

October 7, 2020

Online appendix for “Dynamics of International Giving: How Heuristics Shape Individual Donor Preferences”

Suparna Chaudhry
Lewis & Clark College
schaudhry@lclark.edu

Andrew Heiss
Georgia State University
aheiss@gsu.edu

Contents

Sample	2
Recruitment	2
Sample size	2
Demographics and representativeness	3
Collapsing likelihood variable	6
Experiment	7
Preregistration	7
Treatment assignment balance and CONSORT diagram	7
Priors and models	10
Survey experiment	12
Recruitment and payment	12
Consent	12
Crackdown and private giving	13
Demographics	15
References	17

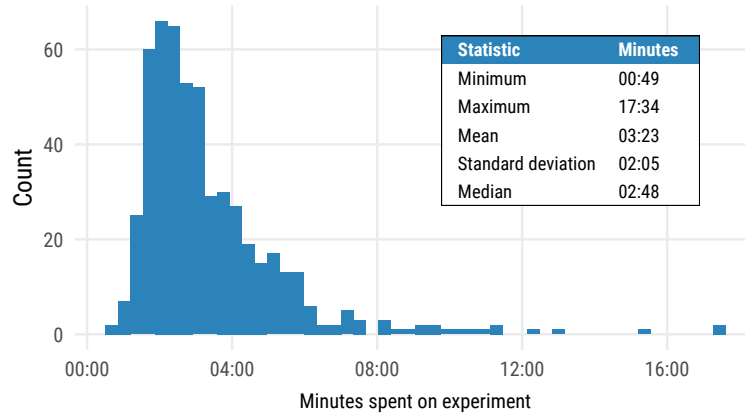


Figure 1: Time spent on experiment

Sample

Recruitment

Experiment participants were recruited via Amazon Mechanical Turk (MTurk), an online crowdsourcing platform that has been extensively used in social science research in recent years. MTurk allows researchers to recruit participants to perform tasks such as filling out surveys and opinion polls, participating in experiments, or coding the content of documents. Researchers advertise their studies as a human intelligence task (HIT) on MTurk, and participants choose only those HITs that interest them, given the promised price and estimated duration of the task. We listed a link to our survey on MTurk on March 22–23, 2018, and offered participants \$0.75 for successfully completing the study. We estimated that the survey would take 5 minutes to complete and paid participants the equivalent of a \$9/hour wage. On average, participants completed the survey in 3 minutes and 20 seconds (see Figure 1).

Sample size

We used power analysis to determine our sample size: with the anticipation of finding a small effect size ($f^2 = 0.02$) at a 0.05 significance level with a power of 0.9, our target sample size was 527, which we then rounded up to 540 in case of error or noncompliance with the survey.

Demographics and representativeness

Amazon’s Mechanical Turk (MTurk) has become an increasingly popular method for fielding survey experiments on semi-representative national samples in an inexpensive manner, though drawing definitive causal inference from convenience samples is not without issues. Previous research has found that MTurk workers tend to be more liberal, younger, and less racially diverse than the general US population (Bernisky et al., 2012; Clifford et al., 2015). The results from our experiment generally mirror these findings, as seen in Table 1, which compares demographic characteristics of our sample with national averages from the US Census’s Current Population Survey (CPS) (Flood et al., 2017). For general demographic information, we use the 2017 Annual Social and Economic Supplement (ASEC) to the CPS. From 2002–2015, the CPS included a Volunteer Supplement every September, so we use 2015 data for data on volunteering and donating to charity. We do not show other respondent demographic details because we do not have good population-level data to compare our sample with. We could theoretically use Pew data to compare political preferences, but Pew collects data on party affiliation, while we collected data about respondents’ ideological positions along a conservative–liberal spectrum, which makes the two variables incomparable.

