Giving social pressure*

Woojin Kim†

July 20, 2021

Abstract

Why do we often avoid confronting someone who infringes a social norm? In this online experiment with over 1000 participants, university students faced the decision whether to send a pressuring email to a campus peer who had yet to register to vote near the deadline for the 2020 U.S. General Election. This experiment elicits incentivized measures of the participants' willingness to pay (WTP) to send the email and their predictions on the effectiveness and on the recipient's (dis)like of the message. On average, senders have a significantly negative WTP, and prefer to pay instead of sending the email. Senders generally believe their emails will be effective, and they value persuading the recipient to register. However, the majority expect the recipient to *dislike* the email, and are willing to pay -$1.4 on average to *avoid* sending the email. In contrast, the minority who expect the recipient to *like* being pressured are willing to pay +$1.6 on average to *send* the email themselves. This sensitivity to the anticipated recipient's (dis)like of the message is significantly lower when the messages can be sent anonymously. Under a simple framework, I estimate that 64% of the motives behind giving social pressure come from self-interest, 29% from the desire to persuade the recipient to follow the norm, and 7% from altruism. Senders also overestimate their social influence over their peers: they predict that their pressuring emails would increase the recipients' registration rate by 14 percentage points on average, but in reality, the messages have no impact.

***JEL* Classification**: D91, D72, P48

*I would like to thank Stefano DellaVigna, Ned Augenblick, Fred Finan, Elizabeth Linos, Dan Moore, Ricardo Perez-Truglia, and Dmitry Taubinsky for their essential guidance on this project, and members of the Berkeley behavioral community for numerous comments and discussions. I gratefully acknowledge funding from the Russell Sage Foundation, the Berkeley Institute for Young Americans, and the Experimental Social Science Laboratory (Xlab) at Berkeley. I also thank Uri Gneezy, Ryan Oprea, and Marta Serra-Garcia for their logistical support, and the Anderson Behavioral Lab at UCLA, the Experimental and Behavioral Economics Lab at UCSB, the Rady Incentives Lab at UCSD, and Xlab at UC Berkeley for facilitating this study. Preregistration: AEARCTR-0006364. IRB: UC Berkeley CPHS #2020-03-13093.

†UC Berkeley, woojin@berkeley.edu

1 Introduction

Many recent studies have used social pressure to increase compliance with norms. Classic examples include utility bills that compare energy usage to the local average, letters that shame tax delinquents, and mass mailings that reveal voting records to neighbors. However, the communications in these studies are typically sent by the researchers or their partnering organizations, and represent only a narrow "top-down" form of social pressure that we might encounter once a month on a utility bill. The bulk of the social pressure that we face everyday instead comes in "decentralized" interactions with our neighbors, colleagues, and friends. We rely on peer interactions to observe one another, confront those who misbehave, and ultimately sustain norms in our communities. Yet, we know very little about when and why we choose to socially pressure our peers. Suppose someone litters on the street—do you tell him to pick it up?

This paper studies the motives behind confronting peers who are not complying with a norm. In this online experiment with over 1000 participants across four University of California campuses, I ask students who had registered to vote ahead of the U.S. 2020 General Election to send pressuring emails to their peers who had yet to register. I find that on average, senders are *un*willing to pressure their peers. Under a simple framework, I estimate that self-interested motives, such as avoiding retaliation or a socially awkward encounter, drives almost two-thirds of the decision to socially pressure others. The desire to persuade the recipient to comply accounts for around one-third, and altruism explains the small remainder. Senders tend to overestimate their social influence over others: they predict that their pressuring emails would have a strong persuasive effect, but in reality, the messages have no impact on the recipients' registration rate.

This experimental sample offers several advantages for a study on social pressure. Most norms are localized forces within a social group, and the closer individuals are in community or in geography, the more inclined they are to agree on the norm, to want the norm to be followed, and to care what others think about them. These conditions are likely to be fulfilled among participants in this sample, who interact in the experiment as students from the same university. In this setting, the participants are clearly members of the same community within a few degrees of separation, but are unlikely to personally know each other. It mimics a host of daily interactions, for example, with people at the local grocery store or a friend of a friend on social media.

The key feature of the experimental design is to measure the preferences and beliefs underlying the psychology of giving social pressure that would be otherwise unobservable. For the main outcome of interest, I elicit the participants' willingness to pay (WTP) to send

a pressuring email to an unregistered peer. The survey incentivizes their choices and uses a binary search procedure to pinpoint the senders' preferences on a multiple price list.

I also capture two important beliefs to distinguish between the senders' motives for (not) pressuring others. I ask senders to forecast (i) how effective the message will be at persuading the recipient to register to vote, and (ii) how much the recipient will like or dislike the message. To incentivize truthful reporting, I award participants a bonus payment if they guess correctly on these predictions, for example, within 5 percentage points of the true rate.

On the first prediction, most senders (77%) believe that the recipient will be more likely to register if they sent a pressuring email, and the correlation between the expected effectiveness of the message and the WTP to send is significant and positive. A increase of 10 percentage points (pp.) in the predicted effect of the message corresponds to an increase of \$0.5 in the WTP to send it. The average expected effect is 14 pp., which translates to +\$0.7 in the average value of pressuring solely from the motive to persuade the recipient.

The willingness to pressure is also highly elastic to the second prediction of how much the senders expect the recipients to like or dislike being pressured. When senders believe the recipient will (strongly) dislike the message, the average WTP is -\$2.3. That is, these senders would pay to avoid sending an email telling an unregistered peer to register to vote. Conversely, the WTP pay switches sign from negative to positive for senders who believe the recipient will (strongly) like the email. These senders would pay +\$1.9 on average to send it themselves.

Senders may seem altruistic in placing such weight on the recipient's (dis)like of the message, but they show much less sensitivity when they can send a pressuring email *anonymously*. When senders believe that the recipient will dislike being pressured, they would much prefer to send an Anonymous Message (which the research team sends on their behalf without listing their name) rather than a Direct Message (which they must send from their own campus email accounts). This disparity in WTP between the Direct and Anonymous Messages disappears when senders expect the recipient either to be indifferent or to like the message. These patterns imply that the senders' elasticity to the anticipated recipient's (dis)like of the message is mostly driven by self-interested concerns, such as the risk of retaliation.

The desire to persuade the recipient and the fear of retaliation are specific features of communicating social pressure. Instead of an email about pressuring the recipient to register, some senders were instead assigned to an email informing the recipient how to search online to identify the local legislators in California. The willingness to send this Info Message is uncorrelated with how effective senders expect it will be at persuading the recipients to look up their local legislators; they value enforcing norms, not just any action. Senders of the Info

Message also display less sensitivity than senders of the Pressure Message to the anticipated recipient's (dis)like of the email under direct communication, as the risk of backlash may be lower without social pressure.

Under a simple framework, I decompose the (un)willingness to socially pressure others into three main motives: persuasion, altruism, and self-interest. The persuasive motive represents the preference for others to follow the norm and explains 29% of the variation. Altruism toward the recipient accounts for only 7%. Self-interested motives, which include time and effort costs, vindictive "cold glow" utility from asserting the norm, and preferences for maintaining privacy and avoiding retaliation and socially awkward interactions, comprise the largest share at 64%.

I also track the effect of the pressuring email on the recipients' voter registration and turnout rates. Participants who had yet to register when they began the study were randomly assigned to either receive a pressuring email (treatment group) or not receive any email (control group). Contrary to the senders' beliefs, I find no effect of the message on electoral participation. The effect on the registration rate is in fact negative and marginally significant, though responses on a survey conducted after the election suggest that this difference is likely due to noise.

