Online Appendix:
Not for publication

Our supplementary material is structured as follows. Appendix A provides proofs of all theoretical results in Section 2. Appendix B provides supporting material for the experiments presented in Section 3. Appendix C provides supporting material for the experiments presented in Section 4. Appendix D discusses the ethical considerations underlying all experimental designs. Finally, Appendix E provides the instruments for all experiments described in the paper.

A Theoretical Results

A.1 Proof of Proposition 1

First, let us prove that for random variable t distributed with c.d.f. $H(\cdot)$ and p.d.f. $h(\cdot)$,

$$\frac{d}{d\tau}\mathbb{E}(t \mid t > \tau) \leq 1.$$

Let $z_\tau = t - \tau$ be a family of random variables indexed by τ; we need to show that

$$\mathbb{E}(z_\tau \mid z_\tau \geq 0)$$

is non-increasing in τ.

Denoting the c.d.f. of z_τ by $F_\tau(\cdot)$ and its p.d.f. by $f_\tau(\cdot)$, we have

$$\mathbb{E}(z_\tau \mid z_\tau \geq 0) = \frac{1}{1 - F_\tau(0)} \int_0^{+\infty} y f_\tau(y)\, dy.$$

The integral may be rewritten as

$$\int_0^{+\infty} y f_\tau(y)\, dy = \int_0^{+\infty} f_\tau(y) \left(\int_0^y 1 dx\right) dy = \int_0^{+\infty} \int_0^y f_\tau(y)\, dx dy$$

$$= \int_0^{+\infty} \int_x^{+\infty} f_\tau(y)\, dy dx = \int_0^{+\infty} (1 - F_\tau(x))\, dx,$$

where we used Fubini's theorem to change the order of integration.

Note that $F_\tau(x) = \Pr(z_\tau \leq x) = \Pr(t \leq x + \tau) = H(x + \tau)$. We therefore have

$$\mathbb{E}(z_\tau \mid z_\tau \geq 0) = \int_0^{+\infty} \frac{1 - F_\tau(x)}{1 - F_\tau(0)} dx = \int_0^{+\infty} \frac{1 - H(x + \tau)}{1 - H(\tau)} dx.$$

The integrand is non-increasing in τ pointwisely (i.e., for any fixed $x \geq 0$), because

$$\frac{d}{d\tau}\left(\frac{1 - H(x + \tau)}{1 - H(\tau)}\right) = \frac{h(\tau)(1 - H(x + \tau)) - h(x + \tau)(1 - H(\tau))}{(1 - H(\tau))^2}$$

$$= \frac{1 - H(x + \tau)}{1 - H(\tau)}\left(\frac{h(\tau)}{1 - H(\tau)} - \frac{h(x + \tau)}{1 - H(x + \tau)}\right) \leq 0, \tag{3}$$

because the first term is positive and the second is nonpositive due to monotone hazard rate property. This proves that $\mathbb{E}(z_\tau \mid z_\tau \geq 0)$ is non-increasing in τ, and thus $\frac{d}{d\tau}\mathbb{E}(t \mid t > \tau) \leq 1$.

Now, for any fixed social cost S, type t_i would choose $d_i = 1$ if $t_i > \frac{1}{\beta}S - w_0$ and would choose $d_i = 0$ if the opposite inequality holds. Thus, every equilibrium is characterized by a threshold τ. This threshold τ satisfies the condition

$$G(\tau) = -w_0, \tag{4}$$

where

$$G(\tau) = \tau - \frac{\gamma}{\beta}\mathbb{E}(t_i \mid t_i > \tau). \tag{5}$$

Since, as we proved, $\frac{d}{d\tau}\mathbb{E}(t_i \mid t_i > \tau) \le 1$ and $\gamma < \beta$, the $G(\tau)$ is strictly increasing in τ, and furthermore

$$\frac{d}{d\tau}G(\tau) \ge 1 - \frac{\gamma}{\beta} > 0.$$

This shows that the equation (4) has a unique solution. This completes the proof. ∎

A.2 Proof of Proposition 2

Since the distributions are normal, the posterior of citizen i is given by the usual formula

$$w_1 = \mathbb{E}(w \mid s) = w_0 \frac{\sigma_\varepsilon^2}{\sigma_w^2 + \sigma_\varepsilon^2} + s\frac{\sigma_w^2}{\sigma_w^2 + \sigma_\varepsilon^2}.$$

We have

$$w_1 - w_0 = \frac{\sigma_w^2}{\sigma_w^2 + \sigma_\varepsilon^2}(s - w_0),$$

so $w_1 > w_0$. From the proof of Proposition 1, the new equilibrium again takes the form of a threshold $\tilde{\tau}$ that satisfies

$$G(\tilde{\tau}) = -w_1,$$

where $G(\cdot)$ is defined in (5). Since $\frac{d}{d\tau}G(\tau) > 0$ and $-w_1 < -w_0$, we have $\tilde{\tau} < \tau$ (and furthermore, since $\frac{d}{d\tau}G(\tau) < 1$, the difference $\tau - \tilde{\tau} > w_1 - w_0$, so the decrease in threshold τ is larger than the increase in w). Now, $\tilde{\tau} < \tau$ implies that the share of citizens choosing $d_i = 1$ has increased: $1 - H(\tilde{\tau}) > 1 - H(\tau)$. Lastly, the social cost is now equal $\gamma\mathbb{E}(t_i \mid t_i > \tilde{\tau}) < \gamma\mathbb{E}(t_i \mid t_i > \tau)$, so it is lower than without the signal s. This completes the proof. ∎

A.3 Proof of Proposition 3

We start by establishing the uniqueness of equilibrium in this case.[31] Let $\bar{S}$ be the social cost of choosing $d_i = 1$ in a hypothetical equilibrium. Then the citizen would choose $d_i = 1$ if $t_i > \frac{1}{\beta}\bar{S} - w_h$ following signal s_h and if $t_i > \frac{1}{\beta}\bar{S} - w_l$ following signal s_l. This implies that there are two thresholds, τ_h and τ_l, that satisfy $\tau_l - \tau_h = w_h - w_l$. Denote $\bar{\tau} = \frac{1}{\beta}\bar{S} - w_0$; then $\tau_h = \bar{\tau} + w_0 - w_h$ and $\tau_l = \bar{\tau} + w_0 - w_l$. From now on we describe the equilibrium in terms of $\bar{\tau}$.

In what follows, we use the following probabilities. We denote

$$p(x, y) = \mu(1 - H(x)) + (1 - \mu)(1 - H(y)),$$

---

[31] Notice first that our assumption of rational expectation of t_i conditional on $d_i = 1$ allows us to bypass the discussion of whether members of the audience get signals s_l, s_h, or both. Rational expectation can be formed in practice if people had prior interactions with those who choose $d_i = 1$ and learned their type, which allows them to make a correct expectation in equilibrium about individuals who choose $d_i = 1$ with a given piece of evidence. An alternative way is to assume that the audience is sophisticated, understands the whole signal structure, but does not know which signal citizen i got, and faces the signal decomposition problem as a result.

so

$$p\left(\bar{\tau}+w_0-w_h,\bar{\tau}+w_0-w_l\right)=p\left(\frac{1}{\beta}\bar{S}-w_h,\frac{1}{\beta}\bar{S}-w_l\right)$$

is the probability of choosing $d_i=1$ if the citizen faces social cost $\bar{S}$. We also let

$$q\left(x,y\right)=\frac{\mu\left(1-H\left(x\right)\right)}{p\left(x,y\right)},$$

so $q\left(\bar{\tau}+w_0-w_h,\bar{\tau}+w_0-w_l\right)$ is the equilibrium conditional probability that citizen i got signal s_h conditional on choosing $d_i=1$.

Define the function

$$\begin{aligned}\bar{S}\left(z\right)&=&\gamma q\left(z+w_0-w_h,z+w_0-w_l\right)\mathbb{E}\left(t_i\mid t_i>z+w_0-w_h\right)\\&&+\gamma\left(1-q\left(z+w_0-w_h,z+w_0-w_l\right)\right)\mathbb{E}\left(t_i\mid t_i>z+w_0-w_l\right).\end{aligned}$$

In equilibrium characterized by $\bar{\tau}$, the social cost of choosing $d_i=1$ equals $\bar{S}\left(\bar{\tau}\right)$. Given the above, thresholds $\tau_h=\bar{\tau}+w_0-w_h$ and $\tau_l=\bar{\tau}+w_0-w_l$ are equilibrium thresholds for choosing $d_i=1$ after getting signals s_h and s_l, respectively, if and only if $\bar{\tau}$ solves the equation

$$\bar{\tau}-\frac{1}{\beta}\bar{S}\left(\bar{\tau}\right)=-w_0.\tag{6}$$

Let us show that $\frac{d}{dz}\frac{1}{\gamma}\bar{S}\left(z\right)\le 1$. Indeed, from the proof of Proposition 1, we have

$$\begin{aligned}\frac{d}{dz}\mathbb{E}\left(t_i\mid t_i>z+w_0-w_h\right)&\le&1;\\\frac{d}{dz}\mathbb{E}\left(t_i\mid t_i>z+w_0-w_l\right)&\le&1.\end{aligned}$$

Furthermore,

$$\mathbb{E}\left(t_i\mid t_i>z+w_0-w_l\right)>\mathbb{E}\left(t_i\mid t_i>z+w_0-w_h\right).$$

Lastly, we have

$$\begin{aligned}q\left(z+w_0-w_h,z+w_0-w_l\right)&=&\frac{\mu\left(1-H\left(z+w_0-w_h\right)\right)}{\mu\left(1-H\left(z+w_0-w_h\right)\right)+\left(1-\mu\right)\left(1-H\left(z+w_0-w_l\right)\right)}\\&=&\frac{1}{1+\frac{1-\mu}{\mu}\frac{1-H\left(z+w_0-w_l\right)}{1-H\left(z+w_0-w_h\right)}}.\end{aligned}$$

Now,

$$\frac{d}{dz}\frac{1-H\left(z+w_0-w_l\right)}{1-H\left(z+w_0-w_h\right)}=\frac{d}{du}\frac{1-H\left(u+\left(w_h-w_l\right)\right)}{1-H\left(u\right)}\le 0,$$

where we denoted $u=z+w_0-w_h$ and used the calculation (3) from the proof of Proposition 1.

This immediately implies that $\frac{d}{dz}q(z + w_0 - w_h, z + w_0 - w_l) \geq 0$. Summing up, we have

$$
\begin{aligned}
\frac{d}{dz}\frac{1}{\gamma}\bar{S}(z) &= q(z + w_0 - w_h, z + w_0 - w_l)\frac{d}{dz}\mathbb{E}(t_i \mid t_i > z + w_0 - w_h) \\
&\quad + (1 - q(z + w_0 - w_h, z + w_0 - w_l))\frac{d}{dz}\mathbb{E}(t_i \mid t_i > z + w_0 - w_l) \\
&\quad + \left(\frac{d}{dz}q(z + w_0 - w_h, z + w_0 - w_l)\right) \\
&\quad \times (\mathbb{E}(t_i \mid t_i > z + w_0 - w_h) - \mathbb{E}(t_i \mid t_i > z + w_0 - w_l)).
\end{aligned}
$$

Notice that the sum of the first two lines does not exceed 1 (since both derivatives do not exceed 1), and term on the third line is positive and the one on the fourth is negative, so their product is negative. This proves that $\frac{d}{dz}\frac{1}{\gamma}\bar{S}(z) \leq 1$. Now, as in the proof of Proposition 1 this implies that the equation (6) has a unique solution $\bar{\tau}$, which proves the uniqueness of equilibrium in this case.

Let us now show that in this solution, $\bar{\tau} < \tau$ and $\bar{S}(\bar{\tau}) < S(\tau)$, where $S(\tau) = \frac{1}{\gamma}\mathbb{E}(t_i \mid t_i > \tau)$ is the equilibrium social cost in the absence of any signal, in the unique solution τ. To do this, it is sufficient to show that $\bar{S}(\tau) < S(\tau)$. Indeed, this would imply that

$$
\tau - \frac{1}{\beta}\bar{S}(\tau) > \tau - \frac{1}{\beta}S(\tau) = -w_0,
$$

and since $\bar{\tau}$ satisfies (6) and the function $x - \frac{1}{\beta}\bar{S}(x)$ is increasing, we would get $\bar{\tau} < \tau$. Then we would get

$$
\bar{S}(\bar{\tau}) = \beta(\bar{\tau} + w_0) < \beta(\tau + w_0) = S(\tau),
$$

as required. So, to complete the proof, we need to show that $\bar{S}(\tau) < S(\tau)$.

In the light of condition (2) and by continuity of $H(\cdot)$, there exists $\hat{w}_h \in (0, w_h)$ such that

$$
\mu(H(\tau) - H(\tau - (\hat{w}_h - w_0))) = (1 - \mu)(H(\tau + (w_0 - w_l)) - H(\tau)).
$$

Let $\hat{S}$ denote the value

$$
\begin{aligned}
\hat{S} &= \gamma q(\tau + w_0 - \hat{w}_h, \tau + w_0 - w_l)\mathbb{E}(t_i \mid t_i > \tau + w_0 - \hat{w}_h) \\
&\quad + \gamma(1 - q(\tau + w_0 - \hat{w}_h, \tau + w_0 - w_l))\mathbb{E}(t_i \mid t_i > \tau + w_0 - w_l);
\end{aligned}
$$

in other words, the expression for $\hat{S}$ is analogous to $\bar{S}(\tau)$, except that w_h is replaced by $\hat{w}_h$.

We now show that $\bar{S}(\tau) < \hat{S} < S(\tau)$. To prove the first inequality, we use some algebra to establish that

$$
\frac{1}{\gamma}\bar{S}(\tau) = (1 - \rho)\frac{1}{\gamma}\hat{S} + \rho\mathbb{E}(t_i \mid t_i \in (\tau + w_0 - w_h, \tau + w_0 - \hat{w}_h)),
$$

where

$$
\rho = q(\tau + w_0 - w_h, \tau + w_0 - w_l)\frac{H(\tau + w_0 - \hat{w}_h) - H(\tau + w_0 - w_h)}{1 - H(\tau + w_0 - w_h)}.
$$

Since $\rho > 0$ and $\frac{1}{\gamma}\hat{S} < \mathbb{E}(t_i \mid t_i \in (\tau + w_0 - w_h, \tau + w_0 - \hat{w}_h))$ as the former is an expectation taken over values to the right of $\tau + w_0 - \hat{w}_h$ while the latter expectation is taken over values to the left

of that point, we get $\bar{S}(\tau) < \hat{S}$.

Let us now prove that $\hat{S} < S(\tau)$. Spelling out $q(\tau + w_0 - \hat{w}_h, \tau + w_0 - w_l)$ and expectations in the definition of $\hat{S}$, we have

$$\frac{1}{\gamma}\left(S(\tau) - \hat{S}\right) = \frac{\int_\tau^\infty xh(x)\,dx}{1 - H(\tau)}$$
$$-\frac{\mu \int_{\tau + w_0 - \hat{w}_h}^\infty xh(x)\,dx + (1-\mu)\int_{\tau + w_0 - w_l}^\infty xh(x)\,dx}{\mu(1 - H(\tau + w_0 - \hat{w}_h)) + (1-\mu)(1 - H(\tau + w_0 - w_l))}.$$

Notice that by the definition of $\hat{w}_h$ the denominators in both terms are equal, hence $S(\tau) - \hat{S}$ has the same sign as

$$\int_\tau^\infty xh(x)\,dx - \left(\mu \int_{\tau + w_0 - \hat{w}_h}^\infty xh(x)\,dx + (1-\mu)\int_{\tau + w_0 - w_l}^\infty xh(x)\,dx\right)$$
$$= (1-\mu)\int_\tau^{\tau + w_0 - w_l} xh(x)\,dx - \mu \int_{\tau + w_0 - \hat{w}_h}^\tau xh(x)\,dx$$
$$= (1-\mu)(H(\tau + w_0 - w_l) - H(\tau))\,\mathbb{E}(t_i \mid t_i \in (\tau, \tau + w_0 - w_l))$$
$$-\mu(H(\tau) - H(\tau + w_0 - \hat{w}_h))\,\mathbb{E}(t_i \mid t_i \in (\tau + w_0 - \hat{w}_h, \tau)).$$

Since the coefficients in front of the expectations in the last two lines are the same (again, by the choice of $\hat{w}_h$), the sign of this expression is the same as the sign of

$$\mathbb{E}(t_i \mid t_i \in (\tau, \tau + w_0 - w_l)) - \mathbb{E}(t_i \mid t_i \in (\tau + w_0 - \hat{w}_h, \tau)),$$

which is positive, because the first term is greater than τ and the second is less than that. Therefore, $\hat{S} < S(\tau)$.

We have thus proved that $\bar{S}(\tau) < \hat{S} < S(\tau)$ which, as we showed earlier, implies the results stated. This completes the proof. ∎

Table B1: Overview of Data Collections

Experiment	Provider	Dates
Panel A: Main Experiments		
Experiment 1: Willingness to post anti-defunding Tweet – Democrats authorizing Twitter access (N=1,122)	Luc.id, Cloudresearch	October 2021
Experiment 2: Interpretation of anti-defunding Tweet – Democrats (N=1,040)	Prolific	November 2021
Experiment 3: Willingness to post pro-deportation Tweet – Republicans authorizing Twitter access (N=1,130)	Luc.id	March 2021
Experiment 4: Interpretation of pro-deportation Tweet – Republicans (N=1,082)	Prolific	November 2021
Panel B: Auxiliary Experiments		
Auxiliary Experiment 1: Persuasiveness of anti-defunding article – Democrats (N=1,008)	Prolific	December 2021
Auxiliary Experiment 2: Placebo: willingness to post pro-conservation Tweet – respondents authorizing Twitter access (N=483)	Luc.id, Cloudresearch	December 2021 and January 2022
Auxiliary Experiment 3: Anticipated persuasiveness of anti-defunding Tweet – Democrats (N=501)	Prolific	November 2021
Auxiliary Experiment 4: Motives underlying the choice – Democrats with Twitter account (N=400)	Prolific	January 2022
Auxiliary Experiment 5: Interpretation of lower-credibility anti-defunding Tweet – Democrats (N=506)	Prolific	November 2021
Auxiliary Experiment 6: Persuasiveness of pro-deportation Tweet – Republicans (N=2,012)	Prolific, Lucid	December 2021
Auxiliary Experiment 7 (wave 1): Willingness to donate to anti-immigrant organization – Conservatives (N=4,457)	Luc.id	January 2020
Auxiliary Experiment 7 (wave 2): Willingness to donate to anti-immigrant organization – Conservatives (N=1,299)	Luc.id	September 2020
Auxiliary Experiment 8: Interpretation of anti-immigrant donation – Liberals (N=3,047)	Luc.id	February 2020

Notes: Reported sample sizes for Experiment 1, Experiment 3, and Auxiliary Experiment 2 include respondents who chose not to join the campaigns and therefore are not included in the sample we analyze.