Table 1: Characteristics of experimental sample

Variable	Sample	National	Δ_{median}	95% HPDI
Female (%) ^a	54.80%	51.0%	3.7%	(-0.7%, 7.7%)
Age (% 35+) ^a	47.27%	53.9%	-6.6%	(-10.7%, -2.2%) [†]
Income (% \$50,000+) ^a	50.39%	27.4%	21.7%	(17.3%, 26%) [†]
Education (% BA+) ^a	46.14%	29.9%	16.2%	(11.9%, 20.2%) [†]
Donated in past year (%) ^b	82.49%	48.8%	33.6%	(30.3%, 36.8%) [†]
Volunteered in past year (%) ^b	54.24%	75.1%	-20.9%	(-25%, -16.7%) [†]

^aAnnual CPS, March 2017
^bMonthly CPS, September 2015
[†]National value is outside the sample highest posterior density interval (HPDI)






As seen in Table 1, our sample is younger, wealthier, and more educated than national averages. This is to be expected, given previous findings about the characteristics of MTurk workers and given that our target population is the portion of Americans willing to give money to charities online, which implies access to technology and disposable income. Our sample self-reports high levels of charitable giving: 82% said they donate at least once a year, while only 49% donate to charity in the previous year nationally. There are several possible explanations for this discrepancy. First, there are differences in the wording of our question and CPS’s question—we ask how often respondents typically donate, while CPS asks if respondents have donated (or not) in the past 12 months. Second, respondents












were primed and knew that the survey was related to NGOs, and thus might be displaying social desirability bias. Finally, some charitably-oriented respondents may have self-selected into the survey, given that the title listed on MTurk mentioned international nonprofit organizations.




We are not overly concerned with these discrepancies, since we aimed our study at those who would be more inclined to donate online. Moreover, in an impressive study to test the validity of experiments conducted on non-representative, internet-based samples, Mullinix et al. (2015) ran 20 identical experiments on both MTurk and on a nationally representative population sample. The results from both types of surveys tracked closely with each other and MTurk-based surveys were able to replicate most of the treatment effects found in national samples. Accordingly, they concluded that internet-based population samples are “useful testing grounds for experimental social science” and that they can still lead to substantial progress in generating knowledge and making inferential claims (Mullinix et al., 2015, p. 124). MTurk-based studies ultimately do not replace the need for more robust nationally representative samples, which still provide the most robust and accurate evidence for causal claims, but they are a cost-effective method for estimating potential population-level effects. Other research corroborates these conclusions—while MTurk samples are not perfectly representative, they are still helpful approximations of representativeness (Bernisky et al., 2012; Clifford et al., 2015; Coppock et al., 2018; Krupnikov & Levine, 2014).

Table 2 provides summary statistics for all the variables we collected in our survey.

Table 2: Descriptive statistics

Variable		Details
Likelihood of donation		Extremely unlikely (46; 8.66%) Somewhat unlikely (110; 20.72%) Neither likely nor unlikely (138; 25.99%) Somewhat likely (192; 36.16%) Extremely likely (45; 8.47%)
Likelihood of donation (binary)		Not likely (294; 55%) Likely (237; 45%)
Amount hypothetically donated (\$)		Median: 10 Mean: 22.4 Std. Dev.: 25.67
Gender		Female (291; 54.80%) Male (237; 44.63%) Other (1; 0.19%) Prefer not to say (2; 0.38%)
Age		Under 18 (1; 0.19%) 18 – 24 (43; 8.10%) 25 – 34 (207; 38.98%) 35 – 44 (129; 24.29%) 45 – 54 (90; 16.95%) 55 – 64 (44; 8.29%) 65 – 74 (16; 3.01%) 75 – 84 (1; 0.19%)

