**THE ‘COMMITMENT TRAP’ REVISITED**

Michal Smetana, Marek Vranka, Ondrej Rosendorf

*Supplementary Appendix*

Contents

[Appendix 1. Sociodemographic composition of the sample 2](#_Toc125111954)

[Appendix 2. Survey items 3](#_Toc125111955)

[Appendix 3. Regression analysis for *H1* and *H2* 2](#_Toc125111956)

[Appendix 4. Regression analysis for *H3* 3](#_Toc125111957)

[Appendix 5. Regression analysis for *H4* 4](#_Toc125111958)

[Appendix 6. Ethical considerations 5](#_Toc125111959)

[Appendix 7. Additional test for *H3* using “explicit threat” treatment 7](#_Toc125111960)

[Appendix 8. Treatment effects in percentage points 9](#_Toc125111961)

[Appendix 9. Selection of the North Korea scenario 13](#_Toc125111962)

[Appendix 10. Re-analysis based on manipulation check failures 15](#_Toc125111963)

[Appendix 11. Additional experiment using an alternative control condition 18](#_Toc125111964)

[Appendix 12. Additional experiment on the likelihood of different responses 22](#_Toc125111965)

[Appendix 13. Additional experiment on information equivalence 27](#_Toc125111966)

[Appendix 14. Additional experiment on preference for nuclear use 30](#_Toc125111967)

[Appendix 15. Alternative dependent variable (impact on credibility) 32](#_Toc125111968)

# Appendix 1. Sociodemographic composition of the sample

**Table 1. Summary statistics for Male**

|  |  |  |  |
| --- | --- | --- | --- |
| Male | Freq. | Percent | Cum. |
| Female | 499 | 49.95 | 49.95 |
| Male | 500 | 50.05 | 100.00 |
| Total | 999 | 100.00 |  |
|  | | | |

**Table 2. Summary statistics for Age**

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| Variable | Obs | Mean | Std. Dev. | Min | Max |
| Age | 997 | 36.863 | 13.81 | 18 | 84 |
|  | | | | | |

**Table 3. Summary statistics for Education (dummy)**

|  |  |  |  |
| --- | --- | --- | --- |
| Education (dummy) | Freq. | Percent | Cum. |
| Less than Bachelor's degree | 453 | 45.99 | 45.99 |
| Bachelor's degree or higher | 532 | 54.01 | 100.00 |
| Total | 985 | 100.00 |  |
|  | | | |

**Table 4. Summary statistics for Income**

|  |  |  |  |
| --- | --- | --- | --- |
| Income | Freq. | Percent | Cum. |
| Less than $10000 | 66 | 6.70 | 6.70 |
| $10000 to $15999 | 57 | 5.79 | 12.49 |
| $16000 to $19999 | 26 | 2.64 | 15.13 |
| $20000 to $29999 | 74 | 7.51 | 22.64 |
| $30000 to $39999 | 89 | 9.04 | 31.68 |
| $40000 to $49999 | 92 | 9.34 | 41.02 |
| $50000 to $59999 | 96 | 9.75 | 50.76 |
| $60000 to $69999 | 85 | 8.63 | 59.39 |
| $70000 to $79999 | 77 | 7.82 | 67.21 |
| $80000 to $89999 | 50 | 5.08 | 72.28 |
| $90000 to $99999 | 61 | 6.19 | 78.48 |
| $100000 to $149999 | 124 | 12.59 | 91.07 |
| More than $150000 | 88 | 8.93 | 100.00 |
| Total | 985 | 100.00 |  |
|  | | | |

**Table 5. Summary statistics for Party**

|  |  |  |  |
| --- | --- | --- | --- |
| Party | Freq. | Percent | Cum. |
| Republican | 327 | 33.10 | 33.10 |
| Democrat | 331 | 33.50 | 66.60 |
| Independent | 330 | 33.40 | 100.00 |
| Total | 988 | 100.00 |  |
|  | | | |

# Appendix 2. Survey items

* **Informed consent**

This survey aims to explore what people think of specific actions regarding geopolitical events. You will not be forced to answer any sensitive questions to complete the survey. It will take approximately 6 minutes to complete the survey.

Filling out the survey is harmless. There will be no detrimental after-effects. Your participation is entirely voluntary. You can terminate participation by simply closing the survey. In that case, all your responses will be discarded.

After evaluating your submission, your Prolific ID will be deleted from the collected data, which will therefore become fully anonymized. Anonymized data will be analyzed, used in scientific publications, and publicly shared with other researchers.

* **Part 1**

*Imagine that about three years from now, there is a very serious crisis on the Korean peninsula involving North Korea, the United States, and U.S. allies in the region. Please, read carefully all the information presented below.*

The United States has a new president James Smith. One day, President Smith makes the following announcement:

*For the control group:*

Graphical user interface, application

Description automatically generated

*For the ambiguous nuclear threat group:*



*For the explicit nuclear threat group:*

Graphical user interface, text, application

Description automatically generated

Imagine that you see this announcement in real life. How strongly do you feel the following emotions?

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
|  | Not at all (1) | A little (2) | Moderately (3) | A lot (4) | Extremely (5) |
| Joy😊 |  |  |  |  |  |
| Sadness 😥 |  |  |  |  |  |
| Anger 😡 |  |  |  |  |  |
| Fear 😨 |  |  |  |  |  |

* **Part 2**

*Now, imagine that this is how the situation then unfolded in the real world:*

* Two days after the U.S. President Smith’s announcement, North Korea used chemical weapons to strike Okinawa, a Japanese island with a large U.S. military presence.
* 1,400 Japanese civilians and 650 U.S. troops stationed in Okinawa died as a result of this chemical attack
* President Smith ordered the U.S. military to prepare a military response to strike military targets in North Korea

*For the nuclear use group:*

President Smith has decided that the U.S. military response will include the use of two nuclear weapons against some targets to send a strong signal to the North Korean leadership.

Graphical user interface, text, application

Description automatically generated

*For the nuclear nonuse group:*

President Smith has decided that the U.S. military response will not include the use of nuclear weapons and that conventional strikes will be enough to send a strong signal to the North Korean leadership.



* **Part 3**

How much do you approve or disapprove of U.S. President Smith’s **overall handling** of this crisis?

* Strongly approve (1)
* Approve (2)
* Somewhat approve (3)
* Neither approve nor disapprove (4)
* Somewhat disapprove (5)
* Disapprove (6)
* Strongly disapprove (7)

If you had to choose between using nuclear weapons and not using nuclear weapons in the scenario you have read, which of these two options would you prefer?

* strongly prefer not to use nuclear weapons (1)
* somewhat prefer not to use nuclear weapons (2)
* somewhat prefer to use nuclear weapons (3)
* strongly prefer to use nuclear weapons (4)
* **Part 4**

To review: in the scenario you have read, what was the content of the U.S. president's first tweet?

* **only** the possibility of a chemical strike by North Korea (1)
* the possibility of a chemical strike by North Korea **and** the nature of the U.S. response to it, but **without explicitly mentioning nuclear weapons** (2)
* the possibility of a chemical strike by North Korea **and** that the U.S. **will use nuclear weapons** in response to it (3)
* the possibility of a chemical strike by North Korea **and** that the U.S. **will not use nuclear weapons** in response to it (4)

To review: in the scenario you have read, did the U.S. president ordered the use of nuclear weapons in the end?

* the U.S. president **ordered** the use of two nuclear weapons (1)
* the U.S. president **did not order** the use of nuclear weapons (2)
* **Survey debriefing**

Thank you very much for participating in this survey!

Note that we are interested in factors that affect how people think about the use of military force in world politics. We do not argue for the use of military force, and we stress that the fictional scenario and options provided in the survey do not cover the full spectrum of aspects and considerations that would be important in reality. The survey is used only for academic purposes and cannot be used to justify the use of military force in the real world. Importantly, any use of military force should be in line with international law, which stipulates under what conditions states are allowed to use military force against other states (*jus ad bellum*) and how the states should behave in war (*jus in bello*).

If you have any additional comments or concerns, please reach out to us at *(e-mail)*

# Appendix 3. Regression analysis for *H1* and *H2*

**Table 1.** Ordinal logistic regression of crisis handling approval

|  |  |  |
| --- | --- | --- |
|  | Model 1 | Model 2 |
| Treatment (control - ambiguity) | 1.076 | 1.102 |
| (0.210) | (0.221) |
| Treatment (explicit - ambiguity) | 0.510\*\*\* | 0.520\*\* |
| (0.102) | (0.107) |
| Gender (male) |  | 1.394\* |
|  |  | (0.233) |
| Age |  | 1.021\*\*\* |
|  |  | (0.00634) |
| Income |  | 0.979 |
|  |  | (0.0231) |
| Education (university degree) |  | 0.754 |
|  | (0.130) |
| Party (Democrat - Republican) |  | 1.090 |
|  | (0.227) |
| Party (Independent - Republican) |  | 0.709 |
|  | (0.144) |
| *N* | 489 | 476 |
| Pseudo *R*2 | 0.010 | 0.024 |

Standard errors in parentheses

\* *p* < 0.05, \*\* *p* < 0.01, \*\*\* *p* < 0.001

# Appendix 4. Regression analysis for *H3*

**Table 1.** Ordinal logistic regression of crisis handling approval

|  |  |  |
| --- | --- | --- |
|  | Model 3 | Model 4 |
| Response (nuclear - non-nuclear) | 0.282\*\*\* | 0.298\*\*\* |
| (0.0574) | (0.0615) |
| Gender (male) |  | 0.908 |
|  |  | (0.180) |
| Age |  | 1.014 |
|  |  | (0.00751) |
| Income |  | 1.056 |
|  |  | (0.0313) |
| Education (university degree) |  | 0.846 |
|  | (0.172) |
| Party (Democrat - Republican) |  | 0.558\* |
|  | (0.138) |
| Party (Independent - Republican) |  | 0.578\* |
|  | (0.140) |
| *N* | 337 | 332 |
| Pseudo *R*2 | 0.033 | 0.046 |

Standard errors in parentheses

\* *p* < 0.05, \*\* *p* < 0.01, \*\*\* *p* < 0.001

# Appendix 5. Regression analysis for *H4*

**Table 1.** Ordinal logistic regression of nuclear weapons use preference

|  |  |  |
| --- | --- | --- |
|  | Model 5 | Model 6 |
| Treatment (ambiguity - control) | 0.864 | 0.945 |
| (0.126) | (0.211) |
| Treatment (explicit - control) | 0.908 | 1.020 |
| (0.134) | (0.232) |
| Response (nuclear - non-nuclear) |  | 2.024\*\* |
|  | (0.436) |
| Treatment (ambiguity) \* Response (nuclear) |  | 0.809 |
|  | (0.245) |
| Treatment (explicit) \* Response (nuclear) |  | 0.755 |
|  | (0.231) |
| Gender (male) |  | 0.680\*\* |
|  |  | (0.0849) |
| Age |  | 0.985\*\* |
|  |  | (0.00448) |
| Income |  | 1.003 |
|  |  | (0.0176) |
| Education (university degree) |  | 1.039 |
|  | (0.135) |
| Party (Democrat - Republican) |  | 0.458\*\*\* |
|  | (0.0716) |
| Party (Independent - Republican) |  | 0.480\*\*\* |
|  | (0.0724) |
| *N* | 1001 | 979 |
| Pseudo *R*2 | 0.000 | 0.035 |

Standard errors in parentheses

\* *p* < 0.05, \*\* *p* < 0.01, \*\*\* *p* < 0.001

# Appendix 6. Ethical considerations

The study was conducted in accordance with the principles of the Declaration of Helsinki, all relevant regulations for conducting academic survey research at Charles University, and the APSA Principles and Guidance for Human Subjects Research. The data collection for this research study also adhered to the general data collection approach used in the Charles University’s research project “Experimental Lab for International Security Studies (ELISS)” that was approved without reservations by the Commission for Ethics in Research of the Faculty of Social Sciences, Charles University (submission #53).

