
Domestic Audience Costs in International Relations: An Experimental Approach

Michael Tomz

Abstract What makes international threats credible? Recent theories point to domestic audience costs—the domestic price a leader would pay for making foreign threats and then backing down. This article provides the first direct evidence of audience costs. The analysis, based on experiments embedded in public opinion surveys, shows that audience costs exist across a wide range of conditions and increase with the level of escalation. The costs are evident throughout the population, and especially among politically active citizens who have the greatest potential to shape government policy. Finally, preliminary evidence suggests that audience costs arise because citizens care about the international reputation of the country or leader. These findings help identify how, and under what conditions, domestic audiences make commitments credible. At the same time, they demonstrate the promise of using experiments to answer previously intractable questions in the field of international relations.

What makes international commitments credible? The answer may lie, in part, at the intersection of foreign affairs and domestic politics. Recent models of international relations assume that leaders would suffer “domestic audience costs” if they issued threats or promises and failed to follow through. Citizens, it is claimed, would think less of leaders who backed down than of leaders who never committed in the first place. In a world with audience costs, the prospect of losing domestic support—or even office—could discourage leaders from making empty threats and promises. The concept of domestic audience costs is now central to theories about military crises, and researchers have incorporated similar ideas into

I thank Time-sharing Experiments for the Social Sciences (TESS), the National Science Foundation (CAREER Grant SES-0548285), and the Center for Advanced Study in the Behavioral Sciences for financial support. Colleagues at Knowledge Networks provided invaluable assistance in fielding the surveys. For helpful comments I am grateful to Jim Fearon, Page Fortna, John Freeman, Jon Krosnick, Skip Lupia, Helen Milner, Diana Mutz, Ken Scheve, Ken Schultz, Jas Sekhon, Alastair Smith, Paul Sniderman, Rob Van Houweling, Jonathan Wand, Jessica Weeks, and the anonymous referees. I also thank seminar participants at Berkeley, CASBS, Columbia, Duke, NYU, Rice, and Yale, and conference participations at the International Studies Association and the TESS meetings at the University of Pennsylvania.

models of alliances, economic sanctions, foreign trade, foreign direct investment, monetary commitments, interstate bargaining, and international cooperation more generally.¹

Despite the prominence of audience costs in international relations theories, it remains unclear whether and when audience costs exist in practice. Most empirical work on the topic is indirect. Fearon conjectured that audience costs are higher in democracies than in autocracies and explained why this gap would cause the two types of regimes to behave differently.² Researchers have, therefore, checked for correlations between democracy and foreign policy.³ Although valuable, these tests do not reveal whether the effects of democracy stem from audience costs or from other differences between political regimes.

One could try to study audience costs directly, perhaps by examining the historical fate of leaders who issued threats and then backed down. The problem, which international relations scholars widely recognize, is strategic selection bias.⁴ If leaders take the prospect of audience costs into account when making foreign policy decisions, then in situations when citizens would react harshly against backing down, leaders would tend to avoid that path, leaving little opportunity to observe the public backlash. It would seem, therefore, that a direct and unbiased measure of audience costs is beyond reach.

This article aims to solve the empirical conundrum. The analysis is based on a series of experiments embedded in public opinion surveys. In each experiment, the interviewer describes a military crisis. Some participants are randomly assigned to a control group and told that the president does not get involved. Others are placed in a treatment condition in which the president escalates the crisis but ultimately backs down. All participants are then asked whether they approve of the way the president handled the situation. By comparing approval ratings in the “stay out” and “back down” conditions, one can measure audience costs directly without strategic selection bias.

In the remainder of this article, I demonstrate that constituents disapprove of leaders who make international threats and then renege. I further explain why many leaders regard disapproval as a political liability. Finally, as a step toward deepening our theoretical as well as empirical understanding of audience costs, I investigate why citizens react negatively to empty threats.

1. The seminal article is Fearon 1994. See also, on military crises, Schultz 2001a; and Smith 1998; on alliances, see Gaubatz 1996; and Smith 1996; on economic sanctions, see Dorussen and Mo 2001; and Martin 1993; on trade, see Mansfield, Milner, and Rosendorff 2002; on foreign direct investment, see Jensen 2003; on monetary commitments, see Broz 2002; on interstate bargaining, see Leventoğlu and Tarar 2005; and on the role of audience costs in international cooperation in general, see Leeds 1999; Lipson 2003; and McGillivray and Smith 2000.

2. Fearon 1994.

3. See, for example, Eyerman and Hart 1996; Gelpi and Griesdorf 2001; Partell and Palmer 1999; and Prins 2003.

4. See Baum 2004; and Schultz 2001b.

Do Audience Costs Exist?

Two questions are fundamental to theories about domestic audience costs: would constituents disapprove if their leader made false commitments, and by what means would disapproving citizens hold their leader accountable? The first question is analytically prior and the focus of this article; complementary work by others examines the secondary question of accountability.⁵ Throughout this article, I use the term audience costs as shorthand for the surge in disapproval that would occur if a leader made commitments and did not follow through. I adopt this terminology with the understanding that changes in approval are more consequential in some political systems than in others.

There is much speculation about whether audience costs exist at all. Some analysts hypothesize that empty commitments would provoke a negative public reaction.⁶ Citizens, it is argued, believe that hollow threats and promises undermine the country's reputation; that empty commitments are dishonorable and embarrassing; or that inconsistency is evidence of incompetence.

Other analysts argue that citizens would not disapprove of committing and backing down. They point out that some citizens pay little attention to foreign policy, and others focus on final outcomes rather than the sequence of threats and promises *in medias res*.⁷ Finally, even citizens who pay careful attention to every jab and parry may forgive leaders for making false commitments. After all, anyone who has played poker understands that bluffing can be an optimal strategy. "Why, then, would constituents punish leaders whose bluffs are sometimes called?"⁸

Citizens may even prefer leaders who try before conceding over leaders who forfeit at the outset. Walt, for example, points out that citizens may "reward a leader who overreaches at first and then manages to retreat short of war. Thus the British and French governments did not suffer domestic audience costs when they backed down during the Rhineland crisis of 1936 or the Munich crisis of 1938, because public opinion did not support going to war."⁹ Walt's historical examples raise an interesting possibility: leaders may gain points at home by escalating before giving up, instead of giving way immediately.

