Can bad norms be good?*

Jesper Akesson, Rena Conti, Robert W. Hahn,
Robert D. Metcalfe and Itzhak Rasooly†

September 2026

Abstract

Descriptive social norm interventions have been used to address a wide range of policy challenges, from encouraging resource conservation to promoting female labor force participation. In this paper, we report the results of a natural field experiment ($n = 5,360$) that randomly exposes participants to different signals about the share of others who have signed up as advocates for a charity. Consistent with prior work, we find that suggesting that everyone else has signed up increases the likelihood that participants themselves sign up as advocates. However, we also show that suggesting that no one else has signed up increases sign-up rates. We replicate these findings in a follow-up experiment ($n = 5,042$) that investigates why it can be effective to reveal 'bad' norms.

KEYWORDS: social norms, field experiment, beliefs

*We would especially like to thank Elisa Weiss, Maria Sae-Hau, Gabriela Gracia, and the team at the Leukemia & Lymphoma Society for helping us design and conduct the first experiment. We would also like to thank Filippo Muzi-Falconi, Hannah Chappell, Ben Harrison, and Ondrej Kacha for excellent research assistance, as well as Nikhil Kalyanpur, Luke Milsom, JingKai Ong, Chris Roth, Toby Shevlane and Jasmine Theilgaard for helpful comments and suggestions. Boston University IRB: A231602CD; AEA RCT registry: AEARCTR-0007938 and AEARCTR-0012528.

†Akesson, King's College London; Conti, Boston University; Hahn, Carnegie Mellon University and Technology Policy Institute; Metcalfe, Columbia University and NBER; Rasooly, City St George's, University of London.

1. Introduction

Descriptive social norm interventions attempt to influence behavior by providing information about the prevalence of an activity. They have been used in a variety of areas and policy domains, including tax compliance (Hallsworth et al., 2017; Perez-Truglia and Troiano, 2018), online reviews (Chen et al., 2010), public health (Borsari and Carey, 2003), voting (Gerber and Rogers, 2009), female labor force participation (Bursztyn et al., 2020), savings (Goldin et al., 2020), waste (Ek and Söderberg, 2024), and resource conservation (Allcott, 2011; Ferraro and Price, 2013; Brent et al., 2015; Brandon et al., 2017, 2019). This body of work reveals that norm-based interventions can exert a powerful influence on behavior, which in turn has encouraged their adoption by policymakers across the globe (Benartzi et al., 2017; OECD, 2017; Sunstein et al., 2018; DellaVigna and Linos, 2022).

Despite the large number of papers on this topic, at least two prominent gaps remain. First, there are no descriptive social norm experiments that examine the effect of suggesting that no one else is doing an activity. Some scholars argue that descriptive norm interventions should not be used in such contexts (*i.e.,* when the norm is 'bad') as they might reinforce undesirable behaviors (see, *e.g.,* Bicchieri and Dimant, 2019). In fact, descriptive social norm interventions in low-compliance environments have been branded a 'big mistake' by prominent academics and policy advisors (see, *e.g.*, Halpern, 2015). However, there are theoretical reasons why it may be beneficial to convey information about low levels of compliance. For example, it is possible that an individual believes that the impact of their action is more important when a low percentage of other individuals is also engaging in an activity. Alternatively, decreasing beliefs regarding the share of individuals that engage in an altruistic activity might trigger feelings of pity or guilt, while also increasing the perceived 'self-image' benefits from engaging in that activity (Bénabou and Tirole, 2006; Falk, 2020).

Second, very few studies examine the relative effects of sending different exogenous signals about the prevalence of a behavior (*e.g.,* that 80% rather than 90% pay their taxes on time). This matters since comparing different norm levels helps distinguish between competing theories. In models of social learning and conformism (Bikhchandani et al., 1992; Banerjee, 1992; Akerlof, 1980), higher signals are expected to increase 'pro-social' behaviour. In models of strategic substitution (Diekmann, 1985; Bergstrom et al., 1986; Duncan, 2004), however, higher signals are expected to do the reverse. Comparing responses across signal levels therefore provides an empirical test between these competing mechanisms. From a practical perspective, it is also informative to test norm interventions that convey different signals (relative to a control group) to identify when to use them. In particular, it can reveal whether descriptive norm interventions are effective only in certain contexts, as some have suggested, or whether they have broader applicability.

In this paper, we address both of the aforementioned gaps. To do this, we conducted a natural field experiment ($n = 5,360$) in partnership with the Leukemia & Lymphoma Society (LLS). Members of the society were randomly placed into one of six treatment groups or a control group. Each of the treatment groups was given a different non-deceptive signal about the share of members who had signed up as advocates for the organization (ranging from 0% to 100% in 20% increments).[1] Subjects were then offered the opportunity to sign up as advocates, which LLS views as key to achieving its goals. Finally, we asked participants whether they believed that they could make a difference if they signed up as advocates. This design allows us to study precisely how sign-up rates vary with the information conveyed by the intervention. By collecting data on individual beliefs both before and after the intervention, we are also able to explore the mechanisms through which the observed effects occur.

This first experiment yields three main findings. First, we find a 'U-shaped' relationship between the signal about prevalence conveyed and the share that signs up. More specifically, as one might expect from the previous literature, we find that the signal suggesting that everyone else signs up increases the share of participants who sign up (the share rises from 22% in the control group to 30%). However, we also find that the signal suggesting that no one else signs up boosts sign-up rates from 22% to 26%. Moreover, while the middling signals (20-80%) fail to change behavior relative to the control group, they do not backfire, even though some revise beliefs downward regarding the share of other participants that signed up. The latter findings contrast with the received wisdom that revealing 'bad' norms is necessarily a mistake.

Second, we use an instrumental variables identification strategy to decompose these treatment effects into: (1) the effect of the signals on beliefs about prevalence, and (2) the effect of beliefs about prevalence on sign-up behavior. We show that higher signals about sign-up rates lead to higher beliefs about sign-up rates, and that this effect operates in a roughly linear way. However, we also show that beliefs about sign-up rates influence actual sign-up rates in a U-shaped manner: just as very high beliefs about prevalence boost sign ups, so do very low beliefs about prevalence. It is thus the combination of the linear effect of treatments on beliefs with the U-shaped effect of beliefs on sign ups that generates the U-shaped treatment effects discussed above.

Next, we examine whether our 'bad' norm works by making individuals believe that their

[1] LLS advocates work with the society to influence policy at federal and state levels. Advocates may be asked to attend events or write letters to their members of Congress. See `https://www.lls.org/policy-advocacy` for more information.

actions are more impactful. This is a very natural conjecture given the familiar logic of diminishing marginal returns: one might think that, the fewer others who sign up, the greater the benefit generated by one's own decision to sign up. This leads to our third finding: the 'bad' norm actually makes people believe that their actions are less efficacious, consistent with a model of social learning (Banerjee, 1992; Bikhchandani et al., 1992). For this reason, an efficacy channel cannot explain the success of our 'bad' norm (although it might help explain the success of the 100% norm).

To study the mechanisms underlying our results in more detail—as well as to verify whether the main effects can be replicated—we conducted a second experiment ($n = 5,042$) using a sample recruited via Prolific. The second experiment closely mirrored the first, with a few important differences. First, subjects were given a series of questions to assess mechanisms after they chose whether to sign up as an advocate. For example, we measured whether the signals had effects on whether they felt bad, the perceived efficacy of signing up, and participants' self-image. This allowed us to examine the impact of sending different signals about prevalence on these (potentially important) mediating variables. Second, to examine whether the impacts are driven by the observability of subjects' decisions—which was a feature of the first experiment—we also varied whether the charity was informed of the participant's decision and provided with their contact details.

We begin by showing that, despite the differences across the two experiments (*e.g.*, in subject pools), the main results from the first experiment also arise in the second. First, we find evidence of a 'U-shaped' relationship between the signal values conveyed and subjects' willingness to sign up as advocates. In particular, the estimates again reveal that 'bad norms' can be good—the 0% intervention increases sign-up rates by 4 percentage points (from 22.6% in the control to 26.5%). Second, the signals have similar effects on participants' beliefs about prevalence. Third, we can again reject that an 'efficacy' channel is underlying the success of the 0% intervention, and we do not find that participants' decision to sign up depends on whether their decision is made public to the charity or not. These results suggest that our field-experimental findings may generalize beyond their original setting.

Next, we find that the 0% signal makes people feel bad that an insufficient number of people have signed up, while the other signals do not affect this outcome. In contrast, we do not find meaningful effects of the 0% signal on several other potentially mediating variables, such as perceived efficacy or the public-image consequences of signing up. We also analyze participants' open-ended explanations for their decisions and find that the 0% signal increases references to feeling bad for the organization. These results suggest that negative emotions are one important mechanism underlying the effect of the 0% signal.[2]

Taken together, these findings suggest that descriptive social norm interventions are both more complicated and more broadly applicable than previous work might suggest. In particular, we show that 'bad' norms can be good, an effect that we observe both in our natural field experiment and in our follow-up study. This result suggests that it is worthwhile to test the impact of norms in low-compliance environments—which are, after all, very common. Based on our results, one may also suspect that such interventions will be most promising in contexts where disclosing very low rates of 'pro-social' activity is likely to make potential supporters feel bad. In contrast, it seems unlikely that such an intervention would work in a setting like tax compliance: individuals are unlikely to feel bad for a tax authority, and may react adversely to the news that many do not pay tax but nonetheless avoid punishment.

Related literature. This paper contributes to the large literature on the impact of descriptive social norm interventions (see Miller and Prentice, 2016 and Wenzel and Woodyatt, 2025 for reviews). In contrast to this literature, we explore the impact of *multiple* different norm levels and include a 0% (*i.e.*, 'bad') norm. Our findings also speak to the theoretical literature on how individual behavior is shaped by beliefs about the behavior of others. Importantly, the U-shaped relationship that we document between signals and behavior is not consistent with many 'single-factor' models of norms. More specifically, theories of social learning (Bikhchandani et al., 1992; Banerjee, 1992) predict an increasing relationship between prevalence and behavior, whereas theories of strategic substitution (Bergstrom et al., 1986; Duncan, 2004) predict a decreasing relationship. While it is clear that neither of these mechanisms alone can explain our results, our results are consistent with a mix of social learning (as evidenced by our results of perceived efficacy), along with negative emotions that are activated by the very low signals.

Several papers are methodologically related to ours and merit closer discussion. In an influential paper, Frey and Meier (2004) examine the impact of informing students that either 64% or 46% of their fellow students contributed to university social funds. They find that the higher signal increases contribution rates, consistent with models of conditional cooperation. Our paper differs from Frey and Meier (2004) in several important ways. Most importantly, they vary the descriptive norm only between two intermediate values, whereas we vary it across the full range from 0 to 100 percent. As a result, unlike Frey and Meier (2004), we can study the effect of an explicitly bad norm and test for non-monotonic responses.

[2]While many psychologists have argued for the role of such emotions in decision-making (Dijker and Koomen, 2003; Lunardo and Bezençon, 2015; Lantos et al., 2020), the role of emotions more generally in economic decision-making has only recently received sustained attention in economics (Loewenstein, 2000; Coricelli et al., 2010; Card and Dahl, 2011; Butera et al., 2022; Bernheim et al., 2024).

Relatedly, Martin and Randal (2008) run a natural field experiment on charitable giving by varying the amount of money visible in a transparent donation box. Most relevant for our purposes is their "empty box" treatment, which is somewhat analogous to our 0% signal if participants interpret an empty box as indicating no prior donations rather than, say, a recently emptied box. However, unlike our design, they do not provide numerical information about the share of previous donors. At best, participants might try to infer these statistics from the visual clues with which they are provided (*e.g.*, the number of coins in the box and the number of bills) along with their beliefs about the frequency with which the box is replaced and the amount that each participant donates. For this reason, the experiment in Martin and Randal (2008) is rather different to ours.

More recently, Linek and Traxler (2021) study the impact of sending different signals about the prevalence of donating to Wikipedia. While many of their treatments vary the framing of a given piece of social information, some treatments instead vary the information itself. Most relevant for our purposes, their Trial 5 compares messages stating that 'less than 0.1%' versus 'less than 1%' of users donate. In contrast to our experiment, these two signals are not compared to a control group or to any higher signals. In addition, the two signals they consider do not appear to have generated different beliefs about donations (possibly because they are so similar). This may explain why they, unlike us, obtain null effects on pro-social behavior.[3]

Outline. The remainder of this paper is structured as follows. Section 2.1 contains the experimental design and Section 2.2 presents the effect of the interventions on the share that sign up as advocates. In Section 2.3, this effect is decomposed into the effect of the interventions on beliefs and the effects of beliefs on sign-ups; Section 2.4 then explores why revising beliefs downward may increase sign-ups. Section 3 replicates our results (and presents further evidence on mechanisms) using a follow-up experiment. Finally, Section 4 reviews our main findings and discusses their policy implications.

[3]There are several other studies that, although less relevant, touch on some themes that we address in the present work. Carattini and Blasch (2024) discuss the difficulty of using descriptive social norms when compliance is low. However, they do not experimentally study the impact of a 'low' descriptive social norm. Krupka and Weber (2009) is also methodologically relevant since it studies how choices in a binary dictator game respond to information about individual choices made by participants in a previous session (see also Kessler et al., 2021). However, this laboratory experiment is not designed to study the kind of explicit descriptive-norm intervention typically used by practitioners in the field. Finally, Croson and Shang (2008) study the impact of informing donors that another donor gave less than they themselves had previously given. This is related to our study insofar as our 0% treatment also conveys 'downward' social information. In contrast to our study, however, Croson and Shang (2008) do not study the impact of different descriptive social norm levels.

2. Experiment 1

2.1 Experimental design

The natural field experiment was conducted in December 2020 in collaboration with the Leukemia & Lymphoma Society (LLS). At that time, LLS was conducting a survey about how COVID-19 had affected the lives of blood cancer patients, survivors, and caregivers in the US. We embedded our experiment within the survey that LLS emailed to its constituents.

In total, 5,360 individuals took part in the experiment, 63% of whom identified as female and 50% of whom reported holding a Bachelor's degree or higher. The average age of respondents was 64. The survey was administered using Qualtrics. Qualtrics was also used to conduct the randomization.

The experiment took place toward the end of the survey, and began by asking individuals if they were familiar with LLS advocacy work (around 69% stated that they were at least slightly familiar with these efforts). Participants were then provided with the following information about LLS 'advocates':

> *Members of the LLS Action Team play a critical role in our efforts to influence health policy at the federal and state levels. Advocates all over the country are sharing their stories and taking action on public policy issues that matter to blood cancer patients and survivors—from improving access to treatment to promoting meaningful insurance coverage for patients to accelerating new cures.*
>
> *LLS sends advocates simple ways to communicate with elected officials via email, phone calls and other actions. Their voice could be the one that sways a decision-maker to do the right thing.*

Participants were subsequently told that we had asked past respondents whether they would like to sign up to the LLS Action Team. They were then asked the following question:

> *Roughly, what percentage of past respondents to this survey do you think signed up and became advocates? Please provide us with your best guess (It's okay if you don't get it right!)*

Participants were able to enter any integer between 0 and 100.[4] We also elicited participants'

[4]We did not pay individuals money for correctly answering this question since this would have seemed out of place within the context of an LLS COVID-19 survey, in part because many respondents were answering the questions to aid cancer research. In fact, providing financial incentives may have crowded out their intrinsic motivation to answer accurately, thus worsening the quality of the data. Moreover, there is little evidence that incentivizing correct guesses makes individuals more likely to reveal their true beliefs, and even some evidence that it makes them less likely to do so (Haaland et al., 2023). Finally, providing incentives may have given away the fact that the survey included an experiment, introducing unwanted Hawthorne effects.

degree of confidence about their response on a 5-point Likert scale (from 'very unconfident' to 'very confident').