We asked our participants for informed and voluntary consent at the beginning of the study, after we explained the general aims, duration, low-risk nature, and other details regarding our study (see the text of the informed consent below):

*Informed consent*

*This survey aims to explore what people think of specific actions regarding geopolitical events. You will not be forced to answer any sensitive questions to complete the survey. It will take approximately 6 minutes to complete the survey.*

*Filling out the survey is harmless. There will be no detrimental after-effects. Your participation is entirely voluntary. You can terminate participation by simply closing the survey. In that case, all your responses will be discarded.*

*After evaluating your submission, your Prolific ID will be deleted from the collected data, which will therefore become fully anonymized. Anonymized data will be analyzed, used in scientific publications, and publicly shared with other researchers.*

*The study is conducted by a research team at the [Redacted]. In case of any questions, please contact [Redacted], the researcher responsible for the survey, at [Redacted].*

*I declare that I have received full detailed information concerning the conditions of my participation in the study. I agree with these conditions and I am willing to participate.*

We did not use any kind of deception and our participants faced no risk of harm or traumatization. In addition, we concluded the study with the following debriefing item to contextualize the presented study materials and preclude creation of possible misconceptions about the use of nuclear weapons:

*Thank you very much for participating in this survey!*

*Note that we are interested in factors that affect how people think about the use of military force in world politics. We do not argue for the use of military force, and we stress that the fictional scenario and options provided in the survey do not cover the full spectrum of aspects and considerations that would be important in reality. The survey is used only for academic purposes and cannot be used to justify the use of military force in the real world. Importantly, any use of military force should be in line with international law, which stipulates under what conditions states are allowed to use military force against other states (jus ad bellum) and how the states should behave in war (jus in bello).*

*If you have any additional comments or concerns, please reach out to us at (e-mail)*

We conducted the data collection on the Prolific platform that offers a relatively diverse subject pool (Newman et al., 2020).[[1]](#footnote-1) Moreover, we used quotas to assure equal gender representation. All participants were paid in accordance with the standards for participants’ payments on the Prolific platform. We did not collect any personal data or data allowing revelation of identity of our participants.

*References:*

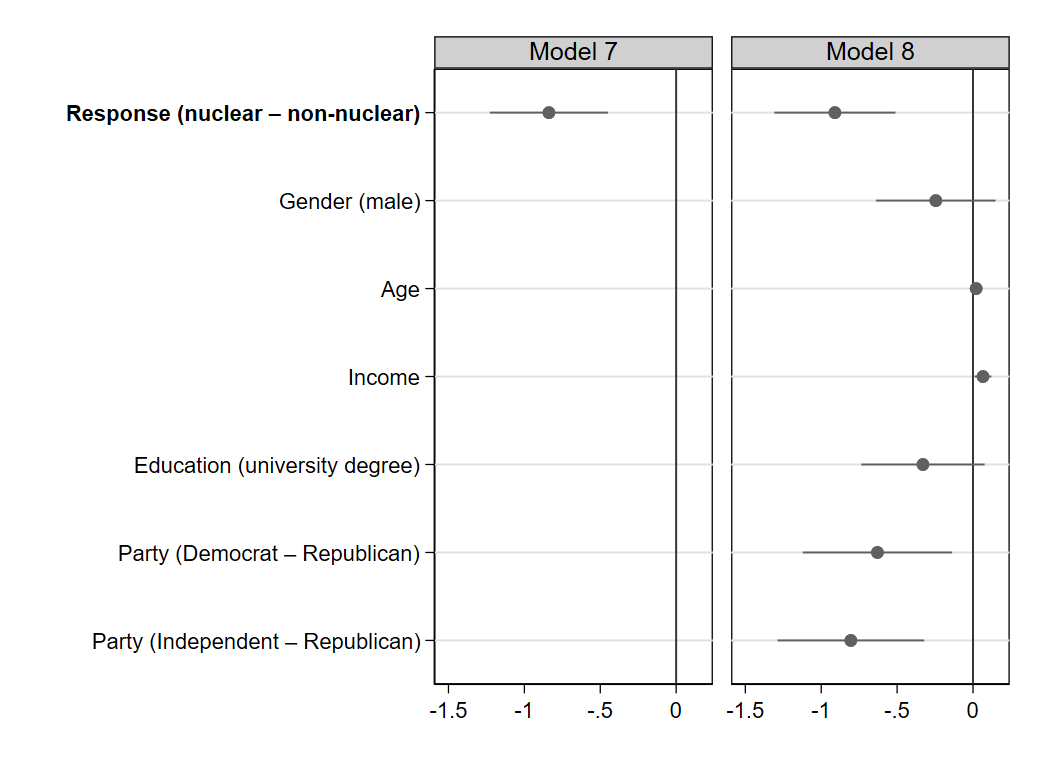
Newman, Alexander, Bavik, Yuen Lam, Mount, Matthew, & Shao, Bo (2021). Data collection via online platforms: Challenges and recommendations for future research. *Applied Psychology*, *70*(3), 1380-1402.

Palan, Stefan, and Christian Schitter. 2018. “Prolific.Ac—A Subject Pool for Online Experiments.” *Journal of Behavioral and Experimental Finance* 17: 22–27.

Peer, Eyal, Laura Brandimarte, Sonam Samat, and Alessandro Acquisti. 2017. “Beyond the Turk: Alternative Platforms for Crowdsourcing Behavioral Research.” *Journal of Experimental Social Psychology* 70: 153–63.

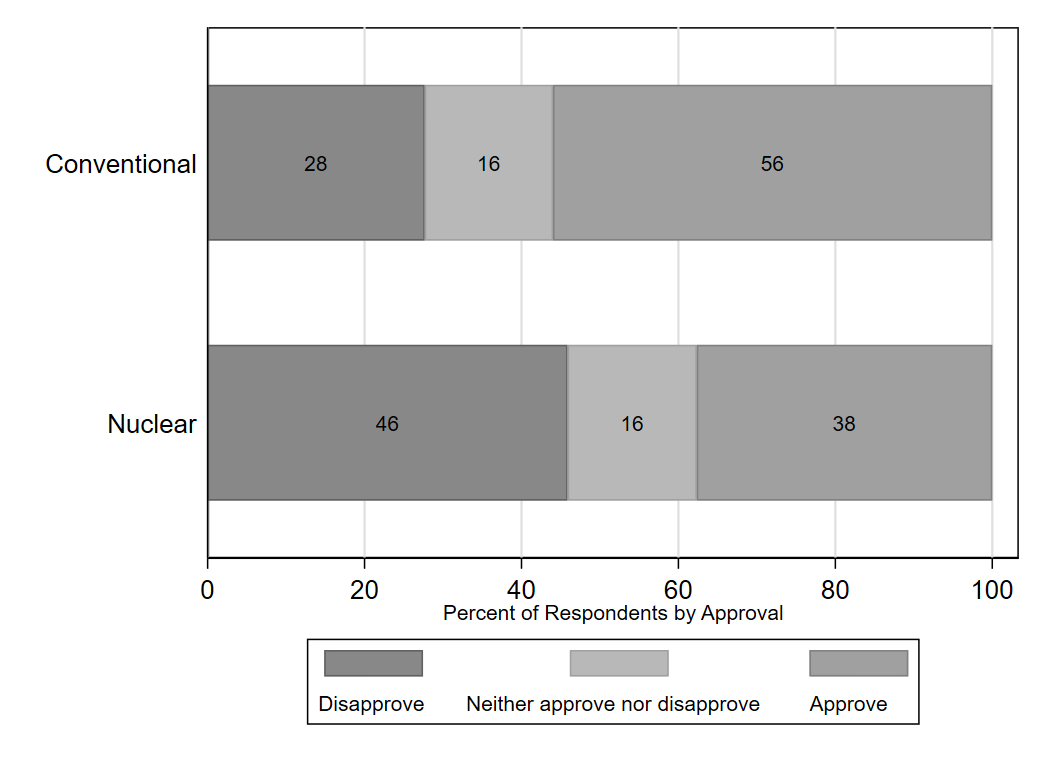
# Appendix 7. Additional test for *H3* using “explicit threat” treatment

**Figure 1.** Ordinal logistic regression estimates



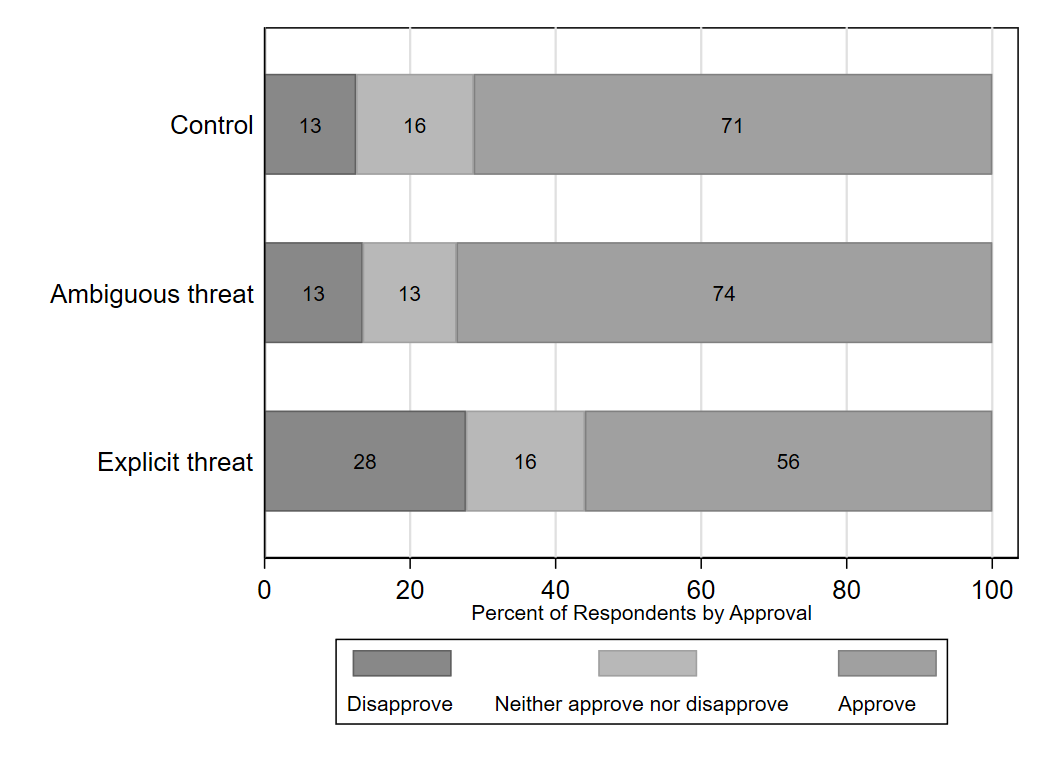
*Note:* Ordinal logistic regression estimates. *N* = 332. 95% CI. Variables whose intervals overlap with the vertical line are statistically indistinguishable from 0. Positive coefficients correspond to a higher level of approval. These results are for participants in the explicit threat treatment (*α2*EXP) with approval as an outcome variable, response *β* as a predictor, and sociodemographic characteristics as control variables. As in the ambiguous threat treatment (*α2*AMB), the nuclear response was negatively associated with approval (*p* < 0.001). Model 7 shows the treatment effect without the inclusion of control variables and Model 8 with control variables included.

