.09)</td>
<td>(0.09)</td>
<td>(0.27)</td>
<td>(0.27)</td>
</tr>
<tr>
<td>US Share</td>
<td>0.76</td>
<td>0.72</td>
<td></td>
<td></td>
</tr>
<tr>
<td>(US Share)^2</td>
<td>(1.18)</td>
<td>(1.24)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Aggr. Rhet.</td>
<td>0.08</td>
<td>0.10</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Respondent Party</td>
<td>0.82***</td>
<td>(0.10)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Respondent Age</td>
<td>0.10**</td>
<td>(0.05)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Respondent Income</td>
<td>0.10***</td>
<td>(0.03)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Respondent Gender</td>
<td>0.03</td>
<td>(0.10)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Respondent Education</td>
<td>0.06</td>
<td>(0.05)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Interceptor</td>
<td>−0.94***</td>
<td>−1.45***</td>
<td>−1.13***</td>
<td>−2.54***</td>
</tr>
<tr>
<td>N</td>
<td>2117</td>
<td>2117</td>
<td>2117</td>
<td>2903</td>
</tr>
</tbody>
</table>

Note: The dependent variable is the binary measure, Presidential Approval. All models were estimated using a logit function.

This forces a simpler functional relationship between share and approval. The estimation produces a globally concave fit to the data, represented by the dashed line in Figure 2. The estimates from this model represent leader preferences in the form they are commonly assumed to take and this therefore serves as an appropriate null hypothesis against which to evaluate the predictions of Model 1. Using a likelihood ratio test, we reject the null hypothesis that the model constrained to concavity is correct ($p < .0001$). Thus, the globally concave model is rejected in favor of the more complex relationship between share and approval, including a high degree of convexity for shares between 0 percent and 50 percent. If the 50 percent dummy variable is excluded from the unconstrained model, we still find a similar, locally convex relationship and the constrained model is again rejected at a similar significance level. Note that the Model 1 and Model 2 predictions for shares above 50 percent are very similar, meaning that the convex area of the parameter space in the unconstrained model is driving the rejection of the constrained model.28

To ensure that no quirk of our econometric specification drives this finding, we employ a second test using

26We also performed a simple test using a model of approval on a dummy variable indicator for the 50 percent share and a dummy variable for share treatments above 50 percent. Confirming the satiation hypothesis, we find that the coefficient on the 50 percent share dummy is larger and a chi-square test rejects the null hypothesis that the two coefficients are equal at the 0.001 level of certainty.

27Log likelihood ratio tests justify the inclusion of the 50 percent share dummy (chi-square distributed ratio test statistic = 25.67, $p = 1.14$ e-06), but do not justify the inclusion of any other share dummy.

28We also estimated a model of approval on US share and the square of US share over the 0 percent to 50 percent share range. Consistent with local convexity, we find that the coefficient on the squared term is positive ($p < .0001$).
non-parametric bootstrapping. At each value of US share, we randomly sample from our data with replacement to generate a new sample of the same size. This produces a new estimate of the functional mapping from share to approval. To generate a distribution of these functions, we repeat this procedure 100,000 times. We find that among this large distribution, not a single mapping is globally concave.\textsuperscript{29} Thus, a variety of tests strongly suggest that the leader utility function implied by voter approval is not weakly risk averse across all fractions of the disputed territory. Further, as we suspected, this appears to be driven by the high levels of approval at the 50 percent share, supporting the Risk Acceptance Hypothesis.

### Aggressive Rhetoric and War

Supporting the Rhetoric and Compromise Hypothesis, we find that aggressive rhetoric lowers approval at compromises below 50 percent shares of the territory. Over this range of peaceful outcomes, presidential approval is 31.5 percent when the Russian leader adopts aggressive rhetoric and 38.2 percent when the leader adopts moderate rhetoric, a difference of 6.7 percent (\(p = 0.028\)). Thus, aggressive foreign rhetoric makes it harder for a leader to offer a dramatic settlement involving substantial concessions to preserve the peace.\textsuperscript{30}

To further investigate how the interaction of foreign rhetoric and the negotiated share of the territory affect presidential approval, we again employ regression analysis. The model specification is similar to those studied above and is listed as Model 3 in Table 2. To estimate the interaction of aggressive rhetoric and the US share, we exclude two of the higher order Taylor series terms. Note that by leaving in the 50 percent dummy variable, we capture much of the convexity around the 50-50 division of the disputed good.