B   Anti-Defunding Experiments: Additional Material

B.1   Experiment 1: Additional Figures and Tables

Table B2: Experiment 1: Balance of covariates

	Overall		Cover	No Cover	p-value
	mean	std.dev.	mean	mean	(R=NR)
Age	39.877	15.130	39.293	40.514	0.355
Black	0.214	0.410	0.228	0.198	0.391
Asian	0.076	0.265	0.072	0.079	0.775
White	0.671	0.470	0.670	0.672	0.968
Hispanic	0.183	0.387	0.159	0.209	0.138
Male	0.580	0.494	0.572	0.589	0.702
High school diploma	0.975	0.155	0.975	0.976	0.903
Bachelors degree	0.435	0.496	0.417	0.455	0.381

Notes: p-values based on robust standard errors reported.

B.2   Auxiliary Experiment 1: Persuasiveness of Defunding Rationale

We conducted this pre-registered experiment in December 2021 with a sample of 1,008 Democrats and Independents recruited from Prolific.[32] After completing a set of demographic questions, respondents assigned to the treatment group read Sharkey's article in the *Washington Post*, while respondents assigned to the control group did not read the article. They then respond to the following two questions: "Do you think that funding for the police should be increased, decreased, or stay the same?" and "How do you think increasing funding for the police would affect violent crime?". We code both questions from -2 ("Decreased a lot" and "Strongly decrease violent crime", respectively) to 2 ("Increased a lot" and "Strongly decrease violent crime", respectively).

Table B3 displays results, with Columns 1–3 corresponding to the first measure and Columns 4–6 corresponding to the second measure. We find a significant effect on both measures, with an effect size of around 0.25 standard deviations for the first outcome and 0.12 standard deviations for the second outcome.

B.3   Auxiliary Experiment 2: Rainforest Placebo

B.3.1   Design and results

We conducted this experiment in December 2021 and January 2022 with a sample of 483 Democrats and Independents recruited from Luc.id and CloudResearch. Respondents logged in to the survey with their Twitter accounts using the same procedure as in Experiment 1. The design is similar to that of Experiment 1, but examines a different (non-stigmatized) context: willingness to post

---

[32]The pre-registration is available in the AEA RCT registry under ID AEARCTR-0008624.

Table B3: Persuasive effects of anti-defunding article

	Belief			Policy preference		
	(1)	(2)	(3)	(4)	(5)	(6)
Provided article	−0.236***	−0.245***	−0.245***	0.135*	0.121*	0.122*
	(0.056)	(0.055)	(0.055)	(0.071)	(0.068)	(0.068)
DV mean	-0.080	-0.080	-0.080	-0.568	-0.568	-0.568
DV std. dev.	0.902	0.902	0.902	1.133	1.133	1.133
Observations	1,008	1,007	1,007	1,008	1,007	1,007
Demographic controls	No	Yes	Yes	No	Yes	Yes
Partisan controls	No	No	Yes	No	No	Yes

Notes: The dependent variable in Columns 1–3 is the respondent's reported belief as to the effect of increasing funding for the police on violence crime, coded between -2 ("Strongly decrease violent crime") and 2 ("Strongly increase violent crime"). The dependent variable in Columns 4–6 is the respondent's reported preference for changing police funding, ranging from -2 ("Decreased a lot") to 2 ("Increased a lot"). Demographic controls include age, age squared, a set of race indicators, a Hispanic indicator, a male indicator, a set of education indicators. Partisan controls include indicators for "Very conservative", "Conservative", "Neither liberal nor conservative" (omitted), "Liberal", and "Very liberal". Robust standard errors are reported.

a Tweet supporting efforts to conserve the Amazon rainforest.[33] Rather than reading an article about the likely effects of defunding the police, respondents read a Reuters article reporting on a study conducted by the Science Panel for the Amazon which finds that over 10,000 species are at risk from deforestation in the Amazon (Science Panel for the Amazon, 2021). The *Cover* Tweet reads:

> I've joined a campaign to immediately stop the destruction of the Amazon rainforest! Before I joined the campaign, I was shown this article about how 10,000 species risk extinction in Amazon: [LINK]. Join the campaign and sign the petition: [LINK].

The *No Cover* Tweet is identical, but replaces "Before I joined the campaign..." with "After I joined the campaign...".

Appendix Table B5 shows no significant difference between posting rates in the *Cover* and *No Cover* conditions, and the difference in effect sizes between the defunding experiment and the placebo experiment is large in magnitude (16 percentage points) and significant at the 5% level, suggesting effects are indeed driven by (anticipated) changes in the stigma associated with dissenting expression rather than some other independent effect of the wording.

B.4 Auxiliary Experiment 3: Anticipated Persuasion Experiment

We conducted this experiment in November 2021 with a sample of 501 Democrats and Independents recruited from Prolific. Only Democrats and Independents with Twitter accounts were eligible to

---

[33]Table B4 shows that our sample is balanced on observables across treatment arms.

Table B4: Rainforest placebo: Balance of covariates

	Overall		Cover	No Cover	p-value
	mean	std.dev.	mean	mean	(R=NR)
Age	39.110	13.244	38.146	40.043	0.201
Black	0.135	0.342	0.172	0.099	0.056
Asian	0.047	0.212	0.025	0.068	0.074
White	0.768	0.423	0.739	0.796	0.226
Hispanic	0.163	0.370	0.166	0.160	0.902
Male	0.473	0.500	0.529	0.420	0.052
High school diploma	0.984	0.124	0.981	0.988	0.628
Bachelors degree	0.392	0.489	0.376	0.407	0.565

Notes: p-values based on robust standard errors reported.

Table B5: Willingness to post pro-conservation Tweet (placebo)

	Dependent variable:		
	Pro-conservation Tweet		
	(1)	(2)	(3)
Cover	−0.044	−0.049	−0.052
	(0.044)	(0.044)	(0.045)
DV mean	0.812	0.812	0.812
DV std. dev.	0.391	0.391	0.391
Observations	319	319	319
Demographic controls	No	Yes	Yes
Partisan controls	No	No	Yes

Notes: The dependent variable is an indicator taking value 1 if the respondent chose to schedule the post. Demographic controls include age, age squared, a set of race indicators, a Hispanic indicator, a male indicator, a set of education indicators. Partisan controls include indicators for "Very conservative", "Conservative", "Neither liberal nor conservative" (omitted), "Liberal", and "Very liberal". Robust standard errors are reported.

take the survey. After completing a set of demographic questions, respondents read Sharkey's article in the *Washington Post*. As in Experiment 1, respondents are asked if they would like to join the campaign to oppose the movement to defund the police, only those who indicate that they would like to join the campaign proceed with the experiment, and those who do proceed are given a chance to re-read the article. They are then randomly shown either the *Cover* or the *No Cover* Tweet from Experiment 1 and are asked: "Suppose you posted the Tweet above on your

account. If you had to guess, what percentage of people who saw your Tweet would choose to join the campaign to oppose defunding the police?"

Table B6 displays results. Reassuringly, we find no significant difference between the anticipated persuasiveness of the Tweets, suggesting that differential posting rates are instead driven by changes in anticipated stigma.

Table B6: Anticipated persuasiveness of Tweet

	Dependent variable:		
	Perceived percentage persuaded		
	(1)	(2)	(3)
Cover	1.897	2.190	2.504
	(2.125)	(2.114)	(2.123)
DV mean	26.309	26.309	26.309
DV std. dev.	23.764	23.764	23.764
Observations	501	501	501
Demographic controls	No	Yes	Yes
Partisan controls	No	No	Yes

Notes: The dependent variable is the respondent's guess as to the percentage of their followers who would join the campaign if they saw the Tweet. Demographic controls include age, age squared, a set of race indicators, a Hispanic indicator, a male indicator, a set of education indicators. Partisan controls include indicators for "Very conservative", "Conservative", "Neither liberal nor conservative" (omitted), "Liberal", and "Very liberal". Robust standard errors are reported.
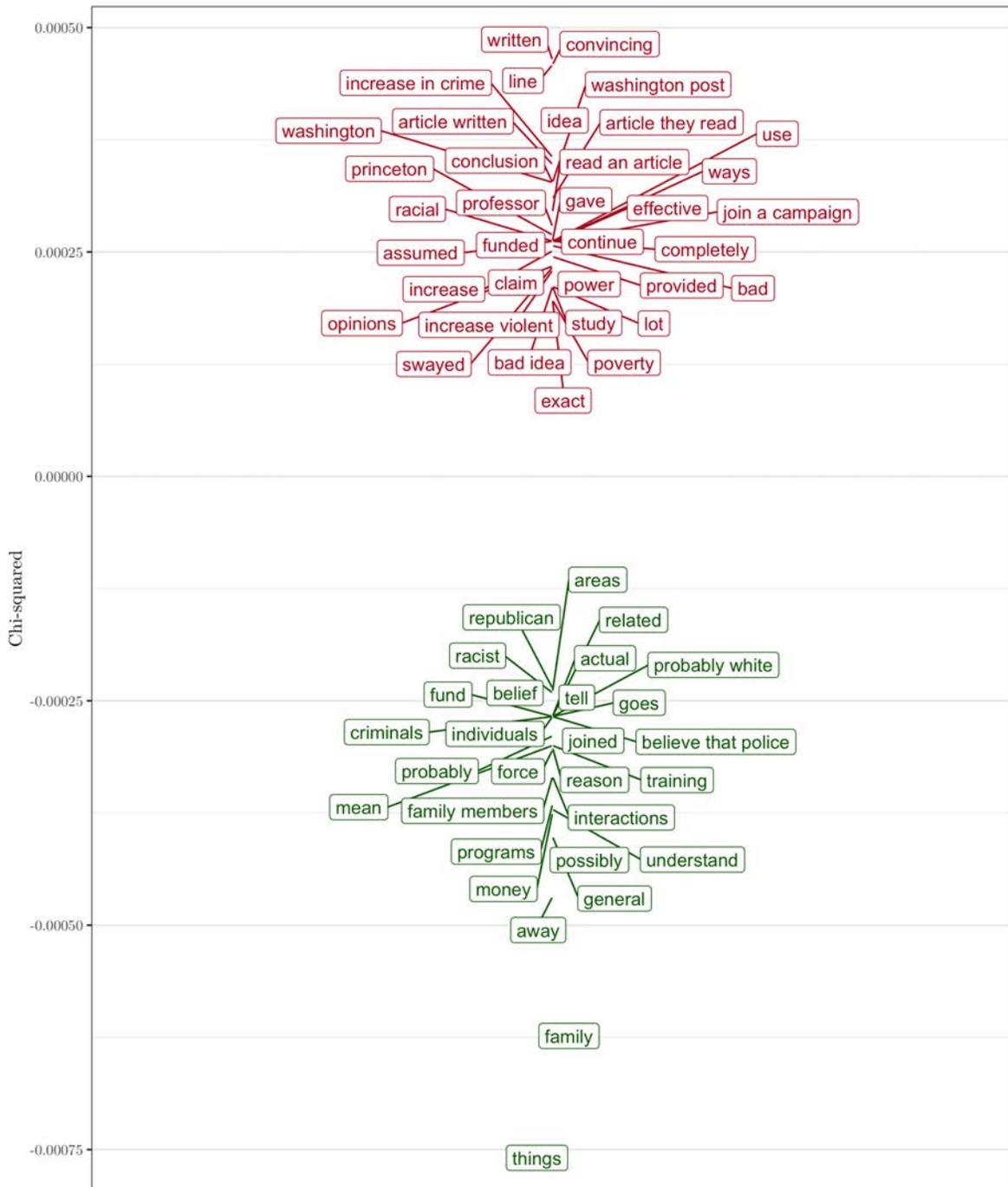
B.5 Experiment 2: Additional Figures and Tables

Table B7: Experiment 2: Balance of covariates

	Overall		Cover	No Cover	p-value
	mean	std.dev.	mean	mean	(R=NR)
Age	30.725	11.258	30.686	30.763	0.912
Black	0.070	0.256	0.086	0.055	0.057
Asian	0.085	0.279	0.089	0.080	0.589
White	0.773	0.419	0.766	0.781	0.563
Hispanic	0.112	0.315	0.093	0.130	0.060
Male	0.374	0.484	0.384	0.365	0.522
High school diploma	0.997	0.054	0.996	0.998	0.552
Bachelors degree	0.572	0.495	0.562	0.582	0.520

Notes: p-values based on robust standard errors reported.

Figure B1: Experiment 2: most distinctive phrases in each condition

Notes: Appendix Figure B1 plots phrases by their associated χ^2 statistic, limiting to the top 100 phrases and multiplying the χ^2 of phrases more characteristic of the "No Cover" condition by -1. The words "article" and "read" have χ^2 values of greater than 0.001 and have been suppressed to facilitate visualization of the remaining phrases.

Table B8: Experiment 2 (lower-credibility): Balance of covariates

| | Overall | | Cover | No Cover | p-value |
	mean	std.dev.	mean	mean	(R=NR)
Age	35.366	14.585	35.275	35.458	0.888
Black	0.053	0.225	0.043	0.064	0.303
Asian	0.132	0.339	0.137	0.127	0.747
White	0.771	0.421	0.773	0.769	0.923
Hispanic	0.107	0.309	0.141	0.072	0.011
Male	0.496	0.500	0.502	0.490	0.789
High school diploma	0.996	0.063	0.992	1.000	0.160
Bachelors degree	0.597	0.491	0.600	0.594	0.884

Notes: p-values based on robust standard errors reported.

C Anti-Immigrant Experiments: Additional Material

C.1 Experiment 3: Additional Tables

Table C1: Experiment 3: Balance of covariates

	Overall		Cover	No Cover	p-value
	mean	std.dev.	mean	mean	(R=NR)
Age	49.422	13.623	48.709	50.094	0.248
Black	0.014	0.116	0.012	0.015	0.762
Asian	0.015	0.124	0.016	0.015	0.934
White	0.952	0.215	0.952	0.951	0.955
Hispanic	0.066	0.248	0.052	0.079	0.214
Male	0.503	0.500	0.490	0.515	0.571
High school diploma	0.994	0.076	0.996	0.992	0.598
Bachelors degree	0.385	0.487	0.343	0.425	0.055

Notes: *p*-values based on robust standard errors reported.

C.2 Auxiliary Experiment 6: Persuasiveness of Deportation Rationale

We conducted a first pre-registered experiment in December 2021 with a sample of 1,008 Republicans recruited from Prolific.[34] After completing a set of demographic questions, respondents assigned to the treatment group viewed the clip from *Tucker Carlson Tonight*, while respondents assigned to the control group did not view the clip. They then indicated their agreement with the following two statements: "Illegal immigrants are not much more likely to commit serious crimes than U.S. citizens" and "The US should immediately deport all illegal Mexican immigrants." We code both questions from -2 ("Strongly disagree") to 2 ("Strongly agree").

Contrary to our pre-registered prediction, we did not estimate a statistically significant effect of viewing the video on either outcome. Two logistical problems complicate interpretation of this result. First, when setting up the survey, we forgot to exclude respondents from some previous experiments which included the video. Thus, some respondents in the *Control* condition had seen the video in previous experiments. Second, there was a highly limited sample of Republicans available on Prolific (fewer than 2000 who met our screening criteria), and we had to pay a higher than usual rate in order to meet our pre-registered sample size. This potentially induced selection into the survey.

We thus ran the same experiment on Luc.id, with the same sample restrictions. Table C2 displays results, with Columns 1–3 corresponding to the first measure and Columns 4–6 corresponding to the second measure. We find a significant effect on both measures, with an effect size of around 0.12 standard deviations for the first outcome and 0.18 standard deviations for the second outcome.

Overall, we take the evidence for the effects of the clip on persuasion as mixed.

---

[34]The pre-registration is available in the AEA RCT registry under ID AEARCTR-0008624.

Table C2: Persuasive effects of *Tucker Carlson Tonight* video

	Dependent variable:					
	Belief			Policy preference		
	(1)	(2)	(3)	(4)	(5)	(6)
Provided article	−0.133*	−0.123*	−0.123*	0.177**	0.179**	0.180**
	(0.070)	(0.071)	(0.071)	(0.074)	(0.073)	(0.073)
DV mean	0.643	0.643	0.643	0.740	0.740	0.740
DV std. dev.	1.112	1.112	1.112	1.173	1.173	1.173
Observations	1,004	1,002	1,002	1,004	1,002	1,002
Demographic controls	No	Yes	Yes	No	Yes	Yes
Partisan controls	No	No	Yes	No	No	Yes

Notes: The dependent variable in Columns 1–3 is the respondent's reported agreement with the statement "Illegal immigrants are more likely to commit serious crimes than US citizens," coded between -2 ("Strongly disagree") and 2 ("Strongly agree"). The dependent variable in Columns 4–6 is the respondent's reported agreement with the statement "The US should immediately deport all illegal Mexican immigrants," ranging from -2 ("Strongly disagree") to 2 ("Strongly agree"). Demographic controls include age, age squared, a set of race indicators, a Hispanic indicator, a male indicator, a set of education indicators. Partisan controls include indicators for "Very conservative", "Conservative", "Neither liberal nor conservative" (omitted), "Liberal", and "Very liberal". Robust standard errors are reported.

C.3 Experiment 4: Additional Tables

Table C3: Experiment 4: Balance of covariates

	Overall		Cover	No Cover	p-value
	mean	std.dev.	mean	mean	(R=NR)
Age	31.729	12.256	32.408	31.046	0.068
Black	0.069	0.254	0.063	0.076	0.389
Asian	0.100	0.300	0.090	0.109	0.297
White	0.767	0.423	0.785	0.750	0.174
Hispanic	0.118	0.323	0.109	0.128	0.325
Male	0.479	0.500	0.492	0.466	0.392
High school diploma	0.995	0.068	0.994	0.996	0.659
Bachelors degree	0.589	0.492	0.590	0.588	0.939

Notes: p-values based on robust standard errors reported.

Figure C1: Experiment 4: most distinctive phrases in each condition

Notes: Appendix Figure C1 plots phrases by their associated χ^2 statistic, limiting to the top 100 phrases and multiplying the χ^2 of phrases more characteristic of the "No Cover" condition by -1.

C.4 Auxiliary Experiment 7: Anti-Immigrant Expression Among More Representative Sample

The social media setting of Experiments 1 and 3 affords a highly natural setting and real-stakes outcome — and is of itself a context of policy relevance — but there are two potential concerns about external validity. First, Twitter users still comprise a relatively small and selected fraction of the population, particularly among Republicans (Wojcik and Hughes, 2019). Second, our requirement that respondents grant our "Tweetability" app permissions to schedule posts on their Twitter account likely induces selection into our experiment. While this selection does not affect the internal validity of Experiments 1 and 3, it might affect the extent to which the results generalize to the broader population. A third, and related, limitation of these experiments is that we are unable to examine heterogeneity in the effect of the rationale based on the composition of a respondent's audience, both because we cannot observe their audience and because we are insufficiently powered to do so. To address these concerns, this section presents an additional experiment that sacrifices some of the naturalness of Experiments 1 and 3 for a large and representative sample (whose geographic location and thus whose local environment we can observe), while retaining a revealed-preference measure of respondents' willingness to publicly express dissent. We discuss ethical considerations in Appendix D.