**Figure 2.** Approval across response



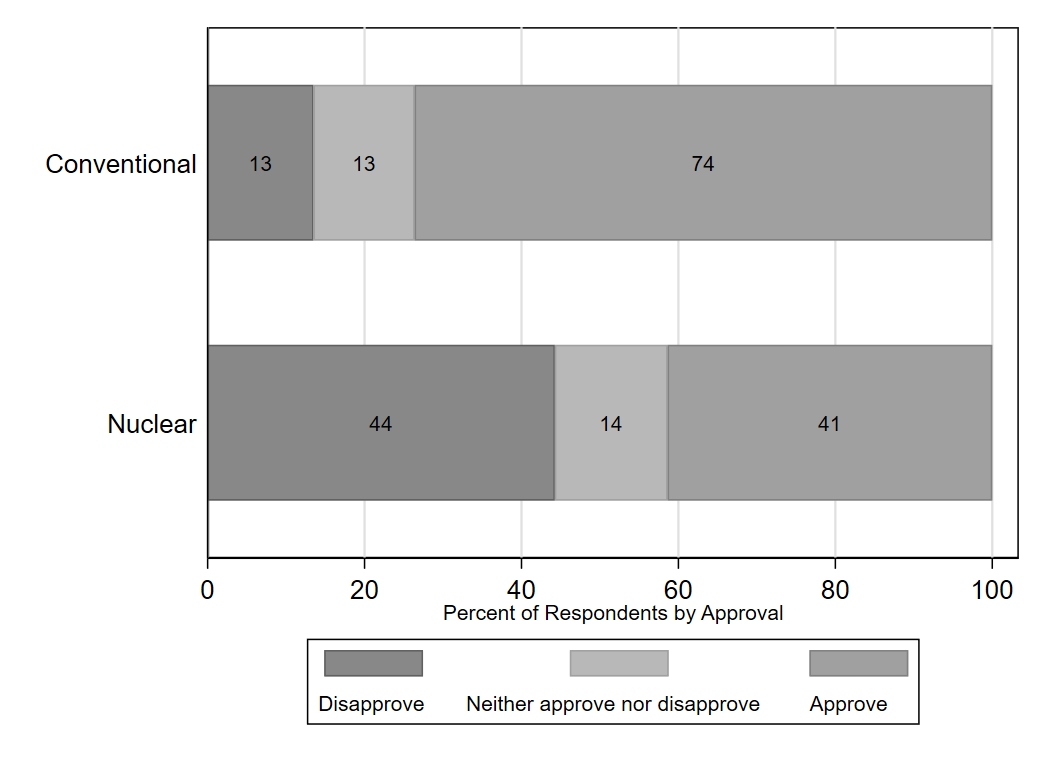
*Note:* An alternative way of calculating domestic audience costs with respect to the test reported above.*.* The figure displays the percentage of respondents who approved, disapproved, and neither approved nor disapproved of the President’s handling of the crisis in treatment *α*EXPunder two different treatments *β1* and *β2*.

# Appendix 8. Treatment effects in percentage points

**Figure 1.** Approval across treatments 

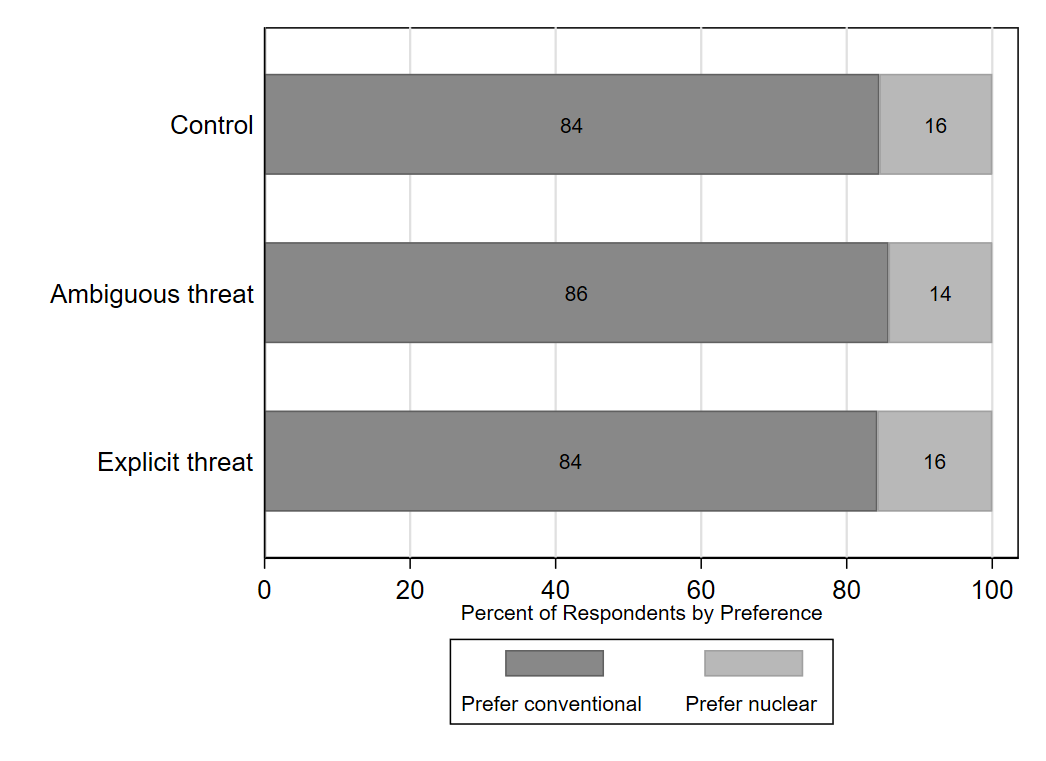
*Note:* An alternative way of calculating domestic audience costs with respect to *H1* and *H2.* The figure displays the percentage of respondents who approved, disapproved, and neither approved nor disapproved of the President’s handling of the crisis in treatment *β1* under three different treatments *α*.

**Figure 2.** Approval across response



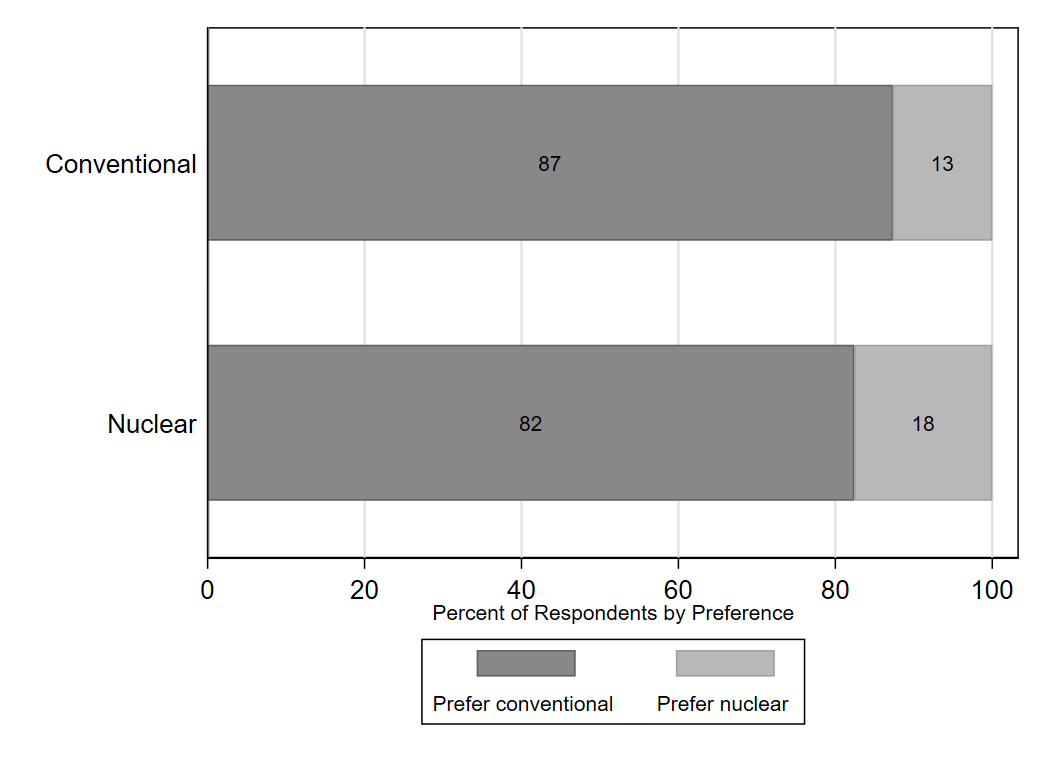
*Note:* An alternative way of calculating domestic audience costs with respect to *H3.* The figure displays the percentage of respondents who approved, disapproved, and neither approved nor disapproved of the President’s handling of the crisis in treatment *α*AMBunder two different treatments *β1* and *β2*.

**Figure 3.** Preference across treatments



*Note:* An alternative way of calculating preference for nuclear use with respect to *H4*. The figure displays the percentage of respondents who preferred a conventional response and those who preferred a nuclear response in three different treatments *α*.

**Figure 4.** Preference across response



*Note:* An alternative way of calculating preference for nuclear use, based on whether the respondents were assigned to a “nuclear response” group or a “conventional response” group. The figure displays the percentage of respondents who preferred a conventional response and those who preferred a nuclear response in treatments *β1* and *β2*.