Do citizens typically respond with scorn, indifference, or praise when their leaders commit without following through? Until this is known, one cannot understand the effects of publicly committing before a domestic audience. If audience costs prove to exist under general conditions, this discovery would provide—for the first time—empirical microfoundations for a broad class of models in

5. For a review and important extension of the literature on accountability and audience costs in democracies and autocracies, see Weeks forthcoming.

6. See, for example, Fearon 1994; Guisinger and Smith 2002; and Smith 1998.

7. Brody 1994, 210.

8. Gowa 1999, 26. See also Desch 2002, 29–32; Ramsay 2004; Schultz 1999, 237; and Slantchev 2006.

9. Walt 1999, 34.

international security and political economy. The discovery would also suggest profitable avenues for new research, especially if the domestic reaction to flip-flopping varied systematically with characteristics of the situation or the audience. If, on the other hand, citizens showed no stronger preference for leaders who avoided commitments than for leaders who committed and subsequently reneged, one would need to rethink how leaders send signals and make commitments in world affairs.

Methods

To study audience costs directly while avoiding the problem of selection bias, I designed and carried out a series of survey experiments. The first experiment was administered to a nationally representative random sample of 1,127 U.S. adults in 2004. (Sampling methods are discussed in the Appendix.) All participants in the Internet-based survey received an introductory script: “You 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 U.S. leaders have taken, and ask whether you approve or disapprove.”¹⁰

Participants then read about a foreign crisis in which “a country sent its military to take over a neighboring country.” To prevent idiosyncratic features of the crisis from driving the results, I randomly varied four contextual variables—regime, motive, power, and interests—that have been shown to be consequential in the international relations literature.¹¹ The country was led by a “dictator” in half the interviews and a “democratically elected government” in the other half. The attacker sometimes had aggressive motives—it invaded “to get more power and resources”—and sometimes invaded “because of a long-standing historical feud.” To vary power, I informed half the participants that the attacker had a “strong military,” such that “it would have taken a major effort for the United States to help push them out,” and told the others that the attacker had a “weak military,” which the United States could have repelled without major effort. Finally, a victory by the attacking country would either “hurt” or “not affect” the safety and economy of the United States.

Having read the background information, participants learned how the U.S. president handled the situation. Half the respondents were told: “The U.S. president said the United States would stay out of the conflict. The attacking country continued to invade. In the end, the U.S. president did not send troops, and the attacking

10. The full text of all experiments is available at (<http://www.stanford.edu/~tomz>). Accessed 16 July 2007.

11. The literature on these four variables is vast. Herrmann and Shannon 2001; and Herrmann, Tetlock, and Visser 1999 discuss the impact of these variables on elite and mass opinion.

country took over its neighbor.” The remaining respondents received a scenario in which the president made a threat but did not carry it out: “The U.S. president said that if the attack continued, the U.S. military would push out the invaders. The attacking country continued to invade. In the end, the U.S. president did not send troops, and the attacking country took over its neighbor.” The language in the experiment was purposefully neutral: it objectively reported the president’s actions, rather than using interpretive phrases such as “backed down,” “wimped out,” or “contradicted himself,” which might have biased the research in favor of finding audience costs.¹²

After displaying bullet points that recapitulated the scenario, I asked: “Do you approve, disapprove, or neither approve nor disapprove of the way the U.S. president handled the situation?” Respondents who approved or disapproved were asked whether they held their view very strongly, or only somewhat strongly. Those who answered “neither” were prompted: “Do you lean toward approving of the way the U.S. president handled the situation, lean toward disapproving, or don’t you lean either way?” The answers to these questions implied seven levels of presidential approval, ranging from very strong disapproval to very strong approval.

By design, the experimental groups differed in only one respect: whether the U.S. president escalated the crisis before letting the attacker take over its neighbor. For this reason, any systematic difference in presidential approval was entirely due to the path the president took, not to variation in background conditions or the outcome of the crisis.

This experimental approach offers distinct advantages, including the opportunity to measure audience costs directly without selection bias. Nonetheless, the approach is not infallible. Indeed, experiments are vulnerable on the dimension where observational data are most compelling: external validity. Do citizens behave differently in interviews than in actual foreign policy crises? If so, do the experiments in this article understate or overstate the magnitude of audience costs? It is hard to say for sure. Ultimately, the best way to make progress on complicated topics is to analyze data from multiple sources. As the first of their kind, the experiments in this article provide new insights to complement what others have found with historical data.

Findings: Direct Evidence of Audience Costs

The experiments described above offer a new way to test competing conjectures in the literature. If audience costs exist, respondents who receive the vignette in which the president stayed out should approve more than respondents who learn

12. The experiment also avoided language that might have reduced audience costs, either by criticizing the president who stayed out or by praising the leader who escalated the crisis.

that the president threatened and yielded. If, on the other hand, citizens do not disparage leaders for getting caught in a bluff, levels of approval should be approximately the same in the two experimental groups. Finally, if leaders score points at home by showing at least some effort abroad, popularity should be higher in the “empty threat” scenario than in the “stay out” scenario.

Which of these conjectures best fits the data? Before answering that question, I confirmed that the treatment and control groups were balanced on baseline covariates that could affect presidential approval. Using a variety of parametric and nonparametric methods, I assessed balance with respect to demographic variables such as gender, age, education, income, and race. I also judged whether the groups were politically balanced by exhibiting similar patterns of party identification, interest in politics, involvement in politics (for example, voter registration, voter turnout, and political activism), and vote choices in the previous two presidential elections. Given that the survey focused on military intervention, I further checked for equality in attitudes toward internationalism and the use of force, and in history of military service. Finally, I looked for balance across contextual variables: the stakes for the United States and the motive, the power, and the domestic political regime of the invading country. Due to randomization, not one of these baseline variables was significantly different in the treatment group than in the control group.¹³

After verifying the existence of balance between the treatment and control groups, I examined how the public responded to each path the president traveled.¹⁴ The results, in Table 1, provide unambiguous evidence of audience costs. For each presidential strategy, the table gives the percentage of respondents who disapproved, approved, or expressed an intermediate view. As the table shows, the president who issued an empty threat (column 1) was significantly less popular than the president who never got involved (column 2). The empty threat caused both strong and moderate disapproval to grow, and it led both moderate and strong approval to shrink. Evidently, backtracking on a verbal threat evoked an adverse public reaction. The political consequences of swings in approval—their effects on presidential power and incentives—are discussed later in the article.

