

Appendix for “Navigating Hostility: The Effect of Nonprofit Transparency and Accountability on Donor Preferences in the Face of Shrinking Civic Space”

Suparna Chaudhry¹, Marc Dotson², and Andrew Heiss³

¹Lewis and Clark College

²Utah State University

³Georgia State University

Author note

Suparna Chaudhry,  <https://orcid.org/0000-0003-3506-7590>

Marc Dotson,  <https://orcid.org/0000-0001-7846-3660>

Andrew Heiss,  <https://orcid.org/0000-0002-3948-3914>

**Appendix for “Navigating Hostility: The Effect of Nonprofit Transparency and
Accountability on Donor Preferences in the Face of Shrinking Civic Space”**

Table of contents

Respondent demographics	3
Table A1: Respondent demographics	3
Table A2: Respondent attitudes toward charity	4
Table A3: Respondent politics, ideology, and religion	5
Table A4: Sample characteristics compared to nationally representative Current Popu- lation Survey (CPS) estimates	6
Model details	7
Model coefficients and estimated marginal means	7
Raw model coefficients	8
Converting coefficients to estimated marginal means	8
Table A5: Original multinomial logistic regression model coefficients	9
Table A6: Excerpt from complete reference grid of all 576 possible combinations of attribute features and levels	10
Table A7: Complete marginal means and AMCEs (see Figure 1 in main paper)	12
Table A8: Marginal means for all combinations of transparency and accountability (see Figure 2 in main paper)	13
Table A9: Marginal means and AMCEs for interaction between transparency, ac- countability, and government relationships (see Figure 3 in paper)	14
Preregistration deviations	15
Condensed preregistration	18
Study information	18

Sampling Plan	21
Design plan	22
Analysis Plan	23
References	25

Respondent demographics

Table A 1: Respondent demographics

Demographics		
Response	N*	%
Gender		
Male	517	50.9%
Female	485	47.7%
Transgender	8	0.8%
Prefer not to say	3	0.3%
Other	3	0.3%
Age		
Less than 2017 national median (36)	500	49.2%
More than median	516	50.8%
Marital status		
Married	403	39.7%
Widowed	21	2.1%
Divorced	104	10.2%
Separated	35	3.4%
Never married	453	44.6%
Education		
Less than high school	25	2.5%
High school graduate	270	26.6%
Some college	287	28.2%
2 year degree	138	13.6%
4 year degree	206	20.3%
Graduate or professional degree	82	8.1%
Doctorate	8	0.8%
Income		
Less than 2017 national median (\$61,372)	585	57.6%
More than median	431	42.4%

* N = 1,016 respondents.

Table A 2: Respondent attitudes toward charity

Attitudes toward charity		
Response	N [*]	%
Frequency of donating to charity		
More than once a month, less than once a year	566	55.7%
At least once a month	450	44.3%
Amount of donations to charity last year		
\$1-\$49	337	33.2%
\$50-\$99	245	24.1%
\$100-\$499	233	22.9%
\$500-\$999	107	10.5%
\$1000-\$4,999	65	6.4%
\$5000-\$9,999	18	1.8%
\$10,000+	11	1.1%
Importance of trusting charities		
Not at all important	7	0.7%
Very unimportant	9	0.9%
Somewhat unimportant	21	2.1%
Neutral	98	9.6%
Somewhat important	168	16.5%
Very important	157	15.5%
Essential	556	54.7%
Level of trust in charities		
No trust at all	14	1.4%
Very little trust	20	2.0%
Little trust	68	6.7%
Neutral	257	25.3%
Some trust	328	32.3%
A lot of trust	169	16.6%
Complete trust	160	15.7%
Frequency of volunteering		
Haven't volunteered in past 12 months	423	41.6%
Rarely	20	2.0%
More than once a month, less than once a year	322	31.7%
At least once a month	251	24.7%

* N = 1,016 respondents.

Table A 3: Respondent politics, ideology, and religion

Politics, ideology, and religion		
Response	N*	%
Frequency of following national news		
Rarely	88	8.7%
Once a week	216	21.3%
At least once a day	712	70.1%
Traveled to a developing country		
No	766	75.4%
Yes	250	24.6%
Voted in last election		
No	274	27.0%
Yes	742	73.0%
Trust in political institutions and the state		
No trust at all	123	12.1%
Very little trust	155	15.3%
Little trust	207	20.4%
Neutral	276	27.2%
Some trust	151	14.9%
A lot of trust	49	4.8%
Complete trust	55	5.4%
Political ideology		
Extremely liberal	87	8.6%
Somewhat liberal	87	8.6%
Slightly liberal	112	11.0%
Moderate	363	35.7%
Slightly conservative	175	17.2%
Somewhat conservative	80	7.9%
Extremely conservative	112	11.0%
Involvement in activist causes		
Not involved	569	56.0%
Involved	447	44.0%
Frequency of attending religious services		
Not sure	11	1.1%
Rarely	600	59.1%
At least once a month	405	39.9%
Importance of religion		
Not important	338	33.3%
Important	678	66.7%

* N = 1,016 respondents.

