SUPPLEMENTARY MATERIAL

THE POWER OF COMPROMISE
Proposal Power, Partisanship, and Public Support in International Bargaining

By Ryan Brutger

World Politics

doi: 10.1017/S0043887120000192

Replication data are available at:

Brutger, Ryan. 2020. "Replication data for: The Power of Compromise: Proposal Power, Partisanship, and Public Support in International Bargaining." Harvard Dataverse, V 1. doi: 10.7910/DVN/SYH7UN

The Power of Compromise: Proposal Power, Partisanship, and Public Support in International Bargaining

Supplementary Appendix

Ryan Brutger[1]

Contents

Replication materials for the paper and appendix are available at:

https://doi.org/10.7910/DVN/SYH7UN

[1]Assistant Professor, Department of Political Science, University of California, Berkeley.
Email: brutger@berkeley.edu. Web: https://sites.google.com/berkeley.edu/brutger/.

1 Sample Demographics and Methodology

Choosing Experiments:

In addition to the discussion in the main text, it is worth noting that an additional advantage of using survey experiments to test the effects of compromise and proposal power, is that survey experiments represent the *most likely* case to observe audience costs, making them a harder case for the ameliorating effects of compromise and proposal power. In traditional audience cost experiments the leader's threat and subsequent inconsistency are temporally proximate and thus audiences should be most likely to punish leaders who back down, whereas in real-world settings audiences may have more difficulty connecting a past threat to the leader's current choice of action. Furthermore, respondents are not presented with competing messages about the efficacy or value of the engagement,[1] which means respondents are more likely to focus on the leader's threat and subsequent inconsistency. For each of these reasons, experiments represent the most likely situation for leaders to be punished, making them an ideal manner in which to test whether compromise and proposal power can overcome the negative effect of leaders' inconsistency.

Sample Sizes and Recruitment Methodology:

Experiment 1: Fielded using Amazon Mechanical Turk in 2014 on a sample of 604 U.S. respondents.

Experiment 2: Fielded using Amazon Mechanical Turk in 2013 on a sample of 1204 U.S. respondents.

Experiment 3: Fielded on a U.S. national sample of 613 respondents using Survey Sampling International.

Experiment 4: Fielded on a U.S. national sample of 543 respondents using Survey Sampling International.

SSI uses an opt-in recruitment method, after which they randomly select panel participants for survey invitations, using population targets rather than quotas to produce broadly representative samples of respondents.

[1] See Levendusky and Horowitz (2012) for an exception, where respondents are presented with new information justifying the change in policy.

Limits and Advantages of mTurk samples:

Samples from mTurk are increasingly widely used in both political science and elsewhere in the social sciences because of the extent to which they can replicate classic experiments on more diverse samples than those traditionally employed in political psychology research, as recent review pieces by Berinksy et al. (2012), Buhrmester, Kwang, and Gosling (2011), Horton, Rand, and Zeckhauser (2011), Paolacci, Chandler, and Iperiotis (2010), and Rand (2012) make clear. As is common in online surveys, the mTurk samples are more male, educated and liberal than the American population as a whole – 56.6% of respondents were male, 58.6% have at least a four-year college degree, and 51.7% identify themselves as being liberals in Experiment 2. However, the samples are particularly useful, since I focus my analysis on specific ideological and partisan subgroups, and thus draw conclusions contingent on the relevant subgroup. Furthermore, across comparable conditions in Experiments 2 and 3, the difference in average approval scores was essentially zero (0.13, $p < 0.65$), highlighting that the more representative SSI sample and mTurk samples generated remarkably consistent results.