C.4.1 Sample and experimental design

Sample composition We conducted Auxiliary Experiment 7 in January 2020 (wave 1) and September 2020 (wave 2) with a sample of Republicans and Independents recruited through Luc.id.[35] Our sample of respondents is broadly representative of Independents and Republicans in the United States (Appendix Table C7) and is well-balanced on observables across treatment arms (Appendix Table C6). The two waves had a largely similar design. The most important differences are that wave 1 included a pure control condition and had a more heavy-handed set of instructions. In the description of the design below, we focus of the leaner set of instructions from wave 2.

Information: Lott study After completing a series of demographic and other background questions, respondents are assigned to two main conditions: *No Cover* and *Cover*.

Specifically, all respondents are first told about the preliminary findings of an unpublished study (Lott, 2018) claiming that immigrants commit more crime than US citizens. Respondents are informed that they will have the opportunity to authorize a $1 donation to Fund The Wall, an organization seeking to construct the proposed US–Mexico border wall, and that we will post their individual donation decision on our website. To vary the availability of a social cover, we tell respondents assigned to the *No Cover* treatment that the web page will state that "all participants were surveyed before Dr. Lott's study was published in an academic journal." In the *Cover* treatment, respondents are instead told that the web page will state that "all participants were shown the the preliminary findings from Dr. Lott's study before deciding whether or not to donate to Fund The Wall." We tell respondents in the *Cover* and *No Cover* conditions about a recent study (Lott, 2018) which argues that undocumented immigrants commit more crimes and more serious crimes than US citizens.[36] This study has been widely covered by the media, including

---

[35]Wave 1 of the experiment was pre-registered in the AEA RCT registry under ID AEARCTR-0005308.

[36]Wave 1 of study also had a pure *Control* condition in which respondents do not learn about the study. We include a discussion of the results from the *Control* condition below.

The Washington Times, *National Review*, and *Fox News*, and has been repeatedly cited by Trump administration officials. We also truthfully tell our respondents that a number of sources (including a researcher affiliated with the Cato Institute, a libertarian think tank) have recently challenged some of the study's methods, claiming that errors in analysis invalidate its results.

Visible donation decisions We ask respondents to consent to us accessing their name, city, and operating system from the survey provider (which confirmed that they would provide us with this data subject to participant consent) and give respondents the option to terminate the survey if they do not consent. Respondents are informed that they will have the opportunity to authorize a $1 donation to Fund The Wall, an organization seeking to construct the proposed US–Mexico border wall, and that we will post the results from the survey, including their individual donation decision, on our study website.[37] To vary the availability of a social cover, we tell respondents assigned to the *No Cover* treatment that the web page will state that "all participants were surveyed before Dr. Lott's study was published in an academic journal." In the *Cover* treatment, respondents are instead told that the web page will state that "all participants were shown the the preliminary findings from Dr. Lott's study before deciding whether or not to donate to Fund The Wall."
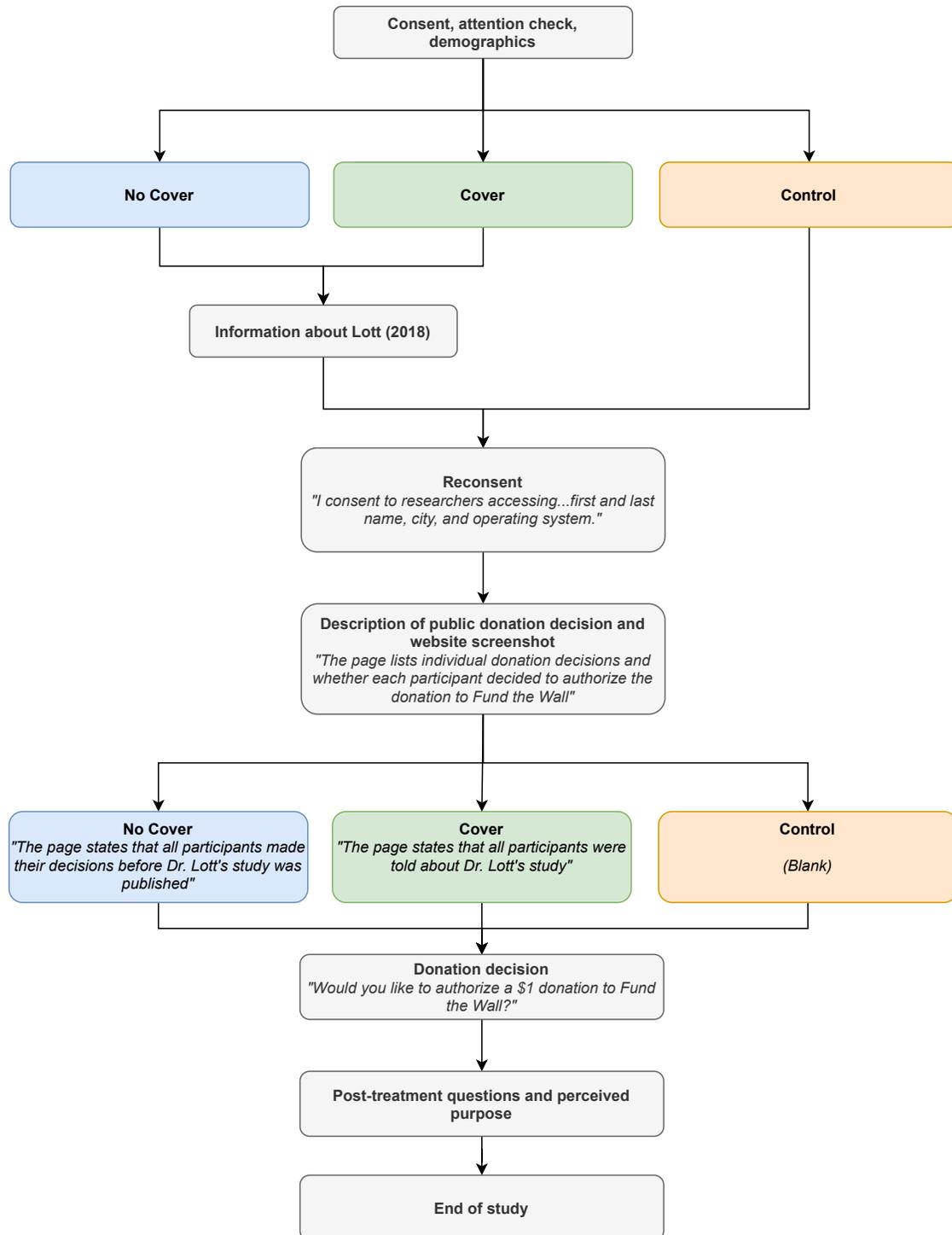
We also inform our respondents that "As researchers, we believe it is important to communicate our findings about political and social attitudes in [City of respondent] to the public".[38] We then inform our respondents that "We will promote our website via Facebook ads to [City of respondent] residents". This generates a plausible social cost for acting in a way that will be stigmatized in the respondent's area. After informing respondents about the content of the website, we ask people whether they "would like to authorize a $ donation to **Fund The Wall**?"

---

[37]To minimize experimenter demand concerns, we in fact informed our respondents that we would randomly select one of two organizations — *Fund the Wall*, or the *Texas Civil Rights Project*, an organization that (among other activities) worked to legally challenge the wall's construction — and that they would have the opportunity to authorize a $1 donation to this organization. In practice, we randomized almost all respondents to Fund the Wall to maximize statistical power for our comparison of interest.

[38]We used participants' IP address to capture and display their current location (i.e. their city).

Figure C2: Auxiliary Experiment 7: Design

Consent, attention check, demographics

No Cover

Cover

Control

Information about Lott (2018)

Reconsent
"I consent to researchers accessing...first and last name, city, and operating system."

Description of public donation decision and website screenshot
"The page lists individual donation decisions and whether each participant decided to authorize the donation to Fund the Wall"

No Cover
"The page states that all participants made their decisions before Dr. Lott's study was published"

Cover
"The page states that all participants were told about Dr. Lott's study"

Control
(Blank)

Donation decision
"Would you like to authorize a $1 donation to Fund the Wall?"

Post-treatment questions and perceived purpose

End of study

C.4.2 Results

Average treatment effects   To identify the joint effects of direct persuasion and anticipated persuasion of the audience (i.e. the direct persuasive effect of learning about the Lott study in addition to the indirect effect of learning that one's audience has learned about the Lott study and may thus be more likely to approve of the donation), we compare the *Control* condition with the *No Cover* condition. To identify the cover effect of rationales, we compare the *No Cover* condition to the *Cover* condition. This design thus allows us to benchmark the cover effect of rationales effect against the joint effect of persuasion and anticipated persuasion.

Table C4 displays mean donation rates by condition. We find a statistically significant effect on respondents' willingness to authorize a donation to Fund the Wall: respondents in the *Cover* condition are 7 percentage points more likely to authorize the donation than respondents in the *No Cover* condition (as shown in Column 6 of Table C4 which pools observations across all conditions). As shown in Panel A of Table C4, effect sizes are almost identical in our pre-specified main study, our pilot study, and a replication several months later. In contrast to the *Cover* vs. *No Cover* comparison, respondents in the *No Cover* condition are only 1 percentage point more likely to authorize a donation than respondents in the *Control* condition, suggesting that the combined effects of persuasion and anticipated persuasion are small. Relatively small persuasion effects are in line with other information provision experiments in the immigration domain, which typically find relatively small or null effects on behavior and stated preferences (Alesina et al., 2019; Hopkins et al., 2019; Grigorieff et al., 2020; Haaland and Roth, 2020), and are also consistent with the mixed evidence on persuasion we find in Experiments 3 and 4.

Heterogeneity by local vote shares   The audience's composition — in this case, the fraction who approve of the decision to donate — should affect the magnitude of the rationale effect. Because we informed respondents that we would promote the website within their geographical area, we might expect that, controlling for respondents' own private views, respondents in areas with a greater fraction of Republicans (who are likely to approve of the decision to donate to Fund the Wall even in the absence of a rationale) should be less sensitive to the availability of a rationale than those in areas with a lower fraction of Republicans. We thus pre-registered investigating heterogeneity by the 2016 Republican vote share of the respondents' county, which we do by interacting our treatment indicators with vote shares (standardized, for ease of interpretation). We flexibly control for differential effect of partisanship by also controlling for the interactions between partisanship and the 2016 Republican vote share. Panel B of Table C4 displays the results, revealing striking heterogeneity by the Republican vote share of respondents' counties. The estimated interaction is large in magnitude: a one standard deviation increase in the Republican vote share of a respondent's county is associated with halving the magnitude of the cover effect of rationales. Of course, these results cannot be interpreted as a causal effect of differences in the composition of respondents' audiences: it may be, for example, that Republicans in Democratic areas feel a greater need to signal their support for the study by publicly donating.

C.4.3 Addressing experimenter demand concerns

One concern is that our instructions in Experiment 7, by directly informing (and reminding) respondents about their audience's information sets, induced experimenter demand effects or otherwise compromised external validity. *A priori*, it seems plausible that showing participants information

Table C4: Willingness to donate to anti-immigrant cause

	Dependent variable:					
	Donated to Fund the Wall					
	(1)	(2)	(3)	(4)	(5)	(6)
Panel A:	*Non-interacted specification*					
Cover	0.063***	0.065***	0.074***	0.066**	0.053**	0.070***
	(0.020)	(0.018)	(0.016)	(0.028)	(0.026)	(0.014)
Control	−0.001	−0.005	−0.001			0.018
	(0.020)	(0.018)	(0.017)			(0.016)
p-value (Cover = Control)	0.0013	< 0.001	< 0.001			0.0012
Panel B:	*Interacted specification*					
Cover	0.061***	0.063***	0.071***	0.042	0.026	0.063***
	(0.020)	(0.019)	(0.016)	(0.030)	(0.028)	(0.014)
Cover × Rep vote share	−0.030	−0.038**	−0.038**	−0.062**	−0.073***	−0.042***
	(0.020)	(0.019)	(0.017)	(0.028)	(0.027)	(0.014)
Control	−0.001	−0.004	−0.001			0.017
	(0.020)	(0.018)	(0.017)			(0.016)
Control × Rep vote share	0.013	0.010	0.010			0.015
	(0.020)	(0.018)	(0.017)			(0.016)
Rep vote share	0.052***	0.0002	−0.0004	0.020	0.019	0.001
	(0.015)	(0.018)	(0.016)	(0.019)	(0.035)	(0.015)
Waves included	1	1	P + 1	2	2	P + 1 + 2
DV mean	0.488	0.488	0.497	0.484	0.484	0.494
DV std. dev.	0.500	0.500	0.500	0.500	0.500	0.500
Observations	3,751	3,751	4,457	1,279	1,279	5,736
Individual controls	No	Yes	Yes	No	Yes	Yes

Notes: The dependent variable is an indicator taking value 1 if the respondent donated to Fund the Wall. Columns 1–2 report results estimated on the sample from the Wave 1; Column 3 pools the sample from the Wave 1 with the sample from the pilot; Columns 4–5 consider only the sample from the Wave 2; and Column 6 pools all waves. In Panel B, the county Republican vote share is from the 2016 US Presidential election and is scaled to a standard normal distribution. The specifications in Columns 2–3 and Columns 5–6 of Panel B include the interactions between our set of partisan controls and the 2016 Republican vote share. Demographic controls include age, age squared, a set of race indicators, a Hispanic indicator, a male indicator, a set of education indicators. Partisan controls include indicators for "Very conservative", "Conservative", "Neither liberal nor conservative" (omitted), "Liberal", and "Very liberal". Robust standard errors are reported.

about the Lott study might induce demand effects and thus affect donation rates, but even if such demand effects are present, they do not bias our main comparison of interest (*Cover* vs. *No Cover*), given that participants in both treatment arms are shown identical information about the study. More concerning is the possibility the treatment manipulation of beliefs about the audience's information set induced differential experimenter demand effects between these two conditions.

Hand-coding of perceived purpose We also ask respondents to respond in open-ended form to the question "If you had to guess, what would you say the purpose of this study was?" We use responses to assess demand effects in two ways. First, we hired two independent research assistants to hand-code the responses. Appendix Table C5 in the Appendix shows that the majority of our respondents believed that we wanted to study the effects of information on anti-immigrant sentiment or participant's willingness to have their decisions posted on the website. Fewer than 1 percent of our sample correctly guessed the true purpose of our experiment (Column 1). The table also shows that on almost all of the dimensions we code, beliefs about the purpose of the study do not significantly differ between the *Cover* and *No Cover* conditions. The exception is Public Image (Column 3): respondents in the *Cover* condition are 2 percentage points more likely than respondents in the *No Cover* condition to believe that the study was about whether people were willing to publicly express political views. Although statistically significant, this difference is small in magnitude and cannot explain our effect sizes. We do find significant differences in perceived purpose between the *Control* condition the other two conditions, likely because we provided respondents in the *No Cover* and *Cover* conditions, but not in the *Control* condition, information about the Lott study. However, these differences do not affect our main comparison of interest (*No Cover* vs. *Cover*).

Table C5: Anti-immigrant donations: Perceived purpose of study

	Cover (1)	Immigration attitudes (2)	Public image (3)	Information (4)	Persuasion (5)	Biased (6)
			Dependent variable:			
Cover	−0.005	0.010	0.019**	0.011	−0.012	−0.0004
	(0.003)	(0.015)	(0.010)	(0.015)	(0.012)	(0.014)
Control	−0.003	0.131***	0.037***	−0.013	−0.080***	−0.041***
	(0.003)	(0.016)	(0.010)	(0.016)	(0.012)	(0.014)
p-value (Cover = Control)	0.62	< 0.001	0.081	0.13	< 0.001	0.0042
DV mean	0.007	0.227	0.084	0.239	0.121	0.176
DV std. dev.	0.084	0.419	0.277	0.426	0.327	0.381
Observations	4,454	4,454	4,454	4,453	4,454	4,452
Individual controls	Yes	Yes	Yes	Yes	Yes	Yes

Notes: The dependent variable in each column is an indicator taking value 1 if the respondent's perceived purpose of the study was coded to fall into the corresponding category. "Cover" takes value 1 if the respondent correctly inferred the study was about whether the ability to link one's donation decision to the Lott study would affect the donation decision. "Immigration attitudes" takes value 1 if the respondent stated the study was about attitudes toward immigration. "Public image" takes value 1 if the respondent stated the study was about whether knowing one's decision will be observable to others would affect the donation decision. "Information" takes value 1 if the respondent stated the study was about disseminating information about immigration. "Persuasion" takes value 1 if the respondent stated the researchers were attempting to persuade them either to donate or not to donate. "Bias" takes value 1 if the respondent stated the researchers were biased. "Other" takes value 1 if the respondent stated a purpose that did not fall into any of the above categories. Categories other than "Other" are not mutually exclusive. All specifications pool the main experiment and the pilot and control for demographics and partisan affiliation. Demographic controls include age, age squared, a set of race indicators, a Hispanic indicator, a male indicator, a set of education indicators. Partisan controls include indicators for "Very conservative", "Conservative", "Neither liberal nor conservative" (omitted), "Liberal", and "Very liberal". Robust standard errors are reported.

Natural language processing   Second, we use natural language processing techniques (in particular, BERT, introduced in Devlin et al. 2019) alongside a neural network classifier to predict treatment status given the participant's response to the open-ended elicitation of perceived purpose. The intuition is simple: if respondents do not perceive differences in the purpose of the study across conditions, there is no scope for experimenter demand effects to affect the results.

At a high level, our exercise proceeds in three steps. After splitting our sample into a training set (75%) and a test set (25%), we create high-dimensional vector representations, or contextual embeddings, of each respondent's answer to the question "If you had to guess, what would you say the purpose of this study was?". These embeddings capture semantic meaning.[39] We then train a neural network classifier on the training set to predict the respondent's treatment condition (*Control*, *Cover*, or *No Cover*) based on their contextual embedding. Finally, we calculate accuracy on the test set.

To generate contextual embeddings for each response, we use DistilBERT (Sanh et al., 2020), a transformer modeled on the BERT (Devlin et al., 2019) architecture that achieves similar performance at substantially lower computational cost. Our implementation is provided by HuggingFace; we use all default parameter values and train the model for the default three epochs.[40] This step outputs a 768-dimensional vector embedding for each respondent. We next train a neural network for sequence classification on the training set, again using the HuggingFace implementation with default parameter values, to predict the respondent's treatment condition based on their embedding. [41] Using this trained model, we predict treatment conditions in the test set and compare predicted conditions with actual conditions.

In order to facilitate comparing classifier accuracy when distinguishing between the *Cover* and *No Cover* condition vs. distinguishing between *Control* and the other two conditions, we repeat the exercise above three times, each time excluding from both the training set and the test set one of the three conditions. Our classifier achieves 51% accuracy when distinguishing between *Cover* and *No Cover*, and we are unable to rule out the null hypothesis of 50% accuracy ($p = 0.48$). In contrast, our classifier achieves 60% accuracy when distinguishing between *Control* and *No Cover* and 58% accuracy when distinguishing between *Control* and *Cover*, with both rates statistically distinguishable from 50% accuracy ($p < 0.001$). Thus, as expected, respondents hold different beliefs about the purpose of the study in *Control* vs. the other two conditions, but not between *Cover* and *No Cover*. Given the differences documented through the hand-coding exercise, we view this result as validation for our method, as it demonstrates that we would likely detect substantial differences in perceived purpose between the *Cover* and *No Cover* conditions if such differences were present.