Prior to conducting the experiment, we had used the results of a pilot survey ($n = 200$) to create six groups of participants, each with ten members.[5] The first group had been chosen so that none of the ten members had signed up; the second group had been chosen so that two of the ten members had signed up; and so on (so that, in every group, X of the 10 had signed up for some $X \in \{0, 2, ..., 10\}$). In the main survey, we then randomly matched each participant with one of these six groups (or instead allocated them to the control group who received no information).[6] Unless in the control group, each participant was then informed of the following:

> *We have matched you with 10 other members of the LLS community who have responded to this survey. X of these 10 decided to sign up and become advocates!*

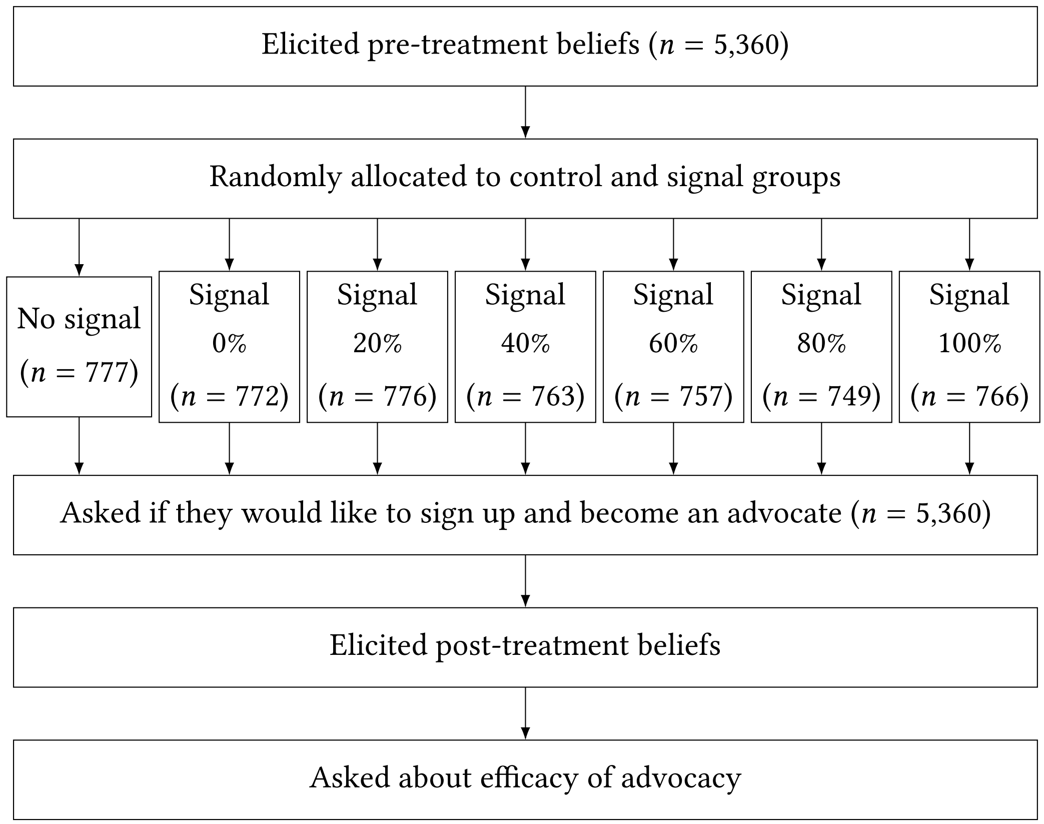
As can be seen in Figure 1, around 760 participants were matched with each of the six groups that received a signal about the share who had signed up (or were instead placed in the control group).[7] We then recorded our main outcome of interest, namely whether individuals signed up and became advocates. We did this by offering them the opportunity to sign up within the survey.

[5]The pilot experiment was identical to the main experiment except for the fact that individuals were not given an 'informational treatment'. Hence, participants in the pilot experiment completed the same survey as those in the control group of the main experiment. We do not use any data from the pilot in our analysis.

[6]Our strategy is thus in a similar spirit to Bergolo et al. (2023), who also generate exogenous variation in perceptions without resorting to deception.

[7]See Table A1 in the Appendix for a balance table. The treatment arms are balanced on observables.

Figure 1. Experimental design



In contrast to several other descriptive social norm interventions that provide information about prevalence within the entire reference population (*e.g.*, Andı and Akesson, 2020; Hallsworth et al., 2017), we only provided information about a small sample of individuals. The main drawback of this approach (relative to that taken by usual descriptive social norm interventions) is that it may limit the extent to which respondents update their beliefs about the share of LLS members who signed up. However, this approach allowed us to provide different, and yet still accurate information, to different respondents. In any case, the signals that we provided did update beliefs in the expected way (see Section 2.3) and in a way similar to other interventions that provide more representative information.

After having decided whether they would like to sign up or not, participants were again asked about their beliefs regarding the share that signed up. More specifically, we asked:

> *Based on what you now know, roughly what percentage of past respondents to this survey do you think signed up and became advocates?*

Respondents could enter any integer between 0 and 100, and we again asked about the extent to which they were confident in their response.

Finally, the experiment concluded by asking:

> *Do you agree with the following statement: I can help shape health policy and improve the lives of patients as a member of the LLS Action Team.*

Respondents provided their answer on a 5-point Likert scale (ranging from 'strongly agree' to 'strongly disagree'). We asked this particular question in an attempt to measure the extent to which the treatments operate by changing beliefs about the value of LLS advocacy.

We consider this experiment to be a *natural* field experiment in line with the taxonomy by Harrison and List (2004) because LLS often surveys members and uses such touch-points to recruit advocates. Further, participants in our experiment were not aware of the randomization of the treatments or indeed that they were part of an experiment. All they knew is that they were taking part in a survey jointly fielded by LLS and Boston University to help LLS understand how COVID-19 has affected its membership. Finally, note that while there was selection into the survey, there was no selection into the experiment.

2.2 Do the interventions encourage people to sign up?

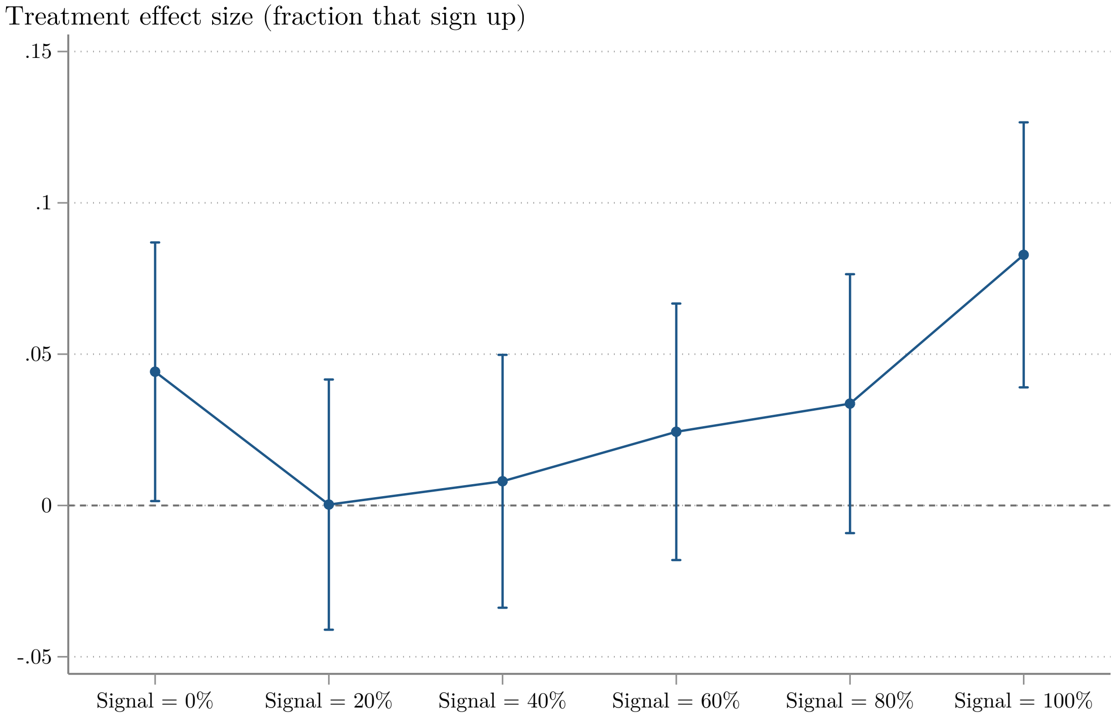
In this section, we examine the effects of the six treatments on the share of participants that sign up and become advocates. In order to do this, we estimate the model

$$y_i = \beta_0 + \sum_{k=1}^{6} \beta_k T_{ki} + u_i, \tag{1}$$

where y_i is a dummy variable indicating whether individual i signs up and the T_{ki} are dummy variables tracking the treatment (or signal) to which they are assigned.

The results are presented in Figure 2. Being shown a signal of 100% has the greatest effect on sign-ups, bringing the share who sign up from 22% (in the control group) to 30% ($p < 0.001$). Perhaps more surprisingly, the second-most effective signal is 0%, with an effect size of 4.4 percentage points relative to the control ($p = 0.04$). While the estimated coefficients for the 40%, 60% and 80% signals are positive and non-negligible, they are not statistically significant. These broad patterns are reproduced by regressions that include baseline controls: see Table A2 for details.

Figure 2. The effects of the signals on sign up rates



Notes: This figure presents the $\hat{\beta}_k$ estimates that result from estimating Equation (1) along with their associated 95% confidence intervals.

The effect of the signal on the share of participants that sign up does not seem to be linear. Instead, at first glance, it appears to be U-shaped. To test this formally, note that linearity is equivalent to the restriction $\beta_{i+1} - \beta_i = \beta_{i+2} - \beta_{i+1}$ for all $i \in \{1, ..., 4\}$. Re-estimating the model under these constraints leads to a substantial worsening of fit; and indeed, an F-test strongly rejects the restrictions ($p < 0.01$). Thus, we can clearly reject the hypothesis of a linear relationship between signals and treatment effects.

This finding raises the question of why the observed relationship is U-shaped. Assuming that the treatments operate via updating participants' beliefs about prevalence, this could be either due to a non-linear effect of the signals on participants' beliefs about the share that signed up; or instead a non-linear effect of beliefs on their behavior (or both). We attempt to answer this question in the following section.

2.3 Decomposing the treatment effects

We now present a decomposition of how the treatments influence sign ups via changes in beliefs. To do this, we first study how the treatments (*i.e.*, signals) alter beliefs about the prevalence of sign ups in the broader LLS population. We then use an instrumental variables strategy to estimate the effect of these beliefs on behaviour (see Haaland et al., 2023; Cullen and Perez-Truglia, 2022; Bottan and Perez-Truglia, 2020; Roth and Wohlfart, 2020; Akesson et al., 2022, 2026 for examples of other studies that use a similar approach). As we later discuss, the resulting decomposition is informative to the extent to which the treatments (*i.e.*, signals) influence sign up rates by changing subjects' beliefs about the sign-up rate in the broader LLS population.

We begin by estimating the effect of treatment assignment on posterior beliefs. To do this, we estimate the equation

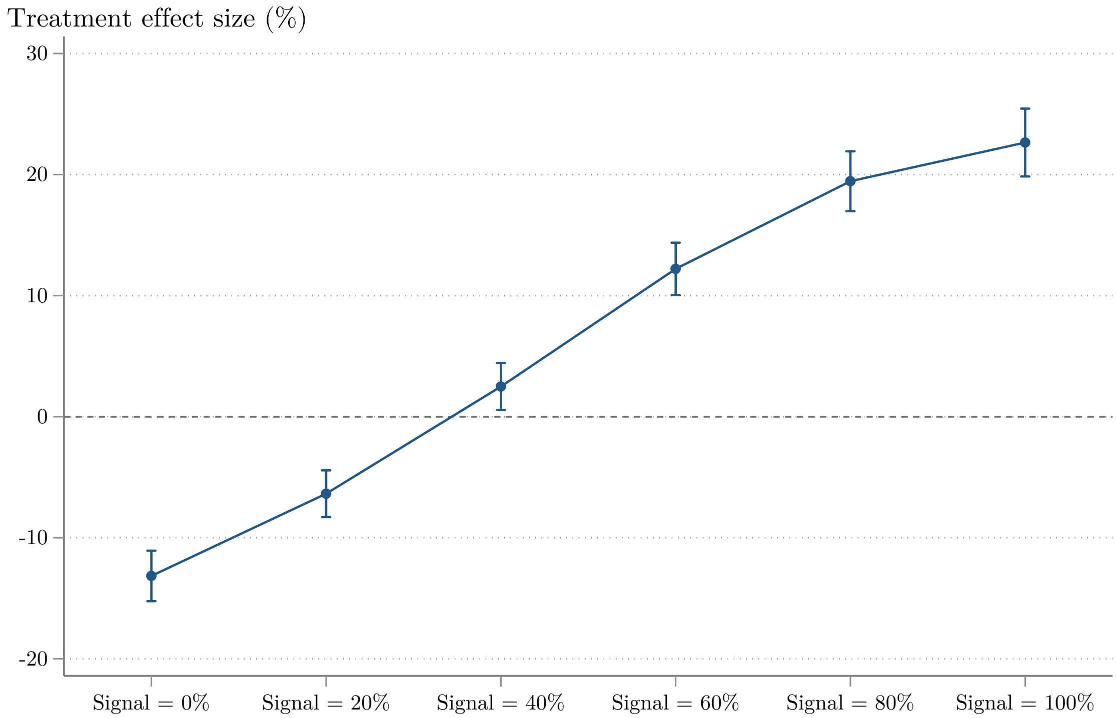
$$b_i = \gamma_0 + \sum_{k=1}^{6} \gamma_k T_{ki} + u_i, \tag{2}$$

where b_i is the 'posterior' belief of individual i (after viewing the signal) and the T_{ki} are dummy variables tracking the treatment to which they were assigned.

The results are displayed in Figure 3 (see also Table A4 which displays the exact numbers both with and without additional demographic control variables).[8] As can be seen, higher signals generally lead to higher beliefs. Indeed, we can reject the hypothesis that average beliefs are the same (with $p < 0.01$) for every 'neighboring' treatment group except for the groups receiving the 80% and 100% signals; and can more weakly ($p < 0.05$) reject the hypothesis that they are the same for the 80% and 100% groups. Moreover, the relationship between signals and posterior beliefs appears to be roughly linear. Indeed, we are not able to reject the hypothesis that $\gamma_{i+1} - \gamma_i = \gamma_{i+2} - \gamma_{i+1}$ for all $i \in \{1, ..., 4\}$ ($p = 0.995$).

[8] In addition, Figures A1–A7 in the appendix display how the treatments alter the entire distribution of beliefs, not just their means.

Figure 3. The effects of the signals on beliefs about sign up rates



Notes: This figure presents the average effects of treatment assignment (relative to the control group) on participants' beliefs about the share of other participants that signed up.

Having examined the effects of treatments on beliefs about prevalence, we now examine how these beliefs influence sign-up rates. Because beliefs may be endogenous, we use random assignment to the treatment arms as instruments for posterior beliefs. More specifically, we estimate the model

$$y_i = \theta_0 + \theta_1 b_i + \theta_2 b_i^2 + u_i, \tag{3}$$

where y_i is a binary variable denoting whether individual i signs up and b_i is individual i's belief (as before). We include the b_i^2 term to allow for possible non-linearity in the effect of beliefs on sign-ups. We then instrument for beliefs using treatment assignment. We omit the control group from this analysis to make the exclusion restriction more plausible (we discuss this issue in more detail later).

We begin by confirming that our instruments are informative (as one might expect in light of Figure 3). To do this, we conduct the standard eigenvalue test (Stock and Yogo, 2005) and can strongly reject the null hypothesis that the instruments are uninformative. We now proceed to consider whether our instruments are valid (*i.e.,* whether the exclusion restriction holds).

It is unclear whether the exclusion restriction would hold (at least exactly) if we compare individuals in the treatment group to those in the control group. For example, providing individuals with information might not just influence their point estimates but may also change the 'salience' of these point estimates, implying that those who receive information (in the treatment groups) are differentially primed to those who do not (in the control). It is more plausible, however, that the exclusion restriction holds when only considering our treatment groups. The only thing that varies across these groups is the signal presented, which means (for instance) that individuals in all groups should be equally 'primed'.

While the exclusion restriction is more plausible when considering only the treatment groups, it is possible that varying signals do not influence just individuals' point estimates but also the confidence that individuals place in these estimates. To test this, we estimate the effect of the treatments on individuals' confidence in their guesses (measured on a five-point Likert scale). We do not find any significant effects and the estimated effects are close to zero (see Columns (1) and (2) of Table A3). This result holds regardless of whether we treat confidence as a continuous variable or collapse it into a binary variable (at any possible cut-off).[9] Moreover, our IV estimates (discussed later) do not change appreciably if confidence is controlled for. Thus, any possible 'confidence effect' is unlikely to be an issue.[10]

Figure 4 displays the results of estimating Equation (3) (see also Table A5). As can be seen, imposing a linear relationship between beliefs and sign-ups reveals a positive effect: increasing point estimates by 1 percentage point is predicted to increase sign-up rates by 0.1 percentage points ($p = 0.035$). However, this result disguises important non-linearities in the effect of beliefs on sign-ups. Indeed, we can clearly reject the hypothesis that the coefficient on the quadratic term is zero ($p < 0.001$). Moreover, Figure 4 suggests that the relationship between beliefs and sign-up rates is U-shaped.