Sample Populations

Table 1: Sample Populations

	Experiment 1	Experiment 2	Experiment 3	Experiment 4
AGE				
Percent Age 18-24	26.8	19.3	11.3	9.3
Percent Age 25-44	62.3	64.0	35.7	32.1
Percent Age 45-64	10.1	15.4	32.8	37.8
Percent Age 65+	0.8	1.3	20.2	20.8
INCOME				
Percent $0-$50,000	57.3	55.2	46.8	43.8
Percent $50,000-$100,000	33.8	35.5	33.4	33.2
Percent $100,000-$150,000	7.1	5.2	11.7	12.0
Percent $150,000-$200,000	0.8	2.3	4.8	8.3
Percent $200,000+	1.0	1.7	3.3	2.7
EDUCATION				
Less than High School	1.1	1.0	2.6	3.1
High School / GED	9.1	11.8	27.5	25.5
Some College	28.1	28.7	11.4	13.1
College Degree	51.0	47.8	39.8	38.6
Masters Degree	8.9	8.6	13.6	15.0
PhD / JD / MD	1.7	2.2	5.2	4.6

Not all percentages add to 100 due to rounding .

2 Results by Partisanship

Figure 1 replicated Figure 2 of the main text, but shows the results broken down by the party of the respondent instead of ideology. Consistent with the theory, support for compromise is strongly conditioned by both ideology and party in the United States.

Figure 1: Average Approval Score for for Compromise by Party

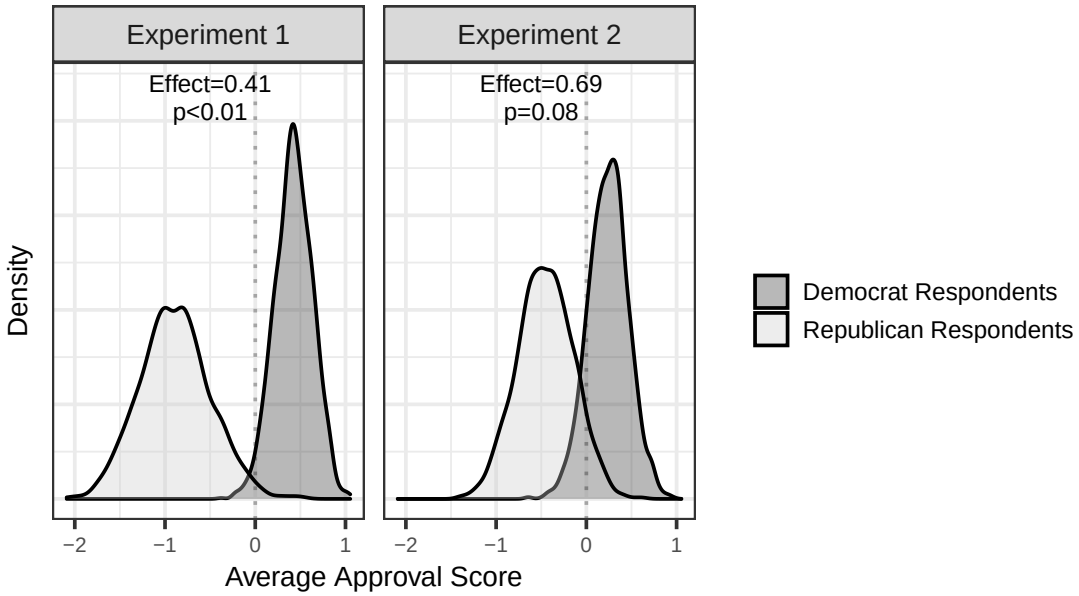

Figure 1 shows the distributions of average approval scores drawn from 2,000 bootstraps based on the self-identified party of the respondents from Experiments 1 and 2. The effects are the difference in approval scores moving from Republican to Democrat respondents; p-values are calculated using a t-test. Outcomes are held constant with 20% territory gained by the attacking country. The approval score is measured on a scale from -3=Strongly Disapprove to 3=Strongly Approve.

3 Experiment 2 - Main Treatments

Experiment 2 followed a similar format as Experiment 1. Each of the treatment conditions varies:

1) The strategy of the president

2) The partisanship of the president

Within the Compromise condition, there is an additional treatment that varies whether the U.S. president or the leader of the foreign country proposed the compromise. The main analysis of Experiment 2 follows Kertzer and Brutger (2016) and uses 20 percent of the territory gained by

the attacking country as the common outcome, but to test the sensitivity to outcomes there is an additional treatment where the attacking country gains an additional 30 percent of the territory in the Compromise and Not Engage treatments.