C.4.4   Additional tables

---

[39]See Liu et al. (2020) for a review of contextual embeddings.
[40]See `https://huggingface.co/docs/transformers/model_doc/distilbert`.
[41]See `https://huggingface.co/docs/transformers/training`.

Table C6: Anti-immigrant donations: Balance of covariates

	Overall		Cover	No Cover	Control	p-values		
	mean	std.dev.	mean	mean	mean	(E=NE)	(E=C)	(NE=C)
Age	45.077	15.724	45.171	44.845	45.209	0.611	0.951	0.560
Black	0.076	0.265	0.070	0.089	0.069	0.075	0.965	0.063
Asian	0.043	0.203	0.041	0.042	0.045	0.886	0.588	0.690
White	0.824	0.381	0.828	0.815	0.830	0.414	0.887	0.330
Hispanic	0.106	0.308	0.112	0.105	0.102	0.580	0.419	0.804
Male	0.498	0.500	0.495	0.503	0.497	0.686	0.925	0.751
High school diploma	0.977	0.151	0.976	0.976	0.977	0.965	0.908	0.944
Bachelors degree	0.376	0.484	0.394	0.367	0.369	0.166	0.196	0.910
Republican	0.425	0.494	0.417	0.436	0.421	0.327	0.825	0.439

Notes: p-values based on robust standard errors reported. Attriters dropped from sample.

Table C7: Anti-immigrant donations: Sample representativeness

	Lott	Pew (Inds and Reps)
Age	45.11	47.17
Black	0.08	0.05
White	0.82	0.77
Asian	0.04	0.03
Hispanic	0.11	0.11
Male	0.49	0.52
High school diploma	0.98	0.93
Bachelors degree or higher	0.37	0.31
Observation	4553	5501

Notes: Table displays mean characteristics, comparing the experimental sample with the 2018 Pew Research Center's American Trends Panel, Wave 39. Attriters are dropped from sample.

C.5   Auxiliary Experiment 8: Interpretation of Lott Rationale

This section presents a pre-registered experiment in which we explore how learning that a previous participant had read the Lott study prior to making their donation decision shapes an audience's inference about the participant's motives and resulting social sanctions.[42] The experimental design is broadly similar to that of Experiments 2 and 4, but in addition to examining inference only about the matched participant's prejudice, we also examine inference about persuadability (or gullibility) — i.e. the extent to which a matched participant might be persuaded by a rationale. This is meant to capture arguments of the form "I have nothing against immigrants...I just believe the study." Of course, persuadability is only one of a set of potential reasons for donating after being exposed to information suggesting immigrants commit more crimes (which is why we choose not to explicitly study it in the main experiments presented in this paper); alternative reasons include lower tolerance for crime, higher levels of risk aversion, etc. We chose to focus on persuadability in this experiment because it is arguably the most natural "second type," because it was the most frequent reason cited in our pilot results, and because it is most consistently and objectively coded.

C.5.1   Sample and Experimental Design

Sample composition   We recruited a sample of 3,047 Democrats through Luc.id in February 2020.[43] Our sample of respondents is broadly representative of Democrats and Independents in the United States (Table C8), though slightly more white and educated, and well-balanced on observables across treatment arms (Table C9).

---

[42]The pre-registration is available in the AEA RCT registry under ID AEARCTR-0005462.

[43]In our pre-registration, we specified that in some specifications, we would pool data from a pilot ($N = 2,019$) with the data from the main experiment. The pilot instrument was virtually identical to the instrument used in the main experiment. We report both unpooled and pooled specifications.

Table C8: Interpretation of Lott rationale: Sample representativeness

	Lott	Pew (Inds and Dems)
Age	41.35	45.86
Black	0.18	0.18
White	0.70	0.59
Asian	0.05	0.05
Hispanic	0.14	0.15
Male	0.45	0.46
High school diploma	0.98	0.89
Bachelors degree or higher	0.45	0.35
Observation	3133	6627

Notes: Table displays mean characteristics, comparing the experimental sample with the 2018 Pew Research Center's American Trends Panel, Wave 39. Attriters are dropped from sample.

Table C9: Interpretation of Lott rationale: Balance of covariates

	Overall		Cover	No Cover	p-value
	mean	std.dev.	mean	mean	(R=NR)
Age	41.376	15.639	41.703	41.048	0.247
Black	0.182	0.386	0.186	0.179	0.612
Asian	0.045	0.208	0.049	0.042	0.386
White	0.710	0.454	0.703	0.716	0.455
Hispanic	0.140	0.347	0.136	0.144	0.561
Male	0.450	0.498	0.451	0.448	0.840
High school diploma	0.983	0.130	0.983	0.983	0.998
Bachelors degree	0.446	0.497	0.454	0.439	0.391

Notes: p-values based on robust standard errors reported.

Information provision Figure C3 outlines the structure of the experiment. We tell all respondents about the Lott study (Lott, 2018) described in Section C.4 and about the fact that the study has been challenged on methodological grounds.[44] We then tell respondents that we conducted a project on political and social attitudes in the United States earlier in the year, and that participants in this previous study were given an opportunity to authorize a $1 donation to Fund the

---

[44]Once again, in order to ensure that our respondents are not misinformed, we debrief them at the end of the study and provide them with a meta-analysis summarizing the work on the effects of immigration on crime (Ousey and Kubrin, 2018).

Wall. We inform participants that we have matched them with one of these respondents, and that this respondent chose to authorize the donation. Respondents in the *Cover* condition are told that their matched respondent was informed about the study before deciding whether or not to authorize the donation to Fund the Wall, while respondents in the *No Cover* condition are told that their matched respondent was not informed about the study before making their donation decision.[45]

---

[45]Because all participants in the donations experiment were informed of the study before making their donation decision, in order to avoid deception, we ran a small auxiliary version of this experiment in which some respondents were *not* informed. These are the participants with whom we match respondents in the experiment described in this section.

Figure C3: Interpretation of Lott rationale: design

Consent, attention check, demographics

↓

Information about Lott (2018)

No Cover
- "Your matched respondent was not informed about Dr. Lott's study"

- Your matched respondent decided to authorize the $1 donation to Fund the Wall

Cover
- "Your matched respondent was informed about Dr. Lott's study"

- Your matched respondent decided to authorize the $1 donation to Fund the Wall

Perceived motive (open-ended)
"Why do you think your matched respondent chose to donate to Fund the Wall?

Gullibility Scale
"If you had to guess, how do you think your matched respondent scored on the Gullibility Scale?"

Foreign Culture Tolerance Scale
"If you had to guess, how do you think your matched respondent scored on the Foreign Culture Tolerance Scale?"

Post-treatment questions, perceived purpose, and feedback

End of study

Measuring inference   After learning whether or not their matched respondent knew about the study, all participants respond to the following open-ended question: "Why do you think your matched respondent chose to donate to Fund the Wall?" These open-ended responses form the raw data for our first measure of inference, which we analyze using the following (pre-registered) procedure. We begin with five "seed words" for each of the two dimensions: for intolerance, we chose *racist*, *biased*, *xenophobic*, *intolerant*, and *prejudiced*; and for persuadability, we chose *convinced*, *persuaded*, *gullible*, *naive*, and *sucker*. We add all "most relevant" synonyms for these words, as classified by the website `www.thesaurus.com`. In order to capture different parts of speech, we stem all words in our list (e.g., xenophobic → xenophob, gullible → gullib), for a total of 23 intolerance-related stems and 30 persuadability-related stems (Gentzkow et al., 2019). We then define two indicator variables, one for the presence of an intolerance-related stem and another for the presence of the persuadability-related stem, and we estimate treatment effects on the probability that the respondent uses at least one word in each list.[46]

For our second measure of inference, participants are then cross-randomized into one of two conditions: "tolerance" and "persuadability".[47] Participants in the "tolerance" condition are told that their matched respondent completed the "Foreign Culture Tolerance Scale," a "short questionnaire measuring tolerance toward foreign values and traditions," before making their donation decisions. Participants in the "persuadability" condition are told that their matched respondent completed the "Gullibility Scale," a "short questionnaire which measures how easily people are manipulated by evidence from untrustworthy sources," before making their donation decisions. All participants are asked to guess their respondent's score; we incentivize this guess by informing them that if they correctly guess the score, they will be entered into a lottery for a $50 Amazon gift card.

C.6   Results

Columns 1–3 of Table C10 displays results for our text-based measure of inference. Participants in the *Cover* condition are 7 percentage points less likely to use a stem related to intolerance when describing their matched respondent's motive, compared to a mean of 17 percent among participants in the *No Cover* condition ($p < 0.001$). These same participants are also 3 percentage points more likely to use words related to persuadability ($p < 0.001$), relative to a mean of 7 percent among participants in the *No Cover* condition.[48] These effect sizes highlight that the availability of a rationale strongly changes people's inference about their matched respondent's motives. That the effect on intolerance is larger than the effect on persuadability is consistent with the fact that persuadability is only one of several possible "second types" to which respondents might be substituting. These results are stable to the inclusion of demographic and partisan controls.

Columns 4–6 display results from our structured belief measures. Respondents in the *Cover* condition rate their matched participant 0.13 standard deviations lower on the intolerance scale ($p < 0.001$) and 0.32 standard deviations higher on the gullibility scale ($p < 0.001$) than partic-

---

[46]Responses that contain both an intolerance-related stem and a persuadability-related stem will have both intolerance and persuadability indicators equal to one, whereas responses that contain neither type of stem will have both indicators equal to zero. Thus, our results are unbiased even if participants perceive a nonzero correlation between intolerance and persuadability.

[47]We measure type inference using a "between" design (in which each respondent is asked only about a single dimension) rather than a "within" design (in which respondents are asked about both dimensions). We employ a between design in order to minimize experimenter demand effects and to avoid order effects (Haaland et al., 2021).

[48]We were intentionally conservative when choosing stem words in order to minimize the rate of false positives.

ipants in the *No Cover* condition. As with the text analysis measure, effects are similar in the pilot and in the pre-registered main experiment, are robust to the inclusion of control variables, and are precisely estimated. Our two measures of type inference are also highly correlated: on average, a respondent who uses a word related to intolerance (persuadability) when describing the matched respondent's motive rates the matched respondent as around half a standard deviation more intolerant (persuadable) than a respondent who does not use such a word.

Table C10: Interpretation of Lott rationale: type inference

	Used word			Inferred score		
	(1)	(2)	(3)	(4)	(5)	(6)
Panel A:	*Inference about intolerance*					
Cover	−0.070***	−0.068***	−0.068***	−0.134***	−0.133***	−0.152***
	(0.012)	(0.012)	(0.012)	(0.051)	(0.051)	(0.039)
Individual controls	No	Yes	Yes	No	Yes	Yes
Observations	3,047	3,047	3,047	1,524	1,524	2,532
DV mean	0.137	0.137	0.137	0.000	0.000	0.000
DV std. dev.	0.344	0.344	0.344	1.000	1.000	1.000
Panel B:	*Inference about persuadability*					
Cover	0.031***	0.032***	0.032***	0.321***	0.310***	0.317***
	(0.010)	(0.010)	(0.010)	(0.050)	(0.050)	(0.039)
Observations	3,047	3,047	3,047	1,523	1,523	2,533
DV mean	0.084	0.084	0.084	0.000	0.000	0.000
DV std. dev.	0.278	0.278	0.278	1.000	1.000	1.000

Notes: The dependent variable in Columns 1–3 is an indicator taking value 1 if the respondent used a key word when describing their matched partner's motive. The dependent variable in Columns 4–6 is the respondent's (standardized) guess about their matched partner's score on the test. Panel A reports results for inference about intolerance; Panel B reports results for inference about persuadability. Demographic controls include age, age squared, a set of race indicators, a Hispanic indicator, a male indicator, a set of education indicators. Partisan controls include indicators for "Very conservative", "Conservative", "Neither liberal nor conservative" (omitted), "Liberal", and "Very liberal". Robust standard errors are reported.

Taken together, our evidence suggests that when judging others' motives, people believe that those who donated with a rationale are more persuadable and less intolerant than those who donated without a rationale.

D Ethical Considerations

Understanding dissenting expression is of great social importance. Identifying the drivers of xenophobic expression is crucial in designing policies best-suited to curbing it, while understanding barriers to dissenting expression in situations where such expression is desirable — for example, speaking out against unjust practices or systems — may help design contexts with lower such barriers.

Nonetheless, ethically conducting revealed-preference experiments on dissenting expression — particularly xenophobic expression — requires balancing three often contradictory objectives: avoiding explicitly deceiving respondents, avoiding compromising respondents' privacy, and avoiding increasing public xenophobic expression. In this section, we provide a more detailed explanation of how our experimental designs balance these objectives. Of course, all experiments obtained approval from multiple Institutional Review Boards.

D.1 Considerations related to information provision (Experiments 3–4 and Lott experiment)

The information on the link between illegal immigration and violent crime we provide to respondents (the clip from *Tucker Carlson Tonight* in Experiments 3–4 and the Lott (2018) study in the robustness experiment) paints an incomplete picture of the academic literature, which generally finds null or negative effects of illegal immigration on violent crime. Although we do not endorse either piece of evidence — indeed, we explicitly inform respondents in the Lott experiment that the study has been challenged by reputable sources — we nonetheless debrief all respondents at the end of the study, providing them with an accessible academic overview of the link between illegal immigration and violent crime (Ousey and Kubrin, 2018) and a list of further readings.[49]

Tucker Carlson Tonight clip In Experiments 3–4, we provide respondents with video clip from *Tucker Carlson Tonight*, the most popular cable news show in the country. While we do not endorse the message, the raw numbers presented in the video clip are taken from the U.S. Sentencing Commission and are factually correct. While we debrief respondents with the above-described meta-study at the end of the survey, this is strictly speaking unnecessary, as the numbers cited in the video clip are not factually wrong.

Lott study In the Lott experiment, treated respondents receive information about a study by John Lott claiming that illegal immigrants commit more crime. Lott holds a PhD in economics from UCLA and has previously held positions at Yale University, the University of Chicago, and the Department of Justice. The working paper we describe — Lott (2018) — has been extensively cited by the Trump administration and is posted on the Social Science Research Network. In other words, it is a real study by an academic economist. The study has been challenged on methodological grounds, but we inform all respondents about this controversy on the very same screen where we present the study.[50] While we again debrief respondents, this is not strictly necessary, as respondents had already been informed that the study's methodology had been challenged.

---

[49]It is common practice to mislead respondents by omitting relevant details from the instructions. Indeed, correspondence studies generally rely on outright deception by sending out fake resumes to employers.

[50]Specifically, at the information page, all respondents are informed about the following: "However, a number of sources (including a researcher affiliated with the Cato Institute, a non-partisan libertarian think tank) have recently

D.2 Considerations related to privacy and deception (Experiments 1, 3, and Lott experiment)

Given that our mechanism examines the effect of perceived social stigma on behavior, it is crucial that respondents in Experiments 1 and 3 and in the Lott experiment believe that their decisions will be visible to others. Although our experiments avoid explicit deception, protecting participants' privacy in this context required us to mislead respondents. We distinguish between the ethical and practical problems associated with deception (the latter relating to concerns about subject pool contamination), addressing the first concern in this section and the second in Section D.3.

D.2.1 Experiments 1 and 3

Twitter login All respondents were required to log in via their Twitter accounts to the "Tweetability" app we created. This app is governed by the Twitter API's terms of service and has the second most restrictive set of permissions among the three application scopes Twitter provides ("Read" and "Write"). That is, the app does not have access to users' passwords, messages, or account settings, but it is able to post Tweets from the users' accounts. We do not use this functionality in any way, and no information that could compromise users' accounts is ever accessed or downloaded. We explicitly inform respondents of the app's permissions in transparent language and give them the option to end the survey if they are uncomfortable granting the app these permissions. We also inform respondents that the app's data, including the tokens that give us access to post on their accounts, will be deleted by no later than August 1, 2021 (Experiment 3) and December 1, 2021 (Experiment 1). Tokens were indeed deleted immediately after collection.

Twitter posts Our key outcome is whether respondents are willing to post a Tweet including a link to a petition to immediately deport all illegal Mexican immigrants. We were not willing to consider designs that asked respondents to actually post such Tweets. We thus asked them to "schedule" their Tweet for the future (using the Tweetability app), to be posted "if/when we have finished surveying people in all US counties". Because we targeted fewer total respondents than the total number of US counties, these posts will never be published. This formulation is therefore misleading, even if it is not explicitly deceptive. Given our desire to avoid leading respondents to publicly post political content (particularly xenophobic content, as in Experiment 3) as part of our survey, we and our Institutional Review Board felt comfortable with this formulation.

D.2.2 Lott experiment

Identifiable donations We asked participants to consent to us accessing their name and city from their survey provider (which confirmed that we could collect this data subject to participant consent). All participants had the option to terminate the survey if they did not consent. We informed those that consented that upon the publication of the Lott study in an academic journal, we would post the results from the survey, including their individual donation decision, on our study website. While we intend to do so should the study be published in an academic journal, this statement is somewhat misleading because it is unlikely that the study will ever be published (given its methodological errors and the fact that Lott has rarely published in peer-reviewed academic

---

challenged some of the study's methods, claiming that errors in analysis invalidate its results. Dr. Lott has responded to this critique with a defense of the study's methods, but the issue remains unresolved."

journals over the past decade). Despite participants consenting to us accessing and publishing their names, and the fact that only a small minority of the Republican and Independent participants are likely to be uniquely identifiable based on their first and last name alone (i.e. absent geographical or other identifiers), we still viewed it as desirable to preserve their anonymity: the formulation of our experiment allows us to do so with high probability.

Cover manipulation   Conceptually, in the *No Cover* condition for the donations experiment, we would like to show respondents a website screenshot stating that "No participants were told about Dr. Lott's study." However, because these participants did in fact learn about the study, such a screenshot would be deceptive. Instead, we exploit the fact that Lott's study had not yet been published in an academic journal (a fact about which we explicitly informed all respondents when describing the website). In particular, we show respondents a website screenshot stating that "We surveyed respondents earlier this year before Dr. Lott's study was published." In the survey, we write that "the website states that you were surveyed before the study was published and does not mention that you were shown an early summary of the study's findings." Respondents in this condition thus believe that their audience will believe that they (the respondents) had no information excusing their decision to donate to Fund the Wall.

This formulation is misleading in that it relies on an academic, rather than commonplace, understanding of the word "published" (that is, "published in an academic journal" rather than "made available for public readership"). However, survey respondents themselves are not misled, as they are fully aware of the study's status and are fully aware of what others reviewing the donation decisions are likely to believe. The group that may be misled is thus the group who visit the website listing donation decisions. Given the low probability that this website will ever be published, we and our Institutional Review Board felt comfortable using this formulation.

D.3   Considerations related to subject pool contamination (Experiments 1, 3, and Lott experiment)

An important concern with deceptive or misleading experiments is that they can contaminate the subject pool by lowering trust in scientists and making respondents less likely to participate in future research studies. Of course, this can only happen if respondents know that they are being misled.