The final two columns of Table 1 summarize the magnitude of the effects. Compared to a baseline condition in which the president stayed out, the decision to threaten and not follow through caused disapproval to swell by 16 points, with

13. The literature on causal inference emphasizes the importance of assessing balance from as many angles as possible. I conducted numerous hypothesis tests, including Fisher’s Exact Test and Beta-Binomial tests for differences in proportions; t-tests for differences in means; and bootstrapped Komolgorov-Smirnov tests for distributional inequality. The p-values associated with these tests were always greater than .10. I also used more subjective methods, including visual inspection of empirical quantile-quantile plots, means, and variances, which again suggested good balance between treatment and control groups. I thank Jas Sekhon and Daniel Ho for helpful discussions about these balance metrics.

14. Due to randomization, there is little need for elaborate statistical models with batteries of control variables. One can obtain unbiased estimates of the treatment effect via cross-tabulation.

TABLE 1. *The domestic political cost of making empty threats*

	Public reaction to empty threat (%)	Public reaction to staying out (%)	Difference in opinion (%)	Summary of differences (%)
<i>Disapprove</i>				
<i>Disapprove very strongly</i>	31 (27 to 35)	20 (17 to 23)	11 (6 to 17)	} 16 (10 to 22)
<i>Disapprove somewhat</i>	18 (14 to 21)	13 (10 to 16)	5 (0 to 9)	
<i>Neither</i>				
<i>Lean toward disapproving</i>	8 (6 to 11)	9 (7 to 11)	0 (-3 to 3)	} -4 (-9 to 2)
<i>Don't lean either way</i>	21 (17 to 24)	21 (18 to 24)	0 (-5 to 4)	
<i>Lean toward approving</i>	8 (6 to 11)	11 (9 to 14)	-3 (-6 to 0)	
<i>Approve</i>				
<i>Approve somewhat</i>	8 (5 to 10)	13 (11 to 16)	-6 (-9 to -2)	} -12 (-17 to -8)
<i>Approve very strongly</i>	6 (4 to 9)	13 (10 to 16)	-7 (-10 to -3)	

Notes: The table gives the percentage of respondents who expressed each opinion. Bayesian 95 percent credible intervals appear in parentheses. Sample size was 477 in the “empty threat” scenario and 650 in the “stay out” scenario. The empty threat scenario involved only a verbal threat; the president did not display or use military force.

almost the entire posterior density being between 10 and 22.¹⁵ At the same time, the percentage of fence-sitters (citizens who neither approved nor disapproved) fell by 4 points, and the share of approvers dropped by 12 points with a posterior probability interval from 8 to 17. Thus, mere threats—without any military deployment or use of force—exposed the president to potentially widespread disapproval.

15. In the tables and in the text, I quantify the level of uncertainty about quantities of interest by reporting 95 percent Bayesian credible intervals (intervals that contain the quantity of interest with probability .95). One can think of the distribution of responses in each experimental group as having a multinomial sampling distribution in which there are k levels of approval, and parameters $\theta_1, \theta_2, \dots, \theta_k$ give the probabilities of falling into each category. A natural distribution for prior beliefs about θ is the k -dimensional Dirichlet, the conjugate prior for the multinomial. With so little prior knowledge about the existence and magnitude of audience costs, I used noninformative priors in which all Dirichlet parameters were set to .5. (These values correspond to the Jeffreys prior, which is noninformative and invariant under transformation. All findings in this article are robust to the use of other diffuse priors, such as the uniform distribution in which all Dirichlet parameters are set to 1.) In this setup, the posterior distribution has a convenient Dirichlet form. By taking random draws from Dirichlet distributions, I simulated the posterior distribution and computed the 95 percent credible interval for each proportion, difference between proportions, and ratio of proportions. See Gelman et al. 2004, 83–84.

Do Audience Costs Increase with the Level of Escalation?

The previous experiment established that even mild acts of escalation—making verbal threats—can set the stage for substantial audience costs. I now investigate whether public sensitivity to renegeing increases with the level of hostility. If so, leaders can send progressively stronger signals by ratcheting crises to higher levels. The literature on militarized interstate disputes (MIDs) distinguishes three levels of escalation prior to war.¹⁶ *Threats* are verbal indications of hostile intent, such as the ultimatum in the first experiment. The next rung on the escalatory ladder is a *display* of force, a military demonstration without combat. Finally, a *use* of force is defined as an active military operation against the foreign target.

Do leaders risk higher audience costs when they display or use force? I investigated this question by expanding the set of presidential responses. In one new scenario, the president “sent troops to the region and prepared them for war. The attacking country continued to invade. In the end, the U.S. president did not send our troops into battle, and the attacking country took over its neighbor.” In another scenario, the president not only threatened and displayed force, but also “ordered U.S. troops to destroy one of the invader’s military bases. U.S. troops destroyed the base, and no Americans died in the operation. The invasion still continued. In the end, the U.S. president did not order more military action, and the attacking country took over its neighbor.” The final scenario was identical, except that “20 Americans died in the operation.” The new scenarios were administered to a random sample of an additional 1,036 U.S. adults.¹⁷

Two features made these scenarios appropriate for testing the hypothesis that audience costs increase with the level of hostility. First, the new scenarios differed only in the approach the president took. In all other respects, including background circumstances and the outcome of the crisis, the extra scenarios were identical to each other and to the “stay out/verbal threat” vignettes discussed earlier. Second, the more hostile scenarios nested the less hostile ones: the vignette about the display of force included a threat to use force, and vignettes about the use of force mentioned previous attempts to threaten and display power. Any extra audience costs should, therefore, be due to layering on higher levels of escalation.