Figure 3 provides an illustration of the predicted effect of foreign rhetoric on presidential approval, mediated by the US share in the negotiated solution. As the figure demonstrates, aggressive foreign rhetoric decreases presidential approval by almost 10 percentage points at low negotiated shares of the territory. Even the upper bound of the 95 percent confidence interval represents a substantially significant 2 to 3 percentage point decrease in presidential approval at low US shares. As the share accorded to the United States increases, the magnitude of the predicted negative effect of aggressive foreign rhetoric on presidential approval diminishes. In fact, aggressive foreign rhetoric is predicted to increase approval at negotiated outcomes involving a 70 percent or greater US share. This predicted increase in approval at outcomes greater than 70 percent is not statistically significant, but the difference in the negative effect of aggressive foreign rhetoric at low negotiated US shares and the positive effect of aggressive rhetoric at high US shares is highly significant. Thus, the analysis demonstrates that aggressive rhetoric provides adversaries with an incentive to negotiate for larger shares of disputed goods. These results are robust to the inclusion of a range of controls for respondent characteristics, such as those shown in Model 4 of Table 2.

Aggressive rhetoric also increases the approval of presidents who prosecute successful wars, supporting the Rhetoric and War Hypothesis. When the Russian leader uses aggressive instead of moderate rhetoric, the expected approval of a successful war increases from 28 percent to 46 percent, a difference of 18 percentage points (\(p = 0.013\)).\textsuperscript{31} The data suggest that approval following an unsuccessful war is 6.3 percent lower when the Russian leader used aggressive rhetoric than when he used

\textsuperscript{29}An appropriate \(p\)-value is difficult to calculate because of the difficulty of specifying the appropriate null hypothesis. To get a sense of the \(p\)-value, however, consider the "sharp" null hypothesis that the share treatments have no effect on responses. Restricting attention to treatments in the 0 percent to 50 percent share range, finding no concave functions in a sample of 2,000 would imply \(p<.01\). We find that not a single mapping is concave over the 0 percent to 50 percent range in the bootstrapped sample of 100,000 mappings. Thus, the \(p\)-value associated with the sharp null is extremely small. The \(p\)-value associated with any reasonable null, corresponding to the hypothesis that the mapping is in fact concave, would be smaller still.

\textsuperscript{30}Online Appendix G illustrates the effect of foreign rhetoric on presidential approval at peaceful settlement outcomes below 50 percent, at 50 percent, and above 50 percent.

\textsuperscript{31}The shift in approval at the successful war outcome appears to be mediated by partisanship (Online Appendix H). Republican approval of the president held at 55-56 percent for a successful war regardless of the type of foreign rhetoric, but Democratic approval soared by more than 40 percentage points (\(p<.0001\)) when the Russian leader switched from moderate to aggressive rhetoric.
moderate rhetoric, but these levels of approval are statistically indistinguishable from each other ($p=0.278$). These effects are illustrated in Figure 4.

Foreign rhetoric therefore incentivizes presidents to escalate conflicts and fight harder, as shown in Figure 4. Surprisingly, given the literature on the effects of success in conflict on presidential approval (Gelpi, Feaver, and Reifler 2009, Mueller 1973, Berinsky 2009), when a foreign leader adopts moderate rhetoric, success in war increases presidential approval by only a statistically insignificant 4 percentage points compared to approval following a military defeat. In contrast, when a foreign leader adopts aggressive rhetoric, a successful war increases approval by a gigantic 28 percentage points compared to approval following a military defeat.

The net effect of aggressive rhetoric at the war and compromise outcomes is a substantial shift in leader incentives toward a militarized dispute. To demonstrate when a leader has an incentive to go to war, we calculated the expected approval of war for different predicted probabilities of victory. Figure 5 provides a comparison of these estimates to the expected approval of different negotiated outcomes. The vertical axis represents the difference between the expected approval from war and the approval at a particular peaceful outcome. For values above 0 percent, therefore, a leader has an incentive to go to war. As the figure illustrates, the US president never has an incentive to go to war when moderate foreign rhetoric is used. Though making large concessions is not popular, approval levels for all negotiated settlements are still higher than the expected approval of engaging in a costly conflict, whatever the probability of victory.