Table A 4: Sample characteristics compared to nationally representative Current Population Survey (CPS) estimates

Variable	Sample	National	Δ^*
Age (% 36+) ^a	50.75%	53.03%	-2.3 [-5.3, 0.9]
Female (%) ^a	47.67%	50.97% [†]	-3.3 [-6.4, -0.2]
Married (%) ^a	39.66%	41.01%	-1.3 [-4.4, 1.6]
Education (% BA+) ^a	29.18%	31.69%	-2.5 [-5.2, 0.4]
Income (% \$61,372+) ^a	42.39%	36.40% [†]	6.0 [3.1, 9.1]
Donated in past year (%) ^b	55.65%	47.40% [†]	8.3 [5.1, 11.3]
Volunteered in past year (%) ^b	58.37%	30.02% [†]	28.3 [25.3, 31.3]
Voted in last November election (%) ^c	73.01%	53.44% [†]	19.6 [16.8, 22.2]

* Values are on the percentage-point scale; single value is posterior median; 95% credible interval in brackets.

† National value is substantially different from the sample; the 95% posterior credible interval for the difference between the sample and national proportions contains 0.

^a Annual Social and Economic Supplement (ASEC) of the Current Population Survey (CPS), March 2019

^b Monthly CPS, September 2019

^c Monthly CPS, November 2018

Model details

We use Stan (Stan Development Team, 2024b, p. v2.32.2; 2024a, p. v2.36) through R (R Core Team, 2024, p. v4.4.2) and {brms} (Bürkner, 2017, p. v2.22) to estimate the model. We simulate 4 MCMC chains with 5,000 draws in each chain, 1,000 of which are used for warmup, resulting in 16,000 ($4,000 \times 4$) draws per model parameter. We assess convergence with visual inspection, and all chains converge.

Complete results from the model, along with posterior predictive checks, goodness-of-fit measures, and model diagnostics—as well as our code and data—are available at a companion statistical analysis compendium at ANONYMIZED_URL.

Model coefficients and estimated marginal means

When working with the results for our multinomial regression model, we rely on estimated marginal means (EMMs) rather than raw regression coefficients because of the complexity of the model. At its core, a “marginal mean” refers to the literal mean in the margins in a contingency table of model predictions, and differences in marginal means are equivalent to marginal effects or regression coefficients.

To find the causal effects defined in each of our estimands, we calculate EMMs by finding the fitted probability-scale values for each cell in a balanced reference grid of all 576 possible combinations of feature levels (2 transparency \times 2 accountability \times 3 government relationships \times 4 organizations \times 4 issues \times 3 funding = 576 rows). We then calculate group averages and contrasts in group averages for each of the features of interest, marginalizing over all other features.

Raw model coefficients

As seen in Table 5, the model returns three sets of coefficients per conjoint level. Each coefficient shows the shift in probability that someone will choose an organization from that appears as the first, second, and third option in the experimental task, represented by μ_1 , μ_2 , and μ_3 . Under experimental conditions where all the feature levels are randomly assigned, it is safe to assume that the cell proportions are equal and then marginalize (i.e. find the average) across the rows or columns. This allows us to take the average of each set of coefficients (e.g. μ_1 , μ_2 , and μ_3 for “Transparency = Yes”) to create a single value per coefficient.

Converting coefficients to estimated marginal means

To convert EMMs and AMCEs to a more interpretable probability scale (rather than the original log odds scale), we generate predicted values (marginalized across the three μ terms) for each of the 576 unique combinations of feature levels. Table 6 provides an excerpt from this grid, showing six rows where accountability, organization, issue area, and funding are identical and held constant, while transparency and government relations vary.

To calculate the marginal mean for a feature, we find the average predicted value across each the levels of that feature. To illustrate, assume that Table 6 represents the full reference grid of all experimental features and levels. The marginal means for transparency would be $(0.486 + 0.396 + 0.347)/3 = 0.410$ when transparency is set to “no”, and $(0.626 + 0.537 + 0.485)/3 = 0.550$ when transparency is set to “yes”. In reality, the marginal mean for transparency reported in the paper reflects the average of 288 rows where transparency is no and 288 rows where transparency is yes.