This paper extends a large and growing literature on the wide-ranging effects of social pressure (e.g., Gerber, Green, and Larimer, 2008; Mas and Moretti, 2009; Funk, 2010; Allcott, 2011; DellaVigna, List, and Malmendier, 2012; Yoeli et al., 2013; Ashraf, 2014; Burszytn and Jensen , 2015; Bradler et al., 2016; Bursztyn, Fujiwara, and Pallais, 2017; DellaVigna et al., 2017; Lagomarsino, 2017; Perez-Truglia and Troiano, 2017; Bursztyn, González, and Yanagizawa-Drott, 2020; Johnson, 2020; Karing, 2021; Butera et al., forthcoming). In many of these studies, people become more likely to comply with the norm when their behavior is made public to one another, but we do not observe their actual interactions. The mechanisms behind these effects often remain bundled in a black box referred to as "social image concerns" or "social pressure". People could behave better because they gain utility merely from being seen by others as "high types" (Benabou and Tirole, 2006), or because they expect others to directly confront and punish them if they violate the norm. This paper instead features actual interactions between participants and finds that a quarter of the sample are confrontational types who would voluntarily confront noncompliers. Therefore, direct peer confrontation is a valid mechanism that may explain some of the effect observed in other social pressure studies.

As discussed in Section 5.1, this paper also informs models on the community enforcement of norms. The theoretical literature extends as far back as Axelrod (1986), Kandori (1992), and Ellison (1994), and remains an active field of research (e.g., Deb, Sugaya, and

Woltzky 2020; Bhaskar and Thomas forthcoming). Although this paper does not derive any deep theoretical contributions, it offers empirical evidence for model-building. As mentioned before, this paper finds that there are a proportion of confrontational types who derive net positive utility from pressuring noncompliers. If there are enough of these confrontational types in the population and there is a system through which they can sanction noncompliers, then a cooperative equilibrium can be generally sustained, even if there are "irrational" types who never cooperate as in Sugaya and Wolitzky (2020). This result aligns with numerous lab studies since Yamagishi (1986), Ostrom, Walker, and Gardner (1992), and Fehr and Gachter (2000) that allow participants to sanction defectors in public-goods games and find that peer punishment is both exercised and successful at maintaining cooperation.

This paper adds to many field and theoretical studies on the calculus of voting, especially on the role of social pressure (e.g., Gerber, Green, and Larimer, 2008; Ali and Lin, 2013; DellaVigna et al., 2017; Levine and Mattozzi, 2020). This experiment finds that a pressuring email sent from a peer to register and vote in a one-shot interaction *before* the election has no effect on voting. In contrast to studies such as Gerber, Green, and Larimer (2008), the senders in this experiment were not informed whether the recipient ultimately registered and voted, and there is no interaction *after* the election. As in DellaVigna et al. (2017), people appear to vote to tell others that they did, not from being told by others to do so.

2 Framework

This section presents a simple model that defines the different motives for pressuring others, and motivates an experimental design to identify each motive. The framework extends the idea from papers such as Kandel and Lazear (1992) and Calvó-Armengol and Jackson (2010) that people derive utility when their peers follow the norm, and may pressure each other typically at a cost to increase the chances of their peers taking the normative action. The model enriches this interaction and incorporates other motives that are likely to influence the decision to give social pressure.

2.1 Model

A potential sender S chooses $p \in \{0, 1\}$, or whether to pressure recipient R to take a binary normative action $x_R \in \{0, 1\}$. The sender's expected utility $u_S(p)$ consists of three main

motives in the following form:

$$u_S(p) = \overbrace{\gamma x_R(p)}^{\text{persuasion}} + \overbrace{\alpha u_R(p)}^{\text{altruism}} + \overbrace{C(p, u_R(p))}^{\text{self-interest}} \tag{1}$$

Persuasion. The first term $\gamma x_R(p)$ represents the persuasive motive (DellaVigna and Gentzkow, 2010) from inducing the recipient to follow the norm. $\gamma \geq 0$ is the amount that the sender values the recipient taking the normative action (i.e., $x_R = 1$). Importantly, x_R is a function of p, and the recipient's action may depend on whether the sender gives pressure.

Altruism. The second term $\alpha u_R(p)$ is the function for "pure" altruism (Andreoni, 1989), which the sender may feel on behalf of the recipient. Altruism arises naturally in this setting since social pressure tends to be costly for the recipient (DellaVigna, List, and Malmendier, 2012; Andreoni, Rao, and Trachtman, 2017; Butera et al., forthcoming). Following papers such as Bandiera, Barankay, and Rasul (2005) and DellaVigna et al. (2021), I assume a parsimonious functional form for altruism that assigns α weight to the recipient's utility $u_R(p)$, which can be affected by pressure. Altruism manifests through a positive weight $\alpha > 0$.

Self-interest. The third term $C(p, u_R(p))$ encompasses various self-interested motives for pressuring others found in the literature. To separate them, I define $C(p, u_R(p))$ as:

$$C(p, u_R(p)) \equiv (-c + v - w)p + \theta u_R(p) \tag{2}$$

The act of communicating social pressure costs the sender time and effort worth c (Elster, 1989; Falk, Fehr, and Fischbacher, 2002),[1] but can also provide vindictive "cold glow" utility v from asserting the norm (Ouss and Peysakhovich, 2015; Falk, Fehr, and Fischbacher, 2005). In many cases, pressuring requires revealing one's identity and also initiating an unsolicited, awkward conversation; w captures these consequences and other social costs from forgoing the anonymity of a bystander and drawing attention as an active norm enforcer. Lastly, the term $\theta u_R(p)$ assigns weight to the recipient's utility, but for self-preserving reasons rather

[1] A richer version of the model could allow the sender to choose an effort level e. For example, more effort could create a more persuasive or emotionally sensitive messages that increase the expected effect of the message or the recipient's utility. In practice, the assumption of a constant effort cost c in this experiment seems reasonable, given that less than 20% of those who sent the pressuring email edited the template message for more than 10 words (out of 56) beyond the greeting and sign-off. Messages were sent unedited from the template 56% of the time. Consequently, the most common cost c seems to be the hassle of opening the email client, copy and pasting the text, and clicking the send button.

than altruism. For instance, senders may be sensitive to the recipient's utility not out of kindness but to avoid conflict, ridicule, or retaliation (Denant-Boemont, Masclet, and Noussair, 2007).

2.2 Empirical strategy

In practice, the functions $x_R(p)$ and $u_R(p)$ are unknown to senders, who must decide whether to pressure based on their own beliefs. Denoting $\Delta_S z \equiv E_S[z|p=1] - E_S[z|p=0]$ as the sender's expectation of the change in z from pressuring, suppose we observe the sender's beliefs on the expected effect of giving pressure on the recipient's action ($\Delta_S x_R$) and utility ($\Delta_S u_R$), and a measure of their willingness to pressure ($\Delta_S u_S$).

Combining Equations 1 and 2, the relationship between these three variables is:

$$\Delta_S u_S = \gamma \Delta_S x_R + \alpha \Delta_S u_R + \theta \Delta_S u_R - w - c + v \tag{3}$$

Regressing $\Delta_S u_S$ on $\Delta_S x_R$ and $\Delta_S u_R$ provides an unbiased estimate of γ, but the observed sensitivity to the recipient's utility ($\alpha+\theta$) overestimates the extent of pure altruism (α). To overcome this omitted variable problem, this experiment introduces an exogenous treatment that disables the motives driving θ.

Consider the act of giving pressure *directly* (e.g., face-to-face or over workplace email) versus *anonymously* (e.g., on an open online forum). In the anonymous case, there are few concerns about conflict or backlash when the sender cannot be traced. On the other hand, pure altruism persists either way. Similarly, some social costs such as relinquishing one's privacy (w) only apply when communication is direct, whereas time and effort costs (c) and cold glow utility (v) are present in both.

Suppose the sender faces the decision to pressure as before, but now with an exogenous mode of delivery: Direct ($d=1$) or Anonymous ($d=0$). The sender's utility takes d as an additional argument, and Equation 3 becomes:

$$\Delta_S u_S(d) = \gamma \Delta_S x_R(d) + \alpha \Delta_S u_R(d) + (\theta \Delta_S u_R(d) - w)d - c + v \tag{4}$$

Now with the experimental variation in d, an ordinary linear regression can separately identify the parameters in Equation 4 (except for c and v, which remain together in the constant), and the willingness to pressure can be decomposed into the persuasive, altruistic, and self-interested motives.[2]

[2]Building on Footnote 1, this identification strategy can be biased if senders choose different levels of effort for composing direct versus anonymous messages. In the experiment, only direct messages were sent, so we cannot compare frequency and intensity of edits between direct and anonymous messages. Intuitively,

3 Experimental design

3.1 Logistics and overview

Based on the discussion in Section 2, the goal of the experimental design was to elicit reliable measures of the willingness to pressure others and senders' forecasts of the effects on the recipients. The study focused on the social norm of registering to vote for the election in the context of the 2020 U.S. General Election.