# Appendix 9. Selection of the North Korea scenario

As noted by one of the anonymous reviewers of our paper, it is important to discuss both the advantages and disadvantages of using North Korea as a U.S. adversary in our scenario. The reviewer adds that “*one potential limitation is that since North Korea is a nuclear weapons power, the dangers of the US using nuclear weapons against them are much greater than if a non-nuclear power (e.g., Syria) was used in the vignette. In that sense, the choice to use North Korea might make this a relatively ‘easy test’ of the authors' argument because members of the public will likely be more hesitant to support the use of nuclear weapons in general and thus may punish leaders less for making an ambiguous nuclear threat and then ultimately deciding not to use nuclear weapons.”*

North Korea was selected as the most likely case where a “commitment trap” scenario could develop in a way discussed by Scott Sagan in his 2000 *International Security* paper. North Korea is estimated to have a sizable chemical weapon program and substantial quantities of prohibited chemical agents, including nerve agents. It is one of the last few countries (with Egypt and Sudan)[[2]](#footnote-2) that have neither signed nor ratified the Chemical Weapons Convention (CWC), and it is, therefore, not subject to the CWC verification mechanisms such as on-site inspections. It is considered a serious security threat to U.S. allies in the region, who rely on the U.S. extended deterrence. Using Syria instead of North Korea in our experiment would certainly address the concern about nuclear retaliation, but it would present us with another set of problems. While there are doubts about the completeness of the Syrian government’s declarations to the CWC and the full extent of the subsequent destruction of its chemical weapons stockpile, it is highly questionable that Damascus would be capable of launching a significant chemical attack against U.S. troops and allies in the region that would warrant a U.S. nuclear response. Moreover, unlike in East Asia, U.S. allies and partners in the Middle East, such as Israel and Saudi Arabia, are significantly stronger in terms of their military power vis-á-vis Syria and could (and most likely would) retaliate militarily against Syria by themselves or with U.S. conventional support (as a nuclear-armed state, Israel could even theoretically issue ambiguous nuclear threats by itself). It is also worth mentioning that, unlike North Korea, Syria – as a non-nuclear member of the NPT – would generally be covered by the U.S. negative security assurances, which could impact the perceived legitimacy of a nuclear response. Overall, it is arguably more difficult to devise a plausible contemporary scenario where the U.S. President – whether explicitly or ambiguously – threatens Syria with nuclear strikes. That is why North Korea (unlike Syria) is explicitly discussed in the U.S. Nuclear Posture Review (NPR) and other strategic documents as a target of U.S. nuclear deterrence posture, not least for the *“the threat posed by its nuclear,* ***chemical****, missile, and conventional capabilities.”* With respect to North Korea and U.S. nuclear response to non-nuclear strategic strikes, the 2022 NPR even states that *“[s]hort of nuclear use, North Korea can also conduct rapid strategic attacks in East Asia. United States nuclear weapons continue to play a role in deterring such attacks”* (U.S. Department of Defense 2022, p. 12).

Moreover, recent survey experiments suggest that the U.S. public is generally less concerned about North Korea’s nuclear retaliation than it may seem. For example, a 2019 experimental study by Haworth, Sagan, and Valentino used scenarios of U.S. conventional and nuclear strikes against North Korea. They found relatively high support for strikes against North Korea even when the participants were informed about a non-trivial chance of North Korea retaliating with nuclear weapons against the U.S. homeland. The study also found that *“[s]eventy-four percent of respondents believe it is highly likely or somewhat likely that, if North Korea were to launch three missiles with nuclear warheads at the United States, ‘current US missile defenses could successfully destroy all the North Korean missiles before they reach their targets’* (Haworth, Sagan, and Valentino 2019, p. 184). This finding suggests that large segments of the U.S. population are overly optimistic about the U.S. prospects of successfully defending against North Korea’s nuclear retaliation.

Another recent experimental study that used a comparable scenario involving North Korea found no statistically significant evidence for a hypothesis that *“a direct threat of retaliation against a subject’s home country will decrease willingness of American, South Korean, or Japanese respondents to use nuclear weapons”* (Allison, Herzog, and Ko 2022, p. 20). To this end, the authors highlight that priming the subjects on the possibility of retaliation actually creates heterogeneous effects. On the one hand, *“North Korean nuclear retaliatory capability likely causes some respondents to become more circumspect”*; on the other hand, *“many respondents actually became more bellicose and inclined to use nuclear weapons when threatened”* (Allison, Herzog, and Ko 2022, p. 20–21).

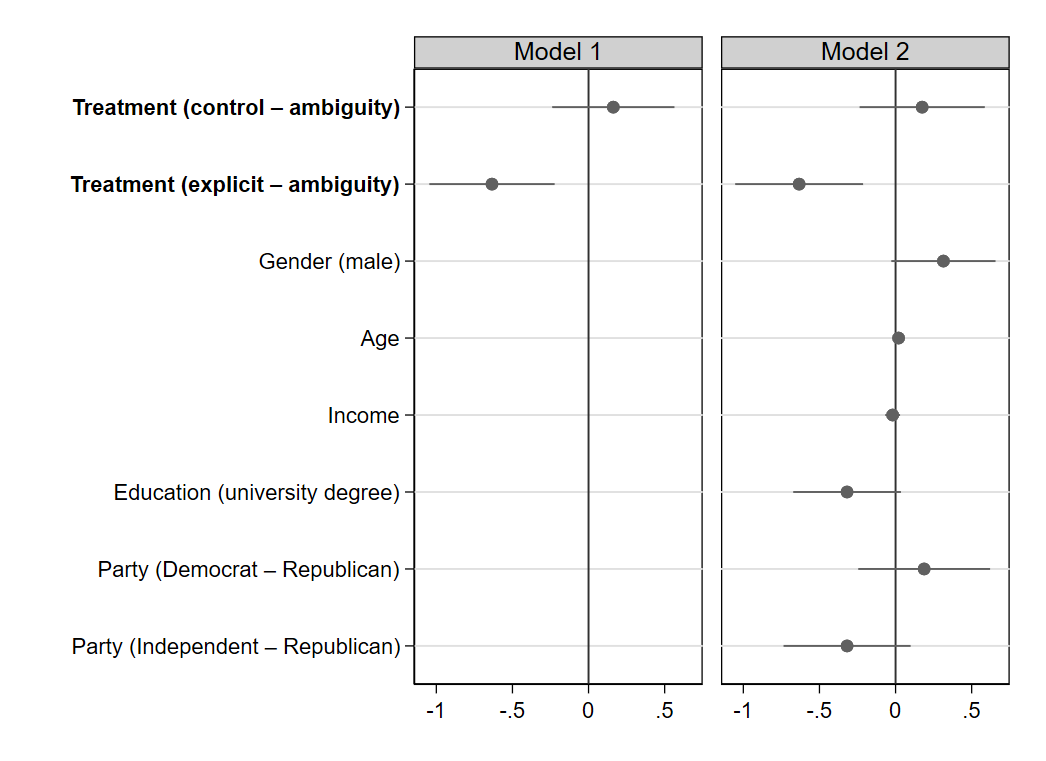
Overall, we found that at least in the context of a chemical threat against U.S. troops and allies, North Korea’s attack is the most suitable scenario for testing Sagan’s claim and the most realistic and policy relevant scenario with respect to U.S. nuclear posture today. Moreover, we believe that the evidence discussed above supports our claim that the risks of seeing significantly lower public support for the nuclear option given North Korea’s retaliatory capability are fairly limited and should not impact our results significantly.

# Appendix 10. Re-analysis based on manipulation check failures

Below is a replication of our main analyses after we removed those participants who failed our main manipulation check.[[3]](#footnote-3) We show that our results are robust, and the interpretation of treatment effects does not change when we used only this specific subset of participants.

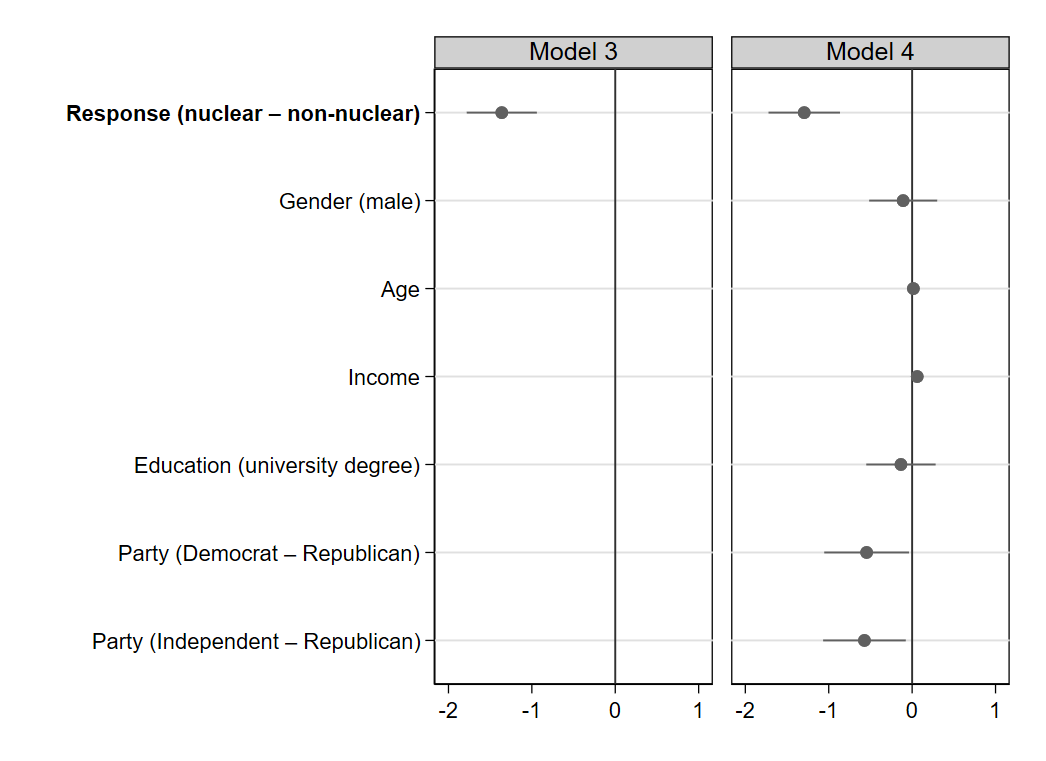
*Hypothesis 1+2 (compare with Figure 4 in the main manuscript)*

**Figure 1.** Ordinal regression estimates

*Note:* Ordinal logistic regression estimates. *N* = 434.95% CI. Variables whose intervals overlap with the vertical line are statistically indistinguishable from 0. Positive coefficients correspond to a higher level of approval. Model 1 shows the treatment effects without the inclusion of control variables and Model 2 with control variables included.

