**Online Appendix**

**Supplemental Information**

**SI-1: Survey Questionnaire**

*Background Information*

As you may know, when there is a dispute about whether a country is abiding by international law, the case is often tried in front of the International Court of Justice (ICJ), the principal judicial organ of the United Nations. It is comprised of fifteen justices from around the world, including one President and one Vice President.

Here is a picture of the current panel of judges.



The Court decides cases by majority rule. When there is a tie vote, the President of the court is entitled to cast the decisive vote to break the tie.

Since 1945, when the ICJ was established, 178 cases have been entered onto the General List for consideration before the court. The United States has appeared, as either applicant or respondent, in 24 of these cases.

*Comprehension/Manipulation Checks*

1. According to what you just read on the previous page, how many judges does ICJ have?

1. 9
2. 12
3. 15

2. When the ICJ votes are tied, what happens next?

1. The case is sent back to the Court for additional deliberation.
2. The President of the Court will cast the deciding vote to break the tie.
3. Don't know.

*[If the answer to either of these questions are incorrect]*

Oops, your answers were incorrect. Please review this information one more time.

[Repeat background information]

*Control*

Suppose that the United States was sued by another country at the ICJ for breaking the international law. The Court vote is tied at 7-7. The President of the Court cast the deciding vote **against** the United States.



*Treatment*

Suppose that the United States was sued by another country at the ICJ for breaking the international law. The Court vote is tied at 7-7. Judge Xue Hanqin from China (pictured below), President of the Court, cast the deciding vote **against** the United States.



*Placebo [Randomly display one of the following five judges]*

Suppose that the United States was sued by another country at the ICJ for breaking the international law. The Court vote is tied at 7-7. Judge Kirill Gevorgian from Russia (pictured below), the President of the Court, cast the deciding vote **against** the United States.



Suppose that the United States was sued by another country at the ICJ for breaking the international law. The Court vote is tied at 7-7. Judge Rosalyn Higgins from the United Kingdom (pictured below), the President of the Court, cast the deciding vote **against** the United States.



Suppose that the United States was sued by another country at the ICJ for breaking the international law. The Court vote is tied at 7-7. Judge Yuji Iwasawa from Japan (pictured below), the President of the Court, cast the deciding vote **against** the United States.



Suppose that the United States was sued by another country at the ICJ for breaking the international law. The Court vote is tied at 7-7. Judge Julia Sebutinde from Uganda (pictured below), the President of the Court, cast the deciding vote **against** the United States.



Suppose that the United States was sued by another country at the ICJ for breaking the international law. The Court vote is tied at 7-7. Judge Patrick Lipton Robinson from Jamaica (pictured below), the President of the Court, cast the deciding vote **against** the United States.



*Legitimacy*

Do you agree or disagree with the following statements?

Agree / Somewhat Agree / Somewhat Disagree / Disagree

(Randomized)

1. I think the ICJ serves an important role in the world.
2. The ICJ should continue to make rulings in the future.
3. I believe the ICJ is necessary.
4. I sympathize with the goals of the ICJ.

*Casual Mechanism*

Please tell us whether you agree or disagree with the following statement.

Agree / Somewhat Agree / Somewhat Disagree / Disagree

(Randomized)

1. Th ICJ should not be allowed to tell the United States what to do.
2. I find the ICJ ruling to be fair.
3. The ICJ is undemocratic.

*Sociodemographic Questions*

*Age*: What year were you born? (Please enter 4 digit year, e.g. 1980)

*Gender*: Are you ...

1. Male
2. Female
3. Prefer not to say

*Education*: What is the highest level of schooling you have completed?

1. Nursery school to 8th grade
2. 9th, 10th or 11th grade
3. 12th grade, no diploma
4. High school graduate - high school diploma or the equivalent (for example: GED)
5. College, CEGEP or other non-university certificate or diploma
6. Some college credit, but less than 1 year
7. Masters
8. Associate degree (for example: AA, AS)
9. Bachelor's degree (for example: BA, AB, BS)
10. Master's degree (for example: MA, MS, MEng, MEd, MSW, MBA)
11. Professional degree (for example: MD, DDS, DVM, LLB, JD)
12. Doctorate degree (for example: PhD, EdD)
13. Prefer not to say

*Income*: What was the range of your household income last year?

