

# Type II Audience Costs

Kai Quek, University of Hong Kong

Traditional audience costs are the political losses a leader incurs for backing down after threatening to fight (type I). Type II audience costs are the losses incurred for entering a conflict after promising *not* to fight. I develop the idea and decompose it experimentally into its constituents: an inconsistency cost plus the loss of a nonbelligerence dividend. Type II audience costs have deep implications, including the reversal of certain microfoundational challenges against type I audience costs in the context of type II audience costs, the credible signaling of a state's resolve *not* to fight, and a reassurance mechanism with attractive properties.

Much attention in international relations is paid to the problem of how leaders can credibly threaten to use force. Less attention is paid to how they can credibly promise *not* to use force. The problem of credible threat may be solved when leaders generate what I will call type I audience costs: the political costs of backing down after threatening to fight (Fearon 1994).<sup>1</sup> The problem of credible reassurance may have a similar solution through what I will call type II audience costs: the political costs of entering a conflict after promising *not* to fight.<sup>2</sup>

Despite a great deal of attention on type I audience costs type II audience costs remain unexplored, except in Levy et al. (2015).<sup>3</sup> I develop the concept and show how type II audience costs change the way we think about audience cost theory and the credible signaling of a state's resolve not to fight. Then, I decompose the concept into its constituents—the cost of inconsistency and the loss of a nonbelligerence dividend—and highlight how these losses tighten the tying-hands constraints on leaders to enhance reassurance credibility. Finally, I compare my results with Levy et al. (2015) and

Tomz (2007) and show how knowledge cumulates when independent experiments cross on a common topic.

## TYPE II AUDIENCE COSTS

Type II audience costs are promise based rather than threat based; they focus on reassurance rather than deterrence; and while both type I and type II audience costs try to shape beliefs, the beliefs they try to shape are orthogonal: type I is about convincing others of one's willingness to bear the costs of war; type II is about convincing others of the converse. These differences have microfoundational implications. In particular, a long-standing issue in audience cost theory is how the consistency mechanism can bear the weight of war (e.g., Slantchev 2006; Snyder and Borghard 2011; Trachtenberg 2012). Why would citizens punish their leader for avoiding a costly war? Can we assume audience costs to be high enough to override the costs of war? Slantchev (2011, 51) argued that “when one talks about things as vague and as amorphous as ‘national honor’ and compares them to the blood, the destruction of lives and property, and the psycho-

Kai Quek (quek@hku.hk) is an assistant professor of politics at the University of Hong Kong.

Data and supporting materials necessary to reproduce the numerical results in the article are available in the JOP Dataverse (<https://dataverse.harvard.edu/dataverse/jop>). An online appendix with supplementary material is available at <http://dx.doi.org/10.1086/693348>. The experiments were approved by the Committee on the Use of Humans as Experimental Subjects at the Massachusetts Institute of Technology. The MIT Political Experiments Research Laboratory funded the primary experiment in this project.

1. Since Fearon (1994), the literature has conceptualized “audience costs” as the costs of backing down from a threat. Developers and critics of audience cost theory have followed the same understanding (e.g., Snyder and Borghard 2011, 437; Tomz 2007, 821; Weeks 2008, 35).

2. I follow the practice of naming orthogonal types as types I and II, e.g., type I and type II errors.

3. Levy et al. (2015) focus on the consistency mechanism in audience cost theory and do not analyze type II audience costs as a distinct phenomenon or decompose the concept into its constituent elements. The experiments here, fielded in 2012 and 2013, were independent from Levy et al.'s study fielded in 2014.

The Journal of Politics, volume 79, number 4. Published online August 29, 2017. <http://dx.doi.org/10.1086/693348>

© 2017 by the Southern Political Science Association. All rights reserved. 0022-3816/2017/7904-0023\$10.00

1438

logical scars a war invariably inflicts on its participants, this assumption becomes heroic indeed.”