When the Russian leader adopts aggressive rhetoric, however, the dynamics of presidential approval may give the American president an incentive to go to war. For US shares below 40 percent of the disputed territory, the president begins to have an incentive to engage in war if the probability of victory is high. For a 30 percent share, this incentive is quite modest and requires a very high degree of confidence in victory, $Pr(\text{Victory}) > 81\%$. The probability of victory required to provide an incentive for war is much lower when conflict is evaluated against lesser negotiated US shares of the territory. For a peaceful division involving a 20 percent US share, the best option for the president is war if the probability of success in the conflict is at least 40 percent. For US shares below 20 percent of the territory, the president’s approval is maximized by bringing the country into a war for nearly any positive probability of victory. Thus, aggressive foreign rhetoric may become a cause of war.

**Conclusion**

Current approaches to crisis bargaining demonstrate that even strictly risk-averse leaders can find themselves fighting an inefficient war due to incomplete information, a commitment problem, or an issue of indivisibility. But, we argue, other factors also influence leader incentives and calculations. The effect of fairness heuristics and leader rhetoric on the construction of conflicts, public preferences, and leader incentives receives comparatively little attention from empirically-minded scholars (but see McDermott et al. 2008). This study demonstrates that such factors will likely have large effects on whether states reach peaceful compromises or go to war.

Contrary to nearly all models of international politics, we find that in situations of the type described in the survey, if leaders prefer to maximize voter approval, then their preferences are not well-represented by insatiable or strictly increasing utility functions. Public opinion shows a strong concern for fairness—rather than simply a greater share of the spoils—especially when foreign leaders use moderate rhetoric in a dispute.

---

32Recall that 1,500 soldiers die on each side, whether or not the war is won. In future work, it would be interesting to analyze the effects of other experimental treatments related to the war outcomes, for example, the impact of differing levels of casualties.

33It is important to note that though leaders may want to maximize their chances of election, they do not always know the best way to achieve it.
These results raise important questions about fairness in IR. Does the public prefer an equitable division in the absence of a fairness prime for moral reasons? Does it believe that perceptibly unfair divisions will lead to more conflict in the future? How does self-interest condition fairness considerations and responses to rhetoric? Studies show that self-interest influences even the perception of fairness (for example, Dawes and Thaler 1988, 195; Rabin 1998, 16; Buchan, Croson, and Johnson 2004). We examined a conflict between two significant powers that each have the ability to harm the other. Perhaps the publics of weaker powers would react differently when faced with aggressive rhetoric from stronger powers. Or, perhaps other publics hold fundamentally different understandings of fairness, providing a different set of incentives to leaders (Henrich, Heine, and Norenzayan 2010). Future research should consider these kinds of questions.

Although the results confirm our theoretical expectations, we should note several caveats. The vignettes did not prime respondents to view the disputed territory as rightfully belonging exclusively to their side. We intentionally anchored fairness around an equal division of the good by introducing ambiguity surrounding each country’s claims to the territory. Yet, fairness does not always imply an equal division. The history of some conflicts may imply different understandings of fairness. Thus, we encourage researchers to consider other scenarios in which leader preferences and incentives may not take commonly assumed forms.

The experimental results suggest that a conflict of interest and a history of dispute are not sufficient to generate non-satiated leader preferences over divisions of the good. The strategic processes and informational asymmetries that have been studied extensively in the field are likely sometimes necessary, but far from sufficient, to generate conflict. Our results suggest that if non-satiated utility functions ever represent leader preferences and incentives, it is more likely the product of particular social processes of conflict construction rather than a direct result of material contexts. These processes may be closely related to those that result in the construction of disputed goods as “indivisible” (Goddard 2009; Hassner 2009), but the results above show that even the notion that “more is better” may not be the norm in international disputes. This is prior to questions about indivisibility. The investigation of these specific social processes is another fruitful avenue for continued research.