- Stay Out: "The U.S. president, who was a [Democrat *or* Republican], said that the United States would stay out of the conflict. The attacking country continued to invade and took control 20 percent of the contested territory. "

Consistent with Experiment 1, in each of the remaining treatment conditions the U.S. escalated with a threat and then the respondents were randomly assigned to one of the following conditions.

Threat: "The U.S. president, who was a [Democrat *or* Republican], said that if the attacking country continued to invade, the United States military would immediately engage and attempt to push out the attacking country. The president sent troops to the region."

- Compromise: "The attacking country continued to invade, but the president did not immediately engage. The [U.S. president *or* leader of the attacking country] proposed a settlement, which was agreed to by all parties, where the attacking country stopped its invasion and took control of [20 *or* 50] percent of the contested territory."
- Not Engage: "The attacking country continued to invade. The U.S. president ordered the military not to engage. The attacking country continued to invade and took control of [20 *or* 50] percent of the contested territory."

Experiment 2 also differs from Experiments 1 and 3, since Experiments 1 and 3 did not include the language "where the attacking country stopped its invasion" in the compromise condition. The language in Experiment 2 adds emphasis to the fact that the settlement stopped the attack, which could be argued to bias results in favor of the compromise condition. To test whether this difference in language biased the results, the comparable results are shown in the main paper in Figure 1. Figure 1 shows that average approval in the compromise conditions is not higher in Experiment 2 than Experiment 1, alleviating concerns that approval for compromise is driven by this difference in language. In aggregate, the results across all four experiments show that the compromise treatment effects are not conditioned by changes in the issue area or modest changes to the wording of the scenario.

4 Experiment 2 - Engage Treatment

To evaluate whether support for a leader's handling of a crisis would be higher if the leader chose to follow through on her threat, Experiment 2 included an additional Engage treatment, where the same threat was made as in the Compromise and Not Engage treatments, but then the leader followed through on the threat.

The text of the Engage treatment read as follows:

- The attacking country continued to invade and the U.S. president ordered the military to engage. The U.S. did not loose any troops in the conflict and the conflict ended with the attacking country taking control of 20 percent of the contested territory.

The treatment follows the text of earlier studies, such as Kertzer and Brutger (2016) who held casualties constant at zero and territorial outcomes constant at 20 percent. In order to isolate the treatment effect of the leader's strategy, it is necessary that casualties are held constant at zero across all conditions, otherwise respondents' approval could shift because of different outcomes as opposed to strategies. In the treatment conditions that do not result in military engagement (Stay Out, Not Engage, and Compromise) casualties are obviously at zero, and thus no mention of casualties is included. However, in the Engage condition respondents may assume higher costs resulting from war,[2] even if outcomes are not explicitly varied (for a related point, see Sher and McKenzie (2006) and Tomz and Weeks (2013) on "information leakage"). To avoid conflating respondent's expectations about the human cost of war with approval for the leader's strategy, I specify that casualties are zero in the engage condition. This inclusion means that the treatment conditions are not *lexically* equivalent, but instead they are *logically* equivalent. This technique was also used by Kertzer and Brutger (2016) who tested the effect of making the conditions *lexically* equivalent by omitting any mention of casualties in the Engage condition. Kertzer and Brutger found that audiences do indeed impose a higher cost when casualties are not mentioned, which means I am choosing a more conservative measure here, since any bias introduced by specifying casualties at zero would make the Engage treatment relatively *more* appealing to audiences (meaning that the Compromise would be relatively less appealing).

[2]This would be expected, given that the crisis bargaining model explicitly includes a cost of war that is separate from the division of the contested prize.