Table 2 summarizes the public reaction associated with each level of escalation. As before, I calculated the percentage of respondents who disapproved either strongly or somewhat when the president escalated and backed down and subtracted

16. Jones, Bremer, and Singer 1996.

17. By design, approximately 40 percent of the fresh sample received the “display of force” vignette and the remaining 60 percent were split evenly between the two “use of force” scenarios. The demographic and political profiles of these new treatment groups, and the contextual information they considered, closely matched the benchmarks in the “stay out” control group. Imbalances arose no more often than implied by chance, and in any case were relatively small. The conclusions in this article remain the same, therefore, after using multivariate logistic or linear regression to adjust for imbalances that might have arisen during the randomization process.

TABLE 2. *Domestic audience costs at four levels of escalation*

<i>Level of escalation</i>	<i>Absolute audience cost (%)</i>	<i>Relative risk of disapproval (%)</i>
<i>Threat of force</i>	16 (10 to 22)	1.5 (1.3 to 1.7)
<i>Display of force</i>	16 (10 to 22)	1.5 (1.3 to 1.7)
<i>Use without U.S. casualties</i>	23 (16 to 29)	1.7 (1.5 to 2.0)
<i>Use with U.S. casualties</i>	32 (26 to 39)	2.0 (1.7 to 2.3)

Notes: The absolute audience cost is the surge in disapproval, expressed in percentage points. The relative risk is the level of disapproval when the president escalated and backed down, divided by the level of disapproval when the president stayed out. Bayesian 95 percent credible intervals appear in parentheses. The sample size was 650 for stay out (the reference category), 477 for threat of force, 420 for display of force, 306 for use without U.S. casualties, and 310 for use with U.S. casualties.

the percentage who disapproved either strongly or somewhat when the president stayed out. This calculation gives the surge in disapproval, or “absolute audience cost,” of committing and not following through. Table 2 also presents the relative risk of disapproval, defined as disapproval in the escalation condition divided by disapproval in the stay-out condition.

The estimates in Table 2 show three clear patterns. First, audience costs unambiguously existed in all four scenarios. When the president escalated and did not follow through, disapproval swelled by between 16 and 32 percentage points.

Second, audience costs did not increase smoothly with the level of escalation. Based on existing models of audience costs, the president who displayed force should have paid a higher price than the president who merely threatened to use it. In our data, though, the costs were similar: disapproval in both scenarios grew by 16 percentage points, with 95 percent of the posterior density between 10 and 22. The experiment, therefore, provides no evidence that audience costs increase as the president moves from threatening to displaying force. This surprising finding, if replicable, would have significant implications for empirical and theoretical work on military crises.

Third, although audience costs did not rise with each level of escalation, they did exhibit a monotonic trend. The use of arms exposed the president to higher audience costs than either threatening or displaying force, and the loss of lives further raised the price of escalating and then backing down. Each level of audience costs was distinguishable from the previous one with probability .95 or better.

TABLE 3. Audience costs as a function of international context

<i>International context</i>	<i>Disapproval if escalate and back down (%)</i>	–	<i>Disapproval if stay out (%)</i>	=	<i>Absolute audience cost (%)</i>
<i>Interests</i>					
<i>Not affect U.S.</i>	51 (48 to 55)		26 (21 to 31)		26 (19 to 31)
<i>Hurt U.S.</i>	56 (52 to 59)		39 (34 to 45)		16 (10 to 23)
<i>Motive</i>					
<i>Historical feud</i>	52 (48 to 55)		29 (24 to 34)		23 (17 to 29)
<i>More power</i>	55 (52 to 59)		36 (31 to 42)		19 (13 to 25)
<i>Regime</i>					
<i>Democracy</i>	53 (50 to 57)		31 (26 to 36)		22 (16 to 28)
<i>Dictatorship</i>	54 (50 to 57)		34 (29 to 40)		19 (13 to 25)
<i>Power</i>					
<i>Strong military</i>	57 (53 to 60)		31 (26 to 36)		26 (20 to 32)
<i>Weak military</i>	50 (47 to 54)		35 (29 to 40)		16 (9 to 22)

Notes: The table gives the percentage of respondents who disapproved in each experimental condition. Bayesian 95 percent credible intervals appear in parentheses. Sample sizes for the escalation and stay-out conditions, respectively, were 746 and 317 for “Not affect the U.S.”; 769 and 313 for “Historical feud”; 744 and 337 for “More power”; 740 and 328 for “Democracy”; 733 and 322 for “Dictatorship”; 767 and 333 for “Would hurt the U.S.”; 728 and 330 for “Strong military”; and 785 and 320 for “Weak military.”

Are Audience Costs Robust to Variation in International Circumstances?

The evidence thus far confirms that empty commitments cause disapproval to surge. Does this finding hold across a wide range of international contexts? Table 3 displays audience costs as a function of four standard international relations variables: material interests, motive, political regime, and military power.¹⁸ The most striking lesson from Table 3 is that, in every scenario, citizens preferred the president who stayed out to the president who escalated and then backed down. The estimated audience costs were at least 16 points and sometimes as high as 26 points, with 95 percent probability intervals that always exceeded 0. Domestic audiences, it seems, disapprove of backtracking against all types of regimes, with varying motivations and military power, whether or not the national interest is at stake.

18. To increase statistical power I pooled the data from all four levels of escalation, but the main findings hold at each step of the escalation ladder.

Although audience costs were always evident, they varied with the material interests of the escalating state. The price of committing and backing down was larger by approximately 10 percentage points when the safety and economy of the United States were not at stake. This difference makes sense. Audience costs depend not only on how the public views empty threats, but also on what the public thinks when the president remains completely aloof. Citizens are naturally less likely to demand military action when their security and livelihood are not at risk. It follows that staying out should be more popular in the “not affect” condition than in the “hurt” condition. Moreover, if much of the public approves when the president stays out, there is more potential for disapproval to grow when the president escalates before backing down. Audience costs should, therefore, be larger when inaction would not threaten the national interest.