*Hypothesis 3 (compare with Figure 5 in the main manuscript)*

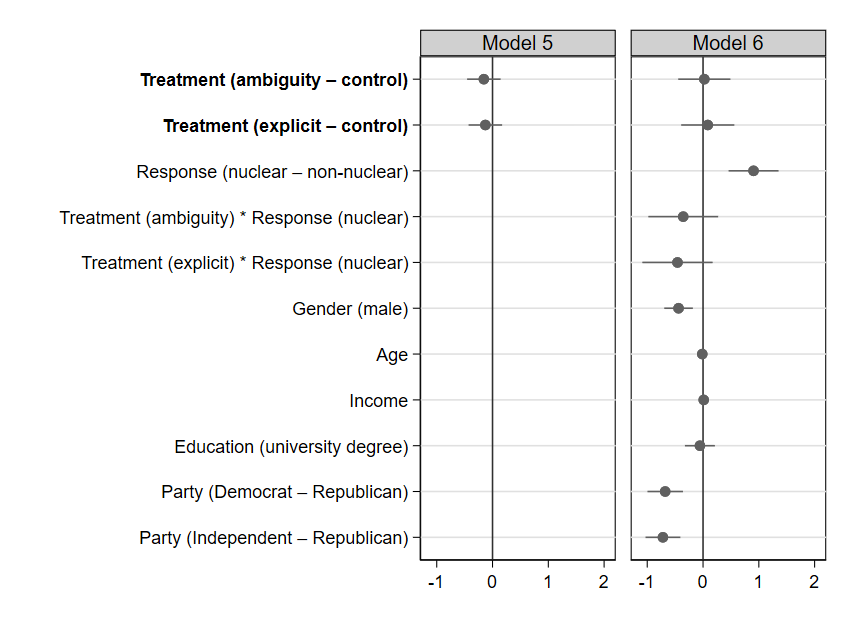
**Figure 2.** Ordinal regression estimates



*Note:* Ordinal logistic regression estimates. *N* = 304. 95% CI. Variables whose intervals overlap with the vertical line are statistically indistinguishable from 0. Positive coefficients correspond to a higher level of approval. Model 3 shows the treatment effect without the inclusion of control variables and Model 4 with control variables included.

*Hypothesis 4 (compare with Figure 6 in the main manuscript)*

**Figure 3.** Ordinal regression estimates



*Note:* Ordinal logistic regression estimates. *N =*901.95% CI. Variables whose intervals overlap with the vertical line are statistically indistinguishable from 0. Positive coefficients correspond to a higher preference for nuclear use. Model 5 shows the treatment effects without the inclusion of control variables and Model 6 with these variables included.

# Appendix 11. Additional experiment using an alternative control condition

During the peer-review process, one of the anonymous reviewers noted that our *α1* treatment (control group) does not fully correspond to the “stay out” treatment in the canonical audience costs literature. In the latter, the leaders explicitly state that they would stay out of intervening militarily in a hypothetical conflict (and then they do not intervene). In our control group / “stay out” treatment, the leader does not explicitly state that he would “stay out” of using nuclear weapons. The reviewer suggested that our operationalization could have led to higher disapproval in the control condition as some respondents could see the decision to use conventional strikes as an escalation if there was no threat issued beforehand (whereas in the traditional “stay out” treatment, the leaders are supposed to be following through on their earlier explicit promise to stay out). A lower approval of the President’s handling of the crisis in the control condition would then impact our overall results, as lower *μ*12 (pay off for the combination of *α1* and *β2* treatments)reduces overall audience costs (*μ*12 – *μ*22).

There were three reasons for our choice of the control condition in the original experiment. First, we felt it was more realistic not to make a threat at all than to state in advance that the U.S. would only use conventional weapons and rule out nuclear use. Second, we were convinced that the decision to engage in conventional strikes in our “stay out” treatment would not be seen as “escalatory”; there has been a recent precedent where the U.S. President responded to Syria’s use of chemical weapons with military strikes even when neither U.S. troops nor allies were among victims of this chemical attack and there was neither explicit nor ambiguous threat issued beforehand. That conventional strike response was generally praised by commentators across the U.S. political spectrum (Doucet 2018). We, therefore, assumed that conventional strikes against North Korea following its chemical attack against U.S. troops and allies would be a minimum baseline expectation by the U.S. public, even in the absence of any preceding threat (and, therefore, such a step would not be seen as “inconsistent” by the participants). Third, and perhaps most importantly, we were concerned that a non-trivial number of respondents would find the “nuclear-stay-out” statement (in which the U.S. President effectively assures a direct U.S. adversary that he would not order the use of nuclear weapons even in response to a WMD attack) more problematic than the “stay out” statement used in the traditional audience costs experiment. Pledging not to intervene on behalf of a generic (supposedly non-allied) country (as described in the traditional audience costs scenarios) is not, by itself, contradicting the U.S. defense policy; in contrast, explicitly pledging not to use nuclear weapons in response to a chemical attack against U.S. allies arguably contradicts the U.S. post-Cold War nuclear posture, which keeps the possibility of nuclear retaliation against WMD attacks deliberately open. The statement threatening a conventional response but ruling out a nuclear response to a chemical strike would correspond to the logic of the no-first-use or sole-purpose doctrines, which have been repeatedly debated in the United States but have not been adopted by any administration to date. Ultimately, we were concerned that the more “traditional” operationalization of our control condition could lower the approval of the President’s handling of the crisis and thereby artificially lower the audience costs when comparing President’s approval in “stay out” and “not engage” conditions (*μ*12 – *μ*22); generally, we opted for what we saw as a more conservative approach for our study and an easier test for Sagan’s (2000) “commitment trap” claim.

To find out whether the concern raised by the reviewer was justified, we have designed an additional survey experiment and fielded it to a sample of 730 U.S. adult citizens via the Prolific online platform. Our aim was to test empirically whether the approval of the U.S. President’s handling of the crisis would be higher if he explicitly said that he would stay out of using nuclear weapons and used conventional weapons after deterrence failed (new *α1* treatment), than if he did not issue any threat at all (old *α1* treatment). If we would find a higher approval in the new *α1* treatment, it would support the reviewer’s argument that we selected an easier test for our claim and a tougher test for Sagan’s claim; if we found higher approval in the old *α1* treatment, it would support our argument that we selected a tougher test for our claim and an easier test for Sagan’s claim; if the approval was statistically indistinguishable, then the two treatments were possibly interchangeable, and we did not err in selecting the original one.

We used a similar setup for this new experiment as for the original one. In the survey vignette, we once again described a development of a crisis on the Korean peninsula involving North Korea, the United States, and U.S. allies in the region (see Appendix 2). We, however, randomly assigned our participants to only two treatments: the “old” *α1* treatment (*α1*OLD; see Figure A1) and the “new” *α1* treatment (*α1*NEW; see Figure A2). The wording of *α1*NEW was made to resemble *α2*EXP (explicit nuclear threat), except the President explicitly ruled out nuclear use and promised a conventional response. The rest of the experiment was identical to the original version, except the participants were assigned to only one response treatment group (non-nuclear response *β2*).

**Figure A1. *α1*OLD treatment**

Graphical user interface, text, application

Description automatically generated

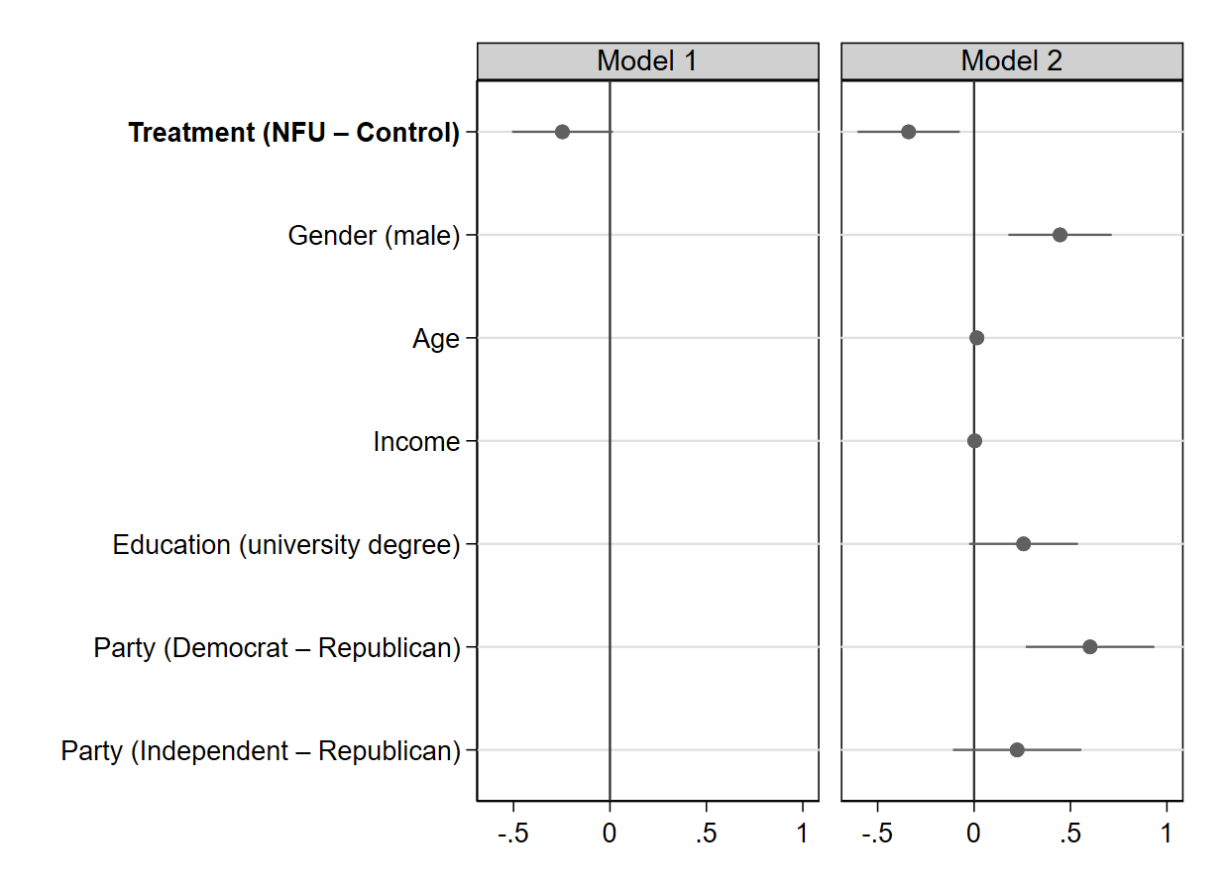
**Figure A2. *α1*NEW treatment**

Text

Description automatically generated

As we show in Figure 1 (and the regression table below), the approval of the President’s handling of the crisis was higher in *α1OLD* than in *α1NEW*. In the Model 1 without control variables, the difference in approval between the two conditions did not pass the test of statistical significance (*p* > 0.05), but it was statistically significant in Model 2 with control variables included (*p* < 0.05). As such, we reject the hypothesis that the President’s approval would be higher under in the new formulation of the control condition.Overall, findings of our follow-up experiment demonstrate that we originally selected a tougher test for our claim and an easier test for Sagan’s “commitment trap” claim – and, therefore, that we originally used a more conservative approach, which should give further credence to the argument advanced in our paper.