5 Comparing Results of Experiments 2 and 3

Figure 2: Experiment 3 - Average Approval Score for Each Strategy

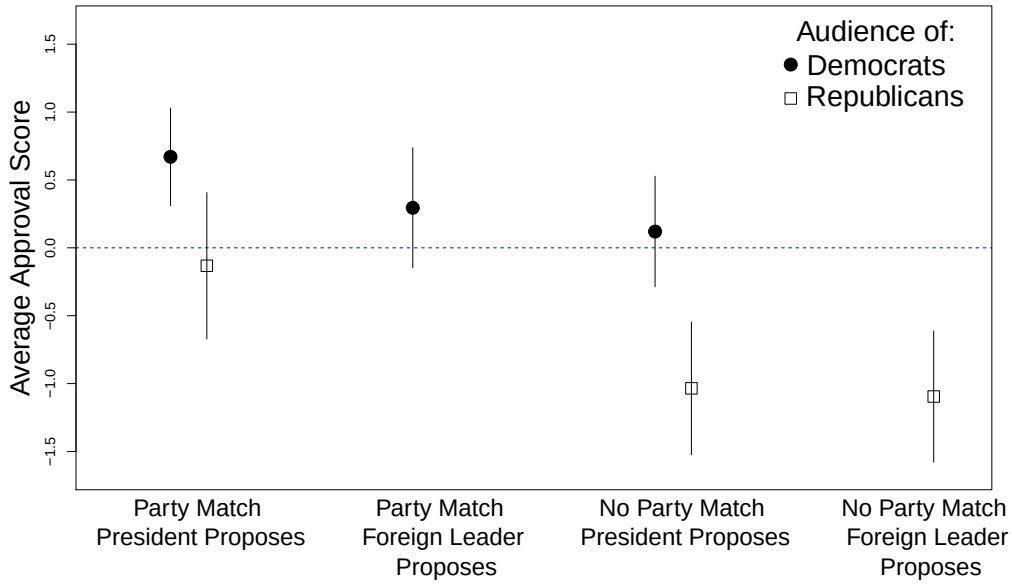

Figure 2 shows the average approval score and 95 percent confidence intervals for audiences of Democrats and Republicans based on whether the US leader or foreign leader proposed the compromise. The approval score is measured on a scale from -3=Srongly Disapprove to 3=Strongly Approve.

Comparing the results displayed in Figure 2 of the appendix to those in Figure 5 of the main paper illustrates that the effects are remarkably consistent across the mTurk and SSI samples. In each study, proposal power and partisanship lead to the highest approval scores for leaders who initiate the compromise among respondents of their own political party. These results highlight that differences in wording, and potential concerns about the wording in Experiments 1 and 2, do not alter the fundamental results of the paper. Due to a randomization error in Experiment 3, no treatment condition with a Republican president *and* the foreign leader proposing the compromise was presented. However, there were still six relevant treatment groups, as shown in Figure 2 of the appendix, that provide a useful comparison of results across samples and treatment wordings, which demonstrate that the main results are consistent across all three security experiments.

6 Results Based on Percent Approving

Figure 3 of the appendix replicates Figure 1 from the main paper, but shows the distribution of percent of respondents who approve of the president's handling of the crisis instead of the seven point approval score. The results are generally consistent regardless of whether the dependent variable is a dichotomous measure of support or the full approval score used in the main analysis.