A similar logic would imply higher audience costs in crises against nonaggressive adversaries. Previous research found Americans less willing to get involved when the invasion was because of a long-standing dispute, rather than a drive for territorial aggrandizement.¹⁹ The same pattern reappeared in this study: dissatisfaction with the “stay out” scenario was less common when the adversary invaded because of a “long-standing historical feud” than when it wanted “more power and resources.” With many citizens approving when the president remained idle during the “feud,” there should have been more potential for audience costs in absolute and relative terms. Indeed, audience costs were slightly higher in the feud scenario, although one cannot affirm the difference with great confidence.²⁰ Finally, Table 3 shows that audience costs increased dramatically with the military power of the adversary, but the opponent’s political regime had little effect.

These findings, though preliminary, suggest that domestic audiences lend more credibility in some international contexts than in others. Threats, for example, may convey more information when issued by leaders who could remain on the sidelines with little risk to their own country. Likewise, threats against status quo states might be more informative than threats against revisionist ones. Finally, although a thorough analysis of the effects of power would require experiments in many countries, threats by a superpower such as the United States may be more revealing when the target is militarily strong than when it is weak.

The Political Consequences of Backing Down

The experiments described in this article establish a necessary and heretofore unproven condition for audience-cost models of international relations: citizens disapprove of empty threats. Would leaders take this disapproval into account when making foreign policy? The answer surely varies across political systems, but in democracies such as the United States, leaders generally view approval as an asset

19. Herrmann, Tetlock, and Visser 1999.

20. With a sample of this size, the probability of a positive difference is only about 8 in 10.

and disapproval as a political cost. Edwards describes the “virtual unanimity” with which presidents, their aides, and participants in the legislative process “assert the importance of the president’s public standing” and regard mass approval as “an important source of presidential power.”²¹

Foreign policy approval ratings affect the power and incentives of the chief executive in several ways. In particular, they shape national elections.²² Citizens consistently list foreign issues among the top problems facing the country, and candidates regularly campaign and speak about foreign policy. This campaign strategy is wise, because judgments about foreign policy are a major component of overall approval, which determines whether sitting leaders can retain office. Aldrich, Sullivan, and Borgida, for example, find that foreign policy issues were just as important as economic ones in predicting how Americans voted in 1980 and 1984, and Gelpi, Reifler, and Feaver offer similarly compelling evidence about the role of foreign policy in the 2004 U.S. election.²³

Public approval also enhances the executive’s influence over the legislature. In the United States, for example, members of the president’s party are more likely to win congressional elections if the president is popular.²⁴ Moreover, holding constant the partisan composition of Congress, high approval ratings help presidents push initiatives through the legislature and increase the likelihood that vetoes will be sustained.²⁵ As Krosnick and Kinder explain: “Presidents who are popular in the country tend to have their way in Washington. Popularity is a vital political resource, perhaps the president’s single most important base of power.”²⁶ Thus, leaders have good reason to view disapproval as a political liability, and to take care not to damage their standing with the public.

The political fallout from making empty threats would be magnified if disapproval were concentrated within the most politically active segments of the population. Table 4 shows precisely this pattern.²⁷ Among respondents who were registered to vote, audience costs averaged 22 percentage points, with nearly all the posterior density between 17 and 27 points. The analogous effect in the unregistered population was 16 percentage points, which had a wider posterior interval but exceeded 0 with probability greater than .99. Table 4 also distinguishes between nonvoters and active voters, who recently cast a ballot in a presidential election.²⁸ Audience costs among active voters averaged 22 percentage points, versus 15 points for those who had not been to the polls in some time. One can, therefore, conclude that empty threats cause disapproval throughout the population, and especially in the group best positioned to apply electoral penalties.

21. Edwards 1997, 113–14.

22. For a literature review, see Aldrich et al. 2006.

23. See Aldrich, Sullivan, and Borgida 1989; and Gelpi, Reifler, and Feaver 2007.

24. Gronke, Koch, and Wilson 2003.

25. For a reconciliation of competing claims in the literature, see Canes-Wrone and de Marchi 2002.

26. Krosnick and Kinder 1990, 497.

27. I obtain statistical power by averaging across international circumstances and levels of escalation.

28. Citizens qualified as active voters if they cast ballots in either 2000 or 2004.

TABLE 4. Audience costs by level of political participation

<i>Level of participation</i>	<i>Disapproval if escalate and back down (%)</i>	–	<i>Disapproval if stay out (%)</i>	=	<i>Absolute audience cost (%)</i>
<i>Registered</i>	56 (53 to 59)		34 (30 to 38)		22 (17 to 27)
<i>Not registered</i>	42 (36 to 49)		26 (18 to 36)		16 (5 to 27)
<i>Voter</i>	57 (54 to 60)		34 (30 to 39)		22 (17 to 28)
<i>Nonvoter</i>	44 (38 to 49)		29 (21 to 37)		15 (5 to 24)
<i>Activist</i>	63 (59 to 68)		29 (23 to 36)		34 (26 to 42)
<i>Nonactivist</i>	50 (47 to 53)		34 (30 to 39)		16 (11 to 21)

Notes: The table gives the percentage of respondents who disapproved in each scenario. Bayesian 95 percent credible intervals appear in parentheses. Sample sizes for the escalation and stay-out conditions, respectively, were 1,120 and 476 for citizens who were “Registered” to vote; 227 and 88 for citizens who were “Not Registered”; 1,026 and 432 for “Voters”; 302 and 126 for “Nonvoters”; 427 and 190 for “Activists”; and 1,031 and 440 for “Nonactivists.”

Empty threats have an even larger effect on citizens who go beyond voting to participate more actively in politics. Following Verba, Schlozman, Brady, and Nie, I classified someone as a political activist if he or she had recently worked for a political campaign, donated money to a campaign, served on a community board, collaborated to solve a community problem, contacted a government official, or attended a political protest or rally.²⁹ Approximately 29 percent of respondents performed at least one of these activities in the previous twelve months. Audience costs among these activists averaged 34 percentage points, more than double the level among citizens who were not in the activist category. Apparently, the most politically active—and possibly the most influential—segments of the population would disapprove at high rates if the executive made threats and did not see them through. These facts bolster the conclusion that escalating and backing down would entail real domestic political costs.

Why Do Citizens Disapprove?