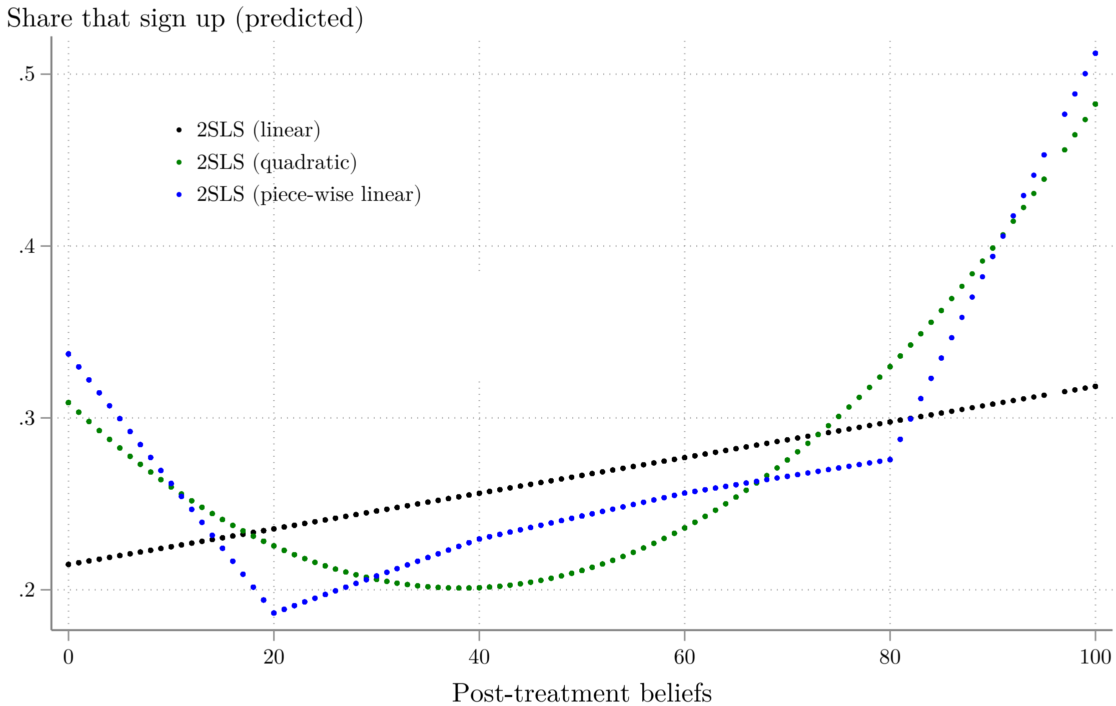
[9] In contrast, we find that the treatments dramatically increase confidence relative to the control (see Columns (3) and (4) of Table A3); providing another reason to exclude the control group from this analysis. This finding emphasizes the importance of including multiple treatment groups in information provision experiments; and suggests that findings that are based on information provision alone should be interpreted with caution.

[10] When considering the plausibility of the exclusion restriction, it is important not to confuse candidate violations with mechanisms through which observed effects occur. For example, we later obtain evidence that the 0% signal may work by triggering negative moral emotions. Presumably, this is because it makes people believe that fewer people are signing up as advocates. In that case, the causal chain would be:

$$\text{0\% signal} \implies \text{beliefs about sign-ups} \downarrow \implies \text{feeling bad} \uparrow \implies \text{sign ups} \uparrow$$

Needless to say, this would not be a violation of the exclusion restriction. Instead, it would be a mechanism through which the observed effects occur. In order to obtain a violation of the exclusion restriction, one would instead need to believe that the 0% signal somehow makes people feel bad, holding constant their beliefs about the share who sign up as advocates. This does not seem very plausible in our setting.

Figure 4. The effects of beliefs on behavior (predicted values)



Notes: In this figure, we present predictions from three instrumental variable regression models (which can be found in Tables A5 and A6). The dots represent the predicted sign-up rate for all individuals in the sample, given their beliefs about the share that sign up.

One might wonder whether the U-shape that we find is an artifact of the model that we specify in equation (3). However, including higher-order polynomials does not substantially change the predicted relationship between beliefs and sign-ups (see Table A5). For example, including a cubic term results in near identical predicted sign-up rates for every value of beliefs. To further study the robustness of this finding, we also estimate the model

$$y_i = \lambda_0 + \sum_{k=0}^{4} \lambda_{k+1}(b_i - 20k)\mathbb{1}[b_i \geq 20k], \tag{4}$$

where $\mathbb{1}[b_i \geq 20k]$ is an indicator function that equals 1 if $b_i \geq 20k$ (and equals zero otherwise). In other words, we assume that the relationship between beliefs and sign-ups is piecewise linear; allowing the effect to vary as beliefs increase but ensuring that the relationship is continuous.[11] Figure 4 displays the results (see also Table A6 in the appendix). As can be seen, the estimated

[11]Since we have six treatment groups, we have five dummy variables indicating treatment assignment—which means that we can instrument for at most 5 explanatory variables. As a result, we allow the relationship between beliefs and sign-ups to vary over 5 intervals, estimating the most flexible possible model whose parameters can be identified using our data.

relationship is broadly similar to before. In particular, the model predicts that increasing beliefs decreases sign-ups when beliefs are low but increases sign-ups when beliefs are high—a non-parametric version of the 'U shape' that the quadratic model estimates.[12]

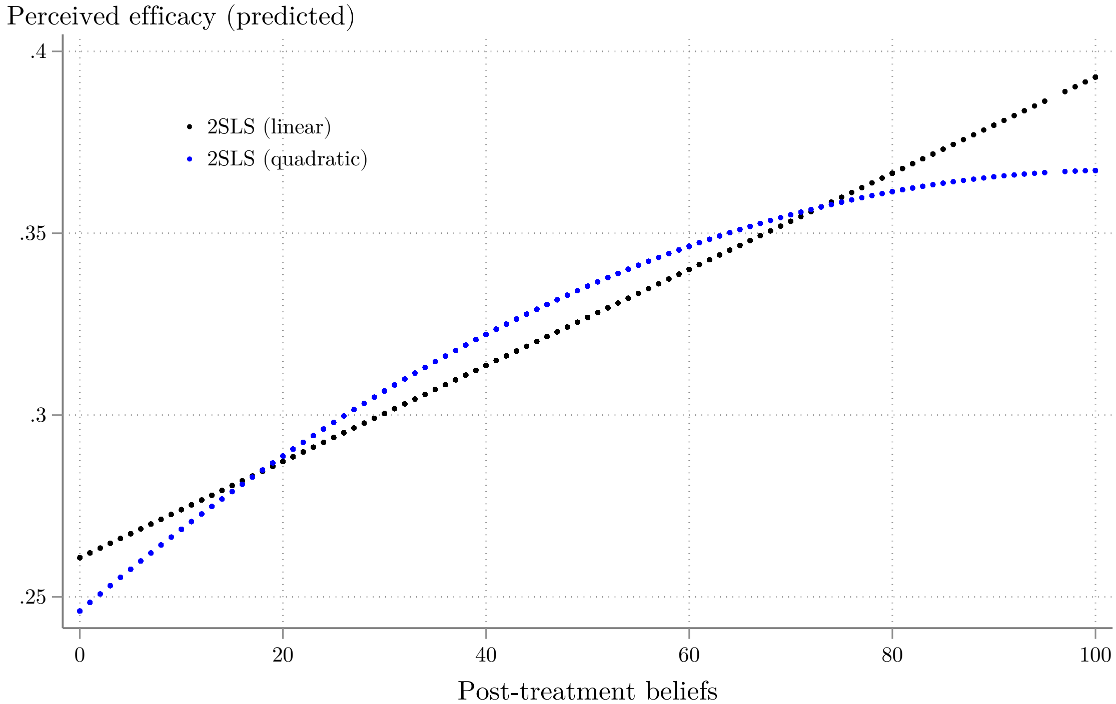
Overall, the decomposition indicates that the non-linearity in treatment effects arises primarily from the way behavior responds to beliefs, rather than from any substantial non-linearity in the effect of signals on beliefs. Higher signals lead respondents to hold higher beliefs about sign-up prevalence in the broader LLS community—a relationship that appears to work in a roughly linear way. However, the data suggest a U-shaped relationship between beliefs about sign-up rates and decisions to sign up. Taken together, this provides an explanation for the pattern of treatment effects observed in Section 2.2.

2.4 Why is there a convex relationship between beliefs and behavior?

The previous section documents that sign-ups can actually rise when beliefs about the share who sign up fall. This may seem surprising: for example, if individuals tend to conform (Akerlof, 1980), or instead view higher sign-up rates as a signal that the charity is effective (Banerjee, 1992), one would expect higher beliefs about sign-up rates to lead to more sign-ups. In this section, we discuss what might explain this rather puzzling result.

[12]As a further robustness check, we estimate a similar model which includes a piece-wise linear relationship over four intervals and an additional dummy that equals one when beliefs are exactly zero. The purpose of this model is to check whether 'zero is special' and whether this might be driving our results. While it appears that zero may be qualitatively different from other beliefs (see Table A6), this specification does not generate a substantially different estimated relationship between beliefs and behavior.

Figure 5. The effects of beliefs on perceived efficacy (predicted values)



Notes: In this figure, we present predictions from two instrumental variable regression models (which can be found in Table A7). The dots represent the (predicted) share of respondents who believe that LLS advocacy is effective, given their level of beliefs about the share that sign up.

One potential explanation for the effect that we observe is the existence of (perceived) 'diminishing returns'. By this, we mean that individuals believe that the impact of their advocacy is lower the greater the number of others who are also advocates. To test this, we examine our measure of participants' perceptions of the efficacy of being an advocate, estimating the exact same equations as in Section 2.3 but changing the dependent variable to the share that believe advocacy is effective. The results of these regressions are presented in Table A7 (and are plotted in Figure 5). As can be seen, perceived efficacy appears to be strictly increasing in beliefs about the share who sign up, as in models of social learning (Banerjee, 1992; Bikhchandani et al., 1992). As a result, perceptions of diminishing returns do *not* seem to explain the results that we observe here: on the contrary, individuals seem to downgrade their beliefs about efficacy when their beliefs about the share who sign up fall.

While we are able to reject the 'diminishing returns' explanation, we are unable to identify precisely what explains the effect we observe; and it is plain that a number of explanations are possible. One possibility is that decreasing the perceived share who sign up raises the perceived 'self-image' benefits from signing up (Bénabou and Tirole, 2006; Falk, 2020). After all, if a person believes that hardly anybody else signs up, then signing up allows the person to tell themselves that they are a better person than nearly everybody else. A second possibility is that individuals believe that the organization is more desperate for their support when the organization sends them a signal that nobody is signing up, thereby increasing the 'social pressure' associated with the signal. It is also possible that the 0% signal encourages individuals to sign up by making them feel bad. Although we are unable to test the importance of these and other explanations using the data from experiment 1, we are able to shed further light on this topic using the data from experiment 2.

3. Experiment 2

3.1 Experimental design

To see if the results replicate in a different context and to provide some additional evidence on mechanisms, we conducted a second experiment in November 2023 with 5,042 US participants. Respondents were recruited to take part in a (Qualtrics) survey via Prolific Academic, and received $1.10 if they completed the survey. On average, the survey took 7 minutes to complete.[13]

At the start of the survey, participants were provided with a list of eight charities (which focused on different causes) along with descriptions of these charities. Participants were told that they needed to read the charity descriptions carefully in order to proceed with the survey. Next, participants were asked about the degree to which they supported the causes on which each of the eight charities focused. We matched participants to the charity whose cause they supported the most (we randomly selected a charity if there was a tie). Participants were also asked to match the charities to their descriptions in a subsequent question, and participants were screened out if they failed to correctly select the description of the charity that they were matched to. Unless otherwise stated, the experiment 2 regressions use the resulting screened analysis sample of 5,031 respondents. Randomization appears balanced on observables in this screened sample (see Table A8).

We recorded participants' gender, age, and income, before randomly allocating them to an experimental condition (participants were allocated to each norm condition with equal probability). All participants were told that they had been matched with a charity and that we were offering them the opportunity to sign up and become advocates for that charity. We then asked participants what proportion of other survey respondents that were matched with this charity they thought had signed up as advocates for that charity.

Participants were then asked if they themselves would like to sign up. Those who said that

[13] The exact wording used in our second experiment can be viewed here.

they would like to become advocates were directed to the charity's sign-up page. Participants allocated to treatment conditions were also told the following just before they were asked whether they would like to sign up:

> *We have matched you with 10 other survey respondents. We asked them if they would like to become advocates and X decided to sign up.*

In the treatment conditions, we allowed for $X \in \{0, 2, 5, 8, 10\}$; meanwhile, participants allocated to the control condition did not see this statement. Broadly, this mirrored the design from Experiment 1; however, we replaced the 'middling' $X = 4, 6$ signals with a single $X = 5$ condition in order to increase statistical power. Further, half of all participants were randomly told that their decision and contact information would be kept confidential, while the other half were randomly told that we would provide the charity with the participants' contact details and that we would inform the charity of their decision so that they could get in touch to discuss their response. The rationale for these public and private conditions was to examine whether revising beliefs downward about the share that signed up increases sign-ups by triggering social image motivations.

After asking participants whether they would like to sign up, we re-elicited their beliefs about the share that signed up. Finally, participants were asked whether they agreed with the following statements (which were presented in random order):

1. If I sign up as an advocate, I will be able to help [the charity] accomplish its goals.
2. [The charity] would benefit from me signing up as an advocate.
3. [The charity] is effective at achieving its goals.
4. I feel bad that not enough people have signed up for [the charity].
5. I will feel guilty if I do not sign up as an advocate for [the charity].
6. [The charity] aligns with my values.
7. [The charity]'s work benefits me or somebody I care about.
8. [The charity] has a positive reputation.
9. I am concerned that people working for [the charity] will judge me if I do not sign up as an advocate.
10. I will feel like a better person if I sign up as an advocate for [the charity].
11. If I sign up as an advocate, people will follow my example and will also sign up.

12. I am familiar with [the charity].

We asked participants whether they agreed with these statements in an attempt to better understand why revising beliefs downward about the share that signed up might increase participants' willingness to sign up.[14] Statements 1–3 were designed to capture efficacy and social-learning considerations, since lower take-up by others may lead participants to infer that advocacy is less effective or that the organization is less impactful (Banerjee, 1992; Bikhchandani et al., 1992; Bursztyn et al., 2014). Statements 4–5 were designed to capture guilt, sympathy, and personal responsibility, which theories of prosocial emotion and bystander intervention suggest may become stronger when fewer others are helping (Darley and Latané, 1968; Fischer et al., 2011; Erlandsson et al., 2016). Statements 6–8 and 10 capture value alignment, personal benefit, reputation, and self-image, which may also vary with norm levels through identity and moral self-signaling channels (Bénabou and Tirole, 2006; Falk, 2020; DellaVigna et al., 2012). Statement 9 captures perceived judgement or social pressure, and statement 11 captures conditional-cooperation or spillover considerations (Frey and Meier, 2004; DellaVigna et al., 2012). Statement 12 captures familiarity with the organization, which may matter if subjects treat others' take-up as a signal about organizational quality or fit (Bursztyn et al., 2014). We also collected an open-ended explanation for the participant's decision.

3.2 Results

First, we investigate whether the 'U shaped' relationship between the signal conveyed and sign up rates replicates. Table 1 displays the treatment effects relative to the control (Column 1) as well as the treatment effects relative to the 0% signal group (Column 2). As can be seen, the 'best' signals are once again the 0% signals and the 100% signals; and the intermediate signals are again statistically indistinguishable from the control. However, the estimated relative efficacy of the 0% and 100% signals is now reversed. Quantitatively, the 0% signal boosts participation by around 4 percentage points relative to the control (a very similar effect to that observed in the previous experiment) whereas the 100% signal boosts participation by around 3 percentage points (a smaller effect than observed previously). Thus, although the general U-shaped pattern from the previous experiment replicates, the exact parameter estimates (notably, the estimates for the 100% signal) are slightly different.[15]

[14] As these questions were asked after participants decided whether to sign up, the mechanism variables may be vulnerable to ex-post rationalization rather than strictly capturing mediators that caused the decision. Asking these questions prior to deciding whether to sign up would, however, likely also have biased participants, as it would prompt them to consider reasons for signing up or not that they might not otherwise have before making their decision.