In the Lott experiment, subjects are told we will publish the website once Lott's study is published in an academic journal. Although we privately believe that the Lott (2018) study is very unlikely to be published in a journal, subjects do not know (and never learn) this. (If the study is published, we are prepared to set up the web page in line with the instructions provided to respondents.) Similarly, in Experiments 1 and 3, subjects are told we will post their Tweets when and if we reach survey respondents on all US counties before August 1, 2021 (Experiment 3) or December 1, 2021 (Experiment 1). Although we privately targeted fewer respondents than the number of US counties, ensuring that this condition would not be met, subjects do not know (and never learn) this is the case. In other words, it is not possible for respondents to know that they have been misled about the implementation of the main outcomes (unless they independently find our working paper). To further substantiate the claim that our experiments had no effect on respondents' trust in social science experiments, we asked Luc.id to calculate the number of studies in which each respondent participated in the following nine months. In Figure D1, we examine

whether this number varies by experimental condition. We find no differences across the three treatment conditions (including the Control condition, in which respondents were not exposed to any information about immigrants).

While it is technically possible that the experimental conditions induced differential trust in social science that was not reflected in the number of studies in which respondents participated, we view this contingency as unlikely in light of the reasoning above. Furthermore, concerns about contaminating the experimental subject pool are most important in an economic lab with clear rules against deception. In online survey marketplaces, where survey participants are expected to regularly participate in studies by psychologists in which explicit deception is common, considerations about contaminating the subject pool are less relevant.

Figure D1: Subsequent survey behavior of respondents
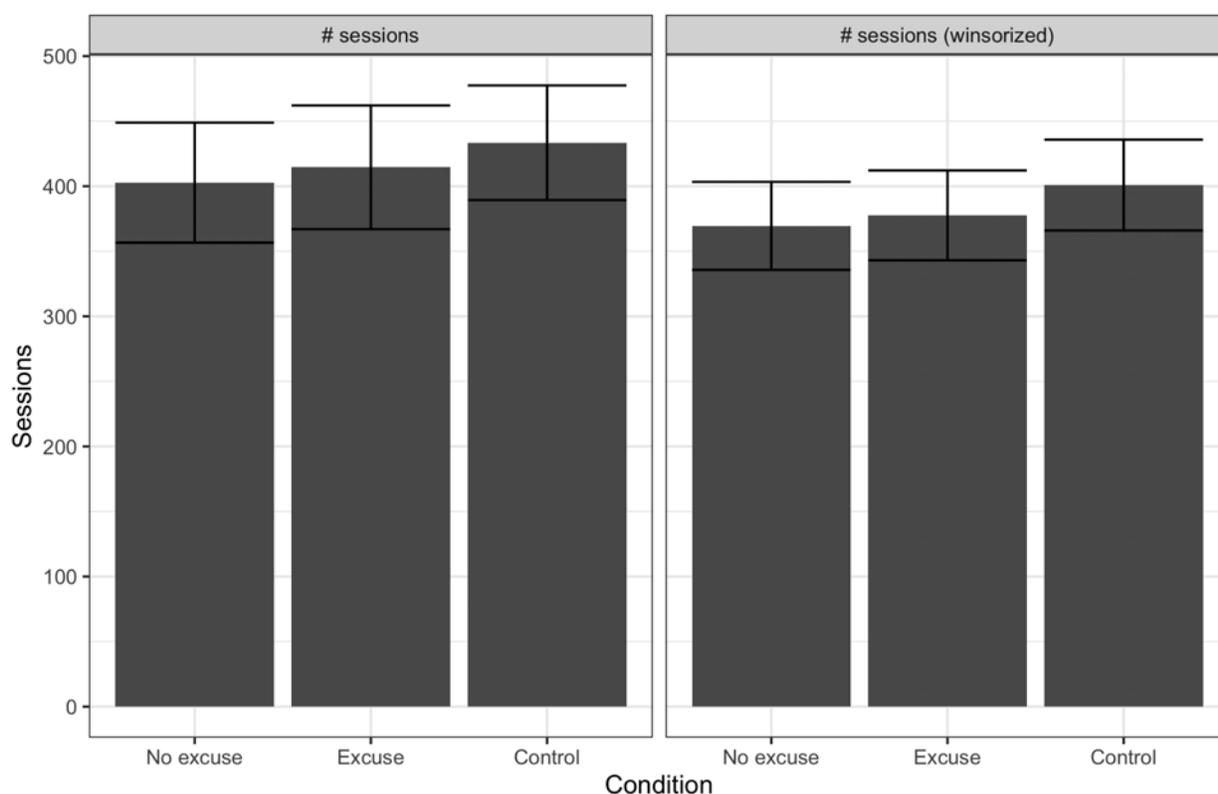
Notes: Figure D1 presents the results of our analysis of the subsequent survey behavior of the respondents in the Lott experiment between the end of collection and December 2020. The figure presents the mean number of surveys (top panel) and the mean of the winsorized number of surveys (bottom panel) in which respondents participated, with the winsorization at the 0.98 quantile.

D.4 Considerations related to starting political Twitter campaigns (Experiments 1 and 3)

As discussed in Appendix D.2, we designed our experiment to ensure that none of the Tweets would ever be posted. It is of course possible that respondents independently posted political content on

85

Twitter as a result of our experiment. This is a concern for Experiment 3, in which respondents were exposed to a clip presenting a misleading narrative about the link between illegal immigration and crime.

To examine whether this was the case, we accessed all Twitter posts made by respondents between the date of experimental collection and August 1, 2021 (the date by which we promised respondents that our access to their accounts and any Twitter-related data would be deleted). We used simple text analysis techniques to identify which posts concern immigrants and quantify the sentiment and content of these posts. The results of this analysis are presented in Figure D2 and Table D1. We find no evidence that respondents in our experiment begin posting more immigrant-related Tweets or more negative content about immigrants after participating (Figure D2). Restricting to the period after the experiment, we find no evidence that respondents in the *Cover* condition post more or fewer Tweets in general, more or fewer Tweets specifically about immigrants, or more or less negative Tweets about immigrants than respondents in our *No Cover* condition (Table D1). This evidence further strengthens our confidence that our experiment did not contribute to anti-immigrant discourse on social media.

Figure D2: Twitter activity of respondents before and after experiment
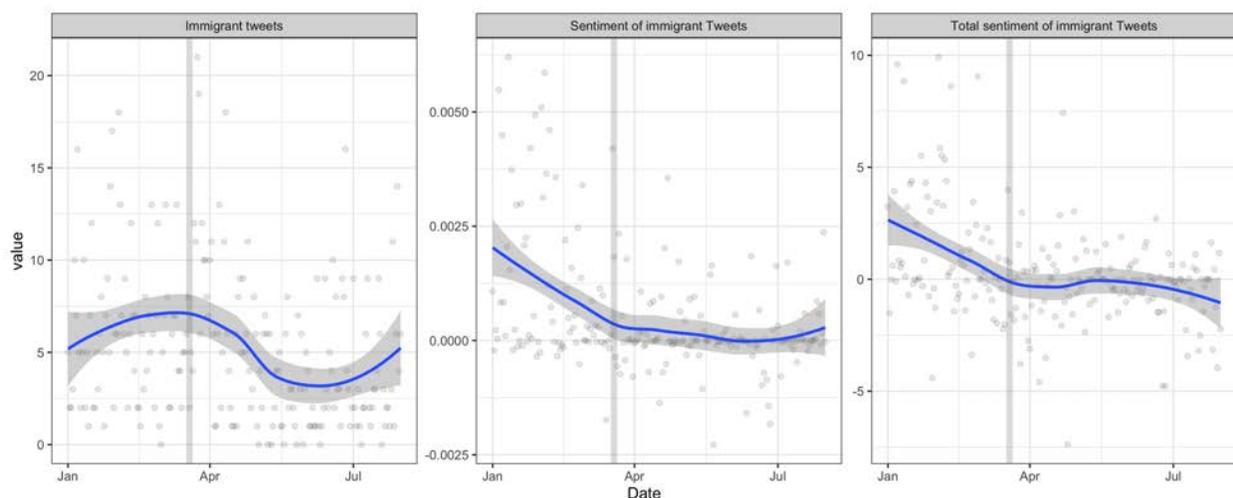
Notes: Figure D2 presents various measures of the Twitter activity of respondents before and after Experiment 3, conducted between March 17 and March 22, 2021 (shaded in a gray rectangle). The left panel of the figure presents the average number of immigrant-related Tweets; the middle panel the average sentiment of immigrant-related Tweets; and the right panel the total expressed sentiment of immigrant-related Tweets.

Table D1: Subsequent Twitter behavior of respondents

	Dependent variable:					
	Tw. (1)	Tw. (w) (2)	Imm. Tw. (3)	Imm. Tw. (w) (4)	Imm. sent. (5)	Tot. imm. sent. (6)
Cover	−44.414	−9.298	−0.583	−0.152	0.005	0.024
	(29.941)	(9.462)	(0.416)	(0.117)	(0.012)	(0.062)
Constant	80.075***	35.951***	0.970***	0.383***	0.003	−0.052
	(20.862)	(6.593)	(0.290)	(0.082)	(0.008)	(0.043)
Observations	517	517	517	517	517	517

Notes: Table D1 presents the results of our analysis of the subsequent Twitter behavior of the respondents in Experiment 3 between the end of our experiment and August 1, 2021. Table presents regressions of various measures of behavior on an indicator for whether the respondent was in the *Cover* condition: Columns 1 and 2 consider the total number of Tweets, Columns 3 and 4 the total number of immigrant-related Tweets, Column 5 the sentiment of immigrant-related Tweets, and Column 6 the sentiment of immigrant-related Tweets multiplied by the number of Tweets. Columns 2 and 4 winsorize the dependent variable at the 0.98 quantile.

E Experimental Instructions

E.1 Experiment 1: Expression of dissent – Democrats

E.1.1 Attention screener

The next question is about the following problem. In questionnaires like ours, sometimes there are participants who do not carefully read the questions and just quickly click through the survey. This means that there are a lot of random answers which compromise the results of research studies. To show that you read our questions carefully, please choose **both** "Extremely interested" and "Not at all interested" on the question below.

Given the text above, how interested are you in sports?

☐ Extremely interested

☐ Very interested

☐ A little bit interested

☐ Very little interested

☐ Not at all interested

›

E.1.2 Twitter information and login

Since our survey is about Twitter and current events, it requires you to grant the system access to your Twitter account through the "Tweetability" app.

Please note that we are **bound by agreement** with the Social and Behavioral Sciences Institutional Review Board at the University of Chicago to adhere to the following terms (in addition to the Twitter terms of service):

- We will **never** use the app to access non-public information from your account (including your posts)
- We will **never** use the app to make posts on your account without your **explicit consent**
- The app **does not give us access to your direct messages or email address**
- All identifying information will be stored on **password-protected directories** secured with **two-factor authentication**, and only **authorized research personnel** will have access
- All identifying information, **including your Twitter handle**, will be deleted by no later than December 1, 2021. Therefore, **the app will lose all access to your account** after this date (if not earlier)

If you have any questions for the researchers, you can contact the researchers at: twitter.study@uchicago.edu

If you have any questions or complaints, you can contact the Social and Behavioral Sciences Institutional Review Board at the University of Chicago at:
The Social & Behavioral Sciences Institutional Review Board, University of Chicago
Phone: (773) 834-7835
E-mail: sbs-irb@uchicago.edu

If you are uncomfortable with these terms in any way, please end the survey now. Otherwise, please click the button below to proceed by signing into Twitter.

Sign in with Twitter

E.1.3 Background questions

Are you Spanish, Hispanic, or Latino or none of these?

○ Yes

○ None of these

What is your year of birth?

[⌄]

What is your sex?

○ Male

○ Female

In politics, as of today, do you consider yourself a Republican, a Democrat, or an Independent?

○ Republican

○ Democrat

○ Independent

››

What is the highest level of school you have completed or the highest degree you have received?

- ○ Less than high school degree
- ○ High school graduate (high school diploma or equivalent including GED)
- ○ Some college but no degree
- ○ Associate degree in college (2-year)
- ○ Bachelor's degree in college (4-year)
- ○ Master's degree
- ○ Doctoral degree
- ○ Professional degree (JD, MD)

Which of the following best describes your race or ethnicity?

- ○ African American/Black
- ○ Asian/Asian American
- ○ Caucasian/White
- ○ Native American, Inuit or Aleut
- ○ Native Hawaiian/Pacific Islander
- ○ Other

Who did you vote for in the 2020 presidential election?

- ○ Donald Trump
- ○ Joe Biden
- ○ Other
- ○ Did not vote

Are you liberal or conservative?

- ○ Very liberal
- ○ Liberal
- ○ Neither liberal nor conservative
- ○ Conservative
- ○ Very conservative

››

E.1.4 Pre-treatment outcomes

On the next page, you will be provided with a recent Washington Post column written by **Princeton Professor of Criminology Patrick Sharkey**, in which he discusses evidence showing that more policing leads to less violent crime.

››

Outlook

Why do we need the police?

Cops prevent violence. But they aren't the only ones who can do it.

By **Patrick Sharkey**
JUNE 12, 2020

The calls to end policing as we know it contain a sort of trap. The best evidence we have makes clear that police are effective in reducing violence, and without designating some group to combat this problem, efforts to weaken them through budget cuts — "defund the police" — are likely to have unanticipated consequences and to destabilize communities. In many cities this is likely to lead to a rise in violence. And research shows that, when violence increases, Americans of all races become more punitive, supporting harsher policing and criminal justice policies. That's how we got to this point.

Patrick Sharkey
@patrick_sharkey is a professor of sociology and public affairs at Princeton University's Woodrow Wilson School of Public and International Affairs. His most recent book is "Uneasy Peace: The Great Crime Decline, the Renewal of City Life, and the Next War on Violence."

Yet none of this means that the police, which have served as an institution of racialized control throughout our nation's history, are the only group capable of reducing violence. Community leaders and residents have proved adept at overseeing their neighborhoods, caring for their populations and maintaining safe streets. Studies show that this work lowers crime, sometimes dramatically. What happens if we put those people *in charge* of containing violence, too?

Over the past 10 years, an expanding body of research has shown just how damaging violence is to community life, children's academic trajectories and healthy child development. We have rigorous, causal evidence that every shooting in a neighborhood affects children's sleep and their ability to focus and learn. When a neighborhood becomes violent, it begins to fall apart, as public spaces empty, businesses close, parks and playgrounds turn dangerous, and families try to move elsewhere. Violence is the fundamental challenge for cities: Nothing works if public space is unsafe.

Those who argue that the police have no role in maintaining safe streets are arguing against lots of strong evidence. One of the most robust, most uncomfortable findings in criminology is that putting more officers on the street leads to less violent crime. We know this from randomized experiments involving "hot spots policing" and natural experiments in which more officers were brought to the streets because of something other than crime — a shift in the terror alert level or the timing of a federal grant — and violent crime fell. After the unrest around the deaths of Freddie Gray in Baltimore and Michael Brown in Ferguson, Mo., police officers stepped back from their duty to protect and serve; arrests for all kinds of low-level offenses dropped, and violence rose. This shouldn't be interpreted to mean that protests against violent policing lead to more violence; rather, it means that when police don't do their jobs, violence often results.

Considered alongside the brutal response to protests over the past few weeks, this evidence forces us to hold two incongruent ideas: Police are effective at reducing violence, the most damaging feature of urban inequality. And yet one can argue that law enforcement is an authoritarian institution that historically has inflicted violence on black people and continues to do so today.

93

Would you like to join a nonpartisan campaign that opposes defunding the police?

○ Yes

○ No

››

You have successfully joined the campaign.

Since you chose to join the campaign, we wanted to give you more time reading the Washington Post column written by **Princeton Professor of Criminology Patrick Sharkey**, where he discusses evidence showing that more policing leads to less violent crime.

The article is available on the next page, and you can spend as much time as you want reading it before you continue with the remaining part of the survey.

››

Outlook

Why do we need the police?

Cops prevent violence. But they aren't the only ones who can do it.

By **Patrick Sharkey**

JUNE 12, 2020

The calls to end policing as we know it contain a sort of trap. The best evidence we have makes clear that police are effective in reducing violence, and without designating some group to combat this problem, efforts to weaken them through budget cuts — "defund the police" — are likely to have unanticipated consequences and to destabilize communities. In many cities this is likely to lead to a rise in violence. And research shows that, when violence increases, Americans of all races become more punitive, supporting harsher policing and criminal justice policies. That's how we got to this point.

Patrick Sharkey
@patrick_sharkey is a professor of sociology and public affairs at Princeton University's Woodrow Wilson School of Public and International Affairs. His most recent book is "Uneasy Peace: The Great Crime Decline, the Renewal of City Life, and the Next War on Violence."

Yet none of this means that the police, which have served as an institution of racialized control throughout our nation's history, are the only group capable of reducing violence. Community leaders and residents have proved adept at overseeing their neighborhoods, caring for their populations and maintaining safe streets. Studies show that this work lowers crime, sometimes dramatically. What happens if we put those people *in charge* of containing violence, too?

Over the past 10 years, an expanding body of research has shown just how damaging violence is to community life, children's academic trajectories and healthy child development. We have rigorous, causal evidence that every shooting in a neighborhood affects children's sleep and their ability to focus and learn. When a neighborhood becomes violent, it begins to fall apart, as public spaces empty, businesses close, parks and playgrounds turn dangerous, and families try to move elsewhere. Violence is the fundamental challenge for cities: Nothing works if public space is unsafe.

Those who argue that the police have no role in maintaining safe streets are arguing against lots of strong evidence. One of the most robust, most uncomfortable findings in criminology is that putting more officers on the street leads to less violent crime. We know this from randomized experiments involving "hot spots policing" and natural experiments in which more officers were brought to the streets because of something other than crime — a shift in the terror alert level or the timing of a federal grant — and violent crime fell. After the unrest around the deaths of Freddie Gray in Baltimore and Michael Brown in Ferguson, Mo., police officers stepped back from their duty to protect and serve; arrests for all kinds of low-level offenses dropped, and violence rose. This shouldn't be interpreted to mean that protests against violent policing lead to more violence; rather, it means that when police don't do their jobs, violence often results.

Considered alongside the brutal response to protests over the past few weeks, this evidence forces us to hold two incongruent ideas: Police are effective at reducing violence, the most damaging feature of urban inequality. And yet one can argue that law enforcement is an authoritarian institution that historically has inflicted violence on black people and continues to do so today.

95

E.1.5 Treatment: "Before" wording (rationale)

This nonpartisan campaign involves signing up people on Twitter **to make a post encouraging their friends and followers to sign a petition** opposing the movement to defund the police.

The posts will be made public if/when we have finished surveying people in all U.S. counties. This strategy is often used to make campaigns "trend" on Twitter. To coordinate these efforts, we will use the *Tweetability* app you signed into earlier to schedule the posts.

I have joined a campaign to oppose defunding the police: bit.ly/3DK3UEr. Before joining, I was shown this article written by a Princeton professor on the strong scientific evidence that defunding the police would increase violent crime:
https://www.washingtonpost.com/outlook/2020/06/12/defund-police-violent-crime/

washingtonpost.com
Perspective | Cops prevent violence. But they aren't the only ones wh...
Communities already know how to police their own. Now put them in charge of it.