Table A 5: Original multinomial logistic regression model coefficients

Feature		Posterior medians		
		μ_1	μ_2	μ_3
Transparency	Transparency × Yes	0.52	0.61	0.59
		[0.45, 0.60]	[0.54, 0.69]	[0.51, 0.67]
Accountability	Accountability × Yes	0.53	0.60	0.56
		[0.45, 0.60]	[0.53, 0.68]	[0.48, 0.64]
Relationship with host government	Criticized	-0.35	-0.42	-0.32
		[-0.44, -0.26]	[-0.51, -0.34]	[-0.41, -0.23]
	Under crackdown	-0.54	-0.62	-0.57
		[-0.63, -0.45]	[-0.71, -0.53]	[-0.66, -0.47]
Organizations	Greenpeace	0.162	0.0973	-0.024
		[0.054, 0.271]	[-0.0083, 0.2028]	[-0.133, 0.086]
	Oxfam	-0.061	-0.136	-0.166
		[-0.174, 0.052]	[-0.247, -0.023]	[-0.280, -0.053]
	Red Cross	0.79	0.60	0.51
		[0.68, 0.89]	[0.50, 0.71]	[0.41, 0.62]
Issue areas	Environment	-0.29	-0.182	-0.22
		[-0.39, -0.18]	[-0.284, -0.079]	[-0.33, -0.11]
	Human rights	-0.028	-0.130	-0.039
		[-0.128, 0.073]	[-0.231, -0.026]	[-0.144, 0.065]
	Refugee relief	-0.38	-0.29	-0.27
		[-0.49, -0.27]	[-0.39, -0.18]	[-0.38, -0.16]
Funding sources	Few wealthy donors	-0.26	-0.123	-0.32
		[-0.35, -0.17]	[-0.211, -0.035]	[-0.41, -0.22]
	Government grants	-0.24	-0.171	-0.165
		[-0.33, -0.16]	[-0.260, -0.083]	[-0.255, -0.076]
Intercept	Intercept	-2.3	-2.4	-2.4
		[-2.5, -2.2]	[-2.5, -2.2]	[-2.5, -2.2]
N		36576		

N = 1,016 respondents. Estimates are median posterior log odds from a multinomial logistic regression model with three possible categories, and the columns for μ_1 , μ_2 , and μ_3 represent estimates for each of the outcomes; 95% credible intervals (equal-tailed quantile intervals) in brackets.

Table A 6: Excerpt from complete reference grid of all 576 possible combinations of attribute features and levels

Organization	Issue	Transparency	Accountability	Funding	Government	EMM
Red Cross	Emergency response	Transparency: No	Accountability: Yes	Funded primarily by many small private donations	Friendly relationship with government	0.486
Red Cross	Emergency response	Transparency: No	Accountability: Yes	Funded primarily by many small private donations	Criticized by government	0.396
Red Cross	Emergency response	Transparency: No	Accountability: Yes	Funded primarily by many small private donations	Under government crackdown	0.347
Red Cross	Emergency response	Transparency: Yes	Accountability: Yes	Funded primarily by many small private donations	Friendly relationship with government	0.626
Red Cross	Emergency response	Transparency: Yes	Accountability: Yes	Funded primarily by many small private donations	Criticized by government	0.537
Red Cross	Emergency response	Transparency: Yes	Accountability: Yes	Funded primarily by many small private donations	Under government crackdown	0.485

To calculate the AMCE for a feature, we find the difference in estimated marginal means. If we again assume that Table 6 contains the full reference grid, the AMCE for transparency would be $0.550 - 0.410$, or 0.140 , or 14 percentage points. Again, this is not actually the true causal effect—the real AMCE for transparency reported in the paper is the difference in marginal means for the 288 rows where transparency is no and the 288 rows where transparency is yes.

In the main paper, we include plots of the marginal means and AMCEs for all experimental features. The tables below correspond to each figure in the paper.