Variable		Details
Income		Less than \$10,000 (32; 6.03%) \$10,000 – \$19,999 (43; 8.10%) \$20,000 – \$29,999 (50; 9.42%) \$30,000 – \$39,999 (78; 14.69%) \$40,000 – \$49,999 (58; 10.92%) \$50,000 – \$59,999 (55; 10.36%) \$60,000 – \$69,999 (43; 8.10%) \$70,000 – \$79,999 (38; 7.16%) \$80,000 – \$89,999 (25; 4.71%) \$90,000 – \$99,999 (30; 5.65%) \$100,000 – \$149,999 (45; 8.47%) More than \$150,000 (21; 3.95%) Prefer not to say (13; 2.45%)
Education		Less than high school (2; 0.38%) High school graduate (47; 8.85%) Some college (128; 24.11%) 2 year degree (68; 12.81%) 4 year degree (212; 39.92%) Graduate or professional degree (66; 12.43%) Doctorate (8; 1.51%)
Frequency of attending religious services		More than once a week (19; 3.58%) Once a week (74; 13.94%) Once or twice a month (50; 9.42%) A few times a year (71; 13.37%) Seldom (98; 18.46%) Never (215; 40.49%) Don't know (4; 0.75%)
Political views		Strong liberal (76; 14.31%) Liberal (150; 28.25%) Independent, leaning liberal (82; 15.44%) Independent (80; 15.07%) Independent, leaning conservative (60; 11.30%) Conservative (61; 11.49%) Very conservative (22; 4.14%)
Frequency of following public affairs		Most of the time (217; 40.87%) Some of the time (213; 40.11%) Only now and then (84; 15.82%) Hardly at all (17; 3.20%)
Frequency of charitable donations		Once a week (37; 6.97%) Once a month (105; 19.77%) Once every three months (105; 19.77%) Once every six months (102; 19.21%) Once a year (89; 16.76%) Once every few years (56; 10.55%) Never (37; 6.97%)
Volunteered in past 12 months		No (288; 54.2%) Yes (243; 45.8%)
Prior favorability towards humanitarian NGOs		Very unfavorable (2; 0.38%) Unfavorable (6; 1.13%) Neutral (39; 7.34%) Favorable (235; 44.26%) Very favorable (249; 46.89%)
Prior favorability towards humanitarian NGOs (binary)		Not favorable (47; 9%) Favorable (484; 91%)
Prior favorability towards human rights NGOs		Very unfavorable (5; 0.94%) Unfavorable (12; 2.26%) Neutral (61; 11.49%) Favorable (226; 42.56%) Very favorable (227; 42.75%)
Prior favorability towards human rights NGOs (binary)		Not favorable (78; 15%) Favorable (453; 85%)

Variable		Details
Prior favorability towards development NGOs		Very unfavorable (5; 0.94%) Unfavorable (8; 1.51%) Neutral (45; 8.47%) Favorable (235; 44.26%) Very favorable (238; 44.82%)
Prior favorability towards development NGOs (binary)		Not favorable (58; 11%) Favorable (473; 89%)
Attention check 2		Correct (531; 100%)

Collapsing likelihood variable

To simplify our analysis, we collapse our likelihood scale from a 1-5 Likert scale to a binary variable:

Original answer	Collapsed answer
Extremely likely	Likely
Somewhat likely	Likely
Neither likely nor unlikely	Unlikely
Somewhat unlikely	Unlikely
Extremely unlikely	Unlikely

To check that the results are internally consistent when collapsed, we ran an ordered probit model (using Stan) to see if the cutpoints follow the distribution of answers, and they do. “Somewhat likely” and above has a cutpoint of > 0.18 , meaning that the likelihood is positive on average for both “Somewhat likely” and “Extremely likely.”

Table 4: Ordered probit regression with donation likelihood as outcome variable

	(1)
Crackdown	0.089 (0.091)
Cutpoint: Extremely unlikely Somewhat unlikely	-1.314 (0.091)
Cutpoint: Somewhat unlikely Neither likely nor unlikely	-0.496 (0.075)
Cutpoint: Neither likely nor unlikely Somewhat likely	0.181 (0.072)
Cutpoint: Somewhat likely Extremely likely	1.416 (0.093)
Observations	531
Posterior sample size	8000.000

Experiment

Preregistration

Prior to launching the survey, we preregistered our hypotheses and research design at the Open Science Framework, and our preregistration protocol is available at <https://osf.io/dx973/>. We made two minor deviations from the original preregistration. First, in the *text* of our preregistration, we inadvertently only specified the non-nested versions of our hypotheses:

1. Donors will give more/be more likely to give to NGOs that face legal crackdowns abroad
2. Donors will give more/be more likely to give to NGOs working on humanitarian issues
3. Donors will give more/be more likely to give to NGOs that do not receive substantial funding from government sources

In our preregistered data analysis *plan*, however, we describe the nested and interacted versions of the hypotheses (i.e. donors will give more to humanitarian NGOs facing legal crackdowns, etc.). This was an oversight—we inadvertently used overly simple (and incorrect) textual hypotheses, but correctly described the full analysis plan (using the correct, unstated, unstated hypotheses).