Do you authorize the *Tweetability* app to schedule the post above to be posted on your account? (If you choose "no," then nothing will be posted on your account.)

○ Yes

○ No

›

This nonpartisan campaign involves signing up people on Twitter **to make a post encouraging their friends and followers to sign a petition** opposing the movement to defund the police.

The posts will be made public if/when we have finished surveying people in all U.S. counties. This strategy is often used to make campaigns "trend" on Twitter. To coordinate these efforts, we will use the *Tweetability* app you signed into earlier to schedule the posts.

I have joined a campaign to oppose defunding the police: bit.ly/3DK3UEr. After joining, I was shown this article written by a Princeton professor on the strong scientific evidence that defunding the police would increase violent crime:
https://www.washingtonpost.com/outlook/2020/06/12/defund-police-violent-crime/

washingtonpost.com
Perspective | Cops prevent violence. But they aren't the only ones wh...
Communities already know how to police their own. Now put them in charge of it.

Do you authorize the *Tweetability* app to schedule the post above to be posted on your account? (If you choose "no," then nothing will be posted on your account.)

○ Yes

○ No

››

E.2 Experiment 2: Interpretation of dissent – Democrats

E.2.1 Attention screener and background questions

The next question is about the following problem. In questionnaires like ours, sometimes there are participants who do not carefully read the questions and just quickly click through the survey. This means that there are a lot of random answers which compromise the results of research studies. To show that you read our questions carefully, please choose both "**Extremely interested**" and "**Not at all interested**" as your answer in the below question.

Given the above, how interested are you in sports?

☐ Extremely interested

☐ Very interested

☐ A little bit interested

☐ Almost not interested

☐ Not at all interested

→

What is your sex?

- ○ Male
- ○ Female

What is your year of birth?

[▼]

In politics, as of today, do you consider yourself a Republican, a Democrat, or an Independent?

- ○ Republican
- ○ Democrat
- ○ Independent

What is the highest level of school you have completed or the highest degree you have received?

- ○ Less than high school degree
- ○ High school graduate (high school diploma or equivalent including GED)
- ○ Some college but no degree
- ○ Associate degree in college (2-year)
- ○ Bachelor's degree in college (4-year)
- ○ Master's degree
- ○ Doctoral degree
- ○ Professional degree (JD, MD)

Are you Spanish, Hispanic, or Latino or none of these?

- ○ Yes
- ○ None of these

Which of the following best describes your race or ethnicity?

- ○ African American/Black
- ○ Asian/Asian American
- ○ Caucasian/White
- ○ Native American, Inuit or Aleut
- ○ Native Hawaiian/Pacific Islander
- ○ Other

Do you lean toward the Republican Party or the Democratic Party?

○ Lean toward the Republican Party

○ Lean toward the Democratic Party

Who did you vote for in the 2020 presidential election?

○ Donald Trump

○ Joe Biden

○ Other

○ Did not vote

Are you liberal or conservative?

○ Very liberal

○ Liberal

○ Neither liberal nor conservative

○ Conservative

○ Very conservative

→

E.2.2 Treatment: "Before" wording (rationale)

We conducted a survey about political and social attitudes in the
United States earlier this year. You have been matched with one
of the participants from that survey.

Your matched participant **joined a campaign to oppose
defunding the police**.

Your matched participant **also agreed to post the Tweet
below** on their Twitter profile to promote the campaign.

I have joined a campaign to oppose defunding the
police: bit.ly/3DK3UEr. Before joining, I was shown this
article written by a Princeton professor on the strong
scientific evidence that defunding the police would
increase violent crime:
https://www.washingtonpost.com/outlook/2020/06/12
/defund-police-violent-crime/

washingtonpost.com
Perspective | Cops prevent violence. But they aren't the only ones wh...
Communities already know how to police their own. Now put them in
charge of it.

Why do you think your matched respondent chose to join the
campaign to oppose defunding the police?

→

Matched Respondent's Donation Decision

We gave your matched respondent the opportunity to donate $10 to the **National Association for the Advancement of Colored People (NAACP)**, America's oldest and largest civil rights organization.

Below, we will ask you to guess whether or not your matched respondent donated $10 to the National Association for the Advancement of Colored People (NAACP).

Reminder: Your matched participant agreed to post the Tweet below on their Twitter account.

I have joined a campaign to oppose defunding the police: bit.ly/3DK3UEr. Before joining, I was shown this article written by a Princeton professor on the strong scientific evidence that defunding the police would increase violent crime:
https://www.washingtonpost.com/outlook/2020/06/12/defund-police-violent-crime/

washingtonpost.com
Perspective | Cops prevent violence. But they aren't the only ones wh...
Communities already know how to police their own. Now put them in charge of it.

Do you think that your matched participant chose to donate $5 to the National Association for the Advancement of Colored People (NAACP)?

○ Yes, I think my matched respondent chose to donate

○ No, I think my matched respondent **did not** choose to donate

→

You now have the opportunity to authorize a $1 bonus payment to your matched respondent. **The bonus payment will not be deducted from your payment.** Your matched respondent did not know you would have the opportunity to decide their bonus.

Reminder: Your matched participant agreed to post the Tweet below on their Twitter account.

I have joined a campaign to oppose defunding the police: bit.ly/3DK3UEr. Before joining, I was shown this article written by a Princeton professor on the strong scientific evidence that defunding the police would increase violent crime:
https://www.washingtonpost.com/outlook/2020/06/12/defund-police-violent-crime/

washingtonpost.com
Perspective | Cops prevent violence. But they aren't the only ones wh...
Communities already know how to police their own. Now put them in charge of it.

Do you want to authorize a $1 bonus to your matched respondent?

○ Yes, I would like to authorize a $1 bonus

○ No, I would not like to authorize a $1 bonus

→

103

E.2.3 Treatment: "After" wording (no rationale)

We conducted a survey about political and social attitudes in the United States earlier this year. You have been matched with one of the participants from that survey.

Your matched participant **joined a campaign to oppose defunding the police**.

Your matched participant **also agreed to post the Tweet below** on their Twitter profile to promote the campaign.

I have joined a campaign to oppose defunding the police: bit.ly/3DK3UEr. After joining, I was shown this article written by a Princeton professor on the strong scientific evidence that defunding the police would increase violent crime:
https://www.washingtonpost.com/outlook/2020/06/12/defund-police-violent-crime/

washingtonpost.com
Perspective | Cops prevent violence. But they aren't the only ones wh...
Communities already know how to police their own. Now put them in charge of it.

Why do you think your matched respondent chose to join the campaign to oppose defunding the police?

Matched Respondent's Donation Decision

We gave your matched respondent the opportunity to donate $10 to the **National Association for the Advancement of Colored People (NAACP)**, America's oldest and largest civil rights organization.

Below, we will ask you to guess whether or not your matched respondent donated $10 to the National Association for the Advancement of Colored People (NAACP).

Reminder: Your matched participant agreed to post the Tweet below on their Twitter account.

I have joined a campaign to oppose defunding the police: bit.ly/3DK3UEr. After joining, I was shown this article written by a Princeton professor on the strong scientific evidence that defunding the police would increase violent crime:
https://www.washingtonpost.com/outlook/2020/06/12/defund-police-violent-crime/

washingtonpost.com
Perspective | Cops prevent violence. But they aren't the only ones wh...
Communities already know how to police their own. Now put them in charge of it.

Do you think that your matched participant chose to donate $5 to the National Association for the Advancement of Colored People (NAACP)?

○ Yes, I think my matched respondent chose to donate

○ No, I think my matched respondent **did not** choose to donate

→

105

You now have the opportunity to authorize a $1 bonus payment to your matched respondent. **The bonus payment will not be deducted from your payment.** Your matched respondent did not know you would have the opportunity to decide their bonus.

Reminder: Your matched participant agreed to post the Tweet below on their Twitter account.

I have joined a campaign to oppose defunding the police: bit.ly/3DK3UEr. After joining, I was shown this article written by a Princeton professor on the strong scientific evidence that defunding the police would increase violent crime:
https://www.washingtonpost.com/outlook/2020/06/12/defund-police-violent-crime/

washingtonpost.com
Perspective | Cops prevent violence. But they aren't the only ones wh...
Communities already know how to police their own. Now put them in charge of it.

Do you want to authorize a $1 bonus to your matched respondent?

○ Yes, I would like to authorize a $1 bonus

○ No, I would not like to authorize a $1 bonus

→

106

E.3 Experiment 3: Expression of dissent – Republicans

E.3.1 Attention screener

The next question is about the following problem. In questionnaires like ours, sometimes there are participants who do not carefully read the questions and just quickly click through the survey. This means that there are a lot of random answers which compromise the results of research studies. To show that you read our questions carefully, please choose **both** "Extremely interested" and "Not at all interested" on the question below.

Given the text above, how interested are you in sports?

☐ Extremely interested

☐ Very interested

☐ A little bit interested

☐ Very little interested

☐ Not at all interested

>>

E.3.2 Twitter information and login

Since our survey is about Twitter and current events, it requires you to grant the system access to your Twitter account through the "Tweetability" app.

Please note that we are **bound by agreement** with the Social and Behavioral Sciences Institutional Review Board at the University of Chicago to adhere to the following terms (in addition to the Twitter terms of service):

- We will **never** use the app to access non-public information from your account (including your posts)
- We will **never** use the app to make posts on your account without your **explicit consent**
- The app **does not give us access to your direct messages or email address**
- All identifying information will be stored on **password-protected directories** secured with **two-factor authentication**, and only **authorized research personnel** will have access
- All identifying information, **including your Twitter handle**, will be deleted by no later than August 1, 2021. Therefore, **the app will lose all access to your account** after this date (if not earlier)

If you have any questions for the researchers, you can contact the researchers at: twitter.study@uchicago.edu

If you have any questions or complaints, you can contact the Social and Behavioral Sciences Institutional Review Board at the University of Chicago at:
The Social & Behavioral Sciences Institutional Review Board, University of Chicago
Phone: (773) 834-7835
E-mail: sbs-irb@uchicago.edu

If you are uncomfortable with these terms in any way, please end the survey now. Otherwise, please click the button below to proceed by signing into Twitter.

Sign in with Twitter

Authorize Tweetability: Schedule Tweets to access your account?

Username or email

Password

☐ Remember me · Forgot password?

Sign In Cancel

This application will be able to:

- See Tweets from your timeline (including protected Tweets) as well as your Lists and collections.
- See your Twitter profile information and account settings.
- See accounts you follow, mute, and block.
- Follow and unfollow accounts for you.
- Update your profile and account settings.
- Post and delete Tweets for you, and engage with Tweets posted by others (Like, un-Like, or reply to a Tweet, Retweet, etc.) for you.
- Create, manage, and delete Lists and collections for you.
- Mute, block, and report accounts for you.

Learn more about third-party app permissions in the Help Center.

Tweetability: Schedule Tweets

This app was created to use the Twitter API.

E.3.3   Demographics

Are you Spanish, Hispanic, or Latino or none of these?

○ Yes

○ None of these

What is your year of birth?

[⌄]

What is your sex?

○ Male

○ Female

In politics, as of today, do you consider yourself a Republican, a Democrat, or an Independent?

○ Republican

○ Democrat

○ Independent

››

What is the highest level of school you have completed or the highest degree you have received?

○ Less than high school degree

○ High school graduate (high school diploma or equivalent including GED)

○ Some college but no degree

○ Associate degree in college (2-year)

○ Bachelor's degree in college (4-year)

○ Master's degree

○ Doctoral degree

○ Professional degree (JD, MD)

Which of the following best describes your race or ethnicity?

○ African American/Black

○ Asian/Asian American

○ Caucasian/White

○ Native American, Inuit or Aleut

○ Native Hawaiian/Pacific Islander

○ Other

Who did you vote for in the 2020 presidential election?

○ Donald Trump

○ Joe Biden

○ Other

○ Did not vote

Are you liberal or conservative?

○ Very liberal

○ Liberal

○ Neither liberal nor conservative

○ Conservative

○ Very conservative

»

111

E.3.4 Video clip

Please see the short video below where Fox News host **Tucker Carlson presents evidence on whether illegal immigrants commit more crime**.

>>

E.3.5 Treatment: "After" wording

Would you like to join a campaign to immediately deport all illegal Mexican immigrants?

○ Yes

○ No

>>

In case you want save the video with Tucker Carlson talking about immigration and crime, here is the link: https://www.youtube.com/watch?v=SDdkkTLCUUQ

>>

You have successfully joined the campaign. This campaign involves signing up people on Twitter **to make a post encouraging their friends and followers to sign a petition** in favor of immediately deporting all illegal Mexican immigrants.

The posts will be made public if/when we have finished surveying people in all U.S. counties. This strategy is often used to make campaigns "trend" on Twitter. To coordinate these efforts, we will use the *Tweetability* app you signed into earlier to schedule the posts.

››

Below, we will ask you if you want to authorize the following Tweet to be posted on your account:

I have joined a campaign to immediately deport all illegal Mexicans! After I joined the campaign, I received a link to this video about how illegals commit more crime: y2u.be/SDdkkTLCUUQ. Sign this petition to immediately deport all illegal Mexicans: bit.ly/2OttyJ6

Tucker: Why didn't we know truth about illegals and crime?
Tucker's Thoughts: For years, we were told illegal immigrants were more law-abiding than American citizens. In fact, the ...
🔗 youtube.com

Do you authorize the *Tweetability* app to schedule the post above to be posted on your account? (If you choose "no," then nothing will be posted on your account.)

○ Yes

○ No

›

E.3.6 Treatment: "Before" wording

In case you want save the video with Tucker Carlson talking about immigration and crime, here is the
link: https://www.youtube.com/watch?v=SDdkkTLCUUQ

>>

Would you like to join a campaign to immediately deport all illegal Mexican immigrants?

○ Yes

○ No

>>

You have successfully joined the campaign. This campaign involves signing up people on Twitter **to make a post encouraging their friends and followers to sign a petition** in favor of immediately deporting all illegal Mexican immigrants.

The posts will be made public if/when we have finished surveying people in all U.S. counties. This strategy is often used to make campaigns "trend" on Twitter. To coordinate these efforts, we will use the *Tweetability* app you signed into earlier to schedule the posts.

››

Below, we will ask you if you want to authorize the following Tweet
to be posted on your account:

I have joined a campaign to immediately deport all
illegal Mexicans! Before I joined the campaign, I
received a link to this video about how illegals commit
more crime: y2u.be/SDdkkTLCUUQ. Sign this petition to
immediately deport all illegal Mexicans: bit.ly/2OttyJ6

Tucker: Why didn't we know truth about illegals and crime?
Tucker's Thoughts: For years, we were told illegal immigrants
were more law-abiding than American citizens. In fact, the ...
🔗 youtube.com

Do you authorize the *Tweetability* app to schedule the post
above to be posted on your account? (If you choose "no," then
nothing will be posted on your account.)

○ Yes

○ No

››

118

E.4 Experiment 4: Interpretation of dissent – Republicans

E.4.1 Attention screener and background questions

The next question is about the following problem. In questionnaires like ours, sometimes there are participants who do not carefully read the questions and just quickly click through the survey. This means that there are a lot of random answers which compromise the results of research studies. To show that you read our questions carefully, please choose both "**Extremely interested**" and "**Not at all interested**" as your answer in the below question.

Given the above, how interested are you in sports?

☐ Extremely interested

☐ Very interested

☐ A little bit interested

☐ Almost not interested

☐ Not at all interested

→

What is your sex?

○ Male

○ Female

What is your year of birth?

[▾]

In politics, as of today, do you consider yourself a Republican, a Democrat, or an Independent?

○ Republican

○ Democrat

○ Independent

What is the highest level of school you have completed or the highest degree you have received?

○ Less than high school degree

○ High school graduate (high school diploma or equivalent including GED)

○ Some college but no degree

○ Associate degree in college (2-year)

○ Bachelor's degree in college (4-year)

○ Master's degree

○ Doctoral degree

○ Professional degree (JD, MD)

Are you Spanish, Hispanic, or Latino or none of these?

○ Yes

○ None of these

Which of the following best describes your race or ethnicity?

○ African American/Black

○ Asian/Asian American

○ Caucasian/White

○ Native American, Inuit or Aleut

○ Native Hawaiian/Pacific Islander

○ Other

→

Who did you vote for in the 2020 presidential election?

○ Donald Trump

○ Joe Biden

○ Other

○ Did not vote

Are you liberal or conservative?

○ Very liberal

○ Liberal

○ Neither liberal nor conservative

○ Conservative

○ Very conservative

→

E.4.2 Treatment: "Before" condition (rationale)

We conducted a survey about political and social attitudes in the United States earlier this year. You have been matched with one of the participants from that survey.

Your matched participant **joined a campaign to immediately deport all illegal Mexican immigrants.**

Your matched participant **also agreed to post the Tweet below** on their Twitter profile to promote the campaign.

I have joined a campaign to immediately deport all illegal Mexicans! Before I joined the campaign, I received a link to this video about how illegals commit more crime: y2u.be/SDdkkTLCUUQ. Sign this petition to immediately deport all illegal Mexicans: bit.ly/2OttyJ6

Tucker: Why didn't we know truth about illegals and crime?
Tucker's Thoughts: For years, we were told illegal immigrants were more law-abiding than American citizens. In fact, the ...
🔗 youtube.com

Why do you think your matched respondent chose to join the campaign to immediately deport all illegal Mexican immigrants?

→

Matched Respondent's Donation Decision

We gave your matched respondent the opportunity to authorize a $5 donation to the US Border Crisis Children's Relief Fund, which delivers humanitarian aid to migrant children and families at the US–Mexico border. The organization is working with local partners to ensure that children and families have necessities such as hygiene kits, diapers and clothing. We told your matched respondent that we would make the donation on their behalf, so the donation did not affect their payment.

Below, we will ask you to guess whether or not your matched respondent authorized the $5 donation to the US Border Crisis Children's Relief Fund.

Reminder: Your matched participant agreed to post the Tweet below on their Twitter profile to promote the campaign.

I have joined a campaign to immediately deport all illegal Mexicans! Before I joined the campaign, I received a link to this video about how illegals commit more crime: y2u.be/SDdkkTLCUUQ. Sign this petition to immediately deport all illegal Mexicans: bit.ly/2OttyJ6

Tucker: Why didn't we know truth about illegals and crime?
Tucker's Thoughts: For years, we were told illegal immigrants were more law-abiding than American citizens. In fact, the ...
🔗 youtube.com

Do you think that your matched participant chose to authorize the $5 donation to the US Border Crisis Children's Relief Fund?

○ Yes, I think my matched respondent chose to authorize the donation

○ No, I think my matched respondent **did not** choose to authorize the donation

→

123

You now have the opportunity to authorize a $1 bonus payment to your matched respondent. **The bonus payment will not be deducted from your payment.** Your matched respondent did not know that you would have the opportunity to decide on their bonus.

Reminder: Your matched participant agreed to post the Tweet below on their Twitter profile to promote the campaign.