Figure 3: Percent Approving for Compromise vs. Not Engage

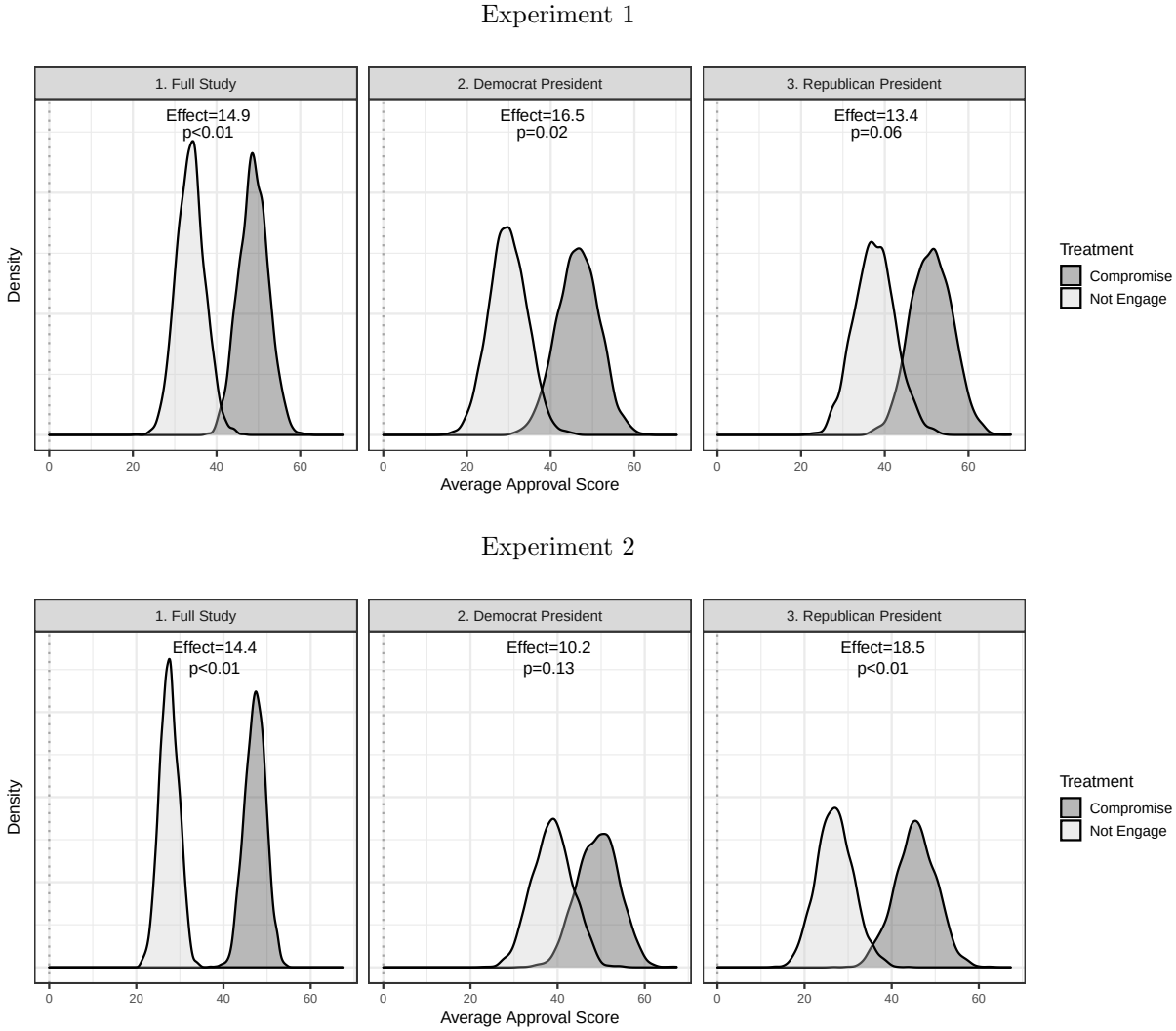

Figure 3 shows the distributions of the percent appriving drawn from 2,000 bootstraps for Experiment 1 and 2 aggregated across presidents' parties and broken down by party of the president. The effects are the change in the percent approving moving from the Not Engage to the Compromise condition and p-values are calculated using a t-test. Outcomes are held constant with 20% territory gained by the attacking country. Approval is measured as those who lean toward approving, somewhat approve, or strongly approve.

Figure 4 of the appendix replicates Figure 5 from the main paper, but shows the percent of respondents who approve of the president's handling of the crisis instead of the seven point approval score. The results are generally consistent regardless of whether the dependent variable is a dichotomous measure of approval or the full approval score used in the main analysis.