[15] We also estimate pooled specifications using the combined data from Experiments 1 and 2. To harmonize treatment arms across studies, the 40% and 60% treatments in Experiment 1 are recoded into a common 50% category before pooling. Results are reported in Table A9.

Table 1. The effect of signals on sign-ups

	(1) Effects relative to control	(2) Effects relative to 0% group
Signal = 0%	0.0394	
	(0.0201)	
Signal = 20%	-0.00707	-0.0465
	(0.0195)	(0.0211)
Signal = 50%	0.00549	-0.0339
	(0.0202)	(0.0215)
Signal = 80%	0.0104	-0.0290
	(0.0199)	(0.0212)
Signal = 100%	0.0340	-0.00541
	(0.0202)	(0.0211)
Constant	0.226	0.265
	(0.0131)	(0.0148)
n	5,031	4,007
R^2	0.002	0.002

Notes. This table presents the effect of signals on sign-up rates, both relative to the control group (column 1) and relative to the 0% group (column 2). Robust standard errors are reported in parentheses.

Having replicated the U-shaped treatment effects documented in the main experiment, we now examine the mechanisms underlying the relative effects of the different signals. In particular, we focus on what explains the surprising efficacy of the 0% signal. To begin, we examine if the 0% signal works by enhancing the image benefits of signing up (for example, because advocates in this condition will be thought of as especially altruistic relative to other participants). To test this, we consider whether signals interact with the public condition by estimating the effects of the signals conditional on being placed (or not being placed) in the public condition. Table 2 displays the results of this exercise. As can be seen, the 0% signal appears to be less effective (if anything) in the public conditions, although the coefficient on the interaction term is not significant. This suggests that the surprising efficacy of the 0% signal is not mediated by visibility.

Table 2. The effect of making sign-ups public

	(1) Private condition	(2) Public condition
Signal = 0%	0.0516	0.0266
	(0.0283)	(0.0286)
Signal = 20%	0.00917	-0.0233
	(0.0279)	(0.0274)
Signal = 50%	-0.0159	0.0255
	(0.0280)	(0.0290)
Signal = 80%	0.0119	0.00878
	(0.0278)	(0.0285)
Signal = 100%	0.0290	0.0385
	(0.0286)	(0.0286)
Constant	0.224	0.228
	(0.0183)	(0.0187)
n	2,522	2,509
R^2	0.002	0.002

Notes. This table shows the effect of signals on sign-up rates in the private condition (column 1) and the public condition (column 2). Robust standard errors are reported in parentheses.

To provide further evidence on mechanisms, we next examine the closed-ended mechanism questions directly. Tables A10 and A11 report regressions of each mechanism item on treatment assignment, once relative to the control group and once relative to the 0% group. This item-level presentation makes it possible to see which specific mechanism questions move with treatment rather than only whether a broader index changes. Consistent with the results of the main experiment, the 0% signal does not appear to enhance the perceived efficacy of the charity, which provides further evidence that efficacy is not the mechanism underlying our result. By contrast, the 0% signal substantially increases the share who state that they would feel bad that not enough people have signed up, while leaving the other mechanism variables largely unchanged.[16]

[16]The fact that the 0% norm affects respondents' reports of "feeling bad" but not guilt or sympathy is perhaps surprising, but it is not necessarily inconsistent. These measures need not capture the same underlying response. "Feeling bad" is a broad measure that may reflect a diffuse negative affective reaction, such as discomfort, unease, or concern, whereas guilt and sympathy are narrower constructs. In particular, guilt typically implies a sense of personal responsibility, while sympathy implies compassionate concern for the charity or its beneficiaries. The 0% norm may therefore shift a general sense that the situation is troubling without strongly affecting either of these more specific emotions. More generally, it is entirely possible to detect an effect on one outcome but not on related outcomes when the treated outcome is a broader and more sensitive measure, while the others are conceptually narrower or measured with more noise.

We complement these closed-ended measures by analyzing respondents' open-ended explanations for their decisions. To do this systematically, we first manually reviewed a subset of responses to build a categorization scheme of underlying motivations. We then used the OpenAI API to classify the full sample of text responses into these predefined categories. This approach follows recent applications of LLM-based classification in economics (Haaland et al., 2024; Celebi and Penczynski, 2024; Ludwig et al., 2025). To maintain consistency with our main results, we restrict this text analysis to the same screened sample of 5,031 respondents used in Tables 1 and 2.

Tables A12 and A13 report the treatment effects on these API-coded motivation indicators, relative either to the control group or to the 0% group. Tables A15–A19 provide randomly sampled examples of the responses assigned to each API-coded category. The main pattern mirrors the closed-ended evidence: respondents in the 0% group are more likely than the control group to say that they signed up because they felt bad, and the 20%, 50%, 80%, and 100% groups are all less likely than the 0% group to cite that motivation. We do not find comparable evidence that the 0% signal increases efficacy-related motives in the text. Taken together, the closed-ended and open-ended responses suggest that the 0% signal works in part by triggering negative moral emotions rather than by increasing perceived efficacy.[17]

4. Conclusions

In this paper, we report the findings from a field experiment on descriptive social norms. The experiment yields three main findings. First, we confirm that a signal suggesting that many others are signing up increases sign-up rates. Second, however, we also find that sign-up rates can be increased by suggesting that no one else is signing up, an effect that we replicate using a

[17] As an additional robustness check, we conduct an exploratory mediation analysis for Experiment 2 that compares the 0% signal to the other non-control norm treatments and asks whether the higher sign-up rate in the 0% arm is statistically associated with respondents reporting that they "felt bad" that too few others had signed up. Following the regression-based mediation framework of Imai et al. (2010, 2011) and the broader discussion of experimental mechanism designs in Imai et al. (2013), we decompose the reduced-form effect of the 0% signal into an indirect component operating through a "feel bad" mediator and a residual direct component, with bootstrap inference for the indirect effect; the results are reported in Appendix Table A14. This approach is useful here because treatment assignment is randomized and the mechanism question of interest is whether the 0% signal shifts an observed mediator that is, in turn, predictive of sign-up. At the same time, the resulting estimates require stronger assumptions than the reduced-form treatment effects. In particular, the mediator is not experimentally manipulated and is measured after the sign-up decision, so the mediator–outcome relationship cannot be interpreted as a clean causal estimate of how sign-up would change if we were to exogenously induce respondents to feel bad. Reported "feeling bad" may partly reflect post-decision rationalization or reverse causality, with respondents inferring or reconstructing their motives after having already decided whether to sign up. For this reason, and consistent with the cautions emphasized by Acharya et al. (2016), we interpret the mediation exercise as an exploratory decomposition that is consistent with "feeling bad" helping to explain the 0% effect, rather than as definitive evidence of causal mediation. In our data, the closed-ended "feel bad" item explains most of the observed sign-up difference between the 0% arm and the other norm arms, whereas the analogous open-ended measure yields a smaller indirect effect.

follow-up experiment in the virtual laboratory. Third, we find that the success of the 0% signal cannot be explained by the familiar logic of diminishing marginal returns; instead, the evidence points to negative moral emotions as the more plausible mechanism in this setting.

While our findings shed light on how organizations like LLS can mobilize support, we do not claim that they will extend to all contexts. In some sense, our setting is rather special: for example, LLS members may care about each other and form a natural 'reference group' (Lindgren et al., 2021), which may then explain why social norms exert such a strong influence in this environment. As a result, it is important that similar studies are conducted to provide us with a better understanding of whether these results generalize.

Nonetheless, we suspect that some of our findings may generalize to other contexts. For example, there appears to be an important qualitative difference between everybody doing an activity and most people doing it, and we conjecture that further studies will also pick this up in their analyses. Similarly, the finding that it can be useful to revise beliefs downward may also generalize, especially in other altruistic or charity-related settings where low levels of compliance may activate feelings of responsibility or concern. Indeed, several non-profit organizations already use this approach, even if it fits poorly with received academic wisdom. For example, the Teenage Cancer Trust currently tries to encourage donations by stating that 'Only 1% of people who visit our website go on to donate' (Teenage Cancer Trust, 2026).

These findings are not only of academic interest, but may also have implications for how policymakers and other individuals should use descriptive social norms. First, they suggest large benefits from shifting beliefs about prevalence from high to very high levels. This might be accomplished either by increasing actual rates of prevalence or alternately by disclosing information about prevalence rates in settings in which prevalence is generally underestimated (see Bursztyn et al., 2020 for a recent example). Second, they suggest that organisations should at least consider revealing low levels of support. While our findings do not imply that such an approach will always work, they do suggest that it should at least be considered (and, when possible, tested). In this sense, our results suggest that descriptive social norm interventions are more broadly applicable than had previously been realized.

Finally, while our paper shows how the efficacy of norm interventions depends on the signal they convey, it is silent on the net welfare effects of different signals. Indeed, a separate line of research shows that the provision of social information can have negative welfare effects, for example by inducing sufficiently large feelings of guilt or shame (Butera et al., 2022). Future work could explore how the types of signals used in this study influence economic welfare and when low norms are likely to trigger negative emotions.

References

Acharya, A., Blackwell, M., and Sen, M. (2016). Explaining causal findings without bias: Detecting and assessing direct effects. *American Political Science Review*, 110(3):512–529.

Akerlof, G. A. (1980). A theory of social custom, of which unemployment may be one consequence. *The Quarterly Journal of Economics*, 94(4):749–775.

Akesson, J., Ashworth-Hayes, S., Hahn, R., Metcalfe, R., and Rasooly, I. (2022). Fatalism, beliefs, and behaviors during the covid-19 pandemic. *Journal of Risk and Uncertainty*, 64(2):147–190.

Akesson, J., Hahn, R. W., Metcalfe, R. D., and Rasooly, I. (2026). Race and redistribution in the united states: An experimental analysis. *Journal of Public Economics*, 255:105555.

Allcott, H. (2011). Social norms and energy conservation. *Journal of Public Economics*, 95(9-10):1082–1095.

Andı, S. and Akesson, J. (2020). Nudging away false news: Evidence from a social norms experiment. *Digital Journalism*, pages 1–21.

Banerjee, A. V. (1992). A simple model of herd behavior. *The Quarterly Journal of Economics*, 107(3):797–817.

Bénabou, R. and Tirole, J. (2006). Incentives and prosocial behavior. *American Economic Review*, 96(5):1652–1678.

Benartzi, S., Beshears, J., Milkman, K. L., Sunstein, C. R., Thaler, R. H., Shankar, M., Tucker-Ray, W., Congdon, W. J., and Galing, S. (2017). Should governments invest more in nudging? *Psychological science*, 28(8):1041–1055.

Bergolo, M., Ceni, R., Cruces, G., Giaccobasso, M., and Perez-Truglia, R. (2023). Tax audits as scarecrows: Evidence from a large-scale field experiment. *American Economic Journal: Economic Policy*, 15(1):110–153.

Bergstrom, T. C., Blume, L., and Varian, H. R. (1986). On the private provision of public goods. *Journal of Public Economics*, 29(1):25–49.

Bernheim, B. D., Kim, K., and Taubinsky, D. (2024). Welfare and the act of choosing. Technical report, National Bureau of Economic Research.

Bicchieri, C. and Dimant, E. (2019). Nudging with care: The risks and benefits of social information. *Public choice*, pages 1–22.

Bikhchandani, S., Hirshleifer, D., and Welch, I. (1992). A theory of fads, fashion, custom, and cultural change as informational cascades. *Journal of Political Economy*, 100(5):992–1026.

Borsari, B. and Carey, K. B. (2003). Descriptive and injunctive norms in college drinking: a meta-analytic integration. *Journal of Studies on Alcohol*, 64(3):331–341.

Bottan, N. L. and Perez-Truglia, R. (2020). Betting on the house: Subjective expectations and market choices. Technical report, National Bureau of Economic Research.

Brandon, A., Ferraro, P. J., List, J. A., Metcalfe, R. D., Price, M. K., and Rundhammer, F. (2017). Do the effects of nudges persist? theory and evidence from 38 natural field experiments. Technical report, National Bureau of Economic Research.

Brandon, A., List, J. A., Metcalfe, R. D., Price, M. K., and Rundhammer, F. (2019). Testing for crowd out in social nudges: Evidence from a natural field experiment in the market for electricity. *Proceedings of the National Academy of Sciences*, 116(12):5293–5298.

Brent, D. A., Cook, J. H., and Olsen, S. (2015). Social comparisons, household water use, and participation in utility conservation programs: Evidence from three randomized trials. *Journal of the Association of Environmental and Resource Economists*, 2(4):597–627.

Bursztyn, L., Ederer, F., Ferman, B., and Yuchtman, N. (2014). Understanding mechanisms underlying peer effects: Evidence from a field experiment on financial decisions. *Econometrica*, 82(4):1273–1301.

Bursztyn, L., González, A. L., and Yanagizawa-Drott, D. (2020). Misperceived social norms: Women working outside the home in saudi arabia. *American Economic Review*, 110(10):2997–3029.

Butera, L., Metcalfe, R., Morrison, W., and Taubinsky, D. (2022). Measuring the welfare effects of shame and pride. *American Economic Review*, 112(1):122–168.

Carattini, S. and Blasch, J. (2024). Nudging when the descriptive norm is low: Evidence from a carbon offsetting field experiment. *Journal of Behavioral and Experimental Economics*, 110:102194.

Card, D. and Dahl, G. B. (2011). Family violence and football: The effect of unexpected emotional cues on violent behavior. *The Quarterly Journal of Economics*, 126(1):103–143.

Celebi, C. and Penczynski, S. P. (2024). Using large language models for text classification in experimental economics. Working Paper 24-01, Centre for Behavioural and Experimental Social Science, University of East Anglia.

Chen, Y., Harper, F. M., Konstan, J., and Li, S. X. (2010). Social comparisons and contributions to online communities: A field experiment on movielens. *American Economic Review*, 100(4):1358–98.

Coricelli, G., Joffily, M., Montmarquette, C., and Villeval, M. C. (2010). Cheating, emotions, and rationality: an experiment on tax evasion. *Experimental Economics*, 13:226–247.

Croson, R. and Shang, J. (2008). The impact of downward social information on contribution decisions. *Experimental Economics*, 11(3):221–233.

Cullen, Z. and Perez-Truglia, R. (2022). How much does your boss make? the effects of salary comparisons. *Journal of Political Economy*, 130(3):766–822.

Darley, J. M. and Latané, B. (1968). Bystander intervention in emergencies: Diffusion of responsibility. *Journal of Personality and Social Psychology*, 8(4):377–383.

DellaVigna, S. and Linos, E. (2022). Rcts to scale: Comprehensive evidence from two nudge units. *Econometrica*, 90(1):81–116.

DellaVigna, S., List, J. A., Malmendier, U., and Rao, G. (2012). Testing for altruism and social pressure in charitable giving. *Quarterly Journal of Economics*, 127(1):1–56.

Diekmann, A. (1985). Volunteer's dilemma. *Journal of Conflict Resolution*, 29(4):605–610.

Dijker, A. J. and Koomen, W. (2003). Extending weiner's attribution-emotion model of stigmatization of ill persons. *Basic and applied social psychology*, 25(1):51–68.

Duncan, B. (2004). A theory of impact philanthropy. *Journal of Public Economics*, 88(9-10):2159–2180.

Ek, C. and Söderberg, M. (2024). Norm-based feedback on household waste: Large-scale field experiments in two swedish municipalities. *Journal of Public Economics*, 238:105191.

Erlandsson, A., Johansson, H., Västfjäll, D., and Slovic, P. (2016). Anticipated guilt for not helping and anticipated warm glow for helping are differently impacted by personal responsibility to help. *Frontiers in Psychology*, 7:1475.

Falk, A. (2020). Facing yourself–a note on self-image. *Journal of Economic Behavior & Organization*.

Ferraro, P. J. and Price, M. K. (2013). Using nonpecuniary strategies to influence behavior: evidence from a large-scale field experiment. *Review of Economics and Statistics*, 95(1):64–73.

Fischer, P., Greitemeyer, T., Pollozek, D., and Frey, A. (2011). The bystander-effect: A meta-analytic review on bystander intervention in dangerous and non-dangerous emergencies. *Psychological Bulletin*, 137(4):517–537.

Frey, B. S. and Meier, S. (2004). Social comparisons and pro-social behavior: Testing" conditional cooperation" in a field experiment. *American Economic Review*, 94(5):1717–1722.

Gerber, A. S. and Rogers, T. (2009). Descriptive social norms and motivation to vote: Everybody's voting and so should you. *The Journal of Politics*, 71(1):178–191.