Table A 7: Complete marginal means and AMCEs (see Figure 1 in main paper)

Feature	Posterior EMM [*]	Contrast	Posterior AMCE [*]	$p_{\text{direction}}^{\dagger}$
Transparency				
Yes	0.307 [0.299, 0.314]	Yes–No	0.103 [0.095, 0.112]	1.00
No	0.204 [0.197, 0.210]	(Reference)	–	–
Accountability				
Yes	0.306 [0.298, 0.313]	Yes–No	0.101 [0.092, 0.110]	1.00
No	0.205 [0.199, 0.211]	(Reference)	–	–
Relationship with host government				
Under government crackdown	0.209 [0.201, 0.216]	Under government crackdown–Friendly relationship with government	–0.104 [–0.115, –0.094]	1.00
Criticized by government	0.244 [0.236, 0.252]	Criticized by government–Friendly relationship with government	–0.069 [–0.080, –0.059]	1.00
Friendly relationship with government	0.313 [0.305, 0.322]	(Reference)	–	–
Organizations				
Red Cross	0.349 [0.339, 0.359]	Red Cross–Amnesty International	0.123 [0.110, 0.136]	1.00
Oxfam	0.206 [0.198, 0.215]	Oxfam–Amnesty International	–0.020 [–0.031, –0.008]	1.00
Greenpeace	0.240 [0.231, 0.249]	Greenpeace–Amnesty International	0.014 [0.002, 0.026]	0.99
Amnesty International	0.226 [0.217, 0.235]	(Reference)	–	–
Issue areas				
Refugee relief	0.227 [0.218, 0.235]	Refugee relief–Emergency response	–0.056 [–0.068, –0.044]	1.00
Human rights	0.270 [0.261, 0.280]	Human rights–Emergency response	–0.012 [–0.025, 0.000]	0.97
Environment	0.241 [0.232, 0.250]	Environment–Emergency response	–0.042 [–0.054, –0.029]	1.00
Emergency response	0.283 [0.273, 0.292]	(Reference)	–	–
Funding sources				
Funded primarily by government grants	0.245 [0.237, 0.253]	Funded primarily by government grants–Funded primarily by many small private donations	–0.035 [–0.046, –0.025]	1.00
Funded primarily by a handful of wealthy private donors	0.239 [0.232, 0.247]	Funded primarily by a handful of wealthy private donors–Funded primarily by many small private donations	–0.041 [–0.052, –0.031]	1.00
Funded primarily by many small private donations	0.281 [0.273, 0.289]	(Reference)	–	–

* N = 576 estimated marginal means collapsed into feature-specific means. Values are on the percentage-point scale; single value is posterior median; 95% credible interval in brackets.

† The probability of direction ($p_{\text{direction}}$) is the probability that the posterior AMCE is strictly positive or negative—

Table A 8: Marginal means for all combinations of transparency and accountability (see Figure 2 in main paper)

Features	Posterior EMM [*]
Transparency: No	
Accountability: No	0.160 [0.154, 0.166]
Accountability: Yes	0.306 [0.298, 0.313]
Transparency: Yes	
Accountability: No	0.307 [0.299, 0.314]
Accountability: Yes	0.364 [0.355, 0.373]

* N = 576 estimated marginal means collapsed into feature-specific means. Values are on the percentage-point scale; single value is posterior median; 95% credible interval in brackets.

Table A 9: Marginal means and AMCEs for interaction between transparency, accountability, and government relationships (see Figure 3 in paper)

Feature	Posterior EMM*	Contrast	Posterior Δ^*	$p_{\text{direction}}^\dagger$
Relationship with government: Under government crackdown				
Transparency: Yes	0.254 [0.245, 0.264]	Yes–No	0.091 [0.083, 0.099]	1.00
Transparency: No	0.163 [0.156, 0.171]	<i>(Reference)</i>	—	—
Accountability: Yes	0.253 [0.244, 0.263]	Yes–No	0.089 [0.081, 0.097]	1.00
Accountability: No	0.164 [0.157, 0.172]	<i>(Reference)</i>	—	—
Relationship with government: Criticized by government				
Transparency: Yes	0.294 [0.284, 0.304]	Yes–No	0.101 [0.092, 0.110]	1.00
Transparency: No	0.193 [0.185, 0.201]	<i>(Reference)</i>	—	—
Accountability: Yes	0.293 [0.283, 0.303]	Yes–No	0.099 [0.090, 0.108]	1.00
Accountability: No	0.194 [0.186, 0.202]	<i>(Reference)</i>	—	—
Relationship with government: Friendly relationship with government				
Transparency: Yes	0.372 [0.362, 0.383]	Yes–No	0.118 [0.108, 0.128]	1.00
Transparency: No	0.254 [0.245, 0.264]	<i>(Reference)</i>	—	—
Accountability: Yes	0.371 [0.360, 0.382]	Yes–No	0.115 [0.105, 0.125]	1.00
Accountability: No	0.256 [0.246, 0.265]	<i>(Reference)</i>	—	—

* N = 576 estimated marginal means collapsed into feature-specific means. Values are on the percentage-point scale; single value is posterior median; 95% credible interval in brackets.