A total of 1,059 students were recruited through behavioral labs from four University of California campuses. The sample was restricted to those who had a valid California Driver License or State ID Card and were U.S. citizens, Californian residents, and over the age of 18. These requirements ensured that all participants were eligible to register to vote online. The experiment was conducted over several rounds from late September until the October 19 deadline for voter registration in California, which was 15 days before the election.

Participants earned between $5 and $15 for completing the study. In addition to the base compensation of $5, participants could earn up to $7 from the willingness to pay (WTP) elicitation, $1 for correct guesses (used as an incentive to report their true beliefs), and $2 for referring their friends to the study. The median total compensation was $12, and the median total participation time was around 15 minutes.

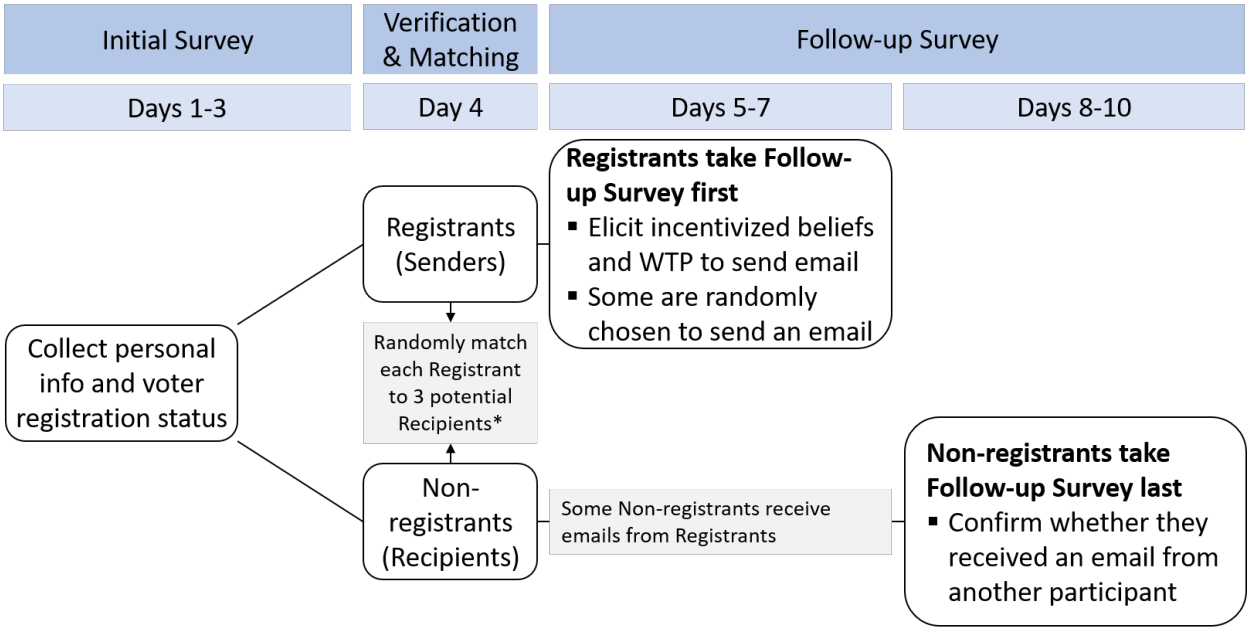
Figure 1 shows the chronological overview of the experiment. Participants took two Qualtrics surveys within 10 days: (i) the Initial Survey and (ii) the Follow-up Survey. On the Initial Survey, participants reported whether they had registered to vote prior to starting the study ("Registrants") or had yet to register ("Non-registrants"). After their self-reported voter registration status was verified, participants were randomly assigned to treatments, and then matched with other participants. The Registrants took the Follow-up Survey first. Near the end of the survey, some randomly selected Registrants were required to send an email pressuring their assigned Non-registrant to register to vote. After the Registrants completed the Follow-up Survey, the Non-registrants were sent a different version of the survey. The surveys are described in chronological order next.

3.2 Initial Survey

The Initial Survey began by asking participants whether they were registered to vote under their current residence (as required under Californian law) on the day *before* the Initial Survey

the effort observed in direct messages should theoretically be an upper bound for the effort in anonymous messages, since for direct messages, there are more reasons to exert effort (e.g., to optimize the terms $\theta\Delta_S u_R, w$). Yet senders put in only a minimal amount of effort to edit the direct messages (see Footnote 1), which implies that there would be a similar, low level of effort on the anonymous messages.

Figure 1: Design overview



* In the Pressure Message condition, potential Recipients are all Non-registrants.
In the Info Message condition, potential Recipients can be Registrants or Non-registrants.

was launched, to prevent participants from registering only after reading the question and stating they had been registered. Next, the survey asked for personal information including their name, campus email address, date of birth, ethnicity or race, gender, entering and graduating year, pursued degree, and field of study.

The survey then elicited two incentivized beliefs. The first was to predict the current and the final (i.e., by the election) voter registration rates among the participants from their campus, as a proxy for their perception of the norm intensity. The second was to forecast in percentage terms the likelihood that the Democratic party would gain a majority in the Senate and that the Republican party would gain a majority in the House.

These predictions were incentivized under a "correct guess bonus" scheme. After the election, one of their guesses across both the Initial and Follow-up Surveys would be randomly selected for each participant, and if it was correct, the participant would be awarded an extra \$1. An answer for the voter registration rate was deemed correct if it was within 5 percentage points of the actual rate; the forecasts for the election outcomes in Congress were graded using the binarized quadratic scoring rule (Hossain and Okui, 2013). The scheme was clearly explained to the participants, although the details of the binarized quadratic scoring rule were intentionally vague given the evidence in Danz, Vesterlund, and Wilson (2020).

Next, the survey informed participants that the goal of the experiment was broadly to study "peer-to-peer communication among students about the election" (without specifying

voter registration), and for this purpose, their full name, voter registration status, and email address could be shared with other participants who may send them an email. The survey was carefully phrased to leave the aim of the study as ambiguous as possible while providing adequate disclosure to mitigate experimenter demand effects (Zizzo, 2010). In fact, the purpose for the prior belief elicitation on the election outcomes in Congress was to draw attention away from voter registration, as well as to proxy for knowledge about the election. The survey ended with an option to refer a campus friend to take the study and a reminder that the link to the Follow-up Survey would be emailed within a week.[3]

3.3 Follow-up Survey

Before the Follow-up Survey, Registrants were verified against the official California voter file to ensure they had indeed registered to vote prior to the study, and then randomly matched with three other participants on their campus who were assigned as their potential recipients. The randomization and matching process was complex, with not only between- and within-subject experimental variation for the Registrants, but also a between-subjects intervention for the Non-registrants that required communication from the Registrants. Online Appendix Section A.2 provides the step-by-step details on the process, which focused on randomizing and balancing the experimental arms on both sides, maintaining incentive-compatibility for all the preference and belief elicitations, and prioritizing power for the main research questions.

3.3.1 Follow-up Survey: Registrants

The Follow-up Survey was first sent to the Registrants. Each Registrant was shown three different names and registration statuses of other participants from their campus. They characterized their relationship with each of them as either Strangers, Acquaintances, Friends, or Close friends.

Registrants were then informed that they could be randomly selected to email one of the three other participants. The topic of the email was randomly assigned, and could be one of two possibilities. In the Pressure Message condition, the email was "a pressuring email message about registering to vote," while in the Info Message condition, it was "an informative email message about local legislative districts." The Info Message condition served as a "control" email to distinguish the costs for sending any kind of unsolicited message, from the costs particular to sending a normative pressuring email.

[3]For the Non-registrants, the survey also elicited their incentivized willingness to pay to not share their information with other participants. Further details are in Online Appendix Section A.1.

Figure 2: Message template and guidelines

(a) Pressure Message

If you're selected to send a message, you'll be provided the **template message** below (no need to copy it now), which you can **edit** within general guidelines. The message must:

- Be addressed to the participant by name
- Mention that the participant hasn't registered to vote yet and suggest that they register
- Include the voter registration link (registertovote.ca.gov) and deadline (Oct. 19)
- Have NO partisan content or personal attacks
- Be within 500 characters

Template:

Subject: You haven't registered to vote yet!