Goldin, J., Homonoff, T., Patterson, R., and Skimmyhorn, W. (2020). How much to save? decision costs and retirement plan participation. *Journal of Public Economics*, 191:104247.

Haaland, I., Roth, C., and Wohlfart, J. (2023). Designing information provision experiments. *Journal of Economic Literature*, 61(1):3–40.

Haaland, I. K., Roth, C., Stantcheva, S., and Wohlfart, J. (2024). Understanding economic behavior using open-ended survey data. NBER Working Paper 32421, National Bureau of Economic Research.

Hallsworth, M., List, J. A., Metcalfe, R. D., and Vlaev, I. (2017). The behavioralist as tax collector: Using natural field experiments to enhance tax compliance. *Journal of public economics*, 148:14–31.

Halpern, D. (2015). How can governments and businesses avoid the 'big mistake?'. https://behavioralscientist.org/how-can-governments-and-businesses-avoid-the-big-mistake/.

Harrison, G. W. and List, J. A. (2004). Field experiments. *Journal of Economic Literature*, 42(4):1009–1055.

Imai, K., Keele, L., and Tingley, D. (2010). A general approach to causal mediation analysis. *Psychological Methods*, 15(4):309–334.

Imai, K., Keele, L., Tingley, D., and Yamamoto, T. (2011). Unpacking the black box of causality: Learning about causal mechanisms from experimental and observational studies. *American Political Science Review*, 105(4):765–789.

Imai, K., Tingley, D., and Yamamoto, T. (2013). Experimental designs for identifying causal mechanisms. *Journal of the Royal Statistical Society: Series A (Statistics in Society)*, 176(1):5–51.

Kessler, J. B., Low, C., and Singhal, M. (2021). Social policy instruments and the compliance environment. *Journal of Economic Behavior & Organization*, 192:248–267.

Krupka, E. and Weber, R. A. (2009). The focusing and informational effects of norms on pro-social behavior. *Journal of Economic Psychology*, 30(3):307–320.

Lantos, N. A., Kende, A., Becker, J. C., and McGarty, C. (2020). Pity for economically disadvantaged groups motivates donation and ally collective action intentions. *European Journal of Social Psychology*, 50(7):1478–1499.

Lindgren, K. P., DiBello, A. M., Peterson, K. P., and Neighbors, C. (2021). Theory-driven interventions: How social cognition can help. *The handbook of alcohol use*, pages 485–510.

Linek, M. and Traxler, C. (2021). Framing and social information nudges at wikipedia. *Journal of Economic Behavior & Organization*, 188:1269–1279.

Loewenstein, G. (2000). Emotions in economic theory and economic behavior. *American Economic Review*, 90(2):426–432.

Ludwig, J., Mullainathan, S., and Rambachan, A. (2025). Large language models: An applied econometric framework. NBER Working Paper 33344, National Bureau of Economic Research.

Lunardo, R. and Bezençon, V. (2015). The neglected ambivalent emotion of pity: Conceptualization and potential (complex) effects on charitable behavior. *ACR North American Advances*.

Martin, R. and Randal, J. (2008). How is donation behaviour affected by the donations of others? *Journal of Economic Behavior & Organization*, 67(1):228–238.

Miller, D. T. and Prentice, D. A. (2016). Changing norms to change behavior. *Annual Review of Psychology*, 67:339–361.

OECD (2017). Behavioural insights and public policy: Lessons from around the world. *OECD Publishing*.

Perez-Truglia, R. and Troiano, U. (2018). Shaming tax delinquents. *Journal of Public Economics*, 167:120–137.

Roth, C. and Wohlfart, J. (2020). How do expectations about the macroeconomy affect personal expectations and behavior? *Review of Economics and Statistics*, 102(4):731–748.

Stock, J. and Yogo, M. (2005). *Asymptotic distributions of instrumental variables statistics with many instruments*, volume 6. Chapter.

Sunstein, C. R., Reisch, L. A., and Rauber, J. (2018). A worldwide consensus on nudging? not quite, but almost. *Regulation & Governance*, 12(1):3–22.

Teenage Cancer Trust (2026). Teenage cancer trust. Accessed March 25, 2026. https://www.teenagecancertrust.org/.

Wenzel, M. and Woodyatt, L. (2025). The power and pitfalls of social norms. *Annual Review of Psychology*, 76:583–606.