† The probability of direction ($p_{\text{direction}}$) is the probability that the posterior AMCE is strictly positive or negative—it is the proportion of the posterior AMCE that is the sign of the median.

Preregistration deviations

We made the following deviations from our preregistered protocol (Willroth & Ather-ton, 2024):

1. Type Analysis

Reason New knowledge

Timing After results known

Original wording “We will examine the aggregate marginal posterior distributions of the attribute levels”

Deviation description This statement was vague and seemed to imply that we would analyze the results of the model by looking only at the raw model coefficients. While it is possible to calculate exact feature contrasts by summing specific combinations of coefficients, we instead calculated estimated marginal means and their contrasts (or AMCEs) using the fitted model.

Reader impact This deviation should improve readers’ interpretation of the findings, since the reported results are no longer on a log-odds or logit scale, and instead are on a more interpretable percentage point scale—estimated marginal means show the percent of respondents who support an NGO given a specific feature, while AMCEs show the percentage point change in support when moving from one feature to another. The risk of bias is minimal as the underlying results are identical whether reported as logit-scale coefficients or marginal means.

2. Type Hypotheses

Reason Stylistic

Timing After results known

Original wording Q5a: “Donors will show increased willingness to donate to NGOs that are financially transparent”

Deviation description We rephrased this as H_{1a} : “If NGOs are financially transparent, then individual private donors will have a higher likelihood of supporting or donating to them.”

Reader impact This deviation has minimal impact on readers’ interpretation of the findings—it is only rephrased to follow an “if... then...” formulation.

3. Type Hypotheses

Reason New knowledge + stylistic

Timing Direction restated after data collection, but before results were known; “if... then...” formulation added after results known

Original wording Q5f: “Donors should be no more or less likely to donate to NGOs that are accountable and hold regular third party audits”

Deviation description We rephrased this as H_{1b} : “If NGOs are accountable, then individual private donors will have a higher likelihood of supporting or donating to them.”

Reader impact This deviation might have some impact on readers’ interpretation of the findings. This deviation was the result of misunderstanding existing work on the effect of nonprofit accountability on donor behavior, and we hypothesized that there would be no effect, contrary to what is predicted by previous research. The risk of bias is low, however—we reversed our prediction after data collection but before we analyzed the data and before the results were known.

4. Type Hypotheses

Reason Stylistic

Timing After results known

Original wording Q2a: “Donors will show increased willingness to donate to NGOs that are facing government crackdown or criticism”

Deviation description We rephrased this as H₂: “If NGOs face legal crackdowns abroad, then individual private donors will have a higher likelihood of supporting or donating to them.”

Reader impact This deviation has minimal impact on readers’ interpretation of the findings—it is only rephrased to follow an “if... then...” formulation.

5. Type Hypotheses

Reason Stylistic

Timing After results known

Original wording Q5b: “Donors will show increased willingness to donate to NGOs that are criticized by the government/under government crackdown when they are also financially transparent”

Deviation description We rephrased this as H₃: “If NGOs face legal restrictions abroad *and* are financially transparent, then individual private donors will have a higher likelihood of supporting or donating to them.”

Reader impact This deviation has minimal impact on readers’ interpretation of the findings—it is only rephrased to follow an “if... then...” formulation.

6. Type Hypotheses

Reason New knowledge

Timing Accountability prediction added after data collection, but before results were known; “if... then...” formulation added after results known

Original wording Q5b: “Donors will show increased willingness to donate to NGOs that are criticized by the government/under government crackdown when they are also financially transparent”

Deviation description We explore the interaction between (1) government crackdown and financial transparency and (2) government crackdown and accountability in the paper, but we only specified the first interaction in the preregistration.

Reader impact This deviation might have some impact on readers’ interpretation of the findings. The omission of a prediction of the relationship between government crackdown and accountability was inadvertent and we had intended to specify it. The risk of bias is low, as we added the new crackdown+accountability hypothesis after data collection and before the results were known.

7. Type Hypotheses

Reason Narrative

Timing After data collection, before results were known

Original wording Q1: Branding; Q3: Issue area; Q4: Funding sources

Deviation description For the sake of narrative simplicity, we do not explicitly test these three predictions as hypotheses. In this paper, our primary interest is crackdown, transparency, and accountability, but we look at branding, issue area, and funding sources to help compare and give context to the magnitude of our main hypotheses.