Hi <Participant>,

I'm also participating in the 2020 Election Study, and I saw that you haven't registered to vote yet as of this week.

Don't you plan on voting? Then you should register ASAP at registertovote.ca.gov.

The deadline to register online is October 19. There's no excuse not to—it only takes 5 minutes!

(b) Info Message

If you're selected to send a message, you'll be provided the **template message** below (no need to copy it now), which you can **edit** within general guidelines. The message must:

- Be addressed to the participant by name
- Inform about local legislative districts
- Include the link to look up their district (findyourrep.legislature.ca.gov)
- Have NO partisan content or personal attacks
- Be within 500 characters

Template:

Subject: Local legislative districts in California

Hi <Participant>,

I'm also participating in the 2020 Election Study, and have some info to pass along about your local legislative district. Many Californians don't actually know which district they're in, or who their local representatives are.

You can look yours up at findyourrep.legislature.ca.gov.

These local reps decide the laws that affect us daily, so it's good to know who they are!

Registrants were shown a template of whichever message they were assigned, and were told that if they were selected to send the message, they could edit it within general guidelines. Figure 2 displays screen captures of the templates and guidelines.

Next, the survey introduced exogenous within-sender variation in the mode of delivery. If the Registrant was selected to send the email message, it could be sent either directly or anonymously, depending on the recipient. For a Direct Message, the Registrant would have to send the email to the recipient from their own campus account, but for an Anonymous Message, the Registrant would submit it to the research team who would then send it to the recipient without mentioning the Registrant's name. For each of the three potential recipients, the Registrant was told whether the email would have to be sent directly or anonymously (see Figure A.1).

The survey then elicited two beliefs for each potential recipient. First, Registrants indicated on a 7-point agree/disagree Likert scale how much they thought the recipient would like to receive the message from them (see Figure A.2). This prediction was incentivized by informing the Registrants that if they were selected to send the message, the recipient would also be asked on the same Likert scale how much he or she actually liked receiving it, and if their answers were the same, the Registrants could earn an extra $1 under the correct guess bonus scheme.

Second, Registrants predicted the likelihood in percentage terms of each potential recipient taking the action recommended by the message. In the Pressure Message, the recommended action was registering to vote by the election, and in the Info Message, it was knowing the names of the local representatives. The Registrants made two predictions for each potential recipient. They predicted the recipient's chances of taking the action without any message from the Registrant, and then with the message from the Registrant (see Figure A.3). The difference in the two predictions measures the Registrant's predicted effectiveness of the message for persuading the recipient.

These predictions were incentivized using the same binary quadratic scoring rule as in the Initial Survey (see Section 3.2). If the Registrant ultimately did not message the potential recipient, then the prediction in the first scenario (without any message) could be chosen for the correct guess bonus scheme, and if the Registrant did eventually message the potential recipient, then the prediction in the second scenario (with the message from the Registrant) could be chosen.[4]

After the predictions, the Registrants stated their willingness to pay to message each

[4]To determine whether the predictions were correct, each Non-registrant's voter registration status was tracked after the election (for potential senders in the Pressure Message condition), and Non-registrants were asked to list the names of their local representatives in the Follow-up Survey (for potential senders in the Info Message condition).

of the potential recipients. For each potential recipient, the Registrants went through a binary choice search procedure and answered four process-of-elimination questions that located their switching point on a 15-row multiple price list. The multiple price list was two-sided, allowing positive and negative WTP amounts between -$7 and $7 in $1 increments.[5]

The survey explained that their WTP choices could be taken into account under the following incentive-compatible program. After they submitted their WTP choices for all three potential recipients, the computer would either (A) randomly select one of the $3 \times 15 = 45$ rows across the three potential recipients and implement the Registrant's preferred option in that row, or (B) randomly decide whether the Registrant must send a message, and if so, to which of the three potential recipients. In Case B, the Registrant would receive an additional $7 regardless, since their WTP choices were not exercised. The survey gave no suggestion whether they should or should not send the message, and stressed that it was in their best interest to answer the questions truthfully. Since this was the main dependent variable in this study, an attention check preceded their choices, and the analysis in Section 4 includes only those who passed. (See Figure A.4 for the full set of instructions for the WTP elicitation.)

After submitting their WTP choices, Registrants were informed of the outcome of the program and whether they had to send a message. Those who were selected to send a message were given the same template they were shown before, with the option to edit the language. Once they had finished editing, the Direct Message senders were required to email their message to the recipient from their campus email account with the research team blind carbon-copied (BCC). All other Registrants were required to send the research team a confirmation email. This confirmation email served to measure the attrition at this stage from inattention or from the hassle of opening the email client and sending an email regardless of the content or recipient. Before closing the survey, Registrants had to tick checkboxes confirming that they had sent the required email, and that if they did not follow all the steps, their participation would be deemed incomplete and they would not receive any compensation.

3.3.2 Follow-up Survey: Non-registrants

After the three-day window for Registrants to take the Follow-up Survey had elapsed, a different version of a Follow-up Survey was sent to the Non-registrants. The survey asked Non-registrants whether they received an email about the election from another participant.

[5]If participants stated a WTP at either bound (-$7 or $7), the survey asked for their hypothetical, unbounded WTP. These uncensored answers are used in robustness checks in Online Appendix Section XXX (in progress).

If so, they were asked what the email was about in a brief phrase, how much they liked or disliked receiving it on a 7-point Likert scale, and whether the sender was a Stranger, Acquaintance, Friend, or Close Friend (or they do not remember). The Non-registrants were allowed to update their guesses from the Initial Survey for the voter registration rates on their campus and the outcomes of the election in Congress. Lastly, they were asked to list the names of any of their local legislators whom they knew.

4 Results

This section covers the results from the experiment. Section 4.1 begins by describing the observable characteristics of the sample. Given the complex design of the study, Section 4.2 provides a simple overview of all the experimental arms. Section 4.3 shows raw statistics on the main variables of interest from the experiment, before Section 4.4 provides the reduced-form results. Building on the reduced-form results, Section 4.5 decomposes the psychology of giving social pressure under the framework from Section 2. Lastly, Section 4.6 explores the effects of receiving social pressure on electoral participation.

4.1 Sample characteristics

In the final sample of 1,059 participants, the majority (84%) are Registrants who had already registered to vote before the study. This rate lies between the 74% of the eligible UC students who registered to vote for the 2018 Midterm Elections[6] and the 88% of all eligible Californians[7] who registered for the 2020 General Election.

Table 1 compares the demographics and beliefs about the norm between the Registrants and Non-registrants. The demographics of the two groups differ in conventional ways. The Non-registrants have a higher proportion of men (40% vs. 31%, $p = 0.03$), just as in the general population, men historically have lower voter registration and turnout rates than women.[8] Since older students have been able to register for longer and are more likely to be eligible to vote in a general election for the second time, Non-registrants are also younger (21.0 vs 21.6 years old, $p = 0.03$) and tend to be pursuing an undergraduate degree (93% vs. 87%, $p = 0.00$).

[6]Source: University of California.
This statistic will be updated to reflect the 2020 General Elections when the data for the UC campuses becomes available from the National Study of Learning, Voting, and Engagement (expected release in late 2021).

[7]Source: California Secretary of State.

[8]Source: Pew Research Center.

Registrants and Non-registrants also hold different beliefs about the intensity of the norm. Compared to the Registrants, Non-registrants predict a 5 percentage point (pp.) lower voter registration rate on their campuses ($p < 0.01$). The relationship between registration status and beliefs could go in either direction: participants could expect others to act like themselves, or Registrants may have registered because they perceived a stronger norm to do so.