This logical liability vanishes in the context of type II audience costs. Type II audience costs assume that the public punishes politicians for a reversal that brings about a costly war—a premise easier to accept than with type I audience costs when the public punishes politicians for a reversal that *avoids* a costly war. And while the logic of type I audience costs traditionally assumes that leaders are able and willing to create audience costs so large such that war is preferable to peace for at least some types (see Fearon 1997, 80; Slantchev 2011, 51),<sup>4</sup> the assumption is unnecessary for type II audience costs. In addition, type II audience costs can shed light on a major debate in audience cost theory: the existence in domestic audiences of a preference for consistency (Fearon 1994; Smith 1998). This is challenged by Snyder and Borghard (2011), who argued that the public cares about policy substance rather than policy consistency. Testing for both type I and type II audience costs in the same experiment can inform this debate. Because respondents are randomly assigned to treatments, and because the two types of audience costs involve mirror forms of policy inconsistency triggered by two opposite policy actions (intervention vs. nonintervention), an intrinsic public desire for consistency is revealed if leaders are punished for their inconsistency regardless of the policy they choose.

Theoretically, type II audience costs offer a mechanism for signaling the resolve to stay out. Much of the extant literature focuses on signaling one’s resolve to fight. Less attention is paid to signaling one’s resolve not to fight.<sup>5</sup> But there are key instances in which a state needs to signal its resolve not to intervene. One example is a bargaining situation whereby one state offers to stay out of another state’s interests in exchange for an irreversible concession.<sup>6</sup> A problem of credible com-

mitment is created because the nonintervention promise can be broken in the future, but the concession—once given—is gone. To make the offer credible, a state can generate type II audience costs by publicizing the promise of nonintervention.

An example of historical significance is the US public non-invasion pledge in the Cuban Missile Crisis. It is interesting that the crisis started with type I audience costs—“Kennedy’s televised speech announcing the presence of Soviet missiles in Cuba” (Fearon 1994, 582)—but ended with an appeal to type II audience costs.<sup>7</sup> In exchange for removing nuclear weapons from Cuba, the Soviet government demanded the United States pledge publicly not to invade Cuba (*New York Times*, October 28, 1962). Many historians now agree that the public pledge not to invade Cuba—rather than the secret promise to remove missiles in Turkey—was a key component of the agreement that ended the crisis and prevented a nuclear war.

#### TYPE II AUDIENCE COSTS DECOMPOSED

Public threats and promises about war and peace will affect public approval if the population is sensitive to the costs and risks of war. If leaders pay a “belligerence cost” (Kertzer and Brutger 2016) for threatening to enter a conflict, then they should also earn a “nonbelligerence dividend” for promising to stay out.<sup>8</sup> The former is a political loss, but the latter is a political gain. Specifically, the belligerence cost is an irrecoverable sunk cost (Kertzer and Brutger 2016), but the nonbelligerence dividend is a retractable gain—a “tying-hands gain” that is lost if the promise is broken. The idea of costly signaling through “tying-hands cost” (Fearon 1997) is well known. But the idea of costly signaling through the loss of the tying-hands gain is not. A conjecture is that the loss of tying-hands gain is at least as salient as tying-hands cost, given an endowment effect whereby leaders assign greater prospective disutility to losing what they already possess (Kahneman and Tversky 1979).

Type II audience costs comprise an inconsistency cost and a loss of nonbelligerence dividend. When a leader breaks a promise to stay out, he or she loses twice—suffering the cost for being inconsistent *and* losing the dividend for nonbelligerence. The dual losses tighten the tying-hands constraints on leaders to strengthen reassurance credibility. Insofar as the inconsistency cost and the loss of nonbelligerence dividend

4. Tarar and Leventoglu (2013) proposed a model in which lesser audience costs can have conflict-reduction effects.

5. There are few works that explicitly analyze how to signal one’s resolve not to enter a war in the first place. The closest literature might be that on war termination, in which belligerents need to persuade each other that they are willing to stop fighting and commit to a peaceful bargain.

6. Another example is a situation when a state has an incentive to stay out of a crisis beyond its borders—insofar as the crisis can be contained by local forces or neighboring states involved in the crisis. But the more the crisis implicates the interests of the distant state, the harder it is to convince others to bear the full costs of containing the crisis on their own. Yet another situation is in times of crisis when states have an incentive to be neutral in the present (e.g., on the eve of war) but also an incentive to opportunistically intervene in the future (e.g., at the close of war). Cursed by the commitment problem, it can be crucial for neutral states to send signals of neutrality to prevent preemptive attacks, by tying their hands as a means of communicating their resolve to stay out.

7. As highlighted in a recent review, “Contrary to long-held belief, JFK’s secret Turkish concession did not influence Khrushchev’s move” (Hershberg 2010, 83).