1. Under $10,000
2. Between $10,000 to $20,000
3. Between $20,000 to $30,000
4. Between $30,000 to $40,000
5. Between $40,000 to $50,000
6. Between $50,000 to $60,000
7. Between $60,000 to $70,000
8. Between $70,000 to $80,000
9. Between $80,000 to $90,000
10. Between $90,000 to $100,000
11. Between $100,000 to $200,00
12. Above $200,000
13. Prefer not to say

**SI-2: IRT Measures of Legitimacy**

In the main analysis of the paper, we use Item Response Theory (IRT) models to calculate individual respondents’ perceived legitimacy. IR models are commonly used to examine the relationship between unobserved latent characteristics such as mathematical aptitude and the probability of correctly answering test questions (items). Applied to this study, we can think of each respondent in our sample having a latent trait in how they perceived legitimacy of the ICJ. From a broader statistical perspective, IRT models are akin to extensions of confirmatory factor analysis models on binary and categorical data and as special cases of generalized linear mixed-effects models. The benefit of IRT is that it allows items to contribute differently to the latent trait unlike other aggregation techniques.

We use a graded response model (GRM) given that responses to the four questions in the survey used to tap perception of ICJ legitimacy are on a four-point ordinal scale (agree, somewhat agree, somewhat disagree, or disagree). The GRM estimates both the difficult and discrimination parameters for each of the four items. The number of difficulty parameters is equal to the number of answer categories minus one (three in this case). The results are presented in Table A1. Here, we can see that four questions have similar difficulty parameters, but the question on whether respondents sympathize with the goals of the ICJ is the least discriminating.

Table A1: Parameter Estimates of the GRM

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Question | Discrimination Parameter | DifficultyParameter 1 | DifficultyParameter 2 | DifficultyParameter 3 |
| Q1: I think the ICJ serves an important role in the world. | 4.866 | -1.974 | -1.31623 | -0.272 |
| Q2: The ICJ should continue to make rulings in the future. | 4.762 | -1.931 | -1.253 | -0.111 |
| Q3: I believe the ICJ is necessary. | 5.596 | -1.977 | -1.236 | -0.215 |
| Q4: I sympathize with the goals of the ICJ. | 2.776 | -2.101 | -1.315 | 0.040 |

The GRM reduces the four legitimacy questions to one latent trait, but we can confirm principle component analysis (PCA). The results, presented in table A2, shows that this is indeed the case—79 percent of the variance is explained by the first factor while the second factor only accounts for 9 percent of the variance. The loadings for the second factor are also consistent with the finding above that the fourth question is the least discriminating.

Table A2: Factor Loadings from the PCA

|  |  |  |
| --- | --- | --- |
| Question | Factor 1 | Factor 2 |
| Q1: I think the ICJ serves an important role in the world. | 0.5070 | -0.3297 |
| Q2: The ICJ should continue to make rulings in the future. | 0.5074 | -0.1139 |
| Q3: I believe the ICJ is necessary. | 0.5122 | -0.3592 |
| Q4: I sympathize with the goals of the ICJ. | 0.4724 | 0.8656 |
| SS Loadings | 3.1663 | 0.3826 |
| Proportion of Variance | 0.7916 | 0.0957 |

**SI-3: Balance Check**

The following tables report balance checks in the main and follow-up surveys on pre-treatment measures on age, gender, education, and income. Pairwise comparisons suggest that in the main survey there are more females in the China and Russia groups relative to the control group, which could have occurred by chance. To correct the imbalance and improve the precision of the estimated effects, therefore, we use covariate adjustment in all of the analyses by including the pre-treatment measures as controls.

Balance is achieved in the follow-up survey as shown in Table A4, though we still employ covariate adjustment using the same pre-treatment measures.

Table A3: Balance Check in the Main Survey

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Group | Age | Female | Education | Income |
| Control | 32.14 | 0.49 | 7.80 | 6.32 |
| China | 31.00 | 0.41 | 7.82 | 6.38 |
| Russia | 32.60 | 0.40 | 7.74 | 6.53 |
| UK | 33.97 | 0.46 | 7.95 | 6.74 |
| Japan | 33.41 | 0.41 | 8.02 | 5.97 |
| Uganda | 34.08 | 0.45 | 7.89 | 6.29 |
| Jamaica | 31.51 | 0.45 | 7.79 | 6.93 |
| F-Stat (ANOVA) | 1.97 | 1.09 | 0.33 | 1.20 |
| P value | 0.067 | 0.367 | 0.921 | 0.303 |

Table A4: Balance Check in the Follow-up Survey

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Group | Age | Female | Education | Income |
| Control | 45.98 | 0.52 | 7.52 | 6.62 |
| China | 45.31 | 0.50 | 7.30 | 6.54 |
| P value (T-test) | 0.612 | 0.665 | 0.272 | 0.802 |

**SI-4: Causal Mediation and Sensitivity Analyses**

We used the STATA package medeff to perform the causal mediation analysis. For each mediator, we include the same pre-treatment covariates (age, gender, education, and income) and specify 1000 simulations to run for the quasi-Bayesian approximation of parameter uncertainty. The results are summarized in Table A5.