Figure 4: Percent Approving for Each Strategy

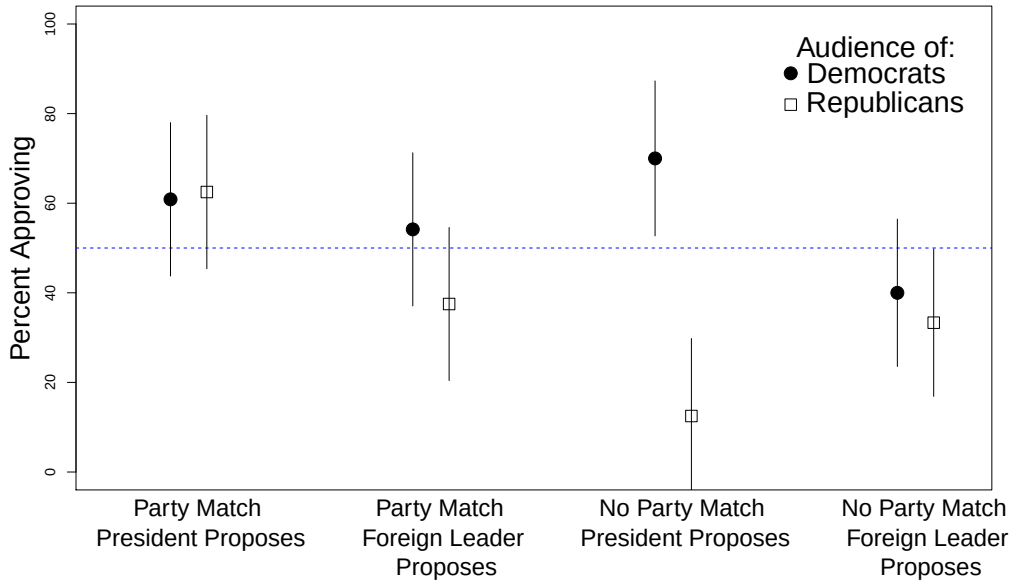

Figure 4 shows the percent of respondents who approve and 95 percent confidence intervals for audiences of Democrats and Republicans based on whether the respondent and the president are from the same party (party-match) interacted with whether the U.S. president or the foreign leader proposed the negotiated settlement. Approval is measured as those who lean toward approving, somewhat approve, or strongly approve.

7 Experiment 4 - Treatments

Experiment 4 presented respondents with an investment dispute, which was presented in the format of a brief news report. The dependent variable is the same as in the first three experiments, where respondents were asked whether they approved, disapproved, or neither, and then how strongly they felt that way. The text of the experiment is as follows:

Company Starts Legal Actions Over Investment Denial

TransCorp., a hypothetical company based in a neighboring country, on Wednesday said it was pursuing legal actions against the United States and the Obama administration

in response to its refusal to issue a border-crossing permit for the company's project.

TransCorp said in a statement that it would initiate an international arbitration case against the U.S. under an international agreement. Through a process known as investor-state dispute settlement (ISDS), companies and investors from one country can challenge the acts of a foreign government and receive compensation if they can show they weren't treated in accordance with international law.

TransCorp said it would attempt to recover more than $15 billion in costs and damages that the company said it has suffered as a result of the U.S. administration's breach of its international obligations.

The final paragraph of the report was randomly assigned, showing either the "U.S. proposes", "foreign proposes", or "U.S. proposes, without threat" treatments, which are shown below.

- U.S. Proposes: The Obama administration originally responded by stating it would fight the challenge until the arbitration panel made its decision. The Obama administration has since proposed a settlement granting TransCorp twenty percent of the value of the suit, and TransCorp accepted the settlement.

- Foreign Proposes: The Obama administration originally responded by stating it would fight the challenge until the arbitration panel made its decision. TransCorp has since proposed a settlement granting TransCorp twenty percent of the value of the suit, and the Obama administration accepted the settlement.