8. Indeed, leaders in modern history have often shown a public desire to appear as the nonbelligerent in a conflict—a behavior that might reflect a belief that nonbelligerence will be rewarded by domestic and international audiences.

Table 1. Type I Audience Costs (%)

	Reaction to Reneged Threat (1)	Reaction to Staying Out (2)	Type I Audience Costs (1 – 2)
Approve	24.3 (17.9 to 30.7)	34.9 (27.8 to 41.9)	–10.6 (–20.1 to –1.1)
Disapprove	47.4 (40.0 to 54.8)	36.0 (28.9 to 43.1)	11.4 (1.1 to 21.7)

Note. Percentages of respondents who approved or disapproved of the president; 95% confidence intervals are in parentheses.

are contingent costs, type II audience costs—unlike type I—constitute a pure tying-hands signal.<sup>9</sup>

## EXPERIMENTS

I fielded two experiments. In May 2012, the primary experiment was implemented on a national sample recruited by Survey Sampling International (SSI) to match the adult US census population on age, gender, income, and geography ( $n = 696$ ).<sup>10</sup> In June 2013, a replication experiment was implemented with subjects recruited through Amazon.com's Mechanical Turk (AMT;  $n = 778$ ). Respondents were linked from SSI or AMT to an Internet survey. They began the experiments by reading about a foreign crisis similar to that in Tomz (2007), which is the standard model for audience cost experiments. A neutral vignette is used to achieve comparability with Kertzer and Brutger (2016), Levy et al. (2015), and Tomz (2007)—all of which used a similar vignette involving an anonymous president without injecting specific domestic cues or frames. Respondents in the primary experiment were randomly assigned to one of four conditions, all of which had a similar crisis scenario in which a foreign country sent its military to take over a neighboring country. The experimental conditions differed only in the way the US president dealt with the crisis.

The control condition and treatment 1 reconstructed Tomz's (2007) test for type I audience costs. In the control condition ("stay-out"), the president said that the United States would stay out of the conflict, and the United States did stay out eventually. In treatment 1 ("reneged-threat"), the president threatened military intervention but eventually backed down. This allows us to isolate the effect of

the treatment (reneging on the threat), while holding constant the same outcome (staying out). Type I audience costs are measured by comparing public approval between the two conditions. Treatments 2 and 3 rescripted the crisis outcome to capture type II audience costs. Two critical twists were introduced: while the president stayed out of the conflict in treatment 1, the president in treatments 2 and 3 did not. In treatment 2, the president said that the United States would intervene *and* it did. In treatment 3, the president said that the United States would not intervene *but* it did. These conditions allow us to isolate the effect of the president's reversal, while holding constant the same outcome (intervention).<sup>11</sup> Type II audience costs are measured by comparing public approval between the two treatments. Appendix 1 (apps. 1 and 2 available online) shows the full experimental text. Appendix 2 (table S1; tables S1–S6 available online) shows the SSI sample is comparable on key demographic variables to the Current Population Survey sample. The AMT sample here is also comparable to the AMT sample in Levy et al. (2015).

## RESULTS

First, I report the results of the primary experiment. Table 1 captures type I audience costs by comparing public approval for the president in the stay-out (control) and the reneged-threat (treatment 1) conditions. The approval for the president who reneged was 11 percentage points lower than that for the president who stayed out. More precisely, the approval rating (seven-point scale from 0 to 6) was lower at 2.38 in treatment 1 compared to 3.02 in the control group ( $p < .01$ ).<sup>12</sup> The results affirm the significance of type I audience costs.

9. Kertzer and Brutger (2016, 236–38) argued that since type I audience costs comprise an inconsistency cost and a belligerence cost, and since the latter is a sunk cost "invoking an alternative signaling mechanism," type I audience costs are "double-barreled" and do not constitute a pure tying-hands signal.

10. SSI samples have been used in many studies in political science, including Kertzer and Brutger (2016).

11. Similar to Levy et al. (2015), postintervention outcomes were suppressed in these treatments to avoid confounding the measurement of audience costs, as respondent support may be influenced by perceptions of battlefield success and the different levels of casualties incurred on both sides.

12. All tests are two-tailed  $t$ -tests. I also analyze the relationship between approval and backing down with ordered logit models with and without sociodemographic controls. The conclusion remains unchanged throughout (table S2 in app. 2).