Reader impact This deviation might have some impact on readers’ interpretation of the findings, as it might appear that we have selectively reported a handful of our predictions. To avoid this, and for the sake of full transparency, we include these results in Figure 1 in the paper and Table 7. The risk of bias is low—we decided on the narrative framing for this paper after collecting the data but before analyzing the results.

Condensed preregistration

Consensed preregistration

This is an anonymized and condensed version of our full OSF preregistration protocol, which is available at <https://doi.org/10.17605/OSF.IO/HSBYD>.

Study information

Title. Why Donors Donate: Disentangling Organizational and Structural Heuristics for International Philanthropy

Research Questions.

OSF question

Please list each research question included in this study.

We use a conjoint survey experiment to examine the impact of organizational features of nongovernmental organizations (NGOs) and the structural factors in target countries in which they operate on donors’ decisions to engage in philanthropy. We explore three research questions in this study:

1. **Do donors rely on structural characteristics of NGOs as heuristics when deciding to donate? How do structural heuristics compare to organizational heuristics?**

Donors rely on shortcuts, signals, and heuristics to determine the trustworthiness of NGOs, since seeking out complete information about an organization’s deservingness and efficiency is costly and time-consuming. Previous research has found that an NGO’s organizational characteristics commonly serve as heuristics for donors. Donors use an organization’s overhead costs, the issues it works on, its transparency and accountability practices, and a host of other organizational practices as signals of an organization’s efficiency and deservingness, which then influences their decision to make a donation. These kinds of heuristics are attributes that organizations can typically control—NGOs can publish annual reports, restructure their management, and engage in other strategies to appear more worthy of donation.

Structural characteristics, such as the political and legal environment an NGO faces in its host country, may also serve as signals to donors of NGO deservingness. We are interested in whether the contentiousness of an NGO’s relationship with its host government influences donor decision making. Do donors care if nonprofits they care about are criticized by, persecuted by, or expelled from the countries they work in?

We are also interested in the effect of organizational characteristics on donor decision making. How do managerial practices (financial transparency and accountability systems), funding sources (private donations and government grants), and issue areas (emergency response, environmental issues, human rights, and refugee relief) compare to structural characteristics when deciding to donate? Which heuristics are more influential?

2. **How do individual-level donor characteristics interact with structural and organizational heuristics? Which kinds of people are more or less likely to consider an NGO’s host country political environment, managerial practices, funding sources, or issue area?**

The decision to donate to an NGO is not determined solely by an organization’s characteristics. Donors themselves have personality traits, preferences, and experiences that make them more or less likely to engage in philanthropy. We are interested in how individual donor characteristics, such as political ideology, political knowledge, religious attendance, involvement in charitable activities, involvement in activism, and demographic attributes interact with organizational- and structural-level factors.

3. **What is the optimal mix of attribute levels for NGOs interested in maximizing donations?**

Finally, given individual donor characteristics and preferences, we are interested in finding the optimal mix of organizational and structural attributes. What might an NGO try to emphasize in its marketing campaigns? Should it highlight its funding sources, managerial practices, issue area, or relationship with its host governments (even if that relationship is negative)?

Hypotheses.

OSF question

For each of the research questions listed in the previous section, provide one or multiple specific and testable hypotheses. Please state if the hypotheses are directional or non-directional. If directional, state the direction. A predicted effect is also appropriate here.

For our first set of questions, we predict that:

1. **Branding**

- Donors will be more likely to donate to Oxfam and Red Cross compared to Amnesty International and Greenpeace [Mechanism: awareness of need and contentiousness of issue area]

2. **Government crackdown**

- Donors will show increased willingness to donate to NGOs that are facing government crackdown or criticism [Mechanism: Governments wouldn’t be cracking down on them if they didn’t perceive a threat from them which means organizations implementing their missions effectively. This perception of efficacy leads to increased donations.]
- Donors will show increased willingness to donate to Oxfam and Red Cross when they are facing government crackdown or criticism compared to when Amnesty or Greenpeace is facing crackdown.

3. Issue area

- Donors will show increased willingness to donate to NGOs working in less contentious issue areas (emergency response and refugee relief) over more contentious issue areas (environment and human rights)
- Donors will show increased willingness to donate to NGOs facing government crackdown/criticism working in less contentious issue areas (emergency response and refugee relief) over more contentious issue areas (environment and human rights) [Mechanisms: Perceptions of deservingness of NGOs dealing with emergency response and refugee relief. Donors are also more likely to donate to programs that are compatible with government preferences and have easily measurable outputs, which environment and human rights programs often lack. NGOs working on more contentious issue areas may be expelled or shut down, which would be a waste of donor resources, make it less likely that they donate to these groups.]