Second, in our preregistration plan, we declared that we would analyze our data with a series of three Bayesian linear regression models with increasing numbers of interaction terms. To find the effect of crackdown within issue area, for instance, we used a crackdown \times issue term, and to find the effect of crackdown conditioned on issue and funding we used a three-way crackdown \times issue \times funding term. We then planned on reassembling the different coefficient and intercept terms to approximate the mean values in each combination of conditions. To simplify our analyses in this version of the paper, we estimated group means directly with Stan (R Core Team, 2016; Stan Development Team, 2016a, 2016b). This approach provides the same results as the regression models with 2- and 3-way interaction terms, but is far simpler to interpret.

Treatment assignment balance and CONSORT diagram

To help ensure the quality of the responses we received and filter out workers who try to get through the HIT as quickly as possible without reading the questions (Berinsky et al., 2014), we included two attention check questions in the experiment. The first check was the second question in the survey (Q1.3) and involved reading several sentences and following instructions to select specific responses. We designed this question to filter out shirking respondents. We included a sec-

ond, simpler check later in the survey, to ensure that participants were still engaged (Q3.8). Only two participants did not pass this check.

Figure 2 and Table 5 show the assignment of participants to the eight different experimental conditions. We excluded participants that (1) did not participate in the experiment through MTurk and (2) failed either attention check.

Table 5: Balance of experimental conditions

Crackdown	Issue	Funding	n
No crackdown	Human rights	Government	68
No crackdown	Human rights	Private	64
No crackdown	Humanitarian assistance	Government	68
No crackdown	Humanitarian assistance	Private	65
Crackdown	Human rights	Government	65
Crackdown	Human rights	Private	65
Crackdown	Humanitarian assistance	Government	68
Crackdown	Humanitarian assistance	Private	68
Total	-	-	531

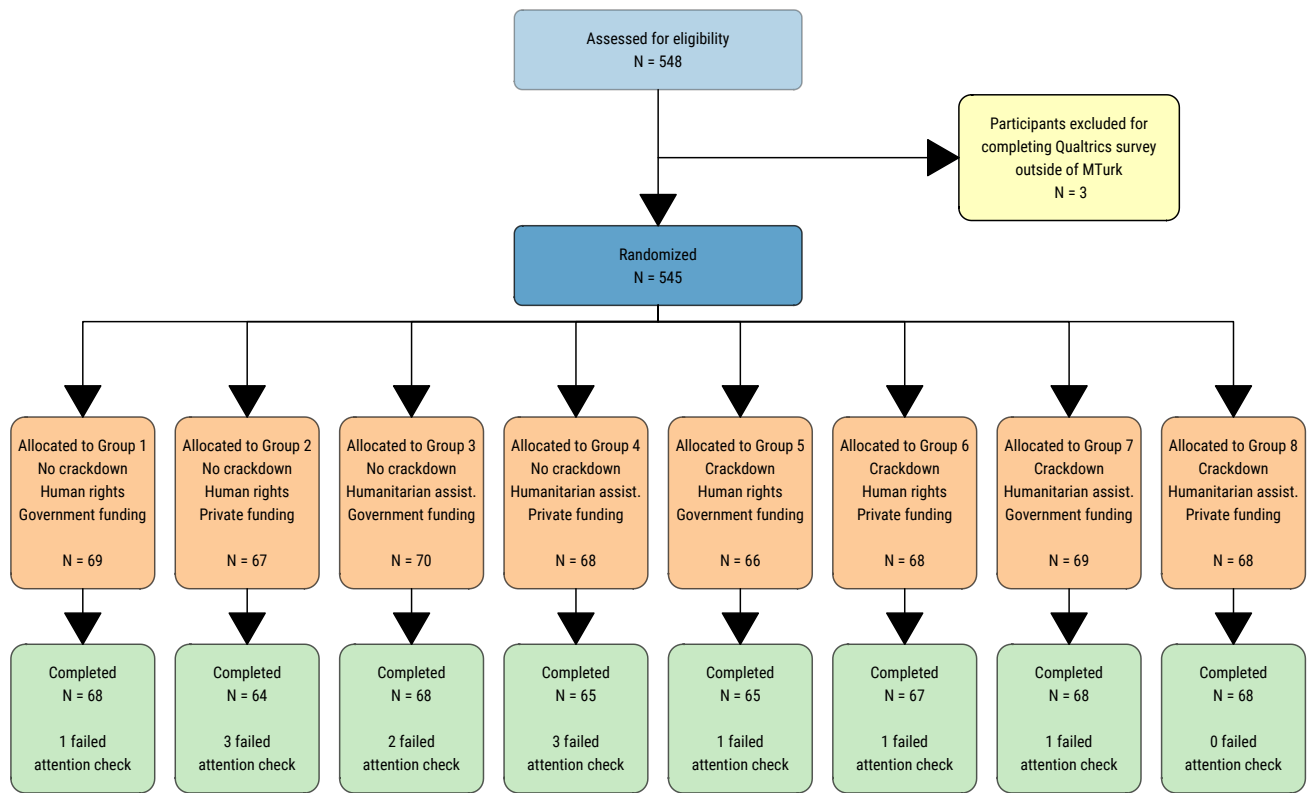
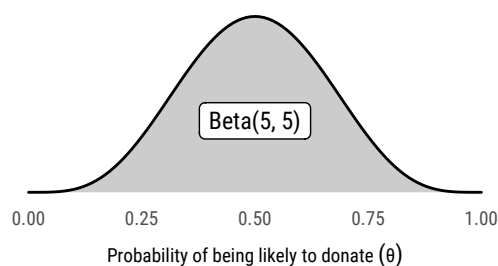