Table 1: Sample characteristics

	Registrants	Non-registrants	Difference (SE)	Diff. p-value
Gender				
Male	0.31 [0.46]	0.40 [0.49]	-0.09 (0.04)	0.03
Female/other	0.69 [0.46]	0.60 [0.49]	0.09 (0.04)	0.03
Age (years)	21.55 [3.09]	21.01 [2.52]	0.54 (0.22)	0.01
Race/ethnicity				
White	0.30 [0.46]	0.32 [0.47]	-0.02 (0.04)	0.65
Asian	0.45 [0.50]	0.41 [0.49]	0.04 (0.04)	0.31
Other	0.25 [0.43]	0.27 [0.45]	-0.02 (0.04)	0.52
Degree				
Masters/doctoral	0.13 [0.34]	0.07 [0.25]	0.07 (0.02)	0.00
Undergraduate	0.87 [0.34]	0.93 [0.25]	-0.07 (0.02)	0.00
STEM major/field	0.70 [0.46]	0.72 [0.45]	-0.03 (0.04)	0.51
Year degree started	2018.09 [1.25]	2018.32 [1.13]	-0.23 (0.10)	0.02
Predictions				
Current reg. rate (%)	60.98 [16.90]	55.62 [16.63]	5.35 (1.41)	0.00
Final reg. rate (%)	74.58 [14.52]	69.69 [16.25]	4.89 (1.35)	0.00
	894	165		

This table shows averages with standard deviations in brackets. The "reg. rate" predictions refer to the forecasts for the voter registration rates on the participant's campus.

4.2 Summary of experimental conditions

Participants were assigned to different experimental conditions depending on whether they were Registrants or Non-registrants. For the Registrants, Table 2a summarizes the experimental arms. Each Registrant considered sending an email message to three other participants. The type of message (Pressure vs. Info) varied between Registrants, but the mode of delivery (Direct vs. Anonymous) varied among the three potential recipients for each Registrant. The assignment to the experimental groups was balanced on all the observable characteristics collected in the study (joint F-test p-value= 0.80, see Table A.1a). Table A.2 summarizes the sample attrition at each stage of the study. Of the 1,118 Registrants who

submitted the Initial Survey, 84% completed the Follow-up Survey. Ninety-six percent of those Registrants finished both surveys, passed the attention check, sent the required email, and form the final sample of Registrants in the analysis.

For the Non-registrants, Table 2b shows the two experimental groups. The assignment to these two groups was also balanced on all the observables (joint F-test p-value= 0.91, see Tables A.1b and A.1c). Out of the 165 Non-registrants who submitted the Initial Survey, 150 (roughly 90%) were randomly selected either to not receive any message or to receive a Direct Pressure Message regardless of their WTP choices (described in Online Appendix Section A.1) or their randomly assigned sender's WTP. Only these 150 Non-registrants who exogenously received or did not receive a Direct Pressure Message are included in the Section 4.6 analysis on the effects of receiving a message.

Table 2a: Experimental arms: Registrants (senders)

	Between-subjects	
Within-subject	Pressure Message ($N = 527$)	Info Message ($N = 367$)
Direct Message	Direct Pressure Msg. (793)	Direct Info Msg. (545)
Anonymous Message	Anon. Direct Msg. (788)	Anon. Info Msg. (555)

This table shows the number of responses in each experimental arm for the Registrants.

Table 2b: Experimental arms: Non-registrants (recipients)

Control	Treatment
No message ($N = 55$)	Receive Direct Pressure Message ($N = 95$)

This table counts only the Non-registrants who were randomly assigned to receive no message or to receive a Direct Pressure Message regardless of their willingness-to-pay (WTP) choices (see Section A.1) or their randomly selected Registrant sender's WTP choices. This table excludes 15 Non-registrants who received or did not receive a message endogenously depending on their or their sender's WTP choices.

4.3 Summary statistics

Before the formal analysis, I provide raw summary statistics on the main outcome (the willingness to send a pressuring message) and the two explanatory variables of interest (the two beliefs elicited from the sender on the effects of the message).

Figure 3 depicts the entire distribution of the senders' WTP for the Direct and Anonymous Messages separately. Most senders are either indifferent or have very strong preferences about pressuring. In the Direct case, around 40% of the responses have a neutral WTP of \$0.

On each of the extremes, there are 12-13% who state the maximum possible WTP to either send ($7) or be exempt from sending the message (-$7). A quarter have a negative WTP in the intermediate range (from -$1 to -$6), and the remaining 11% are in the intermediate positive range. Overall, the average WTP is -$0.4 (s.e.=0.2), indicating that the typical sender would prefer not to pressure the recipient directly.

In the Anonymous case, the general shape of the distribution is similar, but there is a significantly lower fraction on the negative extremity of -$7 (difference=-7.4 pp., s.e.=1.3). With fewer responses in the negative intermediate range as well, this shift in mass pushes the rate of indifference to more than 50% and slightly spills over into the positive WTP values. The average WTP switches sign to +$0.8 (s.e.=0.2), which is $1.2 greater than the Direct case ($p < 0.01$).[9]

Yet the preference for sending the Anonymous rather than the Direct Message may be due to differences in the perceived effectiveness and in the recipient's affinity for the message. Figure 4 compares the senders' beliefs on these two key variables between the Direct and Anonymous Messages. In either case, senders hold optimistic expectations on the persuasive power of their messages. On average, they forecast their messages to increase the recipient's chances of registering by 12.5 pp. (s.e.=0.6). Although the senders predict the Direct Message to be slightly more effective than the Anonymous Message (13.9 vs 11.1 pp., difference $p < 0.01$), they anticipate the recipients would prefer to receive the Anonymous Message instead. In response to the prompt "[Recipient] would like my [Direct/Anonymous] Pressure Message," 47.6% of the responses for the Anonymous Message expect a negative reaction from the recipient, compared to 60.7% (difference $p < 0.01$) of the responses for the Direct Message.[10]

Figure 5 provides a broad overview of the relationships between the WTP to pressure and the sender's two beliefs, and compares the patterns between the Direct and Anonymous Messages. Senders seem somewhat considerate toward the recipient when communication is anonymous, but are much more sensitive when they must send the message directly. When senders predict the recipient to (strongly) *dislike* being pressured, they are willing to pay $2.3 (s.e.=0.3) to *avoid* sending a Direct Message, but their WTP is indistinguishable from $0 when they can send the message anonymously (difference=$2.1, $p < 0.01$). In contrast, when senders predict the recipient to (strongly) *like* the message, the WTP switches sign

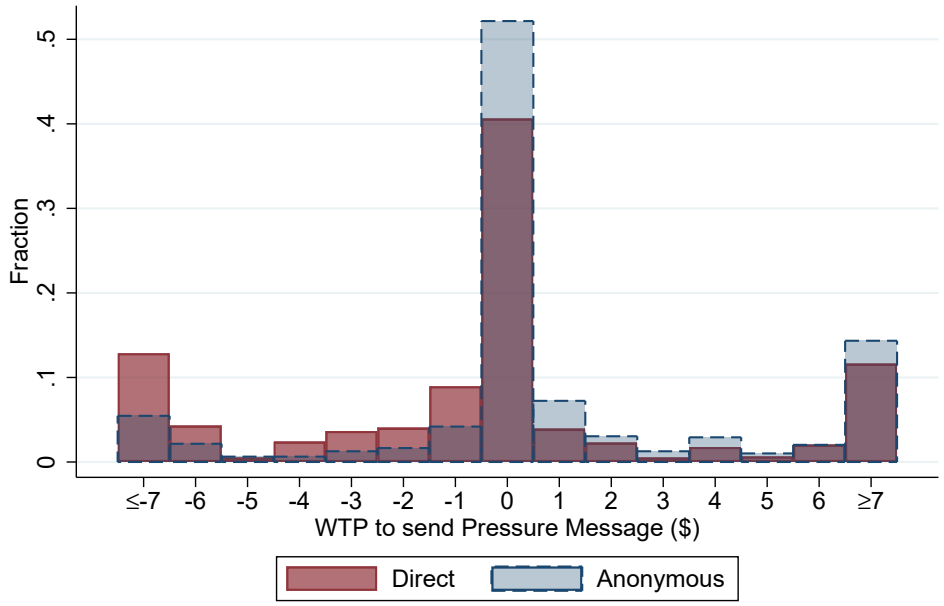
[9]Figure A.5 shows the within-sender differences in the WTP to send a Direct vs. Anonymous Message.