I have joined a campaign to immediately deport all illegal Mexicans! Before I joined the campaign, I received a link to this video about how illegals commit more crime: y2u.be/SDdkkTLCUUQ. Sign this petition to immediately deport all illegal Mexicans: bit.ly/2OttyJ6

Tucker: Why didn't we know truth about illegals and crime?
Tucker's Thoughts: For years, we were told illegal immigrants were more law-abiding than American citizens. In fact, the ...
🔗 youtube.com

Do you want to authorize a $1 bonus to your matched respondent?

○ Yes, I would like to authorize a $1 bonus

○ No, I would not like to authorize a $1 bonus

→

E.4.3 Treatment: "After" condition (no rationale)

We conducted a survey about political and social attitudes in the United States earlier this year. You have been matched with one of the participants from that survey.

Your matched participant **joined a campaign to immediately deport all illegal Mexican immigrants.**

Your matched participant **also agreed to post the Tweet below** on their Twitter profile to promote the campaign.

I have joined a campaign to immediately deport all illegal Mexicans! After I joined the campaign, I received a link to this video about how illegals commit more crime: y2u.be/SDdkkTLCUUQ. Sign this petition to immediately deport all illegal Mexicans: bit.ly/2OttyJ6

Tucker: Why didn't we know truth about illegals and crime?
Tucker's Thoughts: For years, we were told illegal immigrants were more law-abiding than American citizens. In fact, the ...
🔗 youtube.com

Why do you think your matched respondent chose to join the campaign to immediately deport all illegal Mexican immigrants?

→

You now have the opportunity to authorize a $1 bonus payment to your matched respondent. **The bonus payment will not be deducted from your payment.** Your matched respondent did not know that you would have the opportunity to decide on their bonus.

Reminder: Your matched participant agreed to post the Tweet below on their Twitter profile to promote the campaign.

I have joined a campaign to immediately deport all illegal Mexicans! After I joined the campaign, I received a link to this video about how illegals commit more crime: y2u.be/SDdkkTLCUUQ. Sign this petition to immediately deport all illegal Mexicans: bit.ly/2OttyJ6

Tucker: Why didn't we know truth about illegals and crime?
Tucker's Thoughts: For years, we were told illegal immigrants were more law-abiding than American citizens. In fact, the ...
🔗 youtube.com

Do you want to authorize a $1 bonus to your matched respondent?

○ Yes, I would like to authorize a $1 bonus

○ No, I would not like to authorize a $1 bonus

→

126

Matched Respondent's Donation Decision

We gave your matched respondent the opportunity to authorize a $5 donation to the US Border Crisis Children's Relief Fund, which delivers humanitarian aid to migrant children and families at the US-Mexico border. The organization is working with local partners to ensure that children and families have necessities such as hygiene kits, diapers and clothing. We told your matched respondent that we would make the donation on their behalf, so the donation did not affect their payment.

Below, we will ask you to guess whether or not your matched respondent authorized the $5 donation to the US Border Crisis Children's Relief Fund.

Reminder: Your matched participant agreed to post the Tweet below on their Twitter profile to promote the campaign.

I have joined a campaign to immediately deport all illegal Mexicans! After I joined the campaign, I received a link to this video about how illegals commit more crime: y2u.be/SDdkkTLCUUQ. Sign this petition to immediately deport all illegal Mexicans: bit.ly/2OttyJ6

Tucker: Why didn't we know truth about illegals and crime?
Tucker's Thoughts: For years, we were told illegal immigrants were more law-abiding than American citizens. In fact, the ...
🔗 youtube.com

Do you think that your matched participant chose to authorize the $5 donation to the US Border Crisis Children's Relief Fund?

○ Yes, I think my matched respondent chose to authorize the donation

○ No, I think my matched respondent **did not** choose to authorize the donation

→

127

E.5 Auxiliary Experiment 1: Persuasion experiment – Democrats

E.5.1 Pre-treatment beliefs

How do you think decreasing funding for the police, commonly referred to as "defunding the police," would affect violent crime?

○ Strongly increase violent crime

○ Somewhat increase violent crime

○ Neither increase nor decrease violent crime

○ Somewhat decrease violent crime

○ Strongly decrease violent crime

→

E.5.2 Information treatment (treatment group only)

According to a recent article in the Washington Post written by Princeton Professor of Criminology Patrick Sharkey, **one of the most robust findings in criminology is that putting more police officers on the street leads to less violent crime**.

If you want to learn more, you can read the article here: https://www.washingtonpost.com/outlook/2020/06/12/defund-police-violent-crime/

→

E.5.3 Post-treatment outcomes

Do you think that funding for the police should be increased, decreased, or stay the same?

○ Increased a lot

○ Increased a little

○ Stay about the same

○ Decreased a little

○ Decreased a lot

→

How do you think **increasing** funding for the police would affect violent crime?

○ Strongly increase violent crime

○ Somewhat increase violent crime

○ Neither increase nor decrease violent crime

○ Somewhat decrease violent crime

○ Strongly decrease violent crime

→

E.6   Auxiliary Experiment 2: Rainforest placebo

E.6.1   Pre-treatment questions

On the next page, you will be provided with a recent Reuters article reporting about a new landmark study showing that more than 10,000 species are at high risk of extinction due to the destruction of the Amazon rainforest.

››

Over 10,000 species risk extinction in Amazon, says landmark report

By Stephen Eisenhammer and Oliver Griffin

SAO PAULO/BOGOTA, July 14 (Reuters) - More than 10,000 species of plants and animals are at high risk of extinction due to the destruction of the Amazon rainforest - 35% of which has already been deforested or degraded, according to the draft of a landmark scientific report published on Wednesday.

Produced by the Science Panel for the Amazon (SPA), the 33-chapter report brings together research on the world's largest rainforest from 200 scientists from across the globe. It is the most detailed assessment of the state of the forest to date and both makes clear the vital role the Amazon plays in global climate and the profound risks it is facing.

Cutting deforestation and forest degradation to zero in less than a decade "is critical," the report said, also calling for massive restoration of already destroyed areas.

The rainforest is a vital bulwark against climate change both for the carbon it absorbs and what it stores.

Would you like to join a nonpartisan campaign to immediately stop the destruction of the Amazon rainforest?

○ Yes

○ No

›

You have successfully joined the campaign.

Since you chose to join the campaign, we wanted to give you more time reading the Reuters article covering the landmark study showing that more than 10,000 species are at high risk of extinction due to the destruction of the Amazon rainforest.

The article is available on the next page, and you can spend as much time as you want reading it before you continue with the remaining part of the survey.

›

Over 10,000 species risk extinction in Amazon, says landmark report

By Stephen Eisenhammer and Oliver Griffin

SAO PAULO/BOGOTA, July 14 (Reuters) - More than 10,000 species of plants and animals are at high risk of extinction due to the destruction of the Amazon rainforest - 35% of which has already been deforested or degraded, according to the draft of a landmark scientific report published on Wednesday.

Produced by the Science Panel for the Amazon (SPA), the 33-chapter report brings together research on the world's largest rainforest from 200 scientists from across the globe. It is the most detailed assessment of the state of the forest to date and both makes clear the vital role the Amazon plays in global climate and the profound risks it is facing.

Cutting deforestation and forest degradation to zero in less than a decade "is critical," the report said, also calling for massive restoration of already destroyed areas.

The rainforest is a vital bulwark against climate change both for the carbon it absorbs and what it stores.

E.6.2   Treatment: "Before" wording (rationale)

This nonpartisan campaign involves signing up people on Twitter **to make a post encouraging their friends and followers to sign a petition** to immediately stop the destruction of the Amazon rainforest.

The posts will be made public if/when we have finished surveying people in all U.S. counties. This strategy is often used to make campaigns "trend" on Twitter. To coordinate these efforts, we will use the *Tweetability* app you signed into earlier to schedule the posts.

Below, we will ask you if you want to authorize the following Tweet to be posted on your account:

I've joined a campaign to immediately stop the destruction of the Amazon rainforest! Before I joined the campaign, I was shown this article about how 10,000 species risk extinction in Amazon: https://www.reuters.com/business/environment/over-10000-species-risk-extinction-amazon-says-landmark-report-2021-07-14/ Join the campaign and sign the petition: bit.ly/3whrwxT

reuters.com
Over 10,000 species risk extinction in Amazon, says landmark report
More than 10,000 species of plants and animals are at high risk of extinction due to the destruction of the Amazon rainforest - 35% of ...

Do you authorize the *Tweetability* app to schedule the post above to be posted on your account? (If you choose "no," then nothing will be posted on your account.)

○ Yes

○ No

134

E.6.3 Treatment: "After" wording (no rationale)

This nonpartisan campaign involves signing up people on Twitter **to make a post encouraging their friends and followers to sign a petition** to immediately stop the destruction of the Amazon rainforest.

The posts will be made public if/when we have finished surveying people in all U.S. counties. This strategy is often used to make campaigns "trend" on Twitter. To coordinate these efforts, we will use the *Tweetability* app you signed into earlier to schedule the posts.

Below, we will ask you if you want to authorize the following Tweet to be posted on your account:

I've joined a campaign to immediately stop the destruction of the Amazon rainforest! After I joined the campaign, I was shown this article about how 10,000 species risk extinction in Amazon: https://www.reuters.com/business/environment/over-10000-species-risk-extinction-amazon-says-landmark-report-2021-07-14/ Join the campaign and sign the petition: bit.ly/3whrwxT

reuters.com
Over 10,000 species risk extinction in Amazon, says landmark report
More than 10,000 species of plants and animals are at high risk of extinction due to the destruction of the Amazon rainforest - 35% of ...

Do you authorize the *Tweetability* app to schedule the post above to be posted on your account? (If you choose "no," then nothing will be posted on your account.)

○ Yes

○ No

››

135

E.7 Auxiliary Experiment 3: Anticipated persuasion – Democrats

E.7.1 Treatment: "Before" wording (rationale)

This nonpartisan campaign involves signing up people on Twitter **to make a post encouraging their friends and followers to sign a petition** opposing the movement to defund the police.

I have joined a campaign to oppose defunding the police: bit.ly/3DK3UEr. Before joining, I was shown this article written by a Princeton professor on the strong scientific evidence that defunding the police would increase violent crime:
https://www.washingtonpost.com/outlook/2020/06/12/defund-police-violent-crime/

washingtonpost.com
Perspective | Cops prevent violence. But they aren't the only ones wh...
Communities already know how to police their own. Now put them in charge of it.

Suppose you posted the Tweet above on your account. If you had to guess, what percentage of people who saw your Tweet would choose to join the campaign to oppose defunding the police?

| 0 | 10 | 20 | 30 | 40 | 50 | 60 | 70 | 80 | 90 | 100 |

Percentage of people who join

>>

136

E.7.2 Treatment: "After" wording (no rationale)

This nonpartisan campaign involves signing up people on Twitter **to make a post encouraging their friends and followers to sign a petition** opposing the movement to defund the police.

I have joined a campaign to oppose defunding the police: bit.ly/3DK3UEr. After joining, I was shown this article written by a Princeton professor on the strong scientific evidence that defunding the police would increase violent crime: https://www.washingtonpost.com/outlook/2020/06/12 /defund-police-violent-crime/

washingtonpost.com
Perspective | Cops prevent violence. But they aren't the only ones wh...
Communities already know how to police their own. Now put them in charge of it.

Suppose you posted the Tweet above on your account. If you had to guess, what percentage of people who saw your Tweet would choose to join the campaign to oppose defunding the police?

| 0 | 10 | 20 | 30 | 40 | 50 | 60 | 70 | 80 | 90 | 100 |

Percentage of people who join

>>

137

E.8 Auxiliary Experiment 4: Open-ended explanations of preferred anti-defunding Tweet – Democrats

E.8.1 Pre-treatment questions

On the next page, you will be provided with a recent Washington Post column written by **Princeton Professor of Criminology Patrick Sharkey**, in which he discusses evidence showing that more policing leads to less violent crime.

››

Outlook

Why do we need the police?

Cops prevent violence. But they aren't the only ones who can do it.

By **Patrick Sharkey**
JUNE 12, 2020

The calls to end policing as we know it contain a sort of trap. The best evidence we have makes clear that police are effective in reducing violence, and without designating some group to combat this problem, efforts to weaken them through budget cuts — "defund the police" — are likely to have unanticipated consequences and to destabilize communities. In many cities this is likely to lead to a rise in violence. And research shows that, when violence increases, Americans of all races become more punitive, supporting harsher policing and criminal justice policies. That's how we got to this point.

Patrick Sharkey
@patrick_sharkey is a professor of sociology and public affairs at Princeton University's Woodrow Wilson School of Public and International Affairs. His most recent book is "Uneasy Peace: The Great Crime Decline, the Renewal of City Life, and the Next War on Violence."

Yet none of this means that the police, which have served as an institution of racialized control throughout our nation's history, are the only group capable of reducing violence. Community leaders and residents have proved adept at overseeing their neighborhoods, caring for their populations and maintaining safe streets. Studies show that this work lowers crime, sometimes dramatically. What happens if we put those people *in charge* of containing violence, too?

Over the past 10 years, an expanding body of research has shown just how damaging violence is to community life, children's academic trajectories and healthy child development. We have rigorous, causal evidence that every shooting in a neighborhood affects children's sleep and their ability to focus and learn. When a neighborhood becomes violent, it begins to fall apart, as public spaces empty, businesses close, parks and playgrounds turn dangerous, and families try to move elsewhere. Violence is the fundamental challenge for cities: Nothing works if public space is unsafe.

Those who argue that the police have no role in maintaining safe streets are arguing against lots of strong evidence. One of the most robust, most uncomfortable findings in criminology is that putting more officers on the street leads to less violent crime. We know this from randomized experiments involving "hot spots policing" and natural experiments in which more officers were brought to the streets because of something other than crime — a shift in the terror alert level or the timing of a federal grant — and violent crime fell. After the unrest around the deaths of Freddie Gray in Baltimore and Michael Brown in Ferguson, Mo., police officers stepped back from their duty to protect and serve; arrests for all kinds of low-level offenses dropped, and violence rose. This shouldn't be interpreted to mean that protests against violent policing lead to more violence; rather, it means that when police don't do their jobs, violence often results.

Considered alongside the brutal response to protests over the past few weeks, this evidence forces us to hold two incongruent ideas: Police are effective at reducing violence, the most damaging feature of urban inequality. And yet one can argue that law enforcement is an authoritarian institution that historically has inflicted violence on black people and continues to do so today.

Imagine that at this point in the study, you indicated that you wanted to join a campaign that opposes the movement to defund the police.

››

Imagine that you successfully have joined the campaign.

Since you joined the campaign, we wanted to give you more time reading the Washington Post column written by **Princeton Professor of Criminology Patrick Sharkey**, where he discusses evidence showing that more policing leads to less violent crime.

The article is available on the next page, and you can spend as much time (or as little time) as you want reading it before you continue with the remaining part of the survey.

Why do we need the police?

Cops prevent violence. But they aren't the only ones who can do it.

Mixed Card for The Washington Post

By **Patrick Sharkey**

JUNE 12, 2020

The calls to end policing as we know it contain a sort of trap. The best evidence we have makes clear that police are effective in reducing violence, and without designating some group to combat this problem, efforts to weaken them through budget cuts — "defund the police" — are likely to have unanticipated consequences and to destabilize communities. In many cities this is likely to lead to a rise in violence. And research shows that, when violence increases, Americans of all races become more punitive, supporting harsher policing and criminal justice policies. That's how we got to this point.

Patrick Sharkey
@patrick_sharkey is a professor of sociology and public affairs at Princeton University's Woodrow Wilson School of Public and International Affairs. His most recent book is "Uneasy Peace: The Great Crime Decline, the Renewal of City Life, and the Next War on Violence."

Yet none of this means that the police, which have served as an institution of racialized control throughout our nation's history, are the only group capable of reducing violence. Community leaders and residents have proved adept at overseeing their neighborhoods, caring for their populations and maintaining safe streets. Studies show that this work lowers crime, sometimes dramatically. What happens if we put those people *in charge* of containing violence, too?

Over the past 10 years, an expanding body of research has shown just how damaging violence is to community life, children's academic trajectories and healthy child development. We have rigorous, causal evidence that every shooting in a neighborhood affects children's sleep and their ability to focus and learn. When a neighborhood becomes violent, it begins to fall apart, as public spaces empty, businesses close, parks and playgrounds turn dangerous, and families try to move elsewhere. Violence is the fundamental challenge for cities: Nothing works if public space is unsafe.

Those who argue that the police have no role in maintaining safe streets are arguing against lots of strong evidence. One of the most robust, most uncomfortable findings in criminology is that putting more officers on the street leads to less violent crime. We know this from randomized experiments involving "hot spots policing" and natural experiments in which more officers were brought to the streets because of something other than crime — a shift in the terror alert level or the timing of a federal grant — and violent crime fell. After the unrest around the deaths of Freddie Gray in Baltimore and Michael Brown in Ferguson, Mo., police officers stepped back from their duty to protect and serve; arrests for all kinds of low-level offenses dropped, and violence rose. This shouldn't be interpreted to mean that protests against violent policing lead to more violence; rather, it means that when police don't do their jobs, violence often results.

Considered alongside the brutal response to protests over the past few weeks, this evidence forces us to hold two incongruent ideas: Police are effective at reducing violence, the most damaging feature of urban inequality. And yet one can argue that law enforcement is an authoritarian institution that historically has inflicted violence on black people and continues to do so today.

E.8.2 Treatment: "Before" wording (rationale)

As part of the campaign, we plan to ask people **to make a post encouraging their friends and followers to sign a petition** opposing the movement to defund the police.

Imagine that you had joined the campaign. If you were going to post **one** of the following two Tweets on your Twitter account, which would you prefer to post?

Tweet A

I have joined a campaign to oppose defunding the police: https://bit.ly/3DK3UEr.

Tweet B

I have joined a campaign to oppose defunding the police: https://bit.ly/3DK3UEr. Before joining, I was shown this article written by a Princeton professor on the strong scientific evidence that defunding the police would increase violent crime: https://www.washingtonpost.com/outlook/2020/06/12/defund-police-violent-crime/

Which of the above Tweets would you have preferred to post on your account?

○ Tweet A

○ Tweet B

Please explain why you chose this Tweet rather than the other Tweet.

››

142

E.8.3 Treatment: "After" wording (no rationale)

As part of the campaign, we plan to ask people **to make a post encouraging their friends and followers to sign a petition** opposing the movement to defund the police.

Imagine that you had joined the campaign. If you were going to post **one** of the following two Tweets on your Twitter account, which would you prefer to post?

Tweet A

I have joined a campaign to oppose defunding the police: https://bit.ly/3DK3UEr.

Tweet B

I have joined a campaign to oppose defunding the police: https://bit.ly/3DK3UEr. After joining, I was shown this article written by a Princeton professor on the strong scientific evidence that defunding the police would increase violent crime: https://www.washingtonpost.com/outlook/2020/06/12/defund-police-violent-crime/

Which of the above Tweets would you have preferred to post on your account?

○ Tweet A

○ Tweet B

Please explain why you chose this Tweet rather than the other Tweet.

››

143

E.9 Auxiliary Experiments 5: Interpretation of dissent with low-credibility rationale

E.9.1 Treatment: "Before" condition (rationale)

We conducted a survey about political and social attitudes in the United States earlier this year. You have been matched with one of the participants from that survey.

Your matched participant **joined a campaign to oppose defunding the police**.