Figure 2: CONSORT diagram for experiment

1c86b49 on 2020-10-10

Figure 3: Prior θ for binomial models

Priors and models

We use two statistical models for measuring the effect of the crackdown condition on the likelihood of donating and the amount hypothetically donated. Because we measure likelihood of donation with a binary “Likely to donate” vs. “Not likely to donate” variable, we model the proportion of people responding that they would be likely to donate with a binomial distribution. We use a prior distribution of Beta(5, 5) for θ to center the probability of responding positively at 50%. In more formal terms, we model this distribution as follows:

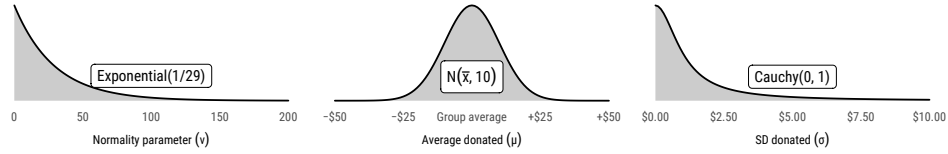
$$n_{\text{group 1, group 2}} \sim \text{Binomial}(n_{\text{total in group}}, \theta_{\text{group}}) \quad [\text{likelihood}]$$

$$\text{Difference} = n_{\text{group 2}} - n_{\text{group 1}} \quad [\text{difference in proportions}]$$

$$n : \text{Number likely to donate}$$

$$\theta_{\text{group 1, group 2}} \sim \text{Beta}(5, 5) \quad [\text{prior prob. of being likely to donate}]$$

We estimate the mean amount donated in each condition (crackdown vs. no crackdown, humanitarian assistance vs. human rights issues, private vs. government funding) using a t distribution. Following Kruschke (2013), we use an exponential distribution with a rate of $1/29$ for the ν parameter; a normal distribution with the group mean and standard deviation of 10 to capture wider variability in how much respondents might donate for the μ parameter; and a Cauchy(0, 1) distribution for the σ parameter. In more formal terms, we use the following model and priors:

Figure 4: Prior ν , μ , and σ for amount models

$$x_{\text{group 1, group 2}} \sim \text{Student } t(\nu, \mu, \sigma) \quad [\text{likelihood}]$$

$$\text{Difference} = x_{\text{group 2}} - x_{\text{group 1}} \quad [\text{difference in means}]$$

x : Mean amount donated

$$\nu \sim \text{Exponential}(1/29) \quad [\text{prior normality}]$$

$$\mu_{\text{group 1, group 2}} \sim \mathcal{N}(\bar{x}_{\text{group 1, group 2}}, 10) \quad [\text{prior donation mean per group}]$$

$$\sigma_{\text{group 1, group 2}} \sim \text{Cauchy}(0, 1) \quad [\text{prior donation sd per group}]$$

We obtain the posterior distribution of each dependent variable with Markov Chain Monte Carlo (MCMC) sampling and simulate values from the joint posterior distribution of the coefficient parameters. We use Stan (R Core Team, 2016; Stan Development Team, 2016a, 2016b) to generate 4 MCMC chains with 4,000 iterations in each chain, 2,000 of which are used for warmup. We use the median values from the posterior distributions as point estimates and calculate credible intervals using the 95% highest posterior density.