A. Additional tables and figures

Figure A1. Beliefs about sign-ups in the control group

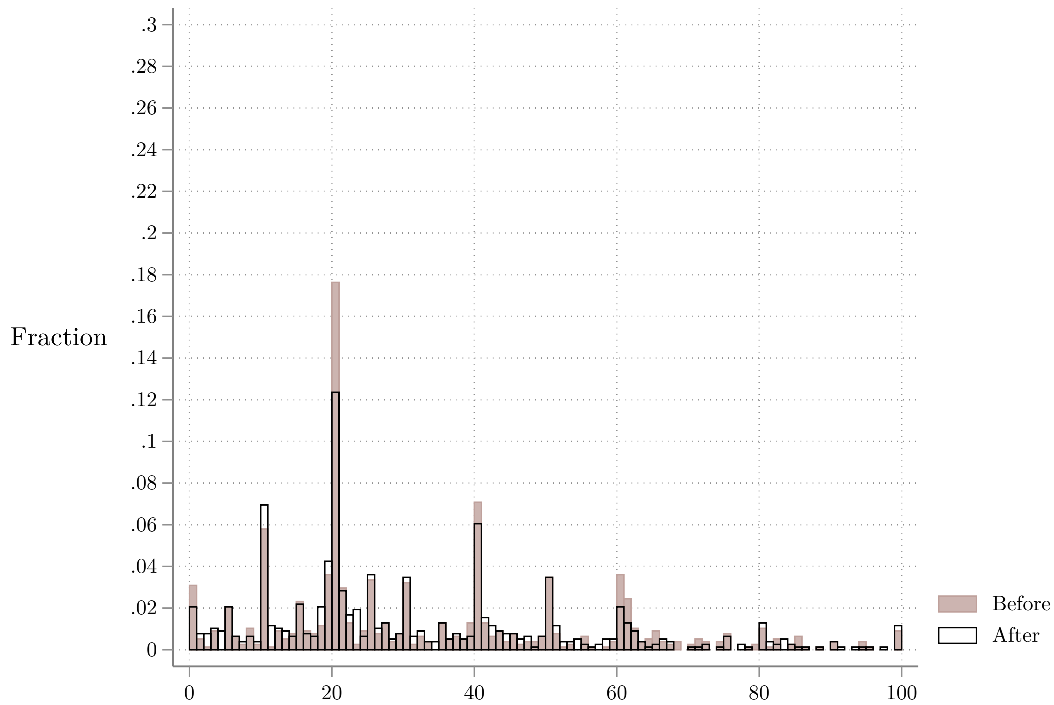


Figure A2. Beliefs about sign-ups in the 0% signal group

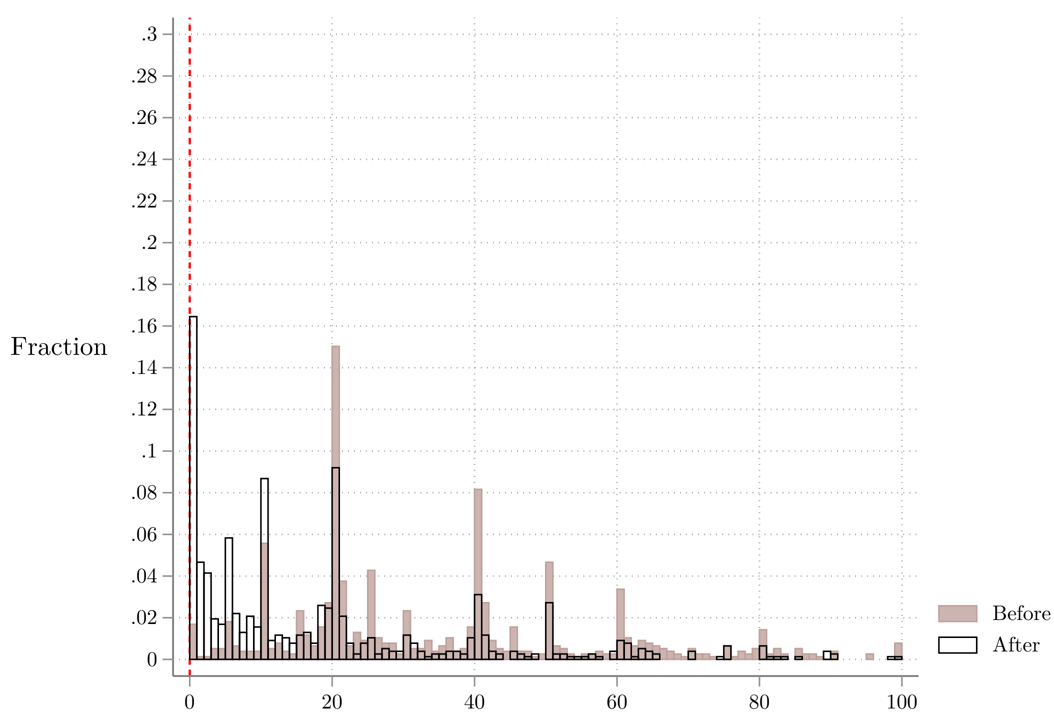


Figure A3. Beliefs about sign-ups in the 20% signal group

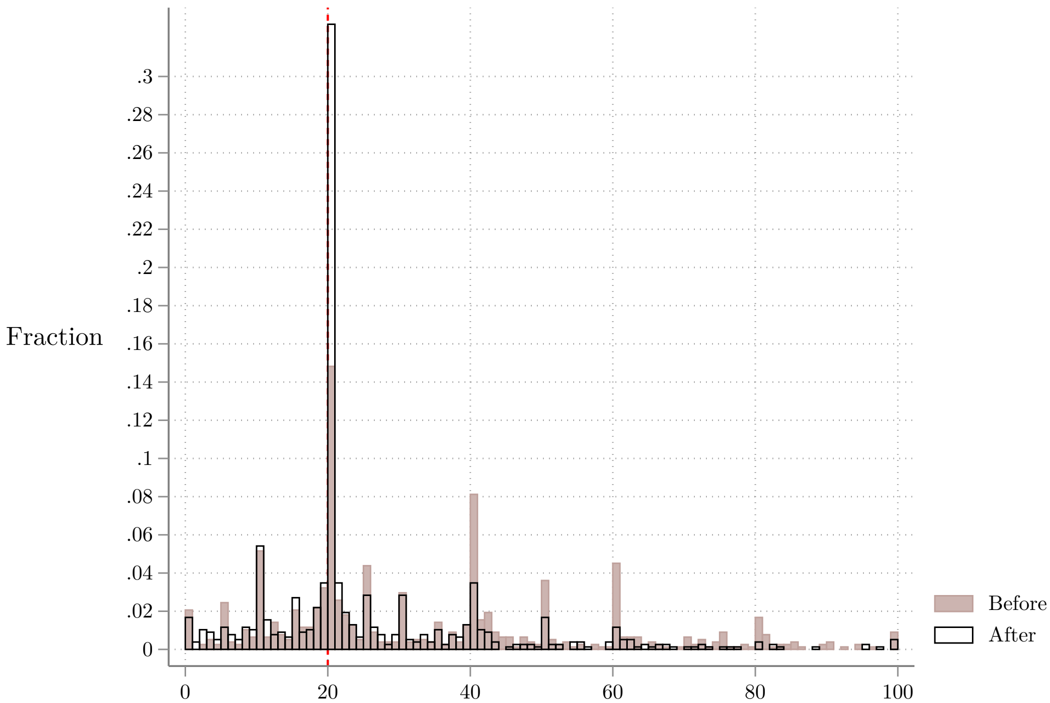


Figure A4. Beliefs about sign-ups in the 40% signal group

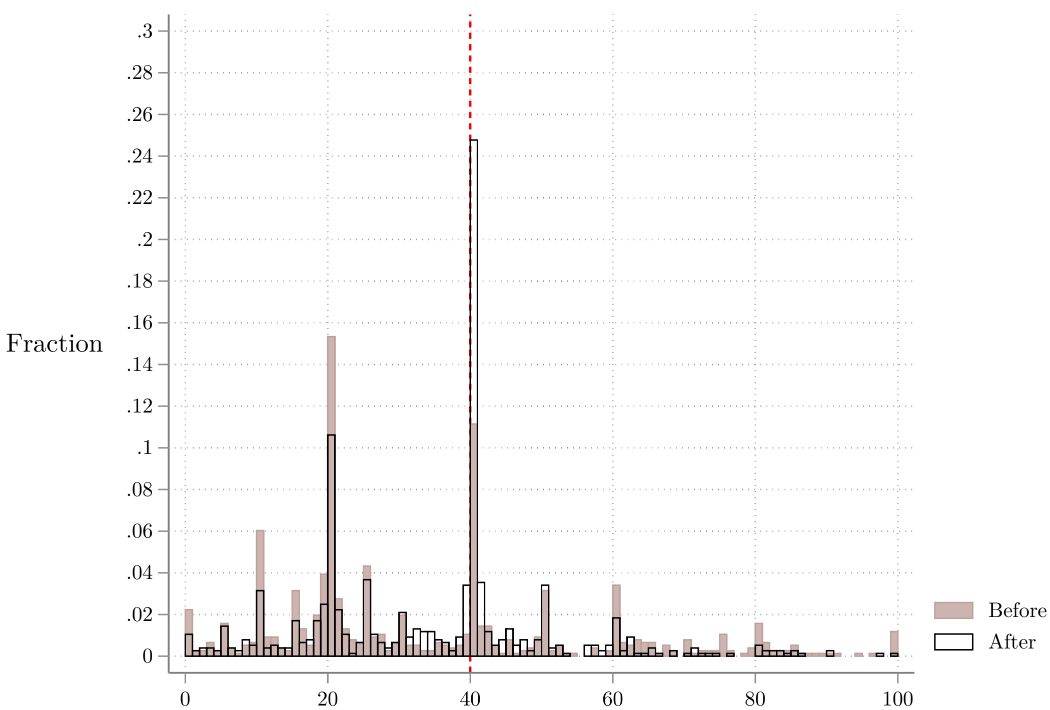


Figure A5. Beliefs about sign-ups in the 60% signal group

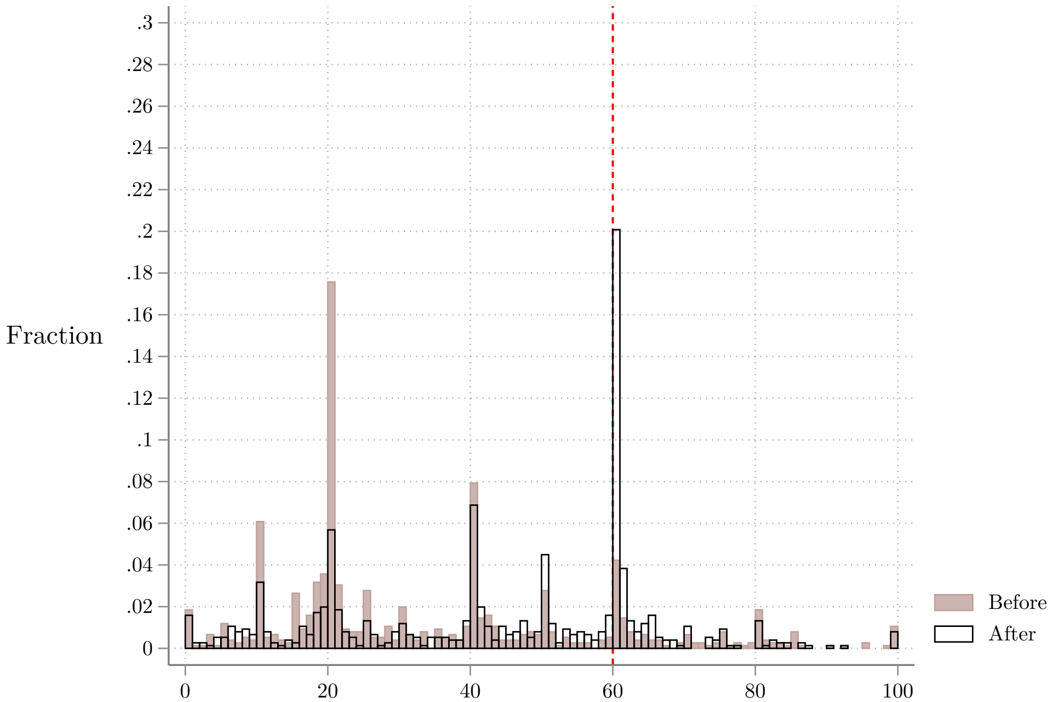


Figure A6. Beliefs about sign-ups in the 80% signal group

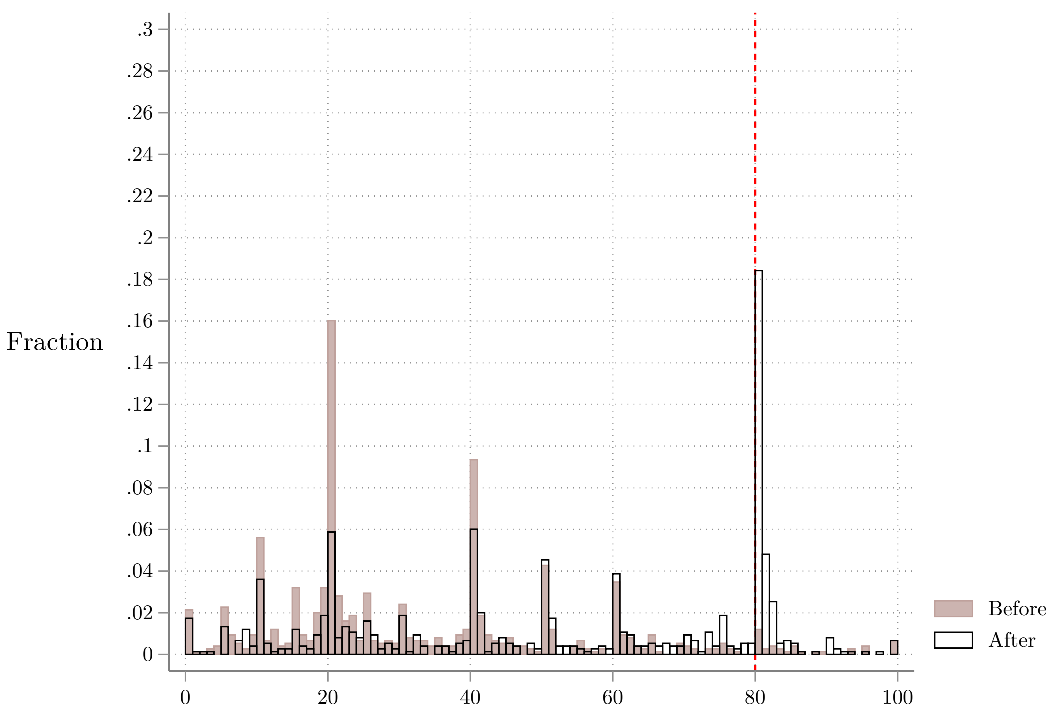


Table A1. Balance table

	Control	Signal 0%	Signal 20%	Signal 40%	Signal 60%	Signal 80%	Signal 100%	p-value
Age (years)	64.0	63.8	64.1	64.2	64.4	64.4	63.6	0.820
Familiar with LLS advocacy	0.663	0.703	0.688	0.702	0.724	0.673	0.689	0.177
BSc or more	0.495	0.509	0.512	0.491	0.484	0.529	0.504	0.740
Pre-treatment beliefs	32.7	34.3	33.2	33.4	33.8	32.1	34.2	0.419
Pre-treatment confidence in beliefs	1.995	1.970	1.912	1.931	1.917	1.917	1.957	0.250
n	777	772	776	763	757	749	766	

Notes. This table presents averages for pre-treatment variables in Experiment 1. Because age and education are missing for a small number of respondents, those rows report means for respondents who answered those questions. The right-most column reports p-values from joint orthogonality tests of treatment assignment.

Table A2. Treatment effect on share that sign up

	(1) Effect on sign-up shares	(2) Effect on sign-up shares
Signal = 0%	0.0442	0.0348
	(0.0218)	(0.0227)
Signal = 20%	0.000285	-0.00177
	(0.0211)	(0.0221)
Signal = 40%	0.00799	-0.00200
	(0.0213)	(0.0221)
Signal = 60%	0.0243	0.00981
	(0.0216)	(0.0225)
Signal = 80%	0.0336	0.0363
	(0.0218)	(0.0232)
Signal = 100%	0.0828	0.0761
	(0.0223)	(0.0234)
Constant	0.221	0.166
	(0.0149)	(0.0445)
Controls	No	Yes
n	5,360	4,368
R^2	0.004	0.068

Notes. In columns (1) and (2), we regress whether participants signed up (binary) on treatment assignment, with the control group omitted. Column (2) additionally controls for age, education, pre-treatment beliefs, pre-treatment confidence in beliefs, and familiarity with LLS advocacy efforts. The sample is smaller in column (2) because demographic information is not observed for all respondents.

Figure A7. Beliefs about sign-ups in the 100% signal group

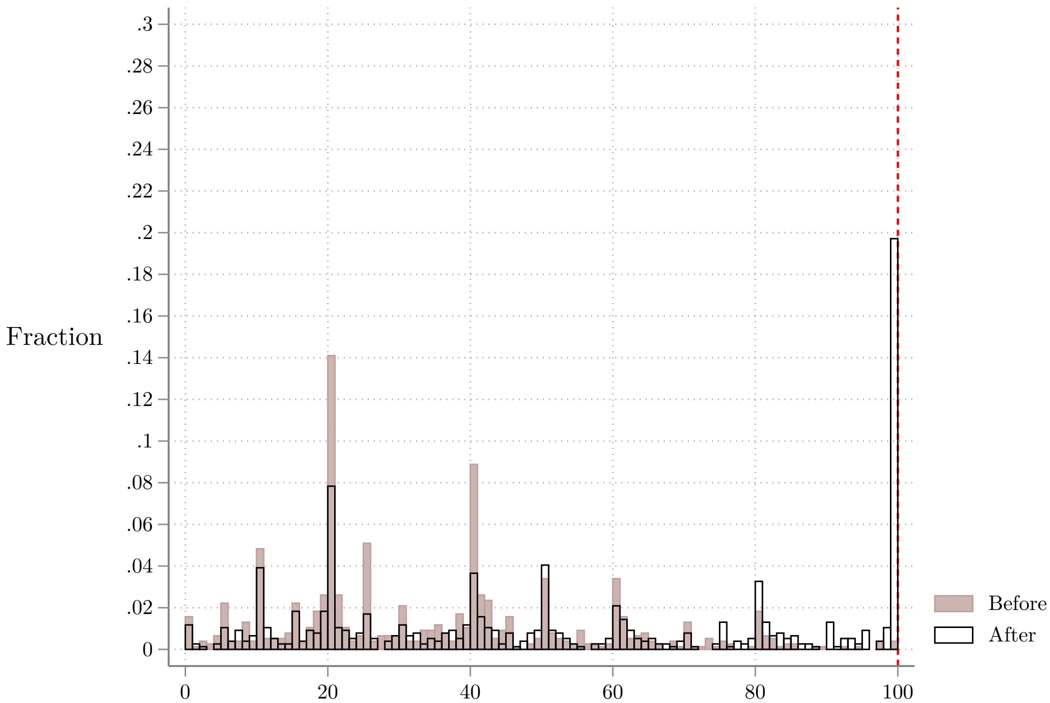


Figure A8. Treatment effects on beliefs by prior belief

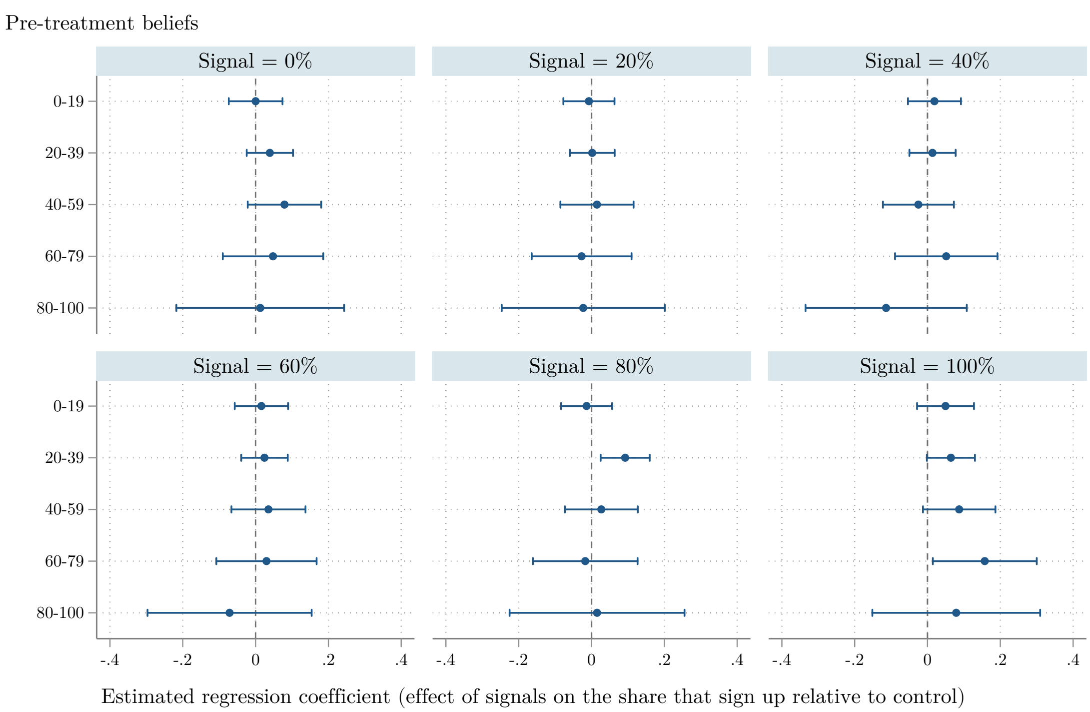


Table A3. Treatment effect on confidence in beliefs

	(1) Confidence in beliefs (Likert)	(2) Confident or very confident (binary)	(3) Confidence in beliefs (Likert)	(4) Confident or very confident (binary)
Signal = 0%			0.185	0.109
			(0.0421)	(0.0210)
Signal = 20%	-0.0330	-0.0130	0.152	0.0956
	(0.0464)	(0.0226)	(0.0421)	(0.0208)
Signal = 40%	-0.00213	-0.0138	0.183	0.0948
	(0.0453)	(0.0227)	(0.0409)	(0.0209)
Signal = 60%	0.0347	0.0108	0.220	0.119
	(0.0450)	(0.0230)	(0.0407)	(0.0213)
Signal = 80%	0.0140	-0.0182	0.199	0.0904
	(0.0452)	(0.0227)	(0.0408)	(0.0209)
Signal = 100%	0.0364	0.0152	0.221	0.124
	(0.0468)	(0.0230)	(0.0426)	(0.0212)
Constant	2.153	0.277	1.968	0.169
	(0.0327)	(0.0161)	(0.0264)	(0.0134)
n	4,583	4,583	5,360	5,360
R^2	0.001	0.001	0.007	0.008

Notes. This table presents four regressions. In columns (1) and (3), the outcome is confidence in beliefs, measured on a five-point Likert scale. In columns (2) and (4), the outcome is an indicator for being confident or very confident in beliefs. Columns (1) and (2) omit the control group, so the 0% signal is the reference category; columns (3) and (4) retain the full sample and use the control group as the reference category.

Table A4. Treatment effect on beliefs about the share that sign up

	(1) Beliefs about sign-ups	(2) Beliefs about sign-ups
Signal = 0%	-13.15	-13.25
	(1.064)	(0.832)
Signal = 20%	-6.363	-5.266
	(0.984)	(0.738)
Signal = 40%	2.489	3.750
	(0.990)	(0.778)
Signal = 60%	12.21	12.04
	(1.106)	(0.971)
Signal = 80%	19.45	21.01
	(1.262)	(1.182)
Signal = 100%	22.65	22.79
	(1.428)	(1.364)
Constant	31.37	16.06
	(0.782)	(2.091)
Controls	No	Yes
n	5,360	4,368
R^2	0.222	0.398

Notes. This table presents two regressions of post-treatment beliefs about the share who sign up on treatment assignment. Column (2) additionally controls for age, education, pre-treatment beliefs, pre-treatment confidence in beliefs, and familiarity with LLS advocacy efforts. The sample is smaller in column (2) because demographic information is not observed for all respondents.

Table A5. Effects of beliefs on sign-ups (1)

	(1) % sign up	(2) % sign up	(3) % sign up	(4) % sign up	(5) % sign up
Belief % sign up	0.00104	-0.00564	-0.00632	-0.0197	-0.0199
	(0.000492)	(0.00204)	(0.00490)	(0.0108)	(0.0268)
Belief ^{2}		7.38e-05	9.21e-05	0.000763	0.000783
		(2.19e-05)	(0.000123)	(0.000500)	(0.00194)
Belief ^{3}			-1.26e-07	-1.11e-05	-1.16e-05
			(8.36e-07)	(7.99e-06)	(5.23e-05)
Belief ^{4}				5.59e-08	6.22e-08
				(4.06e-08)	(5.90e-07)
Belief ^{5}					-0
					(2.35e-09)
Constant	0.215	0.309	0.314	0.365	0.365
	(0.0193)	(0.0343)	(0.0474)	(0.0603)	(0.0829)
n	4,583	4,583	4,583	4,583	4,583
R^2	0.021	0.013	0.014	0.015	0.015

Notes. This table reports five instrumental variables regressions in which the outcome is whether the participant signs up. The control group is omitted, and treatment assignment is used to instrument for beliefs. Column (1) instruments only for beliefs about the share that sign up; columns (2)–(5) additionally instrument for polynomial terms in beliefs. No additional controls are included.

Table A6. Effects of beliefs on sign-ups (2)

	(1)	(2)	(3)
	% sign up	% sign up	% sign up
Belief % sign up	-0.00754	-0.00582	
	(0.00387)	(0.0227)	
" if belief $\geq$ 20	0.00969		
	(0.00694)		
" if belief $\geq$ 40	-0.000818		
	(0.00744)		
" if belief $\geq$ 60	-0.000361		
	(0.00779)		
" if belief $\geq$ 80	0.0108		
	(0.00864)		
" = 0		0.0521	
		(1.032)	
" if belief $\geq$ 25		0.0101	
		(0.0302)	
" if belief $\geq$ 50		-0.00628	
		(0.0120)	
" if belief $\geq$ 75		0.0134	
		(0.00899)	
Belief $>$ 20 & $\leq$ 40			-0.0929
			(0.0631)
Belief $>$ 40 & $\leq$ 60			0.0181
			(0.0650)
Belief $>$ 60 & $\leq$ 80			-0.0609
			(0.0839)
Belief $>$ 80 & $\leq$ 100			0.213
			(0.0723)
Constant	0.337	0.312	0.262
	(0.0497)	(0.427)	(0.0228)
Observations	4,583	4,583	4,583
R^2	0.014	0.012	0.001

Notes. This table reports three instrumental variables regressions in which the outcome is whether the participant signs up. The control group is omitted, and treatment assignment is used to instrument for beliefs.

Table A7. Effects of beliefs on perceived efficacy

	(1) Perceived efficacy	(2) Perceived efficacy	(3) Perceived efficacy	(4) Perceived efficacy	(5) Perceived efficacy
Belief % sign up	0.00132	0.00236	0.0122	0.00234	0.0150
	(0.000513)	(0.00214)	(0.00517)	(0.0116)	(0.0292)
Belief ^{2}		-1.15e-05	-0.000277	0.000215	-0.000754
		(2.31e-05)	(0.000130)	(0.000540)	(0.00212)
Belief ^{3}			1.83e-06	-6.22e-06	2.05e-05
			(8.88e-07)	(8.62e-06)	(5.73e-05)
Belief ^{4}				4.11e-08	-2.62e-07
				(4.38e-08)	(6.45e-07)
Belief ^{5}					1.21e-09
					(2.57e-09)
Constant	0.261	0.246	0.175	0.213	0.183
	(0.0202)	(0.0352)	(0.0487)	(0.0624)	(0.0878)
n	4,583	4,583	4,583	4,583	4,583
R^2	0.015	0.013	0.011	0.010	0.010

Notes. This table reports five instrumental variables regressions in which the outcome is perceived efficacy of LLS advocacy. The control group is omitted, and treatment assignment is used to instrument for beliefs. Column (1) instruments only for beliefs; columns (2)–(5) additionally instrument for polynomial terms in beliefs. No additional controls are included.