Your matched participant **also agreed to post the Tweet below** on their Twitter profile to promote the campaign.

> I have joined a campaign to oppose defunding the police: bit.ly/3DK3UEr. Before joining, I was shown this article that argues that defunding the police would increase violent crime:
> https://www.washingtonpost.com/outlook/2020/06/12/defund-police-violent-crime/
>
>
>
> washingtonpost.com
> Perspective | Cops prevent violence. But they aren't the only ones who can do it. Communities already know how to police their own. Now put them in charge of it.

Why do you think your matched respondent chose to join the campaign to oppose defunding the police?

→

144

Matched Respondent's Donation Decision

We gave your matched respondent the opportunity to donate $5 to the **National Association for the Advancement of Colored People (NAACP)**, America's oldest and largest civil rights organization.

Below, we will ask you to guess whether or not your matched respondent donated $5 to the National Association for the Advancement of Colored People (NAACP).

Reminder: Your matched participant agreed to post the Tweet below on their Twitter account.

I have joined a campaign to oppose defunding the police: bit.ly/3DK3UEr. Before joining, I was shown this article that argues that defunding the police would increase violent crime:
https://www.washingtonpost.com/outlook/2020/06/12/defund-police-violent-crime/

washingtonpost.com
Perspective | Cops prevent violence. But they aren't the only ones who can do it. Communities already know how to police their own. Now put them in charge of it.

Do you think that your matched participant chose to donate $5 to the National Association for the Advancement of Colored People (NAACP)?

○ Yes, I think my matched respondent chose to donate

○ No, I think my matched respondent **did not** choose to donate

→

145

You now have the opportunity to authorize a $1 bonus payment to your matched respondent. **The bonus payment will not be deducted from your payment.** Your matched respondent did not know you would have the opportunity to decide their bonus.

Reminder: Your matched participant agreed to post the Tweet below on their Twitter account.

I have joined a campaign to oppose defunding the police: bit.ly/3DK3UEr. Before joining, I was shown this article that argues that defunding the police would increase violent crime:
https://www.washingtonpost.com/outlook/2020/06/12/defund-police-violent-crime/

washingtonpost.com
Perspective | Cops prevent violence. But they aren't the only ones who can do it. Communities already know how to police their own. Now put them in charge of it.

Do you want to authorize a $1 bonus to your matched respondent?

○ Yes, I would like to authorize a $1 bonus

○ No, I would not like to authorize a $1 bonus

→

E.9.2 Treatment: "After" condition (no rationale)

We conducted a survey about political and social attitudes in the United States earlier this year. You have been matched with one of the participants from that survey.

Your matched participant **joined a campaign to oppose defunding the police**.

Your matched participant **also agreed to post the Tweet below** on their Twitter profile to promote the campaign.

I have joined a campaign to oppose defunding the police: bit.ly/3DK3UEr. After joining, I was shown this article that argues that defunding the police would increase violent crime:
https://www.washingtonpost.com/outlook/2020/06/12/defund-police-violent-crime/

washingtonpost.com
Perspective | Cops prevent violence. But they aren't the only ones who can do it. Communities already know how to police their own. Now put them in charge of it.

Why do you think your matched respondent chose to join the campaign to oppose defunding the police?

→

Matched Respondent's Donation Decision

We gave your matched respondent the opportunity to donate $5 to the **National Association for the Advancement of Colored People (NAACP)**, America's oldest and largest civil rights organization.

Below, we will ask you to guess whether or not your matched respondent donated $5 to the National Association for the Advancement of Colored People (NAACP).

Reminder: Your matched participant agreed to post the Tweet below on their Twitter account.

I have joined a campaign to oppose defunding the police: bit.ly/3DK3UEr. After joining, I was shown this article that argues that defunding the police would increase violent crime:
https://www.washingtonpost.com/outlook/2020/06/12/defund-police-violent-crime/

washingtonpost.com
Perspective | Cops prevent violence. But they aren't the only ones who can do it. Communities already know how to police their own. Now put them in charge of it.

Do you think that your matched participant chose to donate $5 to the National Association for the Advancement of Colored People (NAACP)?

○ Yes, I think my matched respondent chose to donate

○ No, I think my matched respondent **did not** choose to donate

→

148

You now have the opportunity to authorize a $1 bonus payment to your matched respondent. **The bonus payment will not be deducted from your payment.** Your matched respondent did not know you would have the opportunity to decide their bonus.

Reminder: Your matched participant agreed to post the Tweet below on their Twitter account.

I have joined a campaign to oppose defunding the police: bit.ly/3DK3UEr. After joining, I was shown this article that argues that defunding the police would increase violent crime:
https://www.washingtonpost.com/outlook/2020/06/12/defund-police-violent-crime/

washingtonpost.com
Perspective | Cops prevent violence. But they aren't the only ones who can do it. Communities already know how to police their own. Now put them in charge of it.

Do you want to authorize a $1 bonus to your matched respondent?

○ Yes, I would like to authorize a $1 bonus

○ No, I would not like to authorize a $1 bonus

→

149

E.10 Auxiliary Experiment 6: Persuasion experiment – Republicans

E.10.1 Pre-treatment beliefs

Please see the short video below where Fox News host **Tucker Carlson presents evidence on whether illegal immigrants commit more crime**.

E.10.2 Information treatment (only shown to respondents in the treatment group)

To what extent do you agree with the following statement: "The United States should immediately deport all illegal Mexican immigrants."

○ Strongly agree

○ Agree

○ Neither agree nor disagree

○ Disagree

○ Strongly disagree

→

E.10.3 Post-treatment outcomes

To what extent do you agree with the following statement: "Illegal immigrants are not much more likely to commit serious crimes than U.S. citizens."

○ Strongly agree

○ Agree

○ Neither agree nor disagree

○ Disagree

○ Strongly disagree

→

E.11 Auxiliary Experiment 7 (wave 2): Expression of dissent with visible donation

E.11.1 Pre-treatment questions

Do Illegal Immigrants Commit More Crime? Preliminary evidence from Dr. Lott's study

Economist Dr. John R. Lott carried out a study on the relationship between illegal immigration and crime. **The study has not been published in an academic journal, but we summarize some of its findings here.**

Dr. Lott's study uses **detailed data spanning over two decades from the Arizona state prison systems**. He finds that **"undocumented immigrants are at least 142% more likely** to be convicted of a crime than other Arizonans. They also tend to commit **more serious crimes**," such as murder and rape.

Moreover, "young undocumented immigrants commit crime at twice the rate of young U.S. citizens." He calculates that "if undocumented immigrants committed crime nationally as they do in Arizona, in 2016 they would have been responsible for over **1,000 more murders, 5,200 rapes, 8,900 robberies, 25,300 aggravated assaults, and 26,900 burglaries**." The vast majority of the illegal immigrants Lott studies are from Mexico.

However, a number of sources (including a researcher affiliated with the Cato Institute, a non-partisan libertarian think tank) have recently challenged some of the study's methods, claiming that errors in analysis invalidate its results. Dr. Lott has responded to this critique with a defense of the study's methods, but the issue remains unresolved.

Do you think Dr. Lott's study will generate widespread media attention?

○ Yes

○ No

On the next screen, you will have the opportunity to authorize a $1 donation to an organization that seeks to reduce illegal immigration into the United States. We will make the donation on your behalf, so **it will not be deducted from your payment**.

›››

Donation decision

You now have the opportunity to authorize a $1 donation to **Fund the Wall**.

Fund the Wall is a nonprofit organization that seeks to reduce illegal immigration into the United States by raising money for the construction of the US-Mexico border wall. If you wish, you can choose to authorize a $1 donation to Fund the Wall. The donation **will not be deducted from your payment.**

Before proceeding, please confirm your consent to the following:

I consent to the researchers requesting the following information from my study provider: first and last name and **city. We will not have access to any other personally-identifying information, and we will use the information for study purposes only.**

○ Yes, I consent and would like to proceed with the survey.

○ No, I do not consent and would like to terminate the survey now.

»

E.11.2 *No Rationale* **condition**

Publicizing the results of our study

As researchers, we believe it is important to communicate our findings about political and social attitudes in Bergen to the public.

If/when Dr. Lott's study is published in an academic journal, we will post the results from this survey, **including your individual donation decision and the donation decisions of all of the other respondents to this survey who consented on the previous page**, on our website.

We will promote our website via Facebook ads to Bergen residents.

What website visitors will learn

We will post your individual donation decision on our website, which will also include a description of Dr. Lott's study.

- The web page will state that **all participants were surveyed before Dr. Lott's study was published in an academic journal**
- The page lists individual donation decisions: whether or not each participant decided to authorize the donation to Fund the Wall

Donation decision

Would you like to authorize a $1 donation to **Fund the Wall**?

○ Yes, I would like to authorize a $1 donation

○ No, I would not like to authorize a $1 donation

Recall what people will learn when visiting the website:

- The web page will state that **all participants were surveyed before Dr. Lott's study was published in an academic journal**
- The page lists individual donation decisions: whether or not each participant decided to authorize the donation to Fund the Wall

››

E.11.3 *Rationale* condition

Publicizing the results of our study

As researchers, we believe it is important to communicate our findings about political and social attitudes in Bergen to the public.

If/when Dr. Lott's study is published in an academic journal, we will post the results from this survey, **including your individual donation decision and the donation decisions of all of the other respondents to this survey who consented on the previous page**, on our website.

We will promote our website via Facebook ads to Bergen residents.

What website visitors will learn

We will post your individual donation decision on our website, which will also include a description of Dr. Lott's study.

- The web page will state that **all participants were shown the preliminary findings from Dr. Lott's study** before deciding whether or not to donate to Fund the Wall
- The page lists individual donation decisions: whether or not each participant decided to authorize the donation to Fund the Wall

Donation decision

Would you like to authorize a $1 donation to **Fund the Wall**?

○ Yes, I would like to authorize a $1 donation

○ No, I would not like to authorize a $1 donation

Recall what people will learn when visiting the website:

- The web page will state that **all participants were shown the preliminary findings from Dr. Lott's study** before deciding whether or not to donate to Fund the Wall
- The page lists individual donation decisions: whether or not each participant decided to authorize the donation to Fund the Wall

>>

E.12 Auxiliary Experiment 8: Interpretation of dissent with visible donation

E.12.1 Pre-treatment information

Do Illegal Immigrants Commit More Crime? Evidence from Dr. Lott's Study

Dr. John R. Lott, an economist formerly employed at top institutions such as Yale University and the University of Chicago, carried out a study on the relationship between illegal immigration and crime using new high-quality data. **The study has not yet been published in an academic journal, but we obtained an early version and summarize the results below.**

Dr. Lott's study uses **detailed data spanning over two decades from the Arizona state prison systems**. He finds that **"undocumented immigrants are at least 142% more likely** to be convicted of a crime than other Arizonans. They also tend to commit **more serious crimes**," such as murder and rape.

Moreover, "young undocumented immigrants commit crime at twice the rate of young U.S. citizens." He calculates that "if undocumented immigrants committed crime nationally as they do in Arizona, in 2016 they would have been responsible for over **1,000 more murders, 5,200 rapes, 8,900 robberies, 25,300 aggravated assaults, and 26,900 burglaries**. The vast majority of the illegal immigrants Lott studies are from Mexico.

However, a number of sources (including a researcher affiliated with the Cato Institute, a non-partisan think tank) have recently challenged some of the study's methods, claiming that errors in analysis invalidate its results. Dr. Lott has responded to this critique with a defense of the study's methods, but the issue remains unresolved.

›››

160

E.12.2 *No Rationale* condition

We conducted a survey about political and social attitudes in the United States earlier this year. You have been matched with one of the respondents from that survey.

We gave your matched respondent the opportunity to authorize a $1 donation to **Fund the Wall**, a nonprofit organization that seeks to reduce illegal immigration into the United States by helping to fund and construct the US-Mexico border wall. Your matched respondent was told that their donation decision would be posted on our website. The decision on whether to authorize the donation did not have any financial consequences for your matched respondent.

Some respondents were assigned a longer version of the survey and learned about Dr. Lott's study before they decided whether or not to donate. Other respondents were assigned a shorter version of the study and **were not informed** about Dr. Lott's study before they decided whether or not to donate.

Information about your matched respondent

- Your matched respondent **was not informed about Dr. Lott's study**, which finds that illegal immigrants commit more crimes than US citizens
- Your matched respondent decided to authorize the $1 donation to Fund the Wall

››

Why do you think your matched respondent chose to donate to Fund the Wall?

[ ]

Reminder: Information about your matched respondent

- Your matched respondent **was not informed about Dr. Lott's study**, which finds that illegal immigrants commit more crimes than US citizens
- Your matched respondent decided to authorize the $1 donation to Fund the Wall

››

After your matched respondent made their donation decision, they completed the **The Gullibility Scale**, a short questionnaire which measures **how easily people are manipulated by evidence from untrustworthy sources**.

On the next page, we will ask you to guess how your matched respondent scored on this scale. If you guess the correct option, you will be entered into a lottery for a $50 Amazon gift card.

››

The Gullibility Scale

We administered **The Gullibility Scale**, a short questionnaire which measures **how easily people are manipulated by evidence from untrustworthy sources**, to your matched respondent.

The test is scored from 0 to 100, where 0 means "least gullible" and 100 means "most gullible". Thus, a higher score indicates that your matched respondent is more gullible.

Reminder: Information about your matched respondent

- Your matched respondent **was not** informed about Dr. Lott's study, which finds that illegal immigrants commit more crimes than US citizens
- Your matched respondent decided to authorize the $1 donation to Fund the Wall

If you had to guess, how do you think your **matched respondent** scored on **The Gullibility Scale**?

○ Score between 0 and 10 (**Not at all gullible**)

○ Score between 10 and 20

○ Score between 20 and 30

○ Score between 30 and 40

○ Score between 40 and 50

○ Score between 50 and 60

○ Score between 60 and 70

○ Score between 70 and 80

○ Score between 80 and 90

○ Score between 90 and 100 (**Extremely gullible**)

›>

After your matched respondent made their donation decision, they completed the **Foreign Culture Tolerance Scale**, a short questionnaire which measures **tolerance toward foreign values and traditions**.

On the next page, we will ask you to guess how your matched respondent scored on this scale. If you guess the correct option, you will be entered into a lottery for a $50 Amazon gift card.

>>

The Foreign Culture Tolerance Scale

We administered the **Foreign Culture Tolerance Scale**, a short questionnaire which measures tolerance toward **foreign values and traditions**, to your matched respondent.

The test is scored from 0 to 100, where 0 means "least tolerant" and 100 means "most tolerant". Thus, a **higher score indicates that your matched respondent is more tolerant toward foreign values and traditions.**

Reminder: Information about your matched respondent

- Your matched respondent **was not informed about Dr. Lott's study**, which finds that illegal immigrants commit more crimes than US citizens
- Your matched respondent decided to authorize the $1 donation to Fund the Wall

If you had to guess, how do you think your **matched respondent** scored on the **Foreign Culture Tolerance Scale**?

○ Score between 0 and 10 (**Not at all tolerant**)

○ Score between 10 and 20

○ Score between 20 and 30

○ Score between 30 and 40

○ Score between 40 and 50

○ Score between 50 and 60

○ Score between 60 and 70

○ Score between 70 and 80

○ Score between 80 and 90

○ Score between 90 and 100 (**Extremely tolerant**)

››

165

E.12.3 *Rationale* condition

We conducted a survey about political and social attitudes in the United States earlier this year. You have been matched with one of the respondents from that survey.

We gave your matched respondent the opportunity to authorize a $1 donation to **Fund the Wall**, a nonprofit organization that seeks to reduce illegal immigration into the United States by helping to fund and construct the US-Mexico border wall. Your matched respondent was told that their donation decision would be posted on our website. The decision on whether to authorize the donation did not have any financial consequences for your matched respondent.

Some respondents were assigned a longer version of the survey and learned about Dr. Lott's study before they decided whether or not to donate. Other respondents were assigned a shorter version of the study and **were not informed** about Dr. Lott's study before they decided whether or not to donate.

Information about your matched respondent

- Your matched respondent **was informed about Dr. Lott's study**, which finds that illegal immigrants commit more crimes than US citizens
- Your matched respondent then decided to authorize the $1 donation to Fund the Wall

››

Why do you think your matched respondent chose to donate to Fund the Wall?

[text input box]

Reminder: Information about your matched respondent

- Your matched respondent **was informed about Dr. Lott's study**, which finds that illegal immigrants commit more crimes than US citizens
- Your matched respondent then decided to authorize the $1 donation to Fund the Wall

››

After your matched respondent made their donation decision, they completed the **The Gullibility Scale**, a short questionnaire which measures **how easily people are manipulated by evidence from untrustworthy sources**.

On the next page, we will ask you to guess how your matched respondent scored on this scale. If you guess the correct option, you will be entered into a lottery for a $50 Amazon gift card.

››

The Gullibility Scale

We administered **The Gullibility Scale**, a short questionnaire which measures **how easily people are manipulated by evidence from untrustworthy sources**, to your matched respondent.

The test is scored from 0 to 100, where 0 means "least gullible" and 100 means "most gullible". Thus, a higher score indicates that your matched respondent is more gullible.

Reminder: Information about your matched respondent

- Your matched respondent **was informed about Dr. Lott's study**, which finds that illegal immigrants commit more crimes than US citizens
- Your matched respondent then decided to authorize the $1 donation to Fund the Wall

If you had to guess, how do you think your **matched respondent** scored on **The Gullibility Scale**?

○ Score between 0 and 10 (**Not at all gullible**)

○ Score between 10 and 20

○ Score between 20 and 30

○ Score between 30 and 40

○ Score between 40 and 50

○ Score between 50 and 60

○ Score between 60 and 70

○ Score between 70 and 80

○ Score between 80 and 90

○ Score between 90 and 100 (**Extremely gullible**)

››

After your matched respondent made their donation decision, they completed the **Foreign Culture Tolerance Scale**, a short questionnaire which measures **tolerance toward foreign values and traditions**.

On the next page, we will ask you to guess how your matched respondent scored on this scale. If you guess the correct option, you will be entered into a lottery for a $50 Amazon gift card.

The Foreign Culture Tolerance Scale

We administered the **Foreign Culture Tolerance Scale**, a short questionnaire which measures tolerance toward **foreign values and traditions**, to your matched respondent.

The test is scored from 0 to 100, where 0 means "least tolerant" and 100 means "most tolerant". Thus, a **higher score indicates that your matched respondent is more tolerant toward foreign values and traditions.**

Reminder: Information about your matched respondent

- Your matched respondent **was informed about Dr. Lott's study**, which finds that illegal immigrants commit more crimes than US citizens
- Your matched respondent then decided to authorize the $1 donation to Fund the Wall

If you had to guess, how do you think your **matched respondent** scored on the **Foreign Culture Tolerance Scale**?

○ Score between 0 and 10 (**Not at all tolerant**)

○ Score between 10 and 20

○ Score between 20 and 30

○ Score between 30 and 40

○ Score between 40 and 50

○ Score between 50 and 60

○ Score between 60 and 70

○ Score between 70 and 80

○ Score between 80 and 90

○ Score between 90 and 100 (**Extremely tolerant**)

>>