- U.S. Proposes, Without Threat: The Obama administration proposed a settlement granting TransCorp twenty percent of the value of the suit, and TransCorp accepted the settlement.

8 Mediation Analysis - Mechanisms Influencing Support for Compromise

To test what factors are influencing support for compromise, Experiment 2 measured a series of potential mediators that included emotional reactions believed to influence perceptions of negotia-

tions and international relations and their perception of the leader's reputation. Measures of the mediators were as follows:

Reputation: On a scale of 1-5, how much damage do you think there would be to the President's reputation as a result of the President's handling of the situation? [1 (No damage) - 5 (A lot of damage)]

Emotional Mediators: Do any of the following describe your feelings about the president's handling of the situation? [Yes *or* No]

Worried

Proud

Frightened

Angry

Hopeful

Table 2: Experiment 2: Mediation Analysis

	ACME	Proportion Mediated
Mediator:		
President's Reputation	0.38***	0.58***
Proud	0.27***	0.42***
Worried	0.19***	0.30***
Frightened	0.13***	0.20***
Angry	0.27***	0.42***
Hopeful	0.26***	0.40***

Note: *p<0.1; **p<0.05; ***p<0.01

Table 2 shows the average mediation effects (ACME) from Experiment 2 of mediators on audience approval and the proportion of the total effect mediated by the mediator calcualted using the mediation package in R (Tingley et al., 2014). Each mediator is run separately since there is not a clear ordering of which mediators theoretically precede one another in the causal pathway, which is required of multiple mediator analysis. Outcomes are held constant with the attacking country gaining 20% of the territory. The approval score is measured on a scale from -3=Strongly Disapprove to 3=Strongly Approve.

9 Cases of Compromise

Leaders regularly employ compromises in international negotiations. The discussion of the cases in the main text are somewhat brief, so I elaborate further on the Agadir crisis here. The Agadir crisis is a case that is regularly discussed in the literature and was cited by Fearon (1994) as a "prominent example" of inducing audience costs. Specifically, Lloyd George's famous Mansion House speech during the 1911 Agadir Crisis was a clear and public threat that had the potential to induce audience costs (Gartzke and Lupu, 2012; Trachtenburg, 2012). However, a close examination of Britain's involvement in the Agadir Crisis and subsequent events confirms that the crisis took place on the public stage, significant threats were made, and yet the audience chose not to punish its leaders when the government shifted to a softer diplomatic position.

There had been tensions over territorial and economic rights in Morocco for years, but the conflict came to the fore when the German ambassador announced on July 1st that Germany had sent a gunboat, the Panther, to the southern port of Agadir (Barlow, 1940). The British made their position clear when foreign secretary, Edward Grey, informed Germany's Metternich on May 22, 1911 that "England in any case and under all circumstances, would fulfill her obligations to France... [and] in the event of entanglements, all English obligations would become 'operative'" (Barlow, 1940). Although Grey remained optimistic that a peaceful resolution would be negotiated between Germany and France (Grey, 1911), he agreed that Lloyd George should make a public warning to Germany (Trachtenburg, 2012). On the night of July 21st, 1911, Lloyd George delivered his Mansion House speech, where he warned that England would not stand aside and be humiliated (Barlow, 1940). Going public changed the face of the crisis, as the speech invoked the national honor and the press ran with it to the public. The next day the London *Times* published the entire speech and endorsed its message as speaking on behalf of the nation, while the *Daily Chronicle* titled their article "England's Warning to Germany" (Barlow, 1940, 305). One of the most thorough historians on the Agadir Crisis noted in her interpretation of the events that the "[p]opular reception of Lloyd George's speech drove home the full force of its meaning and explains its effect upon the course of the crisis, for just as the foreign offices became more conciliatory, public opinion burst into chauvinistic flames" (Barlow, 1940).