Table A8. Experiment 2 balance table

	Control	Signal 0%	Signal 20%	Signal 50%	Signal 80%	Signal 100%	p-value
Age (years)	41.066	40.978	41.619	40.628	41.596	41.587	0.622
Female (binary)	0.487	0.483	0.470	0.452	0.477	0.483	0.733
Pre-treatment beliefs	30.588	31.463	30.031	31.566	31.216	31.413	0.745
Pre-treatment confidence	0.469	0.484	0.467	0.450	0.478	0.451	0.685
n	1024	834	810	753	801	809	

Notes. This table reports means for pre-treatment characteristics in the screened Experiment 2 analysis sample. The right-most column reports the p-value from a joint orthogonality test of treatment assignment. Confidence is coded as an indicator for being somewhat or very confident in pre-treatment beliefs.

Table A9. Pooled treatment effects on sign-up rates across Experiments 1 and 2

	(1) Baseline specification	(2) + Experiment fixed effect	(3) Baseline specification	(4) + Experiment fixed effect
Signal = 0%	0.042	0.042		
	(0.015)	(0.015)		
Signal = 20%	0.002	0.001	-0.040	-0.040
	(0.014)	(0.014)	(0.015)	(0.015)
Signal = 50%	0.019	0.017	-0.023	-0.026
	(0.013)	(0.013)	(0.014)	(0.014)
Signal = 80%	0.027	0.026	-0.016	-0.016
	(0.015)	(0.015)	(0.016)	(0.016)
Signal = 100%	0.056	0.055	0.014	0.013
	(0.015)	(0.015)	(0.016)	(0.016)
Experiment 2		-0.010		-0.014
		(0.009)		(0.009)
Constant	0.220	0.226	0.263	0.270
	(0.010)	(0.011)	(0.011)	(0.012)
Observations	10,372	10,372	8,593	8,593
R-squared	0.002	0.002	0.002	0.002

Notes. This table pools the main sign-up regressions from Experiments 1 and 2. To harmonize treatment arms across the two studies, the 40% and 60% treatments in Experiment 1 are recoded into a common 50% category before pooling. Columns (1) and (2) use the control group as the omitted category. Columns (3) and (4) omit the control group and use the 0% treatment as the reference category. Robust standard errors are reported in parentheses.

Table A10. Mechanism survey responses relative to the control group

	(1) Accomplish goal	(2) Charity judge	(3) Charity benefit	(4) Better person	(5) Feel bad	(6) Feel guilty	(7) Charity effective	(8) Follow example	(9) Aligns values	(10) Charity benefits me	(11) Positive reputation	(12) Charity familiar
Signal = 0%	0.00741	0.00121	0.0178	0.0198	0.0799	0.00575	-0.0171	0.00899	-0.0189	0.0137	-0.00954	-0.0168
	(0.0142)	(0.00824)	(0.0140)	(0.0147)	(0.0128)	(0.00933)	(0.0176)	(0.00954)	(0.0228)	(0.0182)	(0.0230)	(0.0206)
Signal = 20%	-0.00681	0.00450	-0.0123	-0.00254	-0.00184	-0.00151	-0.0197	0.0103	-0.0306	0.0168	-0.0173	0.00434
	(0.0143)	(0.00830)	(0.0141)	(0.0148)	(0.0129)	(0.00941)	(0.0177)	(0.00962)	(0.0230)	(0.0184)	(0.0232)	(0.0207)
Signal = 50%	0.0238	0.00816	0.0284	0.0223	-0.00809	0.00148	0.00952	0.0216	-0.00132	0.0301	0.0196	-0.0171
	(0.0146)	(0.00848)	(0.0144)	(0.0151)	(0.0132)	(0.00961)	(0.0181)	(0.00982)	(0.0235)	(0.0187)	(0.0237)	(0.0212)
Signal = 80%	0.0129	0.0123	-0.00511	-0.000167	-0.0222	-0.00107	0.00204	0.00577	-0.0175	0.00649	0.0110	-0.0250
	(0.0143)	(0.00833)	(0.0141)	(0.0149)	(0.0130)	(0.00944)	(0.0178)	(0.00965)	(0.0231)	(0.0184)	(0.0233)	(0.0208)
Signal = 100%	0.00689	0.0107	0.0113	0.0211	-0.0104	-0.000225	-0.0109	0.0140	-0.0400	0.00716	0.0389	-0.00521
	(0.0143)	(0.00831)	(0.0141)	(0.0148)	(0.0129)	(0.00941)	(0.0177)	(0.00962)	(0.0230)	(0.0184)	(0.0232)	(0.0207)
Constant	0.0957	0.0264	0.0938	0.103	0.0771	0.0410	0.178	0.0342	0.417	0.176	0.416	0.273
	(0.00948)	(0.00552)	(0.00937)	(0.00986)	(0.00859)	(0.00625)	(0.0118)	(0.00639)	(0.0153)	(0.0122)	(0.0154)	(0.0138)
n	5,031	5,031	5,031	5,031	5,031	5,031	5,031	5,031	5,031	5,031	5,031	5,031
R^2	0.001	0.001	0.002	0.001	0.015	0.000	0.001	0.001	0.001	0.001	0.001	0.001

Notes. This table reports OLS regressions of the closed-ended mechanism indicators on treatment assignment, with the control group omitted. Statements 1–3 capture perceived efficacy, statements 4–5 negative moral emotions, statements 6–8 positive alignment with the charity, and statements 9–12 additional mechanism channels such as judgement, self-image, social spillovers, and familiarity. Robust standard errors are shown in parentheses.

Table A11. Mechanism survey responses relative to the 0% group

	(1) Accomplish goal	(2) Charity judge	(3) Charity benefit	(4) Better person	(5) Feel bad	(6) Feel guilty	(7) Charity effective	(8) Follow example	(9) Aligns values	(10) Charity benefits me	(11) Positive reputation	(12) Charity familiar
Signal = 20%	-0.0142	0.00329	-0.0300	-0.0223	-0.0818	-0.00726	-0.00265	0.00128	-0.0117	0.00314	-0.00771	0.0212
	(0.0151)	(0.00890)	(0.0149)	(0.0157)	(0.0137)	(0.00989)	(0.0185)	(0.0104)	(0.0241)	(0.0194)	(0.0244)	(0.0217)
Signal = 50%	0.0164	0.00695	0.0107	0.00253	-0.0880	-0.00427	0.0266	0.0126	0.0176	0.0164	0.0291	-0.000287
	(0.0154)	(0.00907)	(0.0152)	(0.0160)	(0.0139)	(0.0101)	(0.0189)	(0.0106)	(0.0246)	(0.0197)	(0.0248)	(0.0221)
Signal = 80%	0.00550	0.0111	-0.0229	-0.0199	-0.102	-0.00681	0.0191	-0.00322	0.00142	-0.00718	0.0205	-0.00816
	(0.0151)	(0.00893)	(0.0149)	(0.0158)	(0.0137)	(0.00992)	(0.0186)	(0.0104)	(0.0242)	(0.0194)	(0.0244)	(0.0217)
Signal = 100%	-0.000522	0.00950	-0.00644	0.00131	-0.0903	-0.00597	0.00620	0.00504	-0.0211	-0.00651	0.0484	0.0116
	(0.0151)	(0.00891)	(0.0149)	(0.0157)	(0.0137)	(0.00990)	(0.0185)	(0.0104)	(0.0241)	(0.0194)	(0.0244)	(0.0217)
Constant	0.103	0.0276	0.112	0.122	0.157	0.0468	0.161	0.0432	0.398	0.189	0.406	0.257
	(0.0106)	(0.00625)	(0.0105)	(0.0110)	(0.00959)	(0.00694)	(0.0130)	(0.00727)	(0.0169)	(0.0136)	(0.0171)	(0.0152)
n	4,007	4,007	4,007	4,007	4,007	4,007	4,007	4,007	4,007	4,007	4,007	4,007
R^2	0.001	0.001	0.002	0.001	0.018	0.000	0.001	0.001	0.001	0.000	0.002	0.001

Notes. This table reports OLS regressions of the closed-ended mechanism indicators on treatment assignment after omitting the control group, so the 0% signal is the omitted category. The outcomes correspond to the same mechanism questions as Table A10. Robust standard errors are shown in parentheses.

Table A12. API-coded motivations relative to the control group

	(1) Help / impact	(2) Guilt / obligation	(3) Felt bad	(4) Sympathy	(5) Values alignment	(6) Personal connection	(7) Learn more	(8) Norm / compliance	(9) Signup other	(10) Time / commitment
Signal = 0%	0.015	0.002	0.012	0.002	0.037	0.011	-0.001	0.013	-0.005	-0.027
	(0.016)	(0.004)	(0.004)	(0.004)	(0.016)	(0.009)	(0.006)	(0.006)	(0.005)	(0.021)
Signal = 20%	-0.002	-0.002	0.006	0.006	0.025	-0.002	-0.002	0.001	-0.003	-0.008
	(0.015)	(0.003)	(0.003)	(0.005)	(0.016)	(0.009)	(0.006)	(0.004)	(0.005)	(0.021)
Signal = 50%	0.002	-0.002	0.001	0.005	0.010	-0.012	0.002	-0.001	-0.002	-0.006
	(0.016)	(0.003)	(0.001)	(0.005)	(0.016)	(0.008)	(0.007)	(0.004)	(0.005)	(0.021)
Signal = 80%	-0.004	-0.001	0.002	-0.002	0.019	-0.007	-0.004	-0.003	0.000	-0.018
	(0.015)	(0.003)	(0.002)	(0.004)	(0.016)	(0.008)	(0.006)	(0.004)	(0.005)	(0.021)
Signal = 100%	0.001	-0.001	0.002	-0.003	0.020	-0.004	0.001	0.006	0.005	-0.033
	(0.015)	(0.003)	(0.002)	(0.003)	(0.016)	(0.009)	(0.006)	(0.005)	(0.006)	(0.021)
Constant	0.120	0.005	-0.000	0.007	0.116	0.036	0.018	0.008	0.014	0.277
	(0.010)	(0.002)	(.)	(0.003)	(0.010)	(0.006)	(0.004)	(0.003)	(0.004)	(0.014)
n	5,031	5,031	5,031	5,031	5,031	5,031	5,031	5,031	5,031	5,031
R^2	0.000	0.001	0.004	0.001	0.001	0.001	0.000	0.003	0.001	0.001

Notes. This table reports OLS regressions of API-coded motivation indicators on treatment assignment. Each outcome is a binary indicator equal to one if the response was coded into the relevant category. Robust standard errors are shown in parentheses.

Table A13. API-coded motivations relative to the 0% group

	(1) Help / impact	(2) Guilt / obligation	(3) Felt bad	(4) Sympathy	(5) Values alignment	(6) Personal connection	(7) Learn more	(8) Norm / compliance	(9) Signup other	(10) Time / commitment
Signal = 20%	-0.017	-0.005	-0.006	0.004	-0.013	-0.012	-0.001	-0.012	0.003	0.019
	(0.016)	(0.003)	(0.005)	(0.005)	(0.017)	(0.010)	(0.006)	(0.006)	(0.005)	(0.022)
Signal = 50%	-0.013	-0.005	-0.011	0.004	-0.027	-0.023	0.003	-0.014	0.004	0.020
	(0.017)	(0.003)	(0.004)	(0.005)	(0.017)	(0.009)	(0.007)	(0.006)	(0.005)	(0.022)
Signal = 80%	-0.019	-0.003	-0.009	-0.003	-0.019	-0.018	-0.003	-0.015	0.005	0.009
	(0.016)	(0.004)	(0.004)	(0.004)	(0.017)	(0.009)	(0.006)	(0.005)	(0.005)	(0.022)
Signal = 100%	-0.014	-0.003	-0.010	-0.005	-0.018	-0.015	0.002	-0.007	0.010	-0.006
	(0.017)	(0.004)	(0.004)	(0.004)	(0.017)	(0.010)	(0.007)	(0.006)	(0.006)	(0.021)
Constant	0.135	0.007	0.012	0.008	0.153	0.047	0.017	0.020	0.008	0.251
	(0.012)	(0.003)	(0.004)	(0.003)	(0.012)	(0.007)	(0.004)	(0.005)	(0.003)	(0.015)
n	4,007	4,007	4,007	4,007	4,007	4,007	4,007	4,007	4,007	4,007
R^2	0.000	0.001	0.003	0.002	0.001	0.002	0.000	0.003	0.001	0.001

Notes. This table reports OLS regressions of API-coded motivation indicators on treatment assignment after omitting the control group, so the 0% signal is the omitted category. Each outcome is a binary indicator equal to one if the response was coded into the relevant category. Robust standard errors are shown in parentheses.

Table A14. Exploratory mediation analysis of the 0% signal in Experiment 2

	(1) Closed-ended feel bad	(2) Open-ended API feel bad	(3) Either measure
0% signal → sign-up (total effect)	0.0248	0.0248	0.0248
	(0.0162)	(0.0162)	(0.0162)
0% signal → mediator (a path)	0.0889	0.0089	0.0900
	(0.0132)	(0.0039)	(0.0133)
Mediator → sign-up (b path)	0.3631	0.7170	0.3779
	(0.0273)	(0.0267)	(0.0268)
0% signal → sign-up, controlling for mediator (direct effect)	-0.0075	0.0184	-0.0092
	(0.0155)	(0.0161)	(0.0154)
Indirect effect ($a \times b$)	0.0323	0.0064	0.0340
	(0.0058)	(0.0029)	(0.0060)
n	4,006	4,006	4,006

Notes. This table reports an exploratory mediation analysis for Experiment 2. The sample is restricted to the non-control norm conditions, so the omitted category pools the 20%, 50%, 80%, and 100% signal groups, while the treatment indicator equals one for assignment to the 0% signal. To keep the sample constant across columns, all specifications are estimated on the matched sample of respondents who can also be linked to the saved open-ended-response coding file; one non-control observation in the cleaned Experiment 2 file could not be matched and is therefore dropped. Column (1) uses the closed-ended indicator for agreeing that the respondent would "feel bad that not enough people have signed up." Column (2) uses the API-coded open-ended indicator for stating that the respondent signed up because they "felt bad." Column (3) equals one if either of these two indicators equals one. The total effect is estimated from a linear probability model of sign-up on the 0% treatment indicator and controls. The a path is estimated from a linear probability model of the mediator on the 0% treatment indicator and the same controls. The b path and direct effect are estimated from a linear probability model of sign-up on the 0% treatment indicator, the mediator, and the same controls. All specifications control for age, gender, pre-treatment beliefs, and pre-treatment confidence. Standard errors in the first four rows are heteroskedasticity-robust. The indirect effect is computed as the product of the estimated a and b coefficients, with standard errors obtained by nonparametric bootstrap with 300 replications. Because the mediator is measured after the sign-up decision and is not experimentally manipulated, these estimates should be interpreted as exploratory evidence consistent with mediation, rather than as definitive causal mediation estimates. We do not report the implied proportion mediated because the reduced-form 0% effect is imprecisely estimated in this matched sample, making the ratio of indirect to total effects unstable.

Table A15. API-coded open-ended response examples

Category	Ex.	Signed up	Response text	Coding note
Signed up: help/impact	1	Yes	it would be slightly beneficial to the organization and i would feel guilty to say no	Participant indicated feeling guilty for not signing up, which aligns with guilt and a desire to help.
Signed up: help/impact	2	Yes	I will sign up because I want to help people in need. Is my duty as a human being .	The respondent expressed a desire to help others and mentioned a sense of duty, indicating both a motivation to help and a feeling of obligation.
Signed up: help/impact	3	Yes	to support the needy	Response indicates a desire to help others in need.
Signed up: help/impact	4	Yes	As an anti-Zionist Jew who is watching the Israeli assault on Gaza unfold in shock and horror, I think JVP's vocal demand for a ceasefire and refusal to accept the atrocities of...	Respondent signed up due to a strong alignment with values and emotional reaction to the situation.
Signed up: help/impact	5	Yes	I wanted to see what they have a need for and hope that I can help.	The response indicates a desire to help and learn about needs.
Signed up: help/impact	6	Yes	I did becaus I care about the cause	Caring about the cause aligns with wanting to help.
Signed up: help/impact	7	Yes	The natural environment and all of the species that live on our planet don't have a voice. It is up to humans to protect our planet for later generations. Any cause involving an...	Signed up to help protect the environment and aligns with personal values.
Signed up: help/impact	8	Yes	I would like to show support, because its a meaningful cause.	Respondent expressed desire to support a meaningful cause, indicating a wish to help.
Signed up: help/impact	9	Yes	I did because the world is going backwards on all the strides it was making with income equality and hate. There should be more positive in this world than negative and to me mo...	Signed up to help improve negative social issues and promote positive change.
Signed up: help/impact	10	Yes	I signed up as its the right thing to do they served are country	Signed up due to the belief that it is the right thing to do and appreciation for service.