Survey experiment

Recruitment and payment

Title of Mechanical Turk HIT “Survey on international nonprofit organizations (~ 5 minutes)”

HIT description “We are conducting an academic survey about international nonprofit organizations and want to know your opinion about them. This survey will take roughly five minutes to complete.”

Payment Participants were paid \$0.75 for successfully completing the experiment, commensurate with a \$9/hour wage.

Consent

Q1.1: You have been invited to participate in an academic research survey about international nongovernmental organizations. This study is being conducted by Suparna Chaudhry from Christopher Newport University and Andrew Heiss from Brigham Young University. The experiment should take about 5 minutes to complete. You will be paid \$0.75 for your participation.

Your participation in this study is entirely voluntary and you are free to skip any question or withdraw from the experiment at any time. You will not be asked to provide any personal information, and your answers will remain anonymous. There are no reasonably foreseeable risks or discomforts associated with participating in this study. There will be questions designed to check that you’re paying attention to the details of the experiment. If you answer these incorrectly, the survey will end early and you may not receive compensation.

Beyond your payment, you will receive no direct benefits from participating in this research study, but your responses may reveal important and useful insights about international NGOs and philanthropy.

If you have any questions at any time about the study or the procedures, you can contact Suparna Chaudhry (suparna.chaudhry@cnu.edu; Reiff Center for Human Rights and Conflict Resolution, Christopher Newport University, Newport News, VA 23606) or Andrew Heiss (andrew_heiss@byu.edu; Romney Institute of Public Management, Brigham Young University, Provo, UT 84602) at andrew_heiss@byu.edu. This research has been reviewed by the Institutional Review Board at Christopher Newport University (757-594-7461; IRB@cnu.edu) and the IRB administrator at Brigham Young University (801-422-1461; irb@byu.edu).

By clicking “I agree” below you are indicating that you are at least 18 years old, have read and understood this consent form and agree to participate in this research study. Please print a copy of this page for your records.

Q1.2: I have read the above information, and I consent to take part in the study. Single answer

- Yes
 - No
-

Q1.3: Please read the paragraph below carefully.

What is your favorite color? While we are interested in learning about your preferences on a variety of topics, including colors, we also want to make sure that you are reading the questions we have written. To demonstrate that you've read this much, just go ahead and select both green and yellow from among the alternatives listed below, no matter what your favorite color is. Yes, ignore the first sentence and select both green and yellow.

(End survey without payment if this question is incorrect.) Multiple answers allowed
(*must be green and yellow*)

- Pink
 - Red
 - Green
 - White
 - Yellow
 - Blue
-

Crackdown and private giving

Q2.1: How do you feel about the following types of organizations? Matrix table

	Very unfavorable	Unfavorable	Neutral	Favorable	Very favorable
Nonprofit organizations that provide humanitarian relief	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>
Nonprofit organizations that advocate for human rights	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>
Nonprofit organizations that help with development and alleviate poverty	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>



Q2.2: The **International Rescue Committee (IRC)** focuses on {ISSUE AREA: humanitarian assistance for refugees | human rights for refugees} {MAIN MANIPULATION: and works in countries that have recently passed laws that harshly restrict nonprofit organizations}.

A substantial proportion of IRC's funding comes from {FUNDING SOURCE: government | private donors}.

Q2.3: How likely would you be to donate to IRC? Single answer

- Extremely likely
- Somewhat likely
- Neither likely nor unlikely
- Somewhat unlikely
- Extremely unlikely

Q2.4: If you had an extra \$100, how much would you be willing to donate to IRC right now? (*Total must add to 100*) Numeric slider, options must sum to 100

- Donate to IRC (0–100)
- Keep for yourself (0–100)

Q2.5: Why did you select that amount? Text field

Demographics

Q3.1: Almost done! This is the last page of the study.