Table A5: Causal Mediation Analysis

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Mediator | ACME | Direct Effect | Total Effect | Percent Mediated |
| Mandate | -0.018 | -0.055 | -0.072 | 24.3% |
|  | (-0.035, 0.0001) | (-0.080, -0.029) | (-0.105, -0.040) | (16.7%, 43.5%) |
| Fairness | -0.025 | -0.047 | -0.072 | 34.0% |
|  | (-0.047, 0.003) | (-0.070, -0.024) | (-0.101, -0.042) | (24.3%, 59.1%) |
| Democratic | -0.018 | -0.054 | -0.072 | 24.8% |
|  | (-0.033, 0.003) | (-0.081, -0.027) | (-0.104, -0.040) | (17.3%, 44.8%) |

*Notes:* The control and placebo groups are combined to simplify the interpretation. Numbers in the brackets are 95% confidence intervals.

An important identification assumption for the causal mediation analysis is sequential ignorability, which may or may not hold for the mediators as they were not randomly assigned in the experiment. To test the sequential ignorability assumption, we performed sensitivity analysis proposed by Imai et al (2011), which evaluates the robustness of the findings from causal mediation analysis to the possible existence of unmeasured pretreatment variables that confound the relationship between the mediator and the outcome variable.

Figure A1: Sensitivity Analysis



*Notes:* The control and placebo groups are combined to simplify the interpretation.

Figure A1 presents the result from the sensitivity analysis. In each plot, the monotonic curve represents the different values of the average causal mediated effects at different values of the sensitivity parameter *ρ*. The solid line represents the value of the average causal mediated effects when *ρ* is zero. The further the solid line is from the x-axis and the further the monotonic curve is from the y-axis when it crosses the x-axis, the less likely confounder bias is a problem in the model. This is the case for all three mediators. In fact, the analysis indicates that the original conclusion about the direction of the ACME would be maintained unless *ρ* is greater than 0.5. This implies that the conclusion is plausible given even large departures from the ignorability of the mediator.

**SI-5: Alternative Measures of Legitimacy**

In the main text, we used IRT measures of legitimacy for all the analyses. Alternatively, we can measure legitimacy by simply averaging the responses or by principal component analysis. The Cronbach’s alpha among the three measures is 0.9, indicating high level of internal consistency. This can also be seen in the scatterplot matrix in Figure A2.

Figure A2 Scatterplot Matrix of Different Measures of Legitimacy



Using these two alternative measures of legitimacy yield results similar to those reported in the article (Figure A3).

Figure A3 Treatment Effect on Alternative Measures of ICJ Legitimacy



*Note*: Results based on OLS regression model with robust standard errors adjusted with pre-treatment covariates on age, gender, education, and income. The dependent variable is the additive measure of perceived ICJ legitimacy. Vertical lines indicate 95 percent confidence intervals. Point estimates represent ATEs relative to the control condition. The number of observations in the underlying model is 1,390 due to missing values in the covariates.



*Note*: Results based on OLS regression model with robust standard errors adjusted with pre-treatment covariates on age, gender, education, and income. The dependent variable is the PCA measure of perceived ICJ legitimacy. Vertical lines indicate 95 percent confidence intervals. Point estimates represent ATEs relative to the control condition. The number of observations in the underlying model is 1,390 due to missing values in the covariates.

**SI-6: Reduced Sample Size**

Our main analyses used all respondents from the survey. Here we exclude respondents who failed the comprehension questions. As can be seen in Table A6, 19.7% of the respondents failed the first comprehension question on the number of ICJ judges and 6.2% failed the second comprehension question on the rule for tie-breaking.

Table A6 Result of the Comprehension Check

|  |  |
| --- | --- |
|  | *Comprehension Check 1* |
| *Comprehension Check 2* | Correct | Incorrect |
|  |  |  |
| Correct | 1,036(77.2%) | 222(15.3%) |
| Incorrect | 51(8.9%) | 32(4.4%) |

Figure A4 plots the ATE using the 1,036 respondents that answered both comprehension questions correctly. The results are very similar to Figure 3 in the main text.

Figure A4 Treatment Effect on ICJ Legitimacy (Reduced Sample)



*Note*: Results based on OLS regression model with robust standard errors adjusted with pre-treatment covariates on age, gender, education, and income. The dependent variable is the IRT measure of perceived ICJ legitimacy. Vertical lines indicate 95 percent confidence intervals. Point estimates represent ATEs relative to the control condition. The number of observations in the underlying model is 996 due to missing values in the covariates.

**SI-7: ATE with Post-stratification Weighting**

Table A7 compares a number of key demographic measures between the Prolific sample with the 2019 census data in the United States.