Why, exactly, do citizens disapprove of leaders who escalate crises and then back down? I designed a separate survey of 347 citizens to investigate the micro-mechanisms behind audience costs. As before, citizens considered a situation in

29. Verba et al. 1993.

which a country invaded its neighbor. Some read that the president stayed out; others read that the president escalated the crisis but did not follow through. In all cases, the attacking country ultimately took over its neighbor. This survey went beyond the other experiments, though, by asking citizens to explain the opinions they expressed. After voicing approval or disapproval, participants received a follow-up prompt: "Could you please type a few sentences telling us why you approve/disapprove of the way the U.S. president handled the situation?" Participants entered their answers directly into a text box, making it possible to analyze each respondent's account in his or her own words.

For manageability, the study of motivations contained fewer experimental manipulations than the main instrument. In the category of foreign policy strategy, the president either stayed out or displayed force before backing down. The survey also presented a smaller set of background conditions: the invasion would either hurt or not affect the safety and economy of the United States, but the attacking country was always described as having a strong military, and citizens did not receive information about the motives or political regime of the invader.

The results provided further evidence of audience costs. The president who stayed out received a disapproval score of 32 points, while the president who escalated and backed down got negative ratings from 58 percent of the public. The implied cost of $58 - 32 = 26$ approval points was five times its standard error, and its credible interval ran from 15 to 35 points.³⁰ Thus, the experiment corroborated one of this article's main findings: citizens think more highly of leaders who do nothing than of leaders who commit but do not follow through.

At the same time, the survey provided preliminary evidence about why audience costs exist. In the study, 185 citizens considered a scenario in which the president escalated and backed down. Of these, 105 disapproved either strongly or somewhat of the way the president handled the situation. Why did they view the president's behavior negatively? Some did not say, and a few misunderstood the follow-up question or provided an unclassifiable answer, but eighty-seven of the 105 clearly articulated why they had assigned a negative rating.

The eighty-seven open-ended responses fell into three categories. The first category included people who thought the president should have pushed out the invaders, not because the president had made a prior commitment, but because it was the right thing to do. Some said the United States had a moral obligation to protect the victims of aggression; others pointed out that the safety and economy of the United States would suffer if the invader took over its neighbor. Fourteen of the eighty-seven participants (approximately 16 percent) answered this way. These citizens probably would have objected as much, and for the same reasons, if the president had stayed out. In fact, most disapprovers in the control group (in which the president neither threatened nor showed force) justified their attitudes in similar

30. This estimate exceeds the value for display of force in Table 2. Why the difference? The text was a bit shorter, so backing down might have appeared starker. Moreover, the adversary in this study always had a strong military, a factor that increases audience costs (see Table 3).

terms. Because these reasons apply equally to all scenarios in which the president let the invasion continue, they cannot be a source of audience costs.

A second group of respondents disliked the fact that the president had escalated in the first place. Some contended that it was not the responsibility of the United States to solve other countries' problems ("I do not feel that the U.S.A. should be the police for the world. We should not have sent troops in this situation.") Others argued that the U.S. government should have focused on its own citizens ("The U.S. has enough problems of its own at this time. We have people that are homeless and hungry. We should take care of our own first."). Roughly 12 percent of respondents offered these dovish or isolationist responses, an often overlooked reason for audience costs.

The vast majority of respondents (72 percent) gave a third reason for disapproving: the president behaved inconsistently by saying one thing and doing another. Why did they view inconsistency as problematic? Many complained that waffling would hurt the reputation and credibility of the country. As one citizen explained, "If you say that you are going to do something, you need to do it or else you lose your credibility. It would have been better to ignore the situation completely than to make a public commitment and then not carry it out." Another respondent wrote: "When a President says something, in this case that he will push back the invading country, he must follow through or lose credibility in the world community. He sent troops and when the threat didn't work, he allowed the invasion to continue. That is a terrible precedent to set."

Others criticized the president's inconsistent behavior without citing any consequences. One respondent simply wrote, "the U.S. president did not stick to his word. He should have either stayed out of the other country's business in the first place or attacked as he said he would." Others felt that "this country should not make threats that it does not have the full intention of following through," and that "if the president said he would do something he should have done so. He should not threaten without action."

Passages such as these could reflect a normative preference for honesty, rather than—or in addition to—an instrumental concern for reputation. No respondent who complained about inconsistency impeached the president as "unethical," but moral considerations might have been implicit. Overall, the evidence strongly supports a reputational foundation for audience costs, without ruling out the possibility of a normative foundation as well.

A few respondents disliked inconsistency for reasons potentially distinct from reputation and honesty. Two people complained that the president had wasted money by deploying troops but not using them. In addition, eight judged the president to be incompetent: the president behaved in a puzzling manner ("Why would he have troops there to help and not do anything to help?") or had not shown sufficient foresight ("The United States President must not have truly thought things through"). Some citizens who denounced the president as incompetent may have felt that he failed to weigh the reputational costs of flip-flopping. Nonetheless, even if incompetence is given its own category, the emphasis on reputation and honesty is remarkable: 61 percent of all disapprovers, and 84 percent of those

who complained about inconsistency, denounced the president for breaking his word. By not upholding his commitment to repel the invaders, the president suggested that he and his country could not be trusted.

These responses give preliminary support to a reputation-based theory of audience costs. Early theoretical work proposed that “domestic audiences may provide the strongest incentives for leaders to guard their states’ ‘international’ reputations.”³¹ Guisinger and Smith extend this insight by developing a model in which reputations for honesty help countries and their agents achieve diplomatic success. Knowing this, citizens will punish leaders for “destroying the country’s honest record and thus putting in jeopardy the future benefits of being able to communicate during a crisis.”³² Their model suggests a rational underpinning for audience costs, founded on concerns about reputation.

The evidence in this article is consistent with such a reputational logic. It seems that many citizens value their country’s international reputation and disapprove of leaders who mar it. In countries where citizens can hold leaders accountable, the prospect of a domestic backlash could, therefore, create an added incentive to care about international reputations, and thus an extra reason to avoid making empty commitments.

Conclusions

This article has offered the first direct analysis of audience costs in a way that avoids problems of strategic selection. The research, based on a set of experiments embedded in public opinion surveys, shows that audience costs exist across a wide range of conditions and increase with the level of escalation. The adverse reaction to empty commitments is evident throughout the population, and especially among politically active citizens who have the greatest potential to shape government policy. Finally, preliminary evidence suggests that audience costs arise from concerns about the international reputation of the country and its leaders.