[10]Senders answered this prompt on a 7-point Likert scale (Strongly disagree, Disagree, Somewhat disagree, Neither agree nor disagree, Somewhat agree, Agree, Strongly agree). There were less than 50 responses stating "Strongly disagree" for the Anonymous Message and "Strongly agree" for both messages. As arranged in the pre-registration plan, I combine these categories with the adjacent responses. That is, I merge responses for "Strongly disagree" and "Disagree" together under "(Strongly) disagree", and likewise for the responses on the other end.

to +$1.9 (s.e.=0.5) for the Direct Message and +$1.5 (s.e.=0.4) for the Anonymous case (difference=$0.4, $p = 0.47$).[11]

The difference in WTP between the Direct and Anonymous Messages is only present when senders expect the recipient to have a negative reaction; it disappears as soon as the sender believes the recipient will be indifferent to the message. At the same time, senders predict that a well-received message will be more effective at persuading the recipient to register. The next section takes these intuitive and informative correlations and moves onto a multivariate analysis.

Figure 3: Distribution of willingness to pay (WTP) to send the Pressure Message



Direct: Mean=-$0.42 (SE=0.18), $N_{responses}$=793
Anonymous: Mean=$0.78 (SE=0.16), $N_{responses}$=788
Difference p<0.01, standard errors clustered by sender

[11]Figure A.6 shows the entire distribution of WTP responses for the Presssure Message within each response on the Likert scale for the predicted recipient's (dis)like of the message.

Figure 4: Senders' predictions on the effects of the Pressure Message

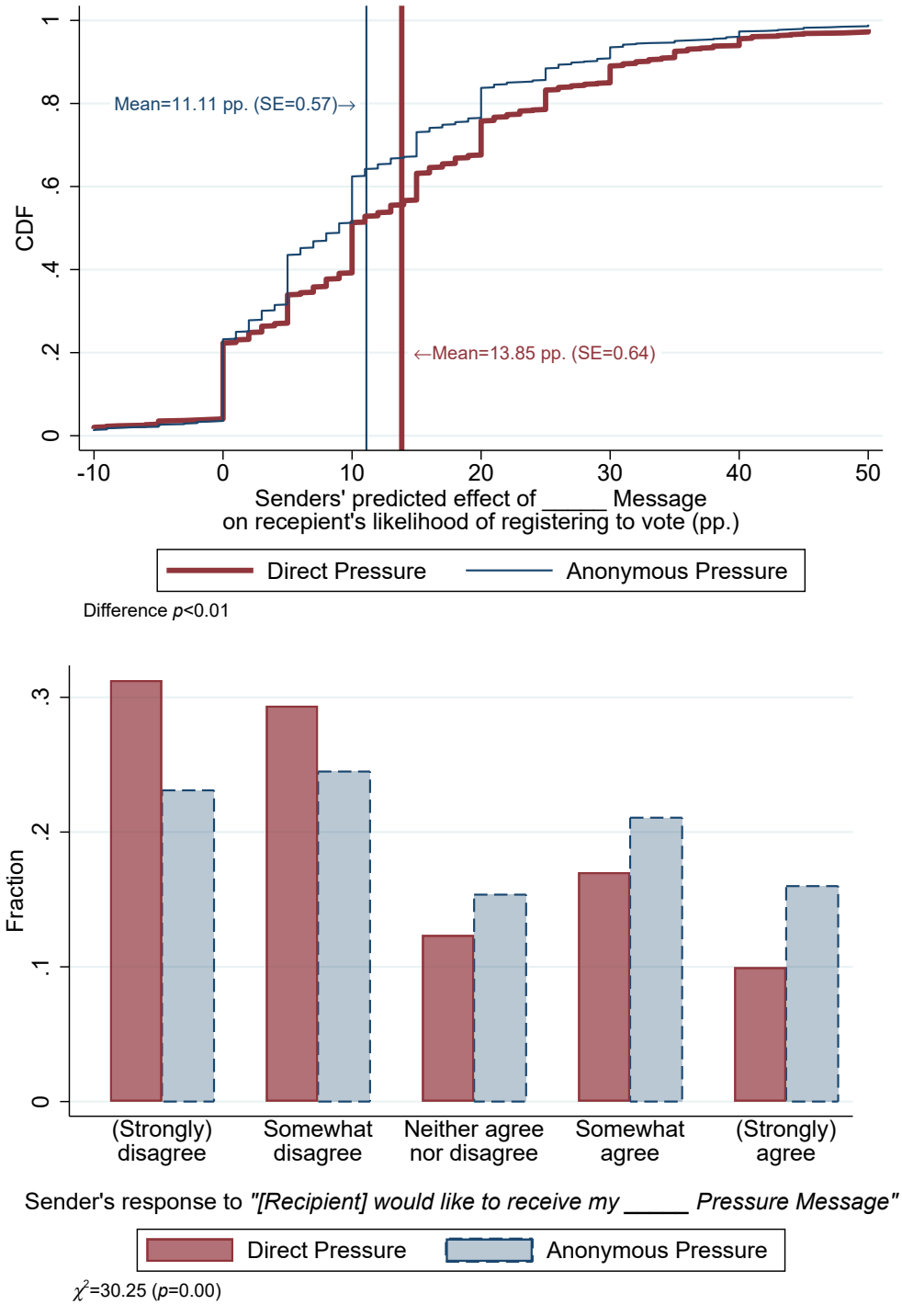
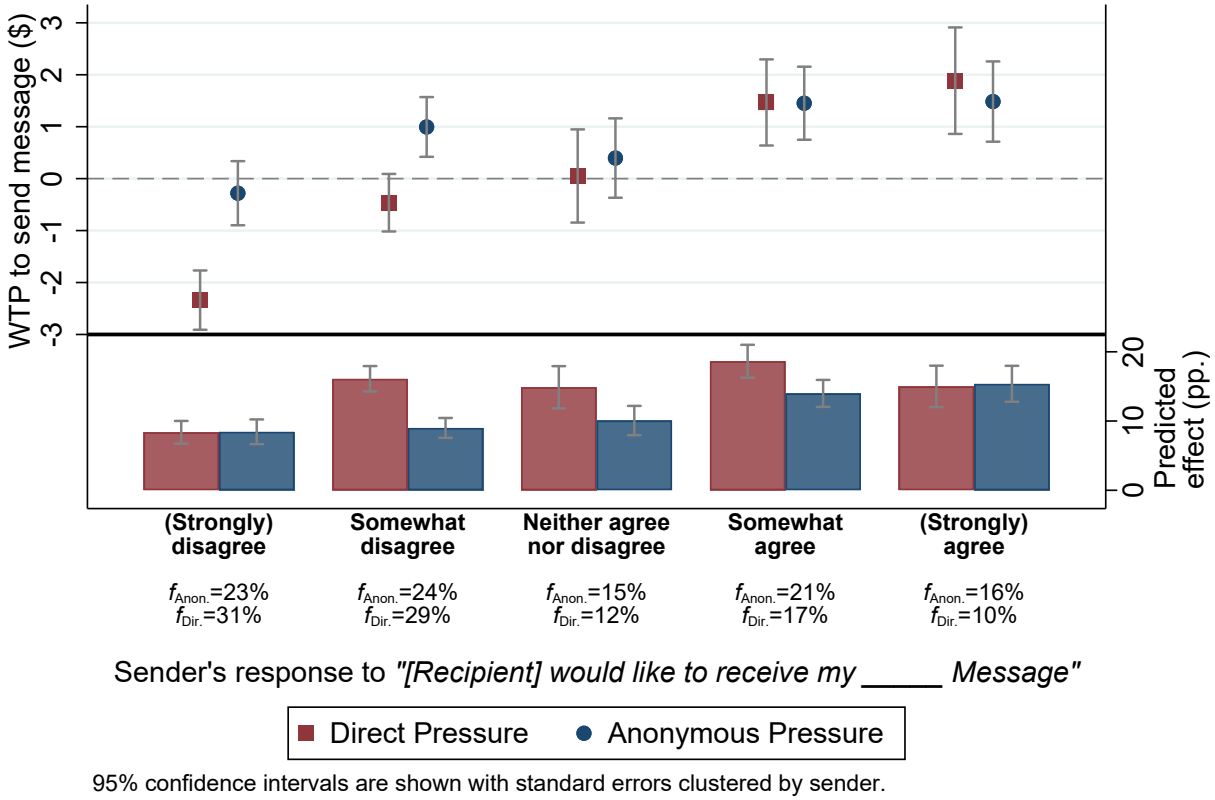


Figure 5: WTP to pressure and senders' beliefs



95% confidence intervals are shown with standard errors clustered by sender.