Table 2. Type II Audience Costs: Primary Experiment (%)

	Reaction to Reneged Commitment (1)	Reaction to Fulfilled Commitment (2)	Type II Audience Costs (1 – 2)
Approve	29.7 (22.9 to 36.5)	47.4 (40.0 to 54.8)	-17.7 (-27.7 to -7.6)
Disapprove	39.4 (32.2 to 46.7)	31.2 (24.3 to 38.1)	8.2 (-1.8 to 18.2)

Note. Percentages of respondents who approved or disapproved of the president; 95% confidence intervals are in parentheses.

Next, I measure type II audience costs by comparing the outcomes between treatment 2 (president intervened after committing to intervene) and treatment 3 (president intervened after committing not to intervene). Table 2 shows that the approval for the president who reneged was 18 percentage points less than that for the president who did not renege.<sup>13</sup> More precisely, the approval rating was 2.71 for the president who reneged. This is significantly lower than the 3.38 received by the president who did not renege ( $p < .01$ ).<sup>14</sup>

These results show that type II audience costs are substantial. The findings also speak to the debate over the existence of a public preference for consistency. It turns out that leaders are punished by the public for their inconsistency whether they choose a policy of intervention (table 2) or the opposite policy of nonintervention (table 1). This suggests that policy consistency is important in the public mind (cf. Snyder and Borghard 2011).<sup>15</sup> The results also speak to Levy et al.'s (2015) experiment, which was fielded on a sample from AMT in 2014.<sup>16</sup> Their experiment found that leaders suffer more for “backing out” than “backing in,” showing the public has a stronger distaste for backing down compared to breaking a peaceful promise, and thus there is greater incentive not to retreat in a crisis.

The experiments here challenge this conclusion. Specifically, Levy et al. (2015, 995) reported a decline in presidential approval by 22 percentage points for backing out compared

to 12 for backing in.<sup>17</sup> I find the converse result: presidential approval declined by 11% for backing out—close to Tomz's (2007, 827) finding of 12%—compared to 18% for backing in (also close to the replication study's 19%; see table 3). The difference in conclusions is due to (i) the higher approval rates found in Levy et al.'s experiment (41%) when respondents learned that the president promised not to use force but reneged and used force, compared to both the SSI experiment (30%) and my later replication on AMT (also 30%), and (ii) the higher type I audience costs found in Levy et al.'s AMT experiment (22% fall in approval), compared to 11% in my SSI experiment and 12% in Tomz's (2007) experiment. Contrary to Levy et al. (2015), my results suggest that we should not downplay the costs of breaking peaceful promises, or overplay the costs of backing down and the political incentives for escalating a crisis.

### Replication tests

Two further tests are conducted using a separate national sample from AMT in June 2013 ( $n = 778$ ). The first test replicated the SSI primary experiment: subjects were told in condition 1 that the president had said the US military would push out the invaders if the attack continued; subjects randomly assigned to condition 2 were told that the president had said the United States would stay out of the conflict. In both conditions, the attacking country continued to invade, and the scenario ended with the president sending troops. The approval rating was 2.57 for the president who reneged on the promise to stay out, which was lower than the 3.47 received by the president who did not renege ( $p < .001$ ). The results are close to those in the primary experiment using the SSI sample (2.71 for the president who reneged and 3.38 for the president who did not). Table 3 shows the president who

13. This is measured by the question: “Do you approve, disapprove, or neither approve nor disapprove of the way the U.S. president handled the situation?” (app. 1). Appendix 2 shows consistent results when we include “leaners” (tables S3 and S4). “Leaners” chose “neither approve nor disapprove” but said they lean toward approving/disapproving when probed further.

14. Table S5 in app. 2 corroborates the result using ordered logit models with sociodemographic controls.

15. The results show that policy consistency matters to the public in a foreign-intervention scenario. They do not imply that consistency is always more important than substance.

16. Levy et al.'s experiment used a similar vignette as the one here and in Tomz (2007) but with a few deviations (see Levy et al. 2015, 992–94).

17. The comparison is based on approval rates following Levy et al.'s reported results. In addition, tables 1–3 here also report the disapproval rates.