4. Funding sources

- Donors will show increased willingness to donate to NGOs that are funded primarily by numerous small private donors compared to NGOs that are funded by a handful of wealthy private donors and government grants [Mechanism: Perception of efficacy - your contribution matters as a small donor. Government funding may also imply lack of independence of government which can reduce the efficiency of an organization.]
- Donors will show increased willingness to donate to NGOs that are facing government crackdown and are funded primarily by numerous small private donors
- Donors will show increased willingness to donate to NGOs that are facing government crackdown and are funded primarily by numerous small private donors and work in less contentious areas (emergency response and refugee relief)

5. Organizational practices

- Donors will show increased willingness to donate to NGOs that are financially transparent [Mechanism: Perception of efficacy]
- Donors will show increased willingness to donate to NGOs that are criticized by the government/under government crackdown when they are also financially transparent
- Donors will show increased willingness to donate to NGOs that are criticized by the government/under government crackdown when they are also financially transparent and are funded primarily by numerous small private donors
- Donors will show increased willingness to donate to NGOs that are criticized by the government/under government crackdown when they are also financially transparent and work in less contentious areas (emergency response and refugee relief)

- Donors will show increased willingness to donate to NGOs that are criticized by the government/under government crackdown when they are also financially transparent and work in less contentious areas (emergency response and refugee relief) and are funded by numerous small donors
- Donors should be no more or less likely to donate to NGOs that are accountable and hold regular third party audits [Mechanism: Donors don’t necessarily seek assurance through third-party programs/audits and charity watchdogs, but rather through word of mouth, personal scrutiny and local networks]

Because of the nature of our statistical methods, we do not have exact hypotheses for the second and third set of questions. We describe how we answer these questions in the “Follow-up analyses” and “Exploratory analysis” sections below.

Sampling Plan

Existing data. Registration prior to creation of data

Explanation of existing data. We will not use any existing data.

Data collection procedures.

OSF question

Please describe the process by which you will collect your data. If you are using human subjects, this should include the population from which you obtain subjects, recruitment efforts, payment for participation, how subjects will be selected for eligibility from the initial pool (e.g. inclusion and exclusion rules), and your study timeline. For studies that don’t include human subjects, include information about how you will collect samples, duration of data gathering efforts, source or location of samples, or batch numbers you will use.

Participants will complete a 10-minute survey on Qualtrics. A static version of the survey is accessible at REDACTED.

Participants of the survey experiment will be recruited through Centiment, a commercial online provider of high quality nonprobability opt-in survey panels. Centiment ensures panel quality by actively recruiting representative samples of the US population and provides monetary incentives and rewards to participants.

To see how varying NGO characteristics influence the decision to donate, our sample will be representative of a population of people who are likely to donate to charity. We ask potential participants a screening question early in the survey (“Q2.5: How often do you donate to charity”). If a participant responds that they give to once every few years or never, they will be disqualified from the study and the survey will end early.

We will provide Centiment with a link to the survey, which is hosted by Qualtrics. Centiment will then distribute the link to their panel. Participants are compensated through Centiment’s internal reward system through cash, points, and other incentives. Centiment does not provide precise details of participant compensation. Centiment states that their compensation is “fair,” and the company’s business model encourages the company to find and maintain high quality panelists. We thus infer that the amount provided is fair and justified. Centiment users receive compensation from the company following the completion of the survey.

Sample size.

OSF question

Describe the sample size of your study. How many units will be analyzed in the study? This could be the number of people, birds, classrooms, plots, interactions, or countries included. If the units are not individuals, then describe the size requirements for each unit. If you are using a clustered or multilevel design, how many units are you collecting at each level of the analysis?

Our target sample size is 1,000 participants.

Sample size rationale.

OSF question

This could include a power analysis or an arbitrary constraint such as time, money, or personnel.

A sample size of at least 500 respondents is typical for estimating a hierarchical Bayesian model based on conjoint data. We double this amount because we are interested in analyzing subpopulations of respondents, which requires a larger sample, and we had sufficient budget to acquire up to 1,000 respondents.

Stopping rule.

OSF question

If your data collection procedures do not give you full control over your exact sample size, specify how you will decide when to terminate your data collection.