**Figure 1.** Ordinal regression estimates



*Note:* Ordinal logistic regression estimates.*N* = 730 (Model 1), 723 (Model 2).95% CI. Variables whose intervals overlap with the vertical line are statistically indistinguishable from 0. Positive coefficients correspond to a higher level of approval. Model 1 shows the treatment effects without the inclusion of control variables and Model 2 with control variables included.

**Table 1.** Ordinal logistic regression of crisis handling approval

|  |  |  |
| --- | --- | --- |
|  | Model 1 | Model 2 |
| Treatment (NFU - control) | 0.782 | 0.712\* |
| (0.104) | (0.0963) |
| Gender (male) |  | 1.562\*\* |
|  |  | (0.214) |
| Age |  | 1.015\*\* |
|  |  | (0.00512) |
| Income |  | 1.003 |
|  |  | (0.0193) |
| Education (university degree) |  | 1.293 |
|  | (0.186) |
| Party (Democrat - Republican) |  | 1.825\*\*\* |
|  | (0.311) |
| Party (Independent - Republican) |  | 1.250 |
|  | (0.212) |
| *N* | 730 | 723 |
| Pseudo *R*2 | 0.001 | 0.018 |

Standard errors in parentheses

\* *p* < 0.05, \*\* *p* < 0.01, \*\*\* *p* < 0.001

# Appendix 12. Additional experiment on the likelihood of different responses

In the main manuscript, we report the overall results of an additional experiment fielded to 151 U.S. adult respondents. Our goal was to examine how the public perceives the message behind ambiguous and explicit nuclear threats. We were interested, in particular, in whether our respondents were less likely to expect the use of nuclear weapons after an ambiguous threat than after an explicit one.

***Experimental design:***

Q1 – condition A:

Imagine that about three years from now, there is a very serious crisis on the Korean peninsula involving North Korea, the United States, and U.S. allies in the region, South Korea and Japan. The United States has a new president James Smith. One day, President Smith makes the following announcement:

**Obsah obrázku text

Popis byl vytvořen automaticky**

Imagine that you see this announcement in real life. Which of these emotions do you feel and how strongly?

TYPE: Matrix, Randomize statements

Joy😊

Sadness 😥

Anger 😡

Fear 😨

1–Not at all

2–A Little

3–Moderately

4–A Lot

5–Extremely

Q1 - condition B:

Imagine that about three years from now, there is a very serious crisis on the Korean peninsula involving North Korea, the United States, and U.S. allies in the region, South Korea and Japan. The United States has a new president James Smith. One day, President Smith makes the following announcement:

**Obsah obrázku text

Popis byl vytvořen automaticky**

**Imagine that you see this announcement in real life. Which of these emotions do you feel and how strongly?**

**TYPE: Matrix**

**Randomize statements**

**Joy😊**

**Sadness 😥**

**Anger 😡**

**Fear 😨**

**1–Not at all**

**2–A Little**

**3–Moderately**

**4–A Lot**

**5–Extremely**

Q2 - condition A:

In the scenario you have just read, President Smith said that *“if North Korea was foolish enough to use chemical weapons, our response will be absolutely overwhelming and devastating, and all military options will be on the table!”* After reading this, what do you expect that the President would actually do if North Korea really used chemical weapons against U.S. troops and allies?

President Smith would respond…

…with a nuclear weapon strike against North Korea

…with a conventional, non-nuclear missile strike against North Korea

…with a cyber strike against North Korea

…by sending U.S. troops to invade North Korea

…by strengthening economic sanctions against North Korea

…by doing nothing

(response on a 0–100 scale slider from very unlikely to very likely)

Q2 - condition B:

In the scenario you have just read, President Smith said that *“if North Korea was foolish enough to use chemical weapons, we will strike back with our powerful nuclear arsenal!”* After reading this, what do you expect that the President would actually do if North Korea really used chemical weapons against U.S. troops and allies?

President Smith would respond…

…with a nuclear weapon strike against North Korea

…with a conventional, non-nuclear missile strike against North Korea

…with a cyber strike against North Korea

…by sending U.S. troops to invade North Korea

…by strengthening economic sanctions against North Korea

…by doing nothing

(response on a 0–100 scale slider from very unlikely to very likely)

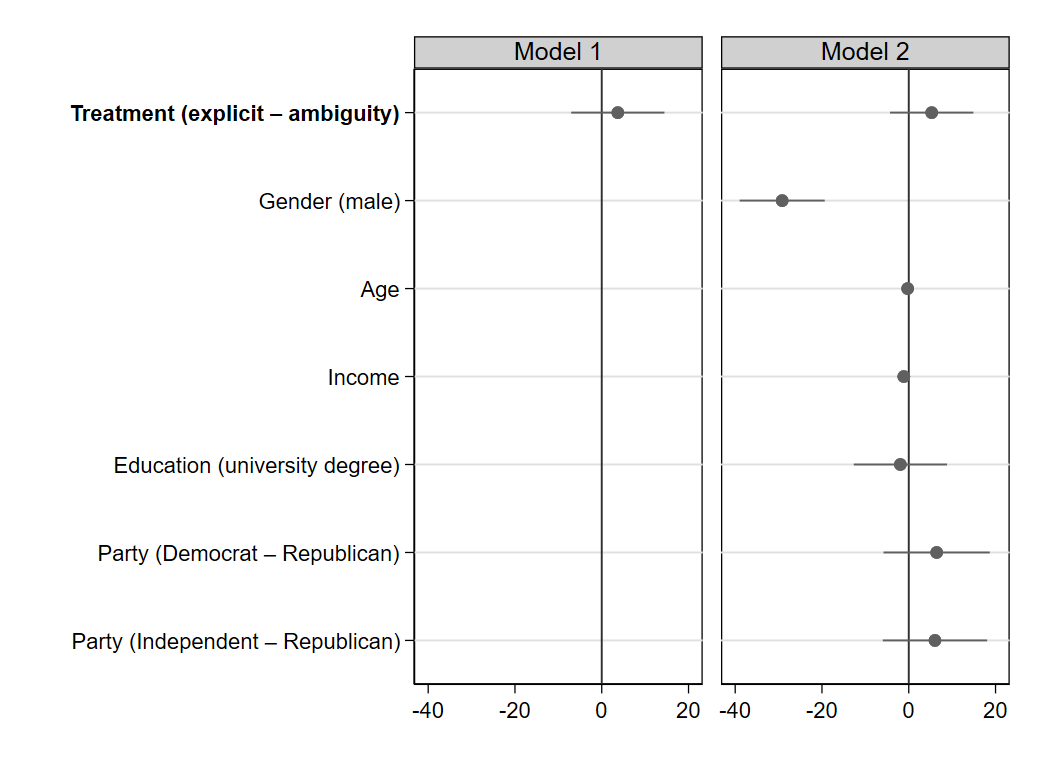
***Results:***

In Table 1, we show the results of independent samples t-test (Welch’s), examining the mean differences for each policy response between the two treatments. The only statistically significant difference was in the expectation of cyber strikes, which was found to be significantly more likely in the ambiguous treatment than in the explicit treatment. Figure 1 and the corresponding regression table show the results of an ordinal logistic regression with the likelihood of a nuclear response as an outcome variable and treatment as a predictor. We show that even after the inclusion of sociodemographic control variables, there was no statistically significant association between the treatment and the estimated likelihood of a nuclear response.

**Table 1: Welch's t-test results**

|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
|  | obs1 | obs2 | Mean1 | Mean2 | dif | St Err | t value | p value |
| nuclear | 75 | 76 | 50.720 | 54.421 | -3.701 | 5.433 | -.7 | .497 |
| conventional | 75 | 76 | 72.133 | 66.145 | 5.989 | 4.593 | 1.3 | .195 |
| cyber | 75 | 76 | 66.507 | 55.829 | 10.678 | 4.558 | 2.35 | .021 |
| invade | 75 | 76 | 69.587 | 61.25 | 8.337 | 4.422 | 1.9 | .061 |
| economic | 75 | 75 | 80.666 | 74 | 6.667 | 4.231 | 1.6 | .117 |
| nothing | 75 | 76 | 20.506 | 19.434 | 1.073 | 4.078 | .25 | .793 |

**Figure 1.** OLS regression estimates



*Note:* OLS regression estimates for the likelihood of a nuclear response. *N* = 151 (Model 1), *N =* 5150 (Model 2). 95% CI. Variables whose intervals overlap with the vertical line are statistically indistinguishable from 0. Positive coefficients correspond to a higher estimated likelihood of nuclear use. Model 3 shows the effects without the inclusion of control variables and Model 4 with control variables included.

**Table 2.** OLS regression of nuclear response likelihood

|  |  |  |
| --- | --- | --- |
|  | Model 1 | Model 2 |
| Treatment (explicit - ambiguity) | 40.49 | 197.8 |
| (220.2) | (961.1) |
| Gender (male) |  | 2.16e-13\*\*\* |
|  |  | (1.07e-12) |
| Age |  | 0.774 |
|  |  | (0.147) |
| Income |  | 0.321 |
|  |  | (0.232) |
| Education (university degree) |  | 0.146 |
|  | (0.792) |
| Party (Democrat - Republican) |  | 621.8 |
|  | (3850.9) |
| Party (Independent - Republican) |  | 420.5 |
|  | (2557.8) |
| Constant | 1.06516e+22\*\*\* | 7.97569e+34\*\*\* |
|  | (4.10884e+22) | (8.58865e+35) |
| *N* | 151 | 150 |
| *R*2 | 0.003 | 0.259 |

Standard errors in parentheses

\* *p* < 0.05, \*\* *p* < 0.01, \*\*\* *p* < 0.001

# Appendix 13. Additional experiment on information equivalence

During the peer-review process, one of the anonymous reviewers raised the issue of information equivalence (Dafoe, Zhang, and Caughey 2018). One possible concern regarding our study was that even though we used a fictional President Smith, the language in the tweet, the type of scenario, and the form of communication could have prime some respondents to see Donald Trump in his stead. Another concern was that we did not control for the U.S. leader’s political identification and some conditions could have been theoretically more indicative of a Democratic president and others of a Republican president.

We designed and fielded an additional experiment to test whether 1) the respondents had different perceptions of the identity of the leader (Trump / non-Trump, Republican / Democrat) across experimental conditions; 2) if these possible differences in perceptions were associated with differences in the preference for nuclear use and could have, therefore, bias our results.

The setup of this additional experiment is described in Appendix 14. To address the aforementioned issue of information equivalence, we added the following single-choice item towards the end of the survey:

**When thinking about President Smith...**

* he reminded me of President Trump.
* he reminded me of President Biden.
* he did not remind me of anyone specific, but I thought about him as a Democrat.
* he did not remind me of anyone specific, but I thought about him as a Republican.
* he did not remind me of anyone specific and I did not think about his partisanship.