Although the Mansion House speech inflamed public opinion, the government was not punished by its audience when it shifted to a more conciliatory strategy. By September, the British had

shifted their focus to alleviating anti-British sentiments in Germany (Barlow, 1940, 375). Rather than standing firmly behind the French on all accounts, the British were eager to put the conflict behind them and backed away from their hard-line policy (Trachtenburg, 2012). In his analysis of the case, Trachtenburg (2012, 21) finds that the government did not experience an audience cost for this shift, and in fact the government "would have paid a price if it had *not* done so." Instead of a hard-line negotiation, Germany was granted lands in the Congo in exchange for the French establishing a protectorate in Morocco (Barlow, 1940, 378-379). In resolving the Agadir Crisis, none of the major players completely backed down and none received their most preferred outcome. As with many international crises, a compromise with mutual concessions brought the crisis to a close and allowed the British leaders to avoid paying audience costs, even though their rhetoric is often cited as a source of such costs. The Agadir Crisis highlights that even in cases where clear and bellicose threats are made, publics often support compromise settlements on the the international stage.

References

Barlow, Christina. 1940. *The Agadir Crisis*. Chapel Hill, NC: University of North Carolina Press. doi: 10.1017/S0002930000229862.

Berinksy, Adam, Gregory Huber, Gabriel Lenz, and Edited by R. Michael Alvarez. 2012. "Evaluating Online Labor Markerts for Experimental Research: Amazon's Mechanical Turk." *Political Analysis* 20 (3): 351-368. doi: 10.1093/pan/mpr057.

Buhrmester, Michael, Tracy Kwang, and Samuel D. Gosling. 2011. "Amazon's Mechanical Turk: A New Source of Inexpensive, Yet High-Quality, Data?" *Perspectives on Psychological Science* 6 (1): 3-5. doi: 10.1177/1745691610393980.

Fearon, James. 1994. "Domestic Political Audiences and the Escalation of International Disputes." *American Political Science Review* 884 (September): 577 - 592. doi: DOI: https://doi.org/10.2307/2944796.

Gartzke, Eric, and Yonatan Lupu. 2012. "Still Looking for Audience Costs." *Security Studies* 21 (3): 391-397. doi: 10.1080/09636412.2012.706486.

Grey, Edward. 1911. *Sir Edward Grey to Sir F. Bertie, July 20, 1911.* HMSO, London chapter in *British Documents on the Origins of War*, 1898 − 1914, edited by G.P. Gooch and Harold Temperley, 1932, p. 382.

Horton, John J., David G. Rand, and Richard J. Zeckhauser. 2011. "The online laboratory: conducting experiments in a real labor market." *Experimental Economics* 14 (3): 399-425.

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): 234–249. doi: 10.1111/ajps.12201.

Levendusky, Matthew S., and Michael C. Horowitz. 2012. "When Backing Down is the Right Decision: Partisanship, New Information, and Audience Costs." *Journal of Politics* 74 (2): 323-338. doi: 10.1017/S002238161100154X.

Paolacci, Gabriele, Jesse Chandler, and Panagiotis G. Iperiotis. 2010. "Running experiments on Amazon Mechanical Turk." *Judgment and Decision Making* 5 (5): 411-419.

Rand, David G. 2012. "The promise of Mechanical Turk: How online labor markets can help theorists run behavioral experiments." *Journal of Theoretical Biology* 299: 172-179.

Sher, Shlomi, and Craig R.M. McKenzie. 2006. "Information leakage from logically equivalent frames." *Cognition* 101: 467-494.

Tingley, Dustin, Teppei Yamamoto, Kentaro Hirose, Luke Keele, and Kosuke Imai. 2014. "mediation: R Package for Causal Mediation Analysis." *Journal of Statistical Software* 59 (5): 1–38.

Tomz, Michael R., and Jessica L. P. Weeks. 2013. "Public Opinion and the Democratic Peace." *American Political Science Review* 107 (4): 849-865.

Trachtenburg, Marc. 2012. "Audience Costs: An Historical Analysis." *Security Studies* 21 (January): 3-42. doi: 10.1080/09636412.2012.650590.