Q3.2: How often do you donate to charity (either cash or in-kind)? Single answer

- Once a week
- Once a month
- Once every three months
- Once every six months
- Once a year
- Once every few years
- Never

Q3.3: Have you volunteered in the past 12 months? Single answer

- Yes
- No

Q3.4: Some people seem to follow what's going on in government and public affairs most of the time, whether there's an election going on or not. Others aren't that interested. Would you say that you follow what's going on in government and public affairs... Single answer

- Most of the time
- Some of the time
- Only now and then
- Hardly at all

Q3.5: How would you describe your political views? Single answer

- Strong liberal
- Liberal
- Independent, leaning liberal
- Independent
- Independent, leaning conservative
- Conservative
- Very conservative

Q3.0: What is the highest degree or level of school you have completed? Single answer

- Less than high school

- High school graduate
- Some college
- 2 year degree
- 4 year degree
- Graduate or professional degree
- Doctorate

Q3.1: Aside from weddings and funerals, how often do you attend religious services?
Single answer

- More than once a week
- Once a week
- Once or twice a month
- A few times a year
- Seldom
- Never
- Don't know

Q3.2: Please select blue from the following list: Single answer

- Red
- Yellow
- Blue
- Green

Q3.3: What is your gender? Single answer

- Female
- Male
- Transgender
- Other:
- Prefer not to say

Q3.10: What is your total household income? Single answer

- Less than \$10,000
- \$10,000 - \$19,999
- \$20,000 - \$29,999
- \$30,000 - \$39,999
- \$40,000 - \$49,999
- \$50,000 - \$59,999

- \$60,000 - \$69,999
- \$70,000 - \$79,999
- \$80,000 - \$89,999
- \$90,000 - \$99,999
- \$100,000 - \$149,999
- More than \$150,000

Q3.11: How old are you? Single answer

- Under 18
- 18 - 24
- 25 - 34
- 35 - 44
- 45 - 54
- 55 - 64
- 65 - 74
- 75 - 84
- 85 or older

References

- Berinsky, A. J., Margolis, M. F., & Sances, M. W. (2014). Separating the shirkers from the workers? making sure respondents pay attention on self-administered surveys. *American Journal of Political Science*, 58(3), 739–53. <https://doi.org/10.1111/ajps.12081>
- Bernisky, A. J., Huber, G. A., & Lenz, G. S. (2012). Evaluating online labor markets for experimental research: Amazon.com's mechanical turk. *Political Analysis*, 20(3), 351–68. <https://doi.org/10.1093/pan/mpo57>
- Clifford, S., Jewell, R. M., & Waggoner, P. D. (2015). Are samples drawn from mechanical turk valid for research on political ideology? *Research & Politics*, 2(4), 1–9. <https://doi.org/10.1177/2053168015622072>
- Coppock, A., Leeper, T. J., & Mullinix, K. J. (2018). *The generalizability of heterogeneous treatment effect estimates across samples* (Working Paper). http://alexandercoppock.com/papers/CLM_heterogeneity.pdf
- Flood, S., King, M., Ruggles, S., & Warren, J. R. (2017). *Integrated public use microdata series, current population survey* [Version 5.0]. Minneapolis: University of Minnesota. Version 5.0. <https://doi.org/https://doi.org/10.18128/Do3o.V5.o>
- Krupnikov, Y., & Levine, A. S. (2014). Cross-sample comparisons and external validity. *Journal of Experimental Political Science*, 1(1), 59–80. <https://doi.org/10.1017/xps.2014.7>
- Kruschke, J. K. (2013). Bayesian estimation supersedes the t test. *Journal of Experimental Psychology*, 142(2), 573–603. <https://doi.org/10.1037/a0029146>
- Mullinix, K. J., Leeper, T. J., Druckman, J. N., & Freese, J. (2015). The generalizability of survey experiments. *Journal of Experimental Political Science*, 2(2), 109–38. <https://doi.org/10.1017/XPS.2015.19>
- R Core Team. (2016). *R: A language and environment for statistical computing* [Version 3.3.0]. Version 3.3.0. R Foundation for Statistical Computing. Vienna, Austria. <https://www.r-project.org/>
- Stan Development Team. (2016a). *RStan: The R interface to Stan* [2.11.0]. 2.11.0. <http://mc-stan.org>

Stan Development Team. (2016b). *Stan modeling language users guide and reference manual* [Version 2.11.0]. Version 2.11.0. <http://mc-stan.org>