Table 3. Type II Audience Costs: Replication Experiment (%)

	Reaction to Reneged Commitment (1)	Reaction to Fulfilled Commitment (2)	Type II Audience Costs (1 – 2)
Approve	30.4 (24.8 to 36.0)	49.8 (43.7 to 55.9)	-19.4 (-27.7 to -11.2)
Disapprove	45.4 (39.3 to 51.4)	30.9 (25.3 to 36.5)	14.5 (6.2 to 22.8)

Note. Percentages of respondents who approved or disapproved of the president; 95% confidence intervals are in parentheses.

reneged obtained 30% approval, 19 percentage points lower than that received by the president who did not renege. Again, these results are consistent with the primary experiment, as shown by comparing tables 2 and 3.<sup>18</sup>

I have measured type II audience costs by comparing public approval between condition 1 (president intervened after committing to intervene) and condition 2 (president intervened after committing not to intervene). Type II audience costs may also be measured by the difference in approval between condition 2 and an alternative baseline—condition 3—when the president did not intervene after committing not to intervene. Hence, I included in the same experiment a second replication test based on condition 3. The approval rating in condition 2 (2.57) was lower than in condition 3 (3.85,  $p < .001$ ). Overall approval for the president who reneged (30%) collapsed by 22 percentage points compared to the president who did not. These results confirm that type II audience costs are significant and large, whichever way we measure them.

### Decomposing type II audience costs

Using conditions 1–3 we can decompose type II audience costs into an inconsistency cost and the loss of the nonbelligerence dividend. Specifically, the loss of the nonbelligerence dividend (0.38 points on the seven-point approval scale) is measured based on the difference in approval when the president promises to stay out and follows through (condition 3) versus promises to intervene and follows through (condition 1). The inconsistency cost (1.28 points) is measured based on the difference in approval when the president promises to stay out and follows through (condition 3) versus promises to stay out and reneges (condition 2). Both differences are statistically significant ( $p < .05$ ). Together, the inconsistency

and belligerence effects represent a loss of public approval of 1.66 points on the seven-point scale.<sup>19</sup>

This decomposition of type II audience costs parallels the decomposition of type I audience costs (1.35 points) in Kertzer and Brutger (2016, 242). In the decomposition of type II audience costs, the loss of the nonbelligerence dividend (0.38 points) is comparable to the belligerence cost (0.45 points) in Kertzer and Brutger, but the inconsistency cost is larger for type II audience costs (1.28 points) than for type I audience costs (0.91 points). This comparison further cautions us against downplaying the costs of breaking peaceful promises and the political incentives of leaders to stay out of war.

### CONCLUSION

This article defines type II audience costs, develops the idea, measures the phenomenon, and decomposes the concept into its constituents. The idea behind type II audience costs is simple but has remained largely unrecognized. It motivates a new way of thinking about audience costs. It extends beyond the type I audience costs traditionally conceived in the literature and deepens our understanding of the microfoundations of audience cost theory. Type II audience costs can help to credibly signal a state's resolve not to fight. To the extent that type II audience costs can be generated by leaders and enforced by publics, leaders will have to think twice before entering a conflict they said they would not, and peaceful reassurances can gain credibility in a way that cheap talk cannot.

Open questions abound. First, what are the determinants of type II audience costs? Given the differences between type I and type II audience costs, it is unknown whether the same factors that move type I audience costs will lead to similar, opposite, or null effects for type II audience costs. Future re-

18. Ordered logit analysis confirms a negative and significant relationship ( $p < .001$ ) between public approval and a reneged commitment to stay out. Table S5 (app. 2) displays the ordered logit estimates side by side with those from the SSI experiment.

19. Breaking the data down by subgroups, we see the inconsistency effect is significant in each demographic (except for Republicans and conservatives). The belligerence effect is generally small and negative (except for Republicans and conservatives) and operates significantly among females, low-income earners, and non-Republicans/conservatives. See table S6 in app. 2.

search should test whether type II audience costs can be moved by these variables in the same way that traditional audience costs are. Second, do the dynamics of type II audience costs vary across different regime types? In terms of reassurance, a regime with a stronger domestic audience should have a stronger ability to signal its cooperative intent via type II audience costs.<sup>20</sup> The greater the type II audience costs generated by the leader, the more likely it is for the leader not to renege. If democratic regimes are indeed more capable of generating type II audience costs, then one may expect to observe higher levels of cooperation and reassurance success among democracies than among nondemocracies, lending support to a “democratic reassurance thesis.”