Notes. This table reports randomly sampled responses from the second experiment. Responses and coding notes are lightly truncated only where necessary to keep the table readable.

Table A16. API-coded open-ended response examples

Category	Ex.	Signed up	Response text	Coding note
Signed up: guilt/obligation	1	Yes	Part was out of guilt (ie I'd feel guilty if I did not sign up) but I have also been wanting to become a better advocate and use my position of power to benefit and advocate for...	Motivation to sign up includes guilt and desire to help others.
Signed up: guilt/obligation	2	Yes	I feel pressured to sign up because it will look bad if I don't	Respondent feels pressured to sign up due to guilt about not doing so.
Signed up: guilt/obligation	3	Yes	it would be slightly beneficial to the organization and i would feel guilty to say no	Participant indicated feeling guilty for not signing up, which aligns with guilt and a desire to help.
Signed up: guilt/obligation	4	Yes	I didn't want to be negatively perceived as a racist or an antisemite.	The response indicates a motivation based on guilt to avoid negative perceptions, aligning with signup_guilt.
Signed up: guilt/obligation	5	Yes	I felt like it was something I should do	The response indicates a sense of obligation to participate.
Signed up: guilt/obligation	6	Yes	I feel guilty and would really like to help.	The response clearly indicates a desire to help coupled with feelings of guilt.
Signed up: guilt/obligation	7	Yes	I've donated to them before, I respect them and if I didn't I would feel bad. I can't offer much but it would boost their advocate number by 1 and the higher the better. Never k...	The response clearly shows motivation to help, feelings of guilt, and respect for the charity.
Signed up: guilt/obligation	8	Yes	I decided to sign up as an advocate partially out of guilt	The response explicitly mentions signing up out of guilt.
Signed up: guilt/obligation	9	Yes	TO STAND BY AND BE SILENT IS AS QUILTY OF DOING IT YOUR SELF	The response indicates a sense of guilt for not taking action.
Signed up: guilt/obligation	10	Yes	I decided to sign up to be an advocate for World Wildlife Found because I believe they are doing really great work and it is my responsibility from moral point of view to suppor...	The response indicates a desire to help and a moral responsibility, supporting signup_help, signup_guilt, and signup_values.
Signed up: felt bad	1	Yes	I feel bad that Vets in our country are so neglected and overlooked	The respondent expressed feeling bad about the neglect of veterans, indicating pity.
Signed up: felt bad	2	Yes	Because I lost my son almost a year ago and I feel like the only way to ease the pain is to be involved with things that help people, I want to do things that have deeper meanin...	The response indicates personal loss and a desire to help others, aligning with values and a sense of deeper meaning.
Signed up: felt bad	3	Yes	The current situation in Gaza (genocide of Palestinians, destruction of buildings and infrastructure, bombing of refugee camps, etc.) saddens and sickens me to an enormous degre...	Response expresses sadness and a desire to help, indicating reasons aligned with multiple signup categories.
Signed up: felt bad	4	Yes	I felt bad seeing that nobody else signed up.	Respondent stated they felt bad, coding 'signup_feelbad'.
Signed up: felt bad	5	Yes	I noticed that no some signe up and it made me feel bad that they werent getting support so i felt if I signed up i can help persuade others.	Participant felt bad about others not getting support and wanted to help persuade others.
Signed up: felt bad	6	Yes	I haven't actually done it, but I feel bad that not enough people are signing up to help. Veterans are such an important part of life	Respondent feels bad about the lack of signups and recognizes the importance of veterans, indicating a desire to help.
Signed up: felt bad	7	Yes	I want to help people when I'm financially able to in whatever amount I can afford to give. It makes me feel better about myself, even if I don't have much to give. It makes me...	Respondent expresses desire to help others and reflect on the importance of helping.
Signed up: felt bad	8	Yes	I felt bad that nobody had signed up and as I've become older, I can appreciate the mission of the American Heart Association and cardiovascular health.	Felt bad about others not signing up and values align with the cause.
Signed up: felt bad	9	Yes	The Jewish community is going through a difficult time with the current war instigated by barbaric terrorist acts. It pains me to see educated Americans voicing slogans in suppo...	The response expresses pain and concern about current events and supports helping the Jewish community, thus fitting 'signup_help,' 'signup_fe...
Signed up: felt bad	10	Yes	I have not signed up yet. This survey is not over, and I have not yet been directed toward their website. I would like to find out a lot more about this group and also research...	Signed up to help advocate for peace and expressed feelings of sadness about conflict.

Notes. This table reports randomly sampled responses from the second experiment. Responses and coding notes are lightly truncated only where necessary to keep the table readable.

Table A17. API-coded open-ended response examples

Category	Ex.	Signed up	Response text	Coding note
Signed up: sympathy	1	Yes	The children need help in the world because of the wars that are going on. It's so sad to know that innocent children are being harmed and family members killed. They need our h...	The response highlights a desire to help children affected by war, indicating both a helping motivation and empathy for their suffering.
Signed up: sympathy	2	Yes	Because in my mind the veterans can never be helped or thanked enough for their service	Signed up to help veterans, indicating a sense of empathy.
Signed up: sympathy	3	Yes	Being a teacher, I advocate for children. I also feel they are the innocent victims of the paths their parents have chosen	The response indicates advocacy to help children and expresses sympathy for their situation.
Signed up: sympathy	4	Yes	I appreciate our service men and women and what they do for us and feel awful that they have suffered for our freedom.	The response expresses appreciation for service members and feelings of sorrow, aligning with signup_help, signup_feelbad, and signup_sympathy.
Signed up: sympathy	5	Yes	Children are helpless. They us to support these precious little ones.	The response emphasizes a desire to help children and support them.
Signed up: sympathy	6	Yes	I feel bad that Vets in our country are so neglected and overlooked	The respondent expressed feeling bad about the neglect of veterans, indicating sympathy.
Signed up: sympathy	7	Yes	Conflict around the world is growing at an extreme pace and the people who suffer most from that are civilians, and especially children. Whatever we can do to assist the displac...	The response indicates a desire to help and expresses concern for civilians and children suffering from conflict, fitting both help and sympathy.
Signed up: sympathy	8	Yes	I love animals and don't believe they have enough of a voice on their own. We need to take care of them, they are being displaced in areas with wildfires, deforestation, etc.	The response clearly indicates a desire to help animals due to their values and empathy towards their suffering.
Signed up: sympathy	9	Yes	As an anti-Zionist Jew who is watching the Israeli assault on Gaza unfold in shock and horror, I think JVP's vocal demand for a ceasefire and refusal to accept the atrocities of...	Respondent signed up due to a strong alignment with values and emotional reaction to the situation.
Signed up: sympathy	10	Yes	Children are precious. They can't help themselves and need us to ensure that they have food, shelter and love. Children grow up in an adult world that adults created and need ou...	Response emphasizes helping children and ensuring their well-being, fitting both help and sympathy categories.
Signed up: values alignment	1	Yes	It's a pretty objectively good association, and I have heard a lot of good things from them with regard to other people's opinions.	Respondent noted a positive perception of the association, indicating alignment with their values.
Signed up: values alignment	2	Yes	They serve a good cause and I probably have the time to do it right now.	The response indicates a desire to help due to the cause and availability.
Signed up: values alignment	3	Yes	It is the right thing to do and would be for a good cause	The response indicates a desire to help for a good cause, aligning with 'signup_help' and 'signup_values'.
Signed up: values alignment	4	Yes	I have been looking for ways to help the jewish community in gaza without feeding in the zionist movement happening in the US.	Respondent signed up to help the Jewish community and aligns with their values.
Signed up: values alignment	5	Yes	aligns with my values and works to help people like me and others I know	Response highlights values alignment and personal connection, thus coded for both signup_values and signup_personal.
Signed up: values alignment	6	Yes	It's a good cause and a cause I support.	Respondent supports the cause, aligning with their values.
Signed up: values alignment	7	Yes	Racial justice is a high priority for me and NAACP is a great organization to help bring that justice to fruition.	Response indicates alignment with racial justice values and intent to help.
Signed up: values alignment	8	Yes	Since the sign up was so low, I felt I should sign up as someone who agrees with the mission and values of the NAACP.	Signed up due to agreement with mission and values, indicating both norm and values motivations.
Signed up: values alignment	9	Yes	I decided to sign up because I think the cause is very important and veteran are the reason the country is safe.	Signed up due to the importance of the cause, aligning with personal values.
Signed up: values alignment	10	Yes	I love nature and care about the environment and I would love to be able to help a cause I care about and that affects everyone on the planet.	The response indicates a desire to help and mentions care for the environment, aligning with both help and values.

Notes. This table reports randomly sampled responses from the second experiment. Responses and coding notes are lightly truncated only where necessary to keep the table readable.

Table A18. API-coded open-ended response examples

Category	Ex.	Signed up	Response text	Coding note
Signed up: personal connection	1	Yes	I have a history of heart complications in my family.. Signing up is an no brainer	Signed up due to personal family history related to the cause.
Signed up: personal connection	2	Yes	I decided to sign up because Jewish people need others to stand with them in solidarity especially right now when there are many people being antisemetic in the world.	Signed up to support Jewish people, showing personal ties to the cause.
Signed up: personal connection	3	Yes	I signed up because I truly believe that we do not take care of our veterans in this country. We give more to illegal aliens than we do to people who have given their lives for...	The respondent signed up to support veterans, citing personal connections and values, which align with the cause.
Signed up: personal connection	4	Yes	I am not Jewish so would likely be rejected from this group as I was when I attempted to sign up for BLM.	Sign-up is influenced by personal identity and previous experiences.
Signed up: personal connection	5	Yes	I love children and I'll be glad to join	The response indicates a personal connection to the cause (love for children).
Signed up: personal connection	6	Yes	My mother passed away unexpectedly from heart disease so I support the heart association.	Respondent has a personal connection to the cause due to losing their mother.
Signed up: personal connection	7	Yes	I am a veteran and I feel it is my duty. I give to Tunnels to Towers and feel they have similar goals.	Motivated by duty as a veteran and alignment with charity goals.
Signed up: personal connection	8	Yes	I was always taught to support a good cause	Response indicates personal upbringing and values aligning with supporting good causes.
Signed up: personal connection	9	Yes	Heart disease runs in my fsmily so this charity fits wih=thin my values	Respondent signed up because the charity aligns with their values and is personally relevant due to family health issues.
Signed up: personal connection	10	Yes	I sincerely do it to support a good cause and in this case an orgranization that has given a lot in the past for the rights of black people and other races like mine.	The respondent signed up to support a cause they value and relate to personally.
Signed up: learn more	1	Yes	I thought it aligned with my values and I'm already familiar with the organization.	Aligned with values and familiarity with the organization.
Signed up: learn more	2	Yes	Yes I am willing to learn more about being an advocate but won't necessarily commit	Respondent is willing to learn more but is not committing.
Signed up: learn more	3	Yes	I am not Jewish, so I would feel like I am intruding. I would sign-up because I like their message, but I would want to do more research on the organization before committing.	Respondent feels they would intrude due to their identity but likes the message and wants to research.
Signed up: learn more	4	Yes	My husband has significant heart disease. He is 20 years older than me. I am a disabled RN and no longer able to work, but wonder if I were more ingaged with my husbands conditt...	Signed up to help others and learn more due to personal experience with husband's heart disease.
Signed up: learn more	5	Yes	I need to volunteer with the organization for a year to two months before I become an advocate. Additionally, I need to make sure what they say and what their mission is in alig...	The response indicates a need for alignment with values and a willingness to learn.
Signed up: learn more	6	Yes	I agreed to sign up because I thought I could get more information about what is expected of an advocate before signing up. I would like to be an advocate, but I would not be su...	The respondent signed up mainly to learn more, but mentioned concerns about not succeeding in fundraising, indicating a reluctance regarding t...
Signed up: learn more	7	Yes	I wanted to see what they have a need for and hope that I can help.	The response indicates a desire to help and learn about needs.
Signed up: learn more	8	Yes	I didn't sign up for advocacy directly, but imhere to help spread awareness and knowledge about issues like wildfires and the importance of organizations like the world wildfire...	Signed up to help spread awareness and knowledge.
Signed up: learn more	9	No	I want to research this more, I have a veteran living under my roof who has PTSD. He was denied benefits from the government after serving his country. This outrages me and I do...	The response shows a desire to research more and expresses outrage about the government's role with veterans, fitting 'nosign_info'.
Signed up: learn more	10	Yes	I decided to sign up because I care about the environment and study it in college.	The response indicates a strong alignment with values and a desire to learn about the environment.

Notes. This table reports randomly sampled responses from the second experiment. Responses and coding notes are lightly truncated only where necessary to keep the table readable.

Table A19. API-coded open-ended response examples

Category	Ex.	Signed up	Response text	Coding note
Norm/compliance	1	Yes	I like to believe that every little bit counts in life. I may be just a single person in a big world, but if millions of single people have the same mindset to do their part - t...	Respondent expresses belief in the collective impact of individual actions, indicating a desire to help and support.
Norm/compliance	2	Yes	I decided to sign up because it may help the charity to have more supporters but wouldn't cost me much, if anything, to help them.	Signed up to help the charity with minimal cost, indicating support and low-cost action.
Norm/compliance	3	Yes	I have seen them promoted quite a bit. I've purchased items from them. I would have to research further before getting any closer to see if there is known corruption, but I don'...	Signed up to help and believes in protecting animal habitats.
Norm/compliance	4	Yes	i support NAACP and am unsure what signing up entails, but figured it would be better if more people signed up	Signed up to learn more and feels it's important for more people to be involved.
Norm/compliance	5	Yes	I chose to sign up to be an advocate because I think it would be a quick and easy process, and I agree with their values. I also know they are trustworthy because they are assoc...	The respondent signed up because it was quick and easy, and they agree with the organization's values.
Norm/compliance	6	Yes	Because other people were doing it.	Respondent signed up because others were participating, fitting the 'signup_norm' category.
Norm/compliance	7	Yes	It's a cause I agree with, and it doesn't really inconvenience me or cost me anything	Aligned with values and convenient to participate.
Norm/compliance	8	Yes	I decided to sign up because it's something that requires little to no effort on my part and it is a cause that is of significant important to others in its benefits.	Coded 'signup_help' and 'signup_values' for the desire to help and alignment with important causes. 'signup_norm' noted for mentioning low eff...
Norm/compliance	9	Yes	It is an organization that aligns with my values. Also, I felt bad that no one else had signed up.	The respondent felt their values aligned and also expressed feeling bad about others not signing up.
Norm/compliance	10	Yes	I noticed that no some signe up and it made me feel bad that they werent getting support so i felt if I signed up i can help persuade others.	Participant felt bad about others not getting support and wanted to help persuade others.
Signed up: other	1	Yes	I have spare time depending on involvement, and the opportunity as presented here was made easy	The response indicates a willingness to engage due to ease and availability, but doesn't fit a specific signup category.
Signed up: other	2	Yes	No real reason not to as long as it is a quick process	Response indicates willingness due to the ease of the process, but no specific motivation is mentioned.
Signed up: other	3	Yes	I'm always looking for new charities to work with!	The response indicates a willingness to engage with charities, but does not fit any specific category beyond general interest.
Signed up: other	4	Yes	Due to the conflict in Gaza there is a rise in antisemitism	Response suggests signing up for awareness rather than a specific motivation.
Signed up: other	5	Yes	I'm very introverted and using my voice to advocate for things makes me scared.	Response indicates a fear of advocating but does not explicitly state a positive motivation for signing up.
Signed up: other	6	Yes	I was debating whether or not to sign up, I saw that 10/10 people signed up for it, and I felt a strong urge to join them. If I could be the inspiration for someone to begin hel...	The response mentions feeling inspired by others and a reputable organization, leading to a signup based on social norms and a general urge to...
Signed up: other	7	Yes	I LIKE THE IDEA	The response indicates general support but does not specify a concrete motivation from the coding options.
Signed up: other	8	Yes	I am already a member of the org and donate to them on a regular basis	Respondent is already a member and donates, suggesting they are engaged but not specifying a new reason for signing up.
Signed up: other	9	Yes	I already take part in some of their compaign.	The response indicates involvement in other campaigns but does not specify a reason fitting the specific categories.
Signed up: other	10	Yes	Trying to be more proactive	The response indicates a general intention to act but does not specify a clear reason for signing up.

Notes. This table reports randomly sampled responses from the second experiment. Responses and coding notes are lightly truncated only where necessary to keep the table readable.