Contingency Table 1 shows the distribution of responses split by treatment. In line with the reviewer’s claim, *“he reminded me of President Trump”* was, overall, a considerably less common response in the control group (10.4%) than in the ambiguous group (37.9%) and explicit group (36%). To see whether this difference could have biased the results of our experiment, we collapsed the associations variable to “Trump” / “not-Trump.” As we shown in Table 2a-d, being reminded of Trump was not associated significantly with a higher preference for nuclear use overall and in any of the three treatments. As such, we conclude that while it is possible that the expressive language in our experimental group made our respondents more likely to be reminded of the former U.S. President Trump, this “Trump effect” was not associated with a statistically significant change in respondents’ preference for the nuclear response; therefore, it was unlikely to bias our findings on preference for nuclear use in a significant way.

**Table 1.** Cross-tabulations of associations across treatments

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| Treatment | President associations | | | | | |
| Trump | Biden | Republican | Democrat | Generic | Total |
| Control | 21 | 10 | 28 | 12 | 131 | 202 |
|  | 10.40 | 4.95 | 13.86 | 5.94 | 64.85 | 100.00 |
| Ambiguous threat | 75 | 6 | 36 | 7 | 74 | 198 |
|  | 37.88 | 3.03 | 18.18 | 3.54 | 37.37 | 100.00 |
| Explicit threat | 72 | 7 | 42 | 4 | 75 | 200 |
|  | 36.00 | 3.50 | 21.00 | 2.00 | 37.50 | 100.00 |
| Total | 168 | 23 | 106 | 23 | 280 | 600 |
|  | 28.00 | 3.83 | 17.67 | 3.83 | 46.67 | 100.00 |
| Pearson Chi2 = 63.54 Prob = 0.0000; First row has *frequencies* and second row has *row percentages* | | | | | | |

**Table 2a.** Cross-tabulations of response preference across associations - Control

|  |  |  |  |
| --- | --- | --- | --- |
| Treatment | President associations | | |
| Not Trump | Trump | Total |
| Prefer conventional | 158 | 17 | 175 |
|  | 87.29 | 80.95 | 86.63 |
| Prefer nuclear | 23 | 4 | 27 |
|  | 12.71 | 19.05 | 13.37 |
| Total | 181 | 21 | 202 |
|  | 100.00 | 100.00 | 100.00 |
| Pearson Chi2 = 0.65 Prob = 0.4189; First row has *frequencies* and second row has *column percentages* | | | |

**Table 2b.** Cross-tabulations of response preference across associations - Ambiguity

|  |  |  |  |
| --- | --- | --- | --- |
| Treatment | President associations | | |
| Not Trump | Trump | Total |
| Prefer conventional | 108 | 65 | 173 |
|  | 87.80 | 86.67 | 87.37 |
| Prefer nuclear | 15 | 10 | 25 |
|  | 12.20 | 13.33 | 12.63 |
| Total | 123 | 75 | 198 |
|  | 100.00 | 100.00 | 100.00 |
| Pearson Chi2 = 0.05 Prob = 0.8151; First row has *frequencies* and second row has *column percentages* | | | |

**Table 2c.** Cross-tabulations of response preference across associations - Explicit

|  |  |  |  |
| --- | --- | --- | --- |
| Treatment | President associations | | |
| Not Trump | Trump | Total |
| Prefer conventional | 114 | 65 | 179 |
|  | 89.06 | 90.28 | 89.50 |
| Prefer nuclear | 14 | 7 | 21 |
|  | 10.94 | 9.72 | 10.50 |
| Total | 128 | 72 | 200 |
|  | 100.00 | 100.00 | 100.00 |
| Pearson Chi2 = 0.07 Prob = 0.7878; First row has *frequencies* and second row has *column percentages* | | | |

**Table 2d.** Cross-tabulations of response preference across associations - Total

|  |  |  |  |
| --- | --- | --- | --- |
| Treatment | President associations | | |
| Not Trump | Trump | Total |
| Prefer conventional | 380 | 147 | 527 |
|  | 87.96 | 87.50 | 87.83 |
| Prefer nuclear | 52 | 21 | 73 |
|  | 12.04 | 12.50 | 12.17 |
| Total | 432 | 168 | 600 |
|  | 100.00 | 100.00 | 100.00 |
| Pearson Chi2 = 0.02 Prob = 0.8762; First row has *frequencies* and second row has *column percentages* | | | |

**Table 3a.** Cross-tabulations of response preference across associations - Control

|  |  |  |  |
| --- | --- | --- | --- |
| Preference | President associations | | |
| Democrat | Republican | Total |
| Conventional | 18 | 37 | 55 |
|  | 81.82 | 75.51 | 77.46 |
| Nuclear | 4 | 12 | 16 |
|  | 18.18 | 24.49 | 22.54 |
| Total | 22 | 49 | 71 |
|  | 100.00 | 100.00 | 100.00 |
| Pearson Chi2 = 0.35 Prob = 0.5563; First row has *frequencies* and second row has *column percentages* | | | |

**Table 3b.** Cross-tabulations of response preference across associations - Ambiguity

|  |  |  |  |
| --- | --- | --- | --- |
| Preference | President associations | | |
| Democrat | Republican | Total |
| Conventional | 12 | 96 | 108 |
|  | 92.31 | 86.49 | 87.10 |
| Nuclear | 1 | 15 | 16 |
|  | 7.69 | 13.51 | 12.90 |
| Total | 13 | 111 | 124 |
|  | 100.00 | 100.00 | 100.00 |
| Pearson Chi2 = 0.35 Prob = 0.5536; First row has *frequencies* and second row has *column percentages* | | | |

**Table 3c.** Cross-tabulations of response preference across associations - Explicit

|  |  |  |  |
| --- | --- | --- | --- |
| Preference | President associations | | |
| Democrat | Republican | Total |
| Conventional | 10 | 104 | 114 |
|  | 90.91 | 91.23 | 91.20 |
| Nuclear | 1 | 10 | 11 |
|  | 9.09 | 8.77 | 8.80 |
| Total | 11 | 114 | 125 |
|  | 100.00 | 100.00 | 100.00 |
| Pearson Chi2 = 0.00 Prob = 0.9716; First row has *frequencies* and second row has *column percentages* | | | |

**Table 3d.** Cross-tabulations of response preference across associations - Total

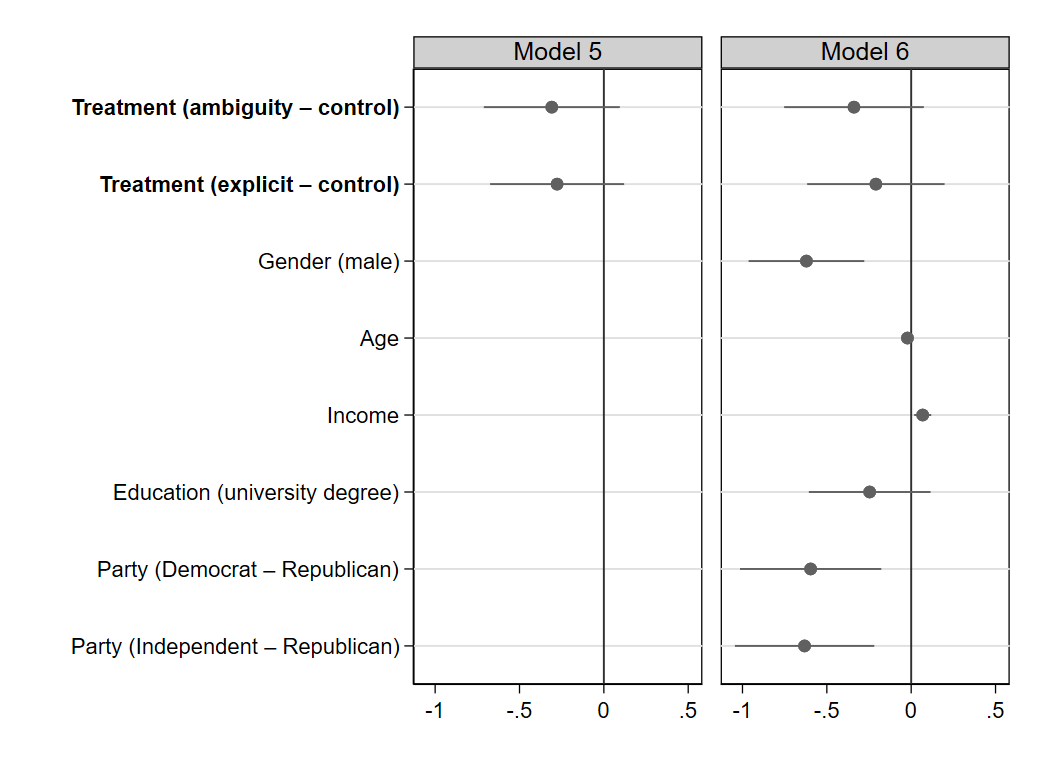
|  |  |  |  |
| --- | --- | --- | --- |
| Preference | President associations | | |
| Democrat | Republican | Total |
| Conventional | 40 | 237 | 277 |
|  | 86.96 | 86.50 | 86.56 |
| Nuclear | 6 | 37 | 43 |
|  | 13.04 | 13.50 | 13.44 |
| Total | 46 | 274 | 320 |
|  | 100.00 | 100.00 | 100.00 |
| Pearson Chi2 = 0.01 Prob = 0.9325; First row has *frequencies* and second row has *column percentages* | | | |

Table 1 also shows that the respondents were more likely to be reminded of a Republican than a Democratic President in the two experimental conditions than in the control condition. To see whether this difference could have biased the results of our experiment, we filtered out the non-partisan responses and collapsed the remaining associations to “Republican” and “Democrat.” As we show in Table 3a-d, being reminded of a Republican President was not associated significantly with a higher preference for nuclear use overall and in any of the three treatments. As such, we conclude that while it is possible that the expressive language in our experimental group made our respondents more likely to be reminded of a Republican, this “partisanship effect” was not associated with a statistically significant change in respondents’ preference for the nuclear response; therefore, it was unlikely to bias our findings on preference for nuclear use in a significant way.