These findings demonstrate the promise of using experiments to address central questions in the field of international relations. By incorporating randomized treatment and control into interviews with masses and elites, researchers can gain new insights about preferences and beliefs that might not be possible through the study of historical data.

These findings also have important substantive implications. In particular, they supply behavioral microfoundations for theories of signaling and commitment in world affairs. It is widely assumed that domestic audiences enhance the credibility of international commitments by punishing leaders who say one thing but do another. I confirm that citizens respond this way, a discovery that was far from preordained. If citizens had focused on foreign policy outcomes rather than

31. Fearon 1994, 581.

32. Guisinger and Smith 2002.

processes, or regarded bluffing as a reasonable strategy, or rewarded leaders for trying before conceding, or cared little about their country's reputation, audience costs would not have emerged. By showing that audience costs arise consistently across a wide range of conditions, this article advances our understanding of signaling and commitment in both international security and political economy.

Finally, this study contributes to an understanding of reputation in world politics. What motivates leaders to protect their international reputations, even at great cost to themselves and others? Domestic audiences may play an important role. Right or wrong, citizens worry that leaders who break commitments will undermine the nation's credibility, and they disapprove when the executive adopts reputation-damaging strategies. Citizens may, therefore, seek to elect leaders who value the national reputation and would be especially competent at preserving it. Once in office, leaders may feel strong domestic pressure to safeguard their foreign reputations. Domestic audiences may, therefore, help explain why many leaders strive to protect the national image and why concerns about reputation shape international relations.

Appendix: Sampling and Interview Methods

The surveys discussed in this article were administered by Knowledge Networks, an Internet-based polling firm, with support from the National Science Foundation. By using random digit dialing to recruit participants, and by providing Internet access to households that do not have it, Knowledge Networks is able to administer questionnaires to a nationally representative sample of the U.S. population. The surveys took place in July and November 2004, and approximately 76 percent of panelists who were invited to complete the surveys actually did so.

Table A1 compares my sample to the U.S. adult population. National population figures came from the U.S. Census Bureau and the Bureau of Labor Statistics, which provide monthly updates of demographic data through the Current Population Survey (CPS). I computed the benchmarks by pooling data from the July and November 2004 CPS studies ($N = 205,580$ adults). The average deviation, in percentage points, between the samples used in this article and the national population was no more than 1.5 percentage points. My sample slightly overrepresented the elderly and residents of the Midwest and the South, while slightly underrepresenting Americans in the highest household income bracket. Even in these categories, though, the deviations were only a few percentage points. Overall, the sample closely matched the population benchmark.³³

33. Was interest in politics higher among respondents than in the nation as a whole? It is hard to know for sure, because the Census Bureau does not collect data on political interest. However, political interest levels in my sample closely matched levels in the General Social Survey (GSS). In my sample, 22 percent of subjects were "very interested" in politics, 40 percent were somewhat interested, 26 percent were slightly interested, and 12 percent were not at all interested. The comparable GSS figures for 2004 were 21, 49, 20, and 10 percent. In any case, the issue is of minor concern because Table 4 indicates large audience costs even among people who do not show much engagement in politics.

TABLE A1. *Characteristics of the national population and the survey sample*

	<i>Adult U.S. population (%)</i>	<i>Audience costs sample (%)</i>	<i>Absolute deviation (%)</i>
<i>Gender</i>			
<i>Male</i>	48	48	0
<i>Female</i>	52	52	0
<i>Average deviation</i>			0
<i>Age</i>			
<i>18–24</i>	12	10	1
<i>25–34</i>	17	15	2
<i>35–44</i>	20	20	0
<i>45–54</i>	20	20	0
<i>55–64</i>	14	15	1
<i>65 or older</i>	17	20	3
<i>Average deviation</i>			1
<i>Education</i>			
<i>No high school diploma</i>	14	14	0
<i>High school diploma</i>	33	31	2
<i>Some college</i>	19	24	5
<i>Associate degree</i>	8	8	1
<i>College degree</i>	26	24	2
<i>Average deviation</i>			2
<i>Income</i>			
<i>Less than \$10,000</i>	7	7	0
<i>\$10,000–\$24,999</i>	17	19	1
<i>\$25,000–\$49,999</i>	29	34	5
<i>\$50,000–\$74,999</i>	20	21	1
<i>\$75,000 or more</i>	26	19	8
<i>Average deviation</i>			3
<i>Marriage</i>			
<i>Married</i>	58	58	0
<i>Not married</i>	42	42	0
<i>Average deviation</i>			0
<i>Race</i>			
<i>White</i>	84	85	1
<i>Nonwhite</i>	16	15	1
<i>Average deviation</i>			1
<i>Region</i>			
<i>Midwest</i>	24	26	2
<i>Northeast</i>	21	18	3
<i>South</i>	30	34	4
<i>West</i>	24	21	3
<i>Average deviation</i>		3	
<i>Total average deviation</i>			1

References

Aldrich, John H., Christopher Gelpi, Peter Feaver, Jason Reifler, and Kristin Thompson Sharp. 2006. Foreign Policy and the Electoral Connection. *Annual Review of Political Science* 9:477–502.