Centiment will monitor how many surveys are successfully completed and will solicit responses until our 1,000 target is met.

Design plan

Study type. Experiment: A researcher randomly assigns treatments to study subjects, this includes field or lab experiments. This is also known as an intervention experiment and includes randomized controlled trials.

Blinding. For studies that involve human subjects, they will not know the treatment group to which they have been assigned.

Study design.

OSF question

Describe your study design. Examples include two-group, factorial, randomized block, and repeated measures. Is it a between (unpaired), within-subject (paired), or mixed design? Describe any counterbalancing required. Typical study designs for observation studies include cohort, cross sectional, and case-control studies.

We use a fractional factorial design. Since no single respondent can possibly see all possible combinations of the attribute levels, we create a number of different versions of the experimental design. We utilize a hierarchical Bayesian model in part to allow for information sharing across like respondents when estimating individual-level preferences for the attribute levels.

Randomization.

OSF question

If you are doing a randomized study, how will you randomize, and at what level?

Every respondent will be randomly assigned a version of the fractional factorial experimental design.

Analysis Plan

Statistical models.

OSF question

What statistical model will you use to test each hypothesis? Please include the type of model (e.g. ANOVA, multiple regression, SEM, etc) and the specification of the model (this includes each variable that will be included as predictors, outcomes, or covariates). Please specify any interactions that will be tested and remember that any test not included here must be noted as an exploratory test in your final article.

We will use a hierarchical Bayesian multinomial logit model with conjugate or otherwise typical priors. The individual-level model is the multinomial logit and the upper-level model of heterogeneity is multivariate normal,

$$\begin{aligned}\beta &\sim \text{Multivariate } \mathcal{N}(Z\Gamma, \xi) \\ y &\sim \text{Multinomiallogit}(X\beta, \varepsilon)\end{aligned}$$

where y = which alternative the respondent chooses to donate, X = design matrix of attribute levels, β = latent individual preferences for the attribute levels, Z = matrix of individual-level covariates, Γ = matrix of coefficients mapping individual-level covariates onto the latent individual-level preferences, and ε and ξ = errors.

Inference criteria.

OSF question

What criteria will you use to make inferences? Please describe the information you'll use (e.g. specify the p-values, Bayes factors, specific model fit indices), as well as cut-off criterion, where appropriate. Will you be using one or two tailed tests for each of your analyses? If you are comparing multiple conditions or testing multiple hypotheses, will you account for this?

We will examine the aggregate marginal posterior distributions of the attribute levels and use 95% credible intervals to establish “significance.” Effects are “significant” if the 95% credible intervals don’t include 0. Similarly, marginal posterior distributions are “significantly” different if the 95% credible intervals don’t overlap.

We will examine the marginal posterior distributions of the following models:

- Organizational and structural attribute levels with an intercept-only distribution of heterogeneity
- Organizational and structural attribute levels with competing sets of covariates in the distribution of heterogeneity

Finally, we will employ the posterior distribution of model parameters to conduct counterfactual analyses via a market simulator to determine optimal policies.

Data exclusion.

OSF question

How will you determine which data points or samples (if any) to exclude from your analyses? How will outliers be handled?

We ask potential participants a screening question early in the survey (“Q2.5: How often do you donate to charity”). If a participant responds that they give to once every few years or never, they will be disqualified from the study and the survey will end early.

We include one question (“Q2.11: Please select blue from the following list:”) to monitor respondent attention. In our analysis we will exclude respondents who fail this question.

Missing data.

OSF question

How will you deal with incomplete or missing data?

Because all survey questions are required, we do not anticipate issues with incomplete or missing data.

References

- Bürkner, P.-C. (2017). brms: An R package for Bayesian multilevel models using Stan. *Journal of Statistical Software*, 80(1), 1–28. <https://doi.org/10.18637/jss.v080.i01>
- R Core Team. (2024). *R: A language and environment for statistical computing* (Version 4.4.2). R Foundation for Statistical Computing. <https://www.r-project.org/>
- Stan Development Team. (2024a). *CmdStan: The shell interface to Stan* (Version 2.36). <https://mc-stan.org/docs/cmdstan-guide/>
- Stan Development Team. (2024b). *Stan modeling language* (Version 2.32.2). <https://mc-stan.org>
- Willroth, E. C., & Atherton, O. E. (2024). Best laid plans: A guide to reporting preregistration deviations. *Advances in Methods and Practices in Psychological Science*, 7(1), 1–14. <https://doi.org/10.1177/25152459231213802>