# Appendix 14. Additional experiment on preference for nuclear use

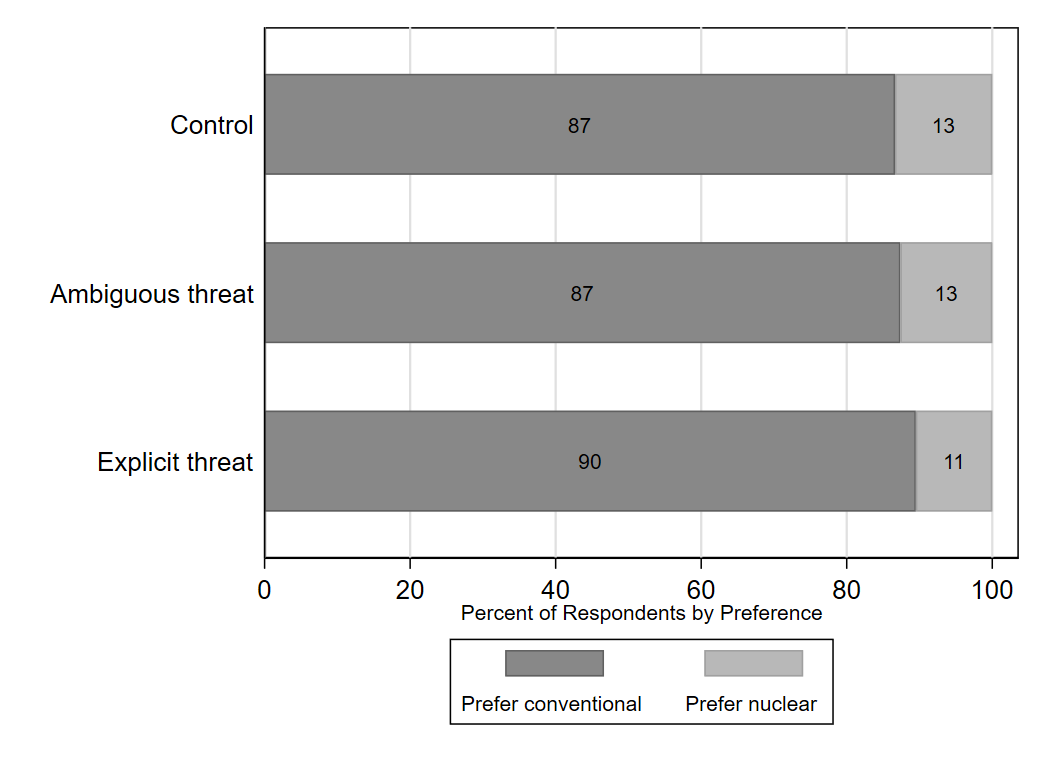
Here, we report the results of an additional experiment where we measured preference for nuclear use before the respondents were informed about the choice made by the U.S. President. This solved the issue of the wording of the non-nuclear strike option raised by one of the reviewers of this paper and makes the connection between threat formulation and preference for nuclear use stronger (the participants are less influenced by the actual choice made by the President). The overall setup of the experiment before we asked about the respondents’ preferences has not changed since the original experiment reported in the manuscript (see Appendix 2). Our results in this additional experiment show, once again, that the nature of the threat was not associated with the change in preference for the use of nuclear weapons.

**Figure 1.** Ordinal logistic regression estimates



*Note:* Ordinal logistic regression estimates.*N* = 600 (Model 5), *N =* 598 (Model 6). 95% CI. Variables whose intervals overlap with the vertical line are statistically indistinguishable from 0. Positive coefficients correspond to a higher preference for nuclear use. Model 3 shows the effects without the inclusion of control variables and Model 4 with control variables included.

**Figure 2.**Preference across treatments



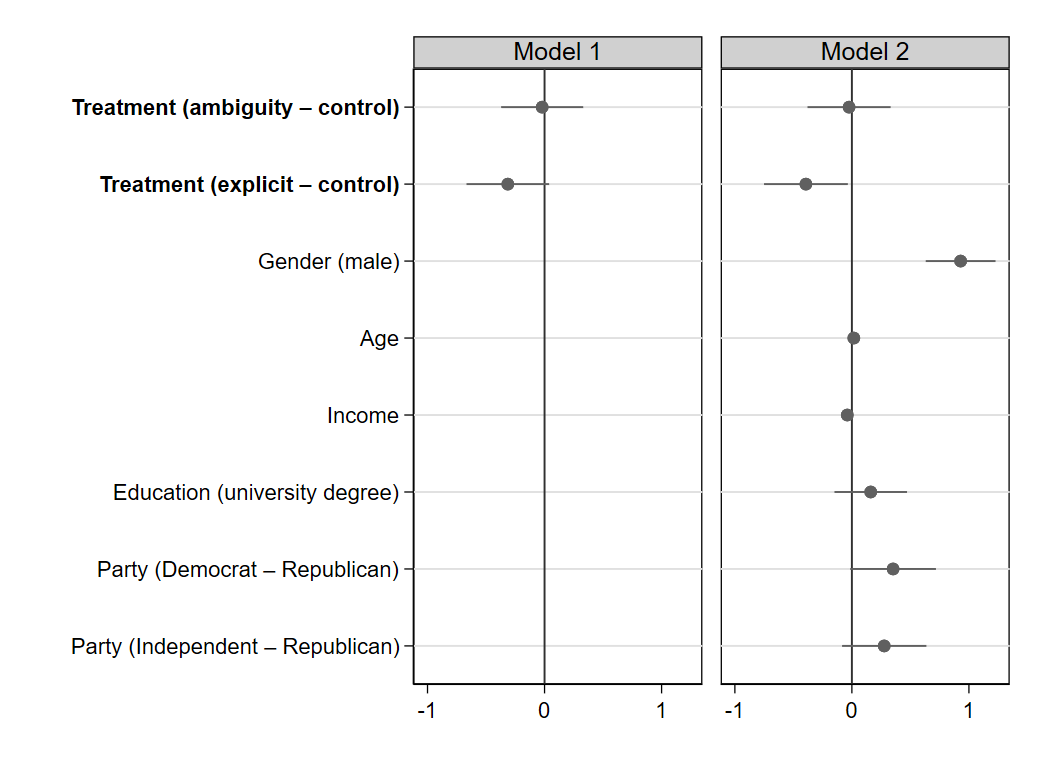
*Note:* An alternative way of calculating preference for nuclear use. The figure displays the percentage of respondents who preferred a conventional response and those who preferred a nuclear response in three different treatments *α*.

# Appendix 15. Alternative dependent variable (impact on credibility)

We took advantage of the additional experiment reported in Appendix 14 to add an alternative dependent variable. After the participants indicated their preference for the type of military response, we asked them how would it impact the U.S. credibility if the President did not order the use of nuclear weapons against military targets in North Korea as part of the U.S. response. The responses were made on a five-point scale from “strongly bolster the U.S. credibility” to “strongly weaken the U.S. credibility.”

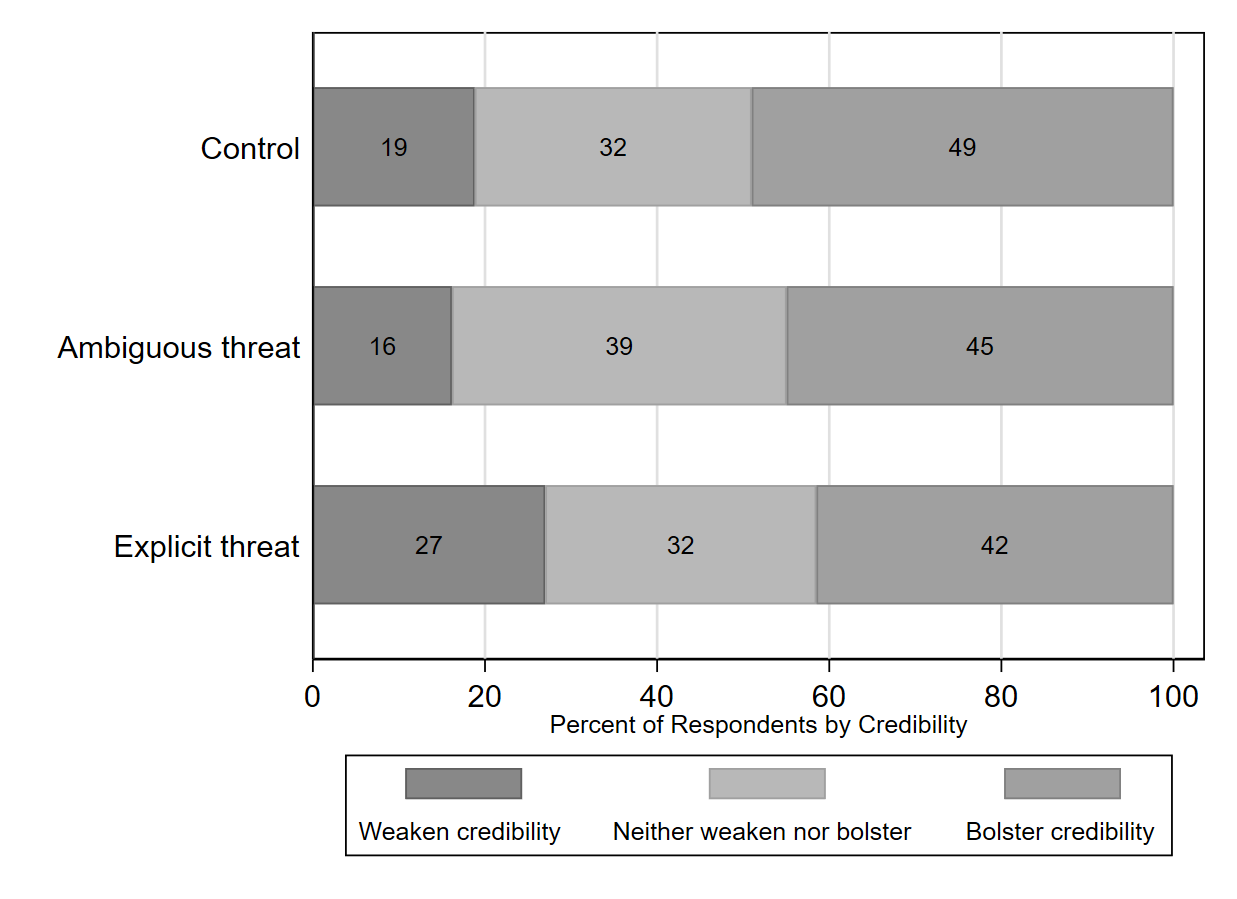
As we show in Figure 1, respondents in the “ambiguous nuclear threat” condition were no more likely to believe that using only conventional strikes against North Korea would weaken the U.S. credibility than those in the control condition. However, those in the “explicit nuclear threat” condition were more likely (*p* < 0.05 in Model 2 after the inclusion of control variables) to believe that “backing down” from nuclear use would weaken U.S. credibility. Figure 2 shows the size of these effects in percentage points. Overall, these findings are in line with the results of our original experiment with respect to *H1* and *H2* .

**Figure 1.** Ordinal logistic regression estimates.



*Note:* Ordinal logistic regression estimates.*N* = 600 (Model 1), *N =* 598 (Model 2). 95% CI. Variables whose intervals overlap with the vertical line are statistically indistinguishable from 0. Positive coefficients correspond to higher credibility. Model 1 shows the effects without the inclusion of control variables and Model 2 with control variables included.

**Figure 2.** Credibility across treatments



*Note:* An alternative way of calculating the perceived impact on credibility. The figure displays the percentage of respondents who believed that the conventional-only response would weaken, bolster, or neither weaken nor bolster the U.S. credibility in three different treatments *α*.

1. See Peer *et al.* (2017) and Palan and Schitter (2018) for a discussion of Prolific as a survey tool. [↑](#footnote-ref-1)
2. Note that Israel has signed but has not ratified the CWC. [↑](#footnote-ref-2)
3. Earlier studies have demonstrated that removing participants who failed post-treatment attention checks or manipulation checks can bias the results (see, e.g., Aronow, Baron, and Pinson 2019 for a thorough discussion) [↑](#footnote-ref-3)