- Aldrich, John H., John L. Sullivan, and Eugene Borgida. 1989. Foreign Affairs and Issue Voting: Do Presidential Candidates "Waltz Before a Blind Audience?" *American Political Science Review* 83 (1):123–41.
- Baum, Matthew A. 2004. Going Private: Public Opinion, Presidential Rhetoric, and the Domestic Politics of Audience Costs in U.S. Foreign Policy Crises. *Journal of Conflict Resolution* 48 (5):603–31.
- Brody, Richard A. 1994. Crisis, War, and Public Opinion: The Media and Public Support for the President. In *Taken by Storm: The Media, Public Opinion, and U.S. Foreign Policy in the Gulf War*, edited by W. Lance Bennett and David L. Paletz, 210–27. Chicago: University of Chicago Press.
- Broz, J. Lawrence. 2002. Political System Transparency and Monetary Commitment Regimes. *International Organization* 56 (4):861–87.
- Canes-Wrone, Brandice, and Scott de Marchi. 2002. Presidential Approval and Legislative Success. *Journal of Politics* 64 (2):491–509.
- Desch, Michael C. 2002. Democracy and Victory: Why Regime Type Hardly Matters. *International Security* 27 (2):5–47.
- Dorussen, Han, and Jongryn Mo. 2001. Ending Economic Sanctions: Audience Costs and Rent-Seeking as Commitment Strategies. *Journal of Conflict Resolution* 45 (4):395–426.
- Edwards, George C., Jr. 1997. Aligning Tests with Theory: Presidential Approval as a Source of Influence in Congress. *Congress & the Presidency* 24 (2):113–30.
- Eyerman, Joe, and Robert A. Hart Jr. 1996. An Empirical Test of the Audience Cost Proposition: Democracy Speaks Louder than Words. *Journal of Conflict Resolution* 40 (4):597–616.
- Fearon, James D. 1994. Domestic Political Audiences and the Escalation of International Disputes. *American Political Science Review* 88 (3):577–92.
- Gaubatz, Kurt Taylor. 1996. Democratic States and Commitment in International Relations. *International Organization* 50 (1):109–39.
- Gelman, Andrew, John B. Carlin, Hal S. Stern, and Donald B. Rubin. 2004. *Bayesian Data Analysis*. 2d ed. New York: Chapman & Hall/CRC.
- Gelpi, Christopher F., and Michael Griesdorf. 2001. Winners or Losers? Democracies in International Crisis, 1918–94. *American Political Science Review* 95 (3):633–47.
- Gelpi, Christopher, Jason Reifler, and Peter Feaver. 2007. Iraq the Vote: Retrospective and Prospective Foreign Policy Judgments on Candidate Choice and Casualty Tolerance. *Political Behavior* 29 (2):151–74.
- Gowa, Joanne S. 1999. *Ballots and Bullets: The Elusive Democratic Peace*. Princeton, N.J.: Princeton University Press.
- Gronke, Paul, Jeffrey Koch, and J. Matthew Wilson. 2003. Follow the Leader? Presidential Approval, Presidential Support, and Representatives' Electoral Fortunes. *Journal of Politics* 65 (36):785–808.
- Guisinger, Alexandra, and Alastair Smith. 2002. Honest Threats: The Interaction of Reputation and Political Institutions in International Crises. *Journal of Conflict Resolution* 46 (2):175–200.
- Herrmann, Richard K., and Vaughn P. Shannon. 2001. Defending International Norms: The Role of Obligation, Material Interest, and Perception in Decision Making. *International Organization* 55 (3):621–54.
- Herrmann, Richard K., Philip E. Tetlock, and Penny S. Visser. 1999. Mass Public Decisions to Go to War: A Cognitive-Interactionist Framework. *American Political Science Review* 93 (3):553–73.
- Jensen, Nathan M. 2003. Democratic Governance and Multinational Corporations: Political Regimes and Inflows of Foreign Direct Investment. *International Organization* 57 (3):587–616.
- 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.
- Krosnick, Jon A., and Donald R. Kinder. 1990. Altering the Foundations of Support for the President Through Priming. *American Political Science Review* 84 (2):497–512.
- Leeds, Brett Ashley. 1999. Domestic Political Institutions, Credible Commitments, and International Cooperation. *American Journal of Political Science* 43 (4):979–1002.
- Leventoğlu, Bahar, and Ahmer Tarar. 2005. Prenegotiation Public Commitment in Domestic and International Bargaining. *American Political Science Review* 99 (3):419–33.

- Lipson, Charles. 2003. *Reliable Partners: How Democracies Have Made a Separate Peace*. Princeton, N.J.: Princeton University Press.
- Mansfield, Edward D., Helen V. Milner, and B. Peter Rosendorff. 2002. Why Democracies Cooperate More: Electoral Control and International Trade Agreements. *International Organization* 56 (3): 477–513.
- Martin, Lisa L. 1993. Credibility, Costs, and Institutions: Cooperation on Economic Sanctions. *World Politics* 45 (3):406–32.
- McGillivray, Fiona, and Alastair Smith. 2000. Trust and Cooperation Through Agent-Specific Punishments. *International Organization* 54 (4):809–24.
- Partell, Peter J., and Glenn Palmer. 1999. Audience Costs and Interstate Crises: An Empirical Assessment of Fearon's Model of Dispute Outcomes. *International Studies Quarterly* 43 (2):389–405.
- Prins, Brandon C. 2003. Institutional Instability and the Credibility of Audience Costs: Political Participation and Interstate Crisis Bargaining, 1816–1992. *Journal of Peace Research* 40 (1):67–84.
- Ramsay, Kristopher W. 2004. Politics at the Water's Edge: Crisis Bargaining and Electoral Competition. *Journal of Conflict Resolution* 48 (4):459–86.
- Schultz, Kenneth A. 1999. Do Democratic Institutions Constrain or Inform? Contrasting Two Institutional Perspectives on Democracy and War. *International Organization* 53 (2):233–66.
- . 2001a. *Democracy and Coercive Diplomacy*. New York: Cambridge University Press.
- . 2001b. Looking for Audience Costs. *Journal of Conflict Resolution* 45 (1):32–60.
- Slantchev, Branislav L. 2006. Politicians, the Media, and Domestic Audience Costs. *International Studies Quarterly* 50 (2):445–77.
- Smith, Alastair. 1996. To Intervene or Not to Intervene: A Biased Decision. *Journal of Conflict Resolution* 40 (1):16–40.
- . 1998. International Crises and Domestic Politics. *American Political Science Review* 92 (3):623–38.
- Verba, Sidney, Kay Lehman Schlozman, Henry Brady, and Norman H. Nie. 1993. Citizen Activity: Who Participates? What Do They Say? *American Political Science Review* 87 (2):303–18.
- Walt, Stephen M. 1999. Rigor or Rigor Mortis? Rational Choice and Security Studies. *International Security* 23 (4):5–48.
- Weeks, Jessica. Forthcoming. Autocratic Audience Costs: Regime Type and Signaling Resolve. *International Organization*.