4.4 Reduced-form results

4.4.1 Pressure Message

Based on the discussion in Section 2, I estimate Equation 4 through the following regression:

$$WTP_{S,R} = \gamma \Delta_S x_R + \alpha \Delta_S \tilde{u}_R + \theta(\Delta_S \tilde{u}_R - w) d_{S,R} + \nu + \varepsilon_{S,R} \tag{5}$$

where $WTP_{S,R}$ is the sender S's willingness to pay to send the Pressure Message to recipient R, $\Delta_S x_R$ is the sender's prediction of the change in the likelihood that the recipient registers to vote upon receiving the message, $\Delta_S \tilde{u}_R$ is the sender's prediction of the recipient's (dis)like of the message, $d_{S,R}$ is an indicator for whether the message is Direct or Anonymous, ν is a constant, and $\varepsilon_{S,R}$ is an error term.

There are two practical differences between Equations 4 and 5. First, WTP is used as a monetary proxy for the sender's utility. While WTP is likely to be riddled with noise, it still remains a widely-used, well-tested method to elicit preferences. Second, the sender's agreement with the statement "[Recipient] would like my [Direct/Anonymous] Pressure Message" takes the place of the recipient's utility from receiving the message as perceived by the sender. Obviously, the responses to the prompt do not directly reflect perceptions of the recipient's utility; nevertheless, it is reasonable to assume that there is a monotonic mapping between the two.

Column 1 of Table 3 presents the regression results. The coefficient on the predicted effectiveness of the message is large and significant: senders care about persuading others to follow the norm. A 10 pp. increase (s.d.=13.6 pp.) in the expected effectiveness of the message in persuading the recipient to register corresponds to an increase of \$0.5 (s.e.=0.1) in the sender's WTP.

The second prediction on how much the recipient would (dis)like the message is included in the regression as a linear term, which seems like a reasonable approximation given the trend in Figure 5.[12] The predictions lie on a condensed 5-point Likert scale captured in integers from -2 to 2, where 0 corresponds to "Neither agree nor disagree". The coefficient on this variable represents the degree of true altruism, α. Senders do display altruistic concerns: even when sending is anonymous, a 1-point increase in the Likert response (e.g., from "Neither agree nor disagree" to "Somewhat agree") increases the WTP by \$0.3 (s.e.=0.1).

[12] This linearity assumption is relaxed in Section 4.5.

Table 3: Sender motives

	Pressure		Info		
Dep. Var.: WTP to send message ($)	(1)	(2)	(3)	(4)	(1)-(3)
Sender's predictions on:					
Effectiveness of message (10 pp.)	0.50	0.07	0.01	-0.06	0.49
	(0.10)	(0.15)	(0.06)	(0.07)	(0.12)
Recipient's (dis)like of message [-2, 2]	0.31	0.25	0.26	0.12	0.05
	(0.11)	(0.10)	(0.09)	(0.08)	(0.14)
Direct Message	-0.85	-0.84	-1.16	-0.97	0.30
	(0.17)	(0.13)	(0.16)	(0.13)	(0.23)
Direct×Predicted recipient's (dis)like	0.68	0.39	0.28	0.05	0.40
	(0.14)	(0.09)	(0.14)	(0.10)	(0.19)
Constant	0.28	0.71	0.15	0.33	0.15
	(0.19)	(0.18)	(0.20)	(0.18)	(0.20)
Sender fixed effects		✓		✓	
Responses	1581	1581	1100	1100	2681
Senders	527	527	367	367	894
R^2	0.14	0.16	0.07	0.11	0.13

Coefficient estimates are from linear regressions with standard errors clustered by sender shown below in parentheses. These regressions estimate the specification in Equation 4 and include two predictions made by the senders. First, senders predicted the effect of the message on the recipient's likelihood of registering to vote by the election (Pressure Message) or knowing the names of their local legislators (Info Message). The coefficient on this prediction corresponds to a 10 percentage point (pp.) increase in the predicted effectiveness. Second, senders indicated on a Likert scale how much they agreed with the statement that the recipient would like to receive their message. The coefficient on this prediction corresponds to a 1-point increase on the Likert scale ranging from -2 (strongly disagree) to 2 (strongly agree).

Yet the bulk of the sensitivity to the anticipated recipient's (dis)like of the message only manifests when sending is direct. In the Direct condition, a 1-point increase in the Likert scale raises the WTP by an additional $0.7 (s.e.=0.1). If we observed only Direct Messages without an Anonymous comparison, we could overstate the extent of true altruism by 220%. In other words, 69% (s.e.=10) of the elasticity with respect to the recipient's affinity for the message is for self-preserving reasons, such as avoiding conflict or retaliation.

There also appear to be large constant social costs from pressuring directly. The coefficient on the Direct Message indicator is -$0.9 (s.e.=0.2), which could arise for various reasons. As discussed in Section 2, students may hold preferences for privacy to not disclose to other students that they are participating in the study. The cost could also reflect the awkwardness of initiating an unsolicited conversation.

In the theoretical framework, the standalone constant term ν in the regression incorporates both negative time and effort costs c and the positive "cold glow" utility v from

asserting the norm to a noncomplier. In this case, the point estimate is positive but statistically indistinguishable from \$0, suggesting that the two components may both be small in magnitude or counteracting one another.

Column 2 adds sender fixed effects to Equation 5. While the coefficient on *Direct*× *Predicted recipient's (dis)like* falls to \$0.4, it still outweighs the coefficient on the uninteracted term *Predicted recipient's (dis)like*. The constant, which is the average of the sender fixed effects, increases significantly to \$0.7 (s.e.=0.2). Contrary to intuition, however, the coefficient on effectiveness becomes insignificant. One major downside in the fixed effects model is that 80% of the total variance in the predicted effectiveness of the message sources from between-sender variation. For this reason, the specification in Column 1 without fixed effects is the preferred specification and is used to estimate the model parameters in the Section 4.5 decomposition.

4.4.2 Pressure Message: Heterogeneity

Table 4 explores the heterogeneity in the willingness to pressure within three observable categories: gender, race, and relationship to the recipient. I run the specification from Column 1 of Table 3 on subgroups within each category and compare the coefficients. For gender, Column 1 of Table 4 shows the results for female and nonbinary senders,[13] and Column 2 for male senders. Compared to male senders, female and nonbinary senders exhibit a significantly higher constant social cost from pressuring directly (difference=-\$0.8, s.e.=0.4). Next, Columns 3 and 4 investigate differences when the sender faces a recipient of the same race versus another race, which could be inferred from the names of the recipients. No differences emerge, which may be due to the lack of salience of race in this online, text-only mode of communication. Lastly, Column 5 displays the results for a small sample of 22 responses in which the sender knows the recipient. Compared to senders facing strangers in Column 6, senders messaging their friends are significantly more sensitive to the anticipated recipient's (dis)like of the pressure and express a higher degree of true altruism (difference=1.6, s.e.=0.6). Although this result seems natural, it must be taken with caution given the small sample size.

[13]There are 7 senders who identify as nonbinary.