Third, are statements of resolve infused with audience costs at the signaler’s end accurately and credibly perceived at the receiver’s end (Quek 2016; Yarhi-Milo, Kertzer, and Renshon 2016)? This question is important to answer because the logic of audience costs may be undercut by the existence of a sender-receiver gap (Quek 2016). Fourth, how do type II audience costs operate across different countries or issue domains? This question is hard to answer unless we can compare the results from different studies conducted across different contexts. To this end, this article hopes to form a foundation for future research to better understand where, and in what ways, the phenomenon of type II audience costs applies.

---

20. Type II audience costs have two defining properties that differentiate the idea from sunk cost reassurance (Kydd 2005) and that make it an attractive signaling mechanism for reassurance purposes. Type II audience costs are based on tying hands. They are contingent costs. Whether they are incurred depends on whether the promise is renege. Unlike the sunk cost signaler who incurs a cost regardless of whether the promise is kept, our signaler here does not incur a cost unless the promise is broken. At the same time, type II audience costs are ex post costs. Thus, our signaler does not face the same resource constraint that the sunk cost signaler faces. The sunk cost signaler must be willing and able to pay the cost at the point of signaling, but our signaler here does not. It is the unwillingness and inability to suffer type II audience costs ex post that make them ex ante compelling as a reassurance device.

## ACKNOWLEDGMENTS

I thank Adam Berinsky, Wilfred Chow, Mike Sances, Jim Snyder, and my anonymous reviewers for their useful feedback. A previous version of this article was delivered at the 2012 American Political Science Association annual meeting under the title “Do Domestic Audience Costs Really Matter?”

## REFERENCES

- Fearon, James D. 1994. “Domestic Political Audiences and the Escalation of International Disputes.” *American Political Science Review* 88 (3): 577–92.
- Fearon, James D. 1997. “Signaling Foreign Policy Interests: Tying Hands versus Sinking Costs.” *Journal of Conflict Resolution* 41 (1): 68–90.
- Hershberg, James. 2010. “The Cuban Missile Crisis.” In Melvyn P. Leffler and Odd Arne Westad, eds., *Cambridge History of the Cold War*, vol. 2. New York: Cambridge University Press.
- Kahneman, Daniel, and Amos Tversky. 1979. “Prospect Theory: An Analysis of Decision under Risk.” *Econometrica* 42 (2): 263–92.
- Kertzer, Joshua D., and Ryan Brutger. 2016. “Decomposing Audience Costs: Bringing the Audience Back into Audience Cost Theory.” *American Journal of Political Science* 60 (1): 234–49.
- Kydd, Andrew. 2005. *Trust and Mistrust in International Relations*. Princeton, NJ: Princeton University Press.
- Levy, Jack S., Michael K. McKoy, Paul Poast, and Geoffrey P. R. Wallace. 2015. “Backing Out or Backing In? Commitment and Consistency in Audience Costs Theory.” *American Journal of Political Science* 59 (4): 988–1001.
- Quek, Kai. 2016. “Are Costly Signals More Credible? Evidence of Signaler-Receiver Gaps.” *Journal of Politics* 78 (3): 925–40.
- Slantchev, Branislav L. 2006. “Politicians, the Media, and Domestic Audience Costs.” *International Studies Quarterly* 50:445–77.
- Slantchev, Branislav L. 2011. *Military Threats: The Costs of Coercion and the Price of Peace*. New York: Cambridge University Press.
- Smith, Alastair. 1998. “International Crises and Domestic Politics.” *American Political Science Review* 92 (3): 623–39.
- Snyder, Jack, and Erica D. Borghard. 2011. “The Cost of Empty Threats: A Penny, Not a Pound.” *American Political Science Review* 105 (3): 437–56.
- Tarar, Ahmer, and Bahar Leventoglu. 2013. “Limited Audience Costs in International Crises.” *Journal of Conflict Resolution* 57 (6): 1065–89.
- Tomz, Michael. 2007. “Domestic Audience Costs in International Relations: An Experimental Approach.” *International Organization* 61 (4): 821–40.
- Trachtenberg, Marc. 2012. “Audience Costs: An Historical Analysis.” *Security Studies* 21:3–42.
- Weeks, Jessica. 2008. “Autocratic Audience Costs: Regime Type and Signaling Resolve.” *International Organization* 62 (1): 35–64.
- Yarhi-Milo, Keren, Joshua D. Kertzer, and Jonathan Renshon. 2016. “Tying Hands, Sinking Costs, and Leader Attributes.” Working paper.