A related caveat derives from the fact that international disputes often involve more than one issue dimension. Peaceful agreements often involve complex packages of compensation involving tradeoffs, logrolls, and side-payments (Fearon 1995, 382; Trager 2011). In such contexts, publics may have more difficulty perceiving what fairness implies. Expanding the analysis to include such scenarios and exploring which factors affect the standards of fairness in a variety of international contexts is an important direction for further research.

Figure 5. Aggressive rhetoric paving the way for war. (Note. Values above 0 indicate when the US president has an incentive to go to war. The shaded area provides 95 percent confidence intervals.)
The experimental results also demonstrate that leader preferences may not always be well-represented by risk-averse utility functions. For US shares below 50 percent, the reactions of the American public suggest presidential incentives consistent with a strongly convex utility function. Similarly, mass opinion in the United States draws little distinction between 0 percent and 30 percent shares, giving the president comparably low marks for any division in this range. In such contexts, there may not be a negotiated solution that both sides prefer to conflict.

Lastly, we demonstrate the profound effect foreign rhetoric can have on leader incentives in international crises. If the Russian leader adopts aggressive rhetoric, mass opinion creates incentives for war below a 40 percent offer when the prospects of victory are sufficiently high. In contrast, when the Russian leader adopts moderate rhetoric, the US president never has an incentive to engage in a costly war. Thus, the rhetorical framing of disputes should be considered alongside material factors in understanding why some disputes result in war.

The findings on the role of aggressive rhetoric indicate that leaders are faced with a dilemma that is central to the dynamics of international crisis bargaining. On the one hand, leaders have an incentive to moderate their rhetoric to better enable their adversaries to compromise. On the other hand, leaders motivate their own publics in ways that challenge the status and rights of their adversaries. Leaders must also convince adversaries of their resolve to take actions, and this too may involve public statements that increase the risks of conflict. Kurizaki (2007) points out, this dynamic provides an incentive for leaders to communicate in private, and private communications in such contexts will be informative. These factors may also relate to Downes and Secher’s (2012) findings that democracies issue few threats and that these threats are no more successful than threats from autocratic regimes. The sorts of rhetoric that rally domestic democratic constituencies may be opposite to the sorts of rhetoric that most effectively influence adversaries.

The assumptions that utility functions representing leader preferences over a disputed good are continuously increasing and weakly concave remain central to formal models and IR scholarship. But no study to date has systematically investigated the marginal utility of bargaining outcomes in the context of international crisis behavior. In this article, we addressed these issues through a randomized survey experiment administered on a sample of 2,500 American adults that closely matches the US population. We find that for leaders who prefer to maximize voter approval, these assumptions do not hold; their preferences are better characterized by satiation and local risk acceptance. Further, aggressive foreign rhetoric dramatically increases the expected utility of war and decreases the utility of substantial compromise. If leaders do have an incentive to maximize their state’s share of a disputed good, this incentive either does not directly derive from public reactions to leader actions, or it results from particular processes of conflict construction rather than directly from material contexts.

These results have important implications for our understanding of the current international security environment and the factors that influence conflict behavior in general. If India or Pakistan returns to bellicose rhetoric over Jammu and Kashmir, or if India or China, Russia or Japan, the United Kingdom or Spain, or Ethiopia or Eritrea, among many others, employs violent rhetoric in their disputes, the effects of such rhetoric on the preferences of foreign publics imply that mutually preferred peaceful settlements will be harder to find. More generally, very different factors from those pointed to in current rationalist explanations of conflict may be central to understanding why some disputes turn into wars. Preference-based explanations for behavior are sometimes thought to be lacking and, indeed, to verge on tautology. For instance, it appears unsatisfying to explain wars with the argument that the actors involved preferred war. Yet, when systematic determinants of the preferences and incentives of leaders are uncovered, such explanations are of great interest. Our findings show that leader incentives are strongly influenced by factors that do not derive directly from material contexts. A leader’s rhetoric not only influences her own future incentives and the beliefs of foreign governments about her intentions, but also directly influences the incentives of foreign leaders as well.

References


54 Of course, a foreign leader might prefer conflict, and his or her choice of aggressive rhetoric could reflect that preference.