Table A7: Survey Sample Comparison with Census

|  |  |  |
| --- | --- | --- |
|  | Survey | Census |
| Age 18-24 | 30.9% | 12.1% |
| Age 25-34 | 36% | 18% |
| Age 35-44 | 17.6% | 16.3% |
| Age 45-54 | 7.9% | 16.8% |
| Age 55-74 | 7.4% | 28.4% |
| Female | 44.3% | 51.3% |
| Bachelor’s degree | 34.3% | 29.1% |
| White | 67.9% | 72.5% |

Figure A5 reports the average treatment effects using post-stratification weighting. The weights are calculated using an iterative proportional fitting algorithm (i.e. raking) to achieve known population margins on age, gender, education, and race. The results are similar to those reported in Figure 3 in the main text—the Chinese judge reduces respondents perceived ICJ legitimacy.

Figure A5: Average Treatment Effects Using Weighted Sample



*Note*: Results based on OLS regression model with robust standard errors adjusted with pre-treatment covariates on age, gender, education, and income. The dependent variable is the additive measure of perceived ICJ legitimacy. Vertical lines indicate 95 percent confidence intervals. Point estimates represent ATEs relative to the control condition. The number of observations in the underlying model is 1,390 due to missing values in the covariates.

**SI-8: ATE with Tobit Model for Censored DV**

Our IRT measure of legitimacy is bounded by zero and one and left-skewed, raising concerns for the use of OLS regression as the skewness can lead to biased coefficient estimates, non-normal residuals, and heteroscedasticity. In all of our analysis, we use robust standard errors, which is often employed to mitigate the influence of outliers and reduce the impact of skewness on coefficient estimates.

Additionally, as part of the robustness check, we estimated a limited dependent variable model (Tobit) for the average treatment effect in order to account for the censoring and truncation in the dependent variable. The results, shown below in Figure A6, are very similar to the main findings reported in Figure 2, and thus our conclusion remains unchanged.

Figure A6: Average Treatment Effects Using Tobit Model

****

*Note*: Results based on Tobit regression model adjusted with pre-treatment covariates on age, gender, education, and income. The dependent variable is the IRT measure of perceived ICJ legitimacy. Vertical lines indicate 95 percent confidence intervals. Point estimates represent ATEs relative to the control condition. The number of observations in the underlying model is 1,390 due to missing values in the covariates.

**SI-9: UK Experiment Wording**

In the UK replication study, we maintained the exact wording with the original U.S. questionnaire, while replacing “United States” with “United Kingdom” (and vice versa), e.g.:

Suppose that the *United Kingdom* was sued by another country at the ICJ for breaking the international law. The Court vote is tied at 7-7. Judge Rosalyn Higgins from the *United States* (pictured below), the President of the Court, cast the deciding vote **against** the *United Kingdom*.



Please tell us whether you agree or disagree with the following statement.

Agree / Somewhat Agree / Somewhat Disagree / Disagree

1. Th ICJ should not be allowed to tell the *United Kingdom* what to do.
2. I find the ICJ ruling to be fair.
3. The ICJ is undemocratic.

**SI-10: Summary Statistics**

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| Variable | Obs | Mean | Std. Dev. | Min | Max |
| Legitimacy Q1 | 1,451 | 3.45 | 0.77 | 1.00 | 4.00 |
| Legitimacy Q2 | 1,451 | 3.37 | 0.78 | 1.00 | 4.00 |
| Legitimacy Q3 | 1,451 | 3.41 | 0.78 | 1.00 | 4.00 |
| Legitimacy Q4 | 1,451 | 3.29 | 0.80 | 1.00 | 4.00 |
| Legitimacy (IRT) | 1,451 | 0.72 | 0.26 | 0.00 | 1.00 |
| Legitimacy (PCA) | 1,451 | 0.00 | 1.78 | -6.08 | 1.57 |
| Legitimacy (Additive) | 1,451 | 3.38 | 0.70 | 1.00 | 4.00 |
| Mandate | 1,451 | 3.10 | 0.92 | 1.00 | 4.00 |
| Fairness | 1,451 | 1.81 | 0.76 | 1.00 | 4.00 |
| Democratic | 1,451 | 3.17 | 0.85 | 1.00 | 4.00 |
| Validate | 1,451 | 0.71 | 0.45 | 0.00 | 1.00 |
| Age | 1,451 | 32.35 | 12.15 | 18.00 | 78.00 |
| Female | 1,451 | 0.44 | 0.50 | 0.00 | 1.00 |
| White | 1,451 | 0.68 | 0.47 | 0.00 | 1.00 |
| Education | 1,443 | 7.84 | 2.12 | 1.00 | 12.00 |
| Income | 1,392 | 6.43 | 3.41 | 1.00 | 12.00 |
| Republican | 1,451 | 0.15 | 0.35 | 0.00 | 1.00 |
| Democrat | 1,451 | 0.48 | 0.50 | 0.00 | 1.00 |