Table 4: Heterogeneity in willingness to pressure

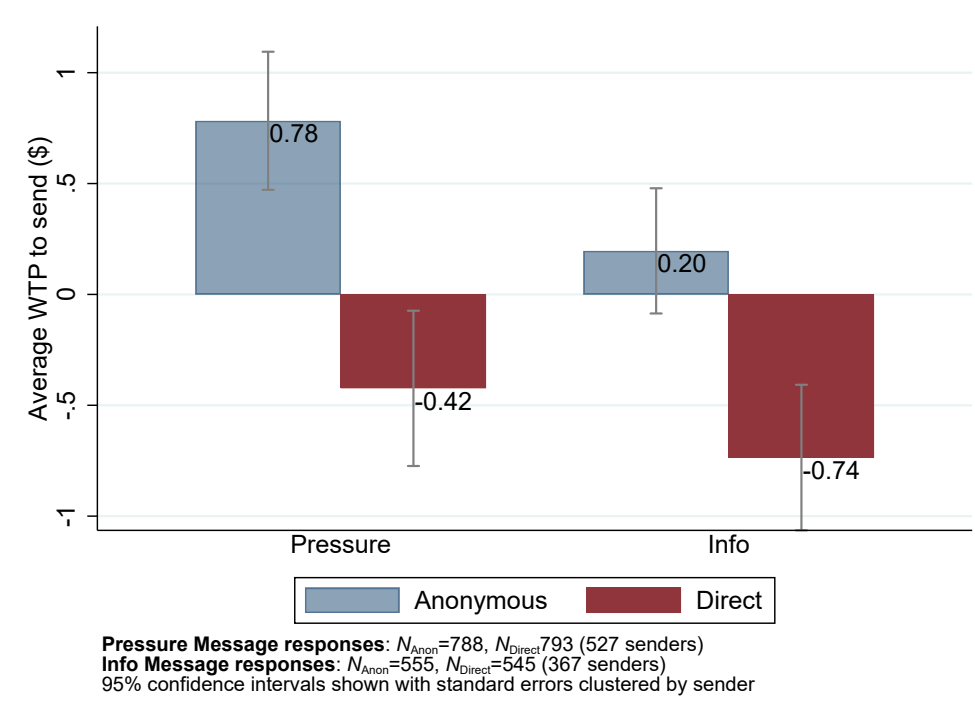
	Sender's gender			Recipient's race			Relationship		
Dep. Var.: WTP to send Pressure Msg. ($)	Fem./other (1)	Male (2)	(1)-(2)	Same (3)	Different (4)	(3)-(4)	Friends (5)	Strangers (6)	(5)-(6)
Sender's predictions on:									
Effectiveness of message (10 pp.)	0.53	0.40	0.13	0.48	0.52	-0.04	0.96	0.50	0.46
	(0.12)	(0.18)	(0.22)	(0.14)	(0.11)	(0.14)	(0.73)	(0.10)	(0.64)
Recipient's (dis)like of message [-2, 2]	0.19	0.53	-0.34	0.27	0.33	-0.05	1.91	0.30	1.61
	(0.13)	(0.19)	(0.23)	(0.18)	(0.12)	(0.20)	(0.62)	(0.11)	(0.56)
Direct Message	-1.09	-0.27	-0.83	-1.03	-0.75	-0.28	0.79	-0.87	1.66
	(0.19)	(0.36)	(0.41)	(0.29)	(0.22)	(0.38)	(1.89)	(0.18)	(1.68)
Direct×Predicted recipient's (dis)like	0.71	0.72	-0.01	0.64	0.71	-0.07	-0.86	0.69	-1.55
	(0.16)	(0.27)	(0.31)	(0.22)	(0.17)	(0.27)	(1.14)	(0.14)	(1.02)
Constant	0.39	0.11	0.27	0.44	0.19	0.25	-1.62	0.29	-1.91
	(0.24)	(0.34)	(0.41)	(0.27)	(0.22)	(0.29)	(1.54)	(0.20)	(1.37)
Responses	1086	495	1581	565	1016	1581	22	1559	1581
Senders	362	165	527	386	499	527	20	527	527
R^2	0.14	0.17	0.15	0.14	0.15	0.15	0.46	0.14	0.15

Coefficient estimates are from linear regressions with standard errors clustered by sender shown below in parentheses. These regressions replicate the specification in Column 1 of Table 3 on the subsample indicated in the column headers.

4.4.3 Info Message

How does social pressure differ from non-normative communication? Figure 6 compares the average WTP to send the Pressure and Info Messages. In both cases, the average WTP to send a Direct Message is significantly negative and around $1 lower than the WTP to send an Anonymous Message. The average WTP to send an Anonymous *Info* Message, however, is statistically indistinguishable from $0, whereas the average WTP to send an Anonymous *Pressure* Message is significantly positive at $0.8 (s.e.=0.2).

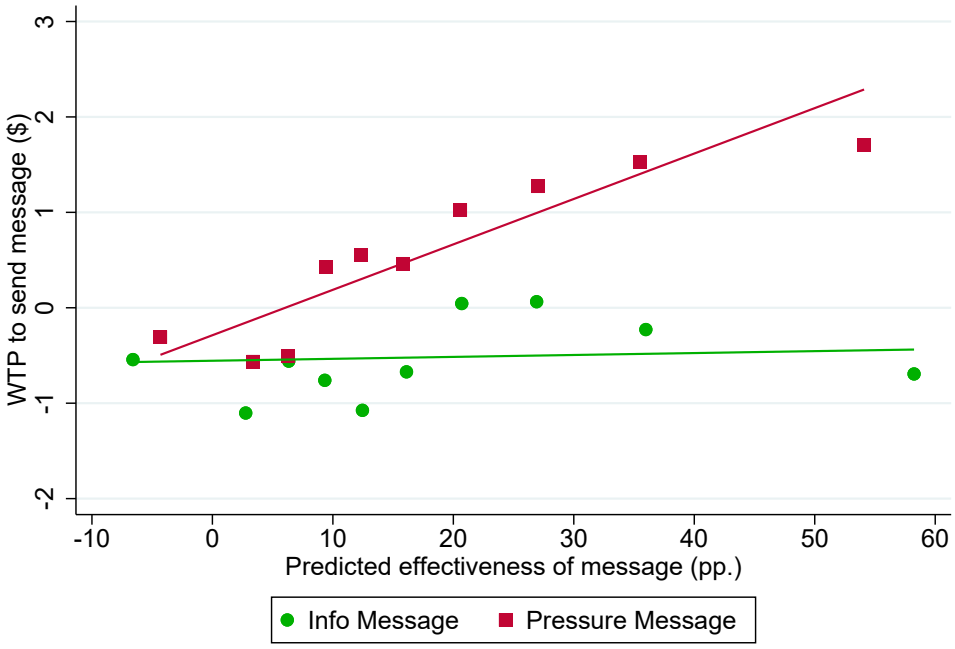
Figure 6: Average WTP for Pressure and Info Messages



To better understand the differences in motives, Column 3 of Table 3 estimates the same regression of Equation 5 but on the WTP responses for the Info Message instead. The motives for sending the Pressure and Info Messages contrast in two ways. First, senders do not care about the effectiveness of the Info Message: the slope on the predicted effect of the Info Message is zero. Figure 7 plots the bin scatter of the senders' WTP and their predicted effectiveness of the message. As described in Section 3, the purpose of the Info Message is to inform the recipient how to search online for their local legislators. The experiment elicits the predicted effectiveness by asking senders to forecast the percentage point effect of the message on the chances that the recipients could name their local legislators, and senders generally expect the Info Message to be very effective (average=26.4 pp., s.e.=1.0, see Figure A.7). Yet knowing one's local legislators—unlike registering to vote—is not a social norm.

Hence, the effectiveness of the Info Message does not factor into the senders' decision.

Figure 7: Bin scatter of WTP and predicted effectiveness



The WTP values in this bin scatter have been residualized on d, $\Delta_S u_R$, and the interaction term.

The second contrast is that senders of the Pressure Message seem to be much warier of retaliation. The coefficient on *Direct*×*Predicted recipient's (dis)like* is 0.4 higher (s.e.=0.2) for the Pressure Message than for the Info Message. While the degree of true altruism is similar in both, senders exhibit heightened sensitivity to the anticipated recipient's (dis)like when sending a Pressure Message directly.

These comparisons deliver two distinguishing features of social norms. First, norms are actions that people want others to do. Hence, in Figure 6, the average WTP to send the Anonymous Pressure Message is significantly higher than the WTP to send an Anonymous Info Message. This difference, however, is subdued by the second distinguishing feature of norms when the sender must confront the recipient directly: people perceive a higher risk of retaliation from socially pressuring others than from relaying non-normative communication.[14]

[14]For the Pressure Message, senders only saw Non-registrants as potential recipients, but for the Info Message, senders encountered both Registrants and Non-registrants as potential recipients. Table A.3 affirms that the differences between the Pressure and Info Messages are not due to whether the recipient is a Registrant or Non-registrant.

4.5 Decomposition of the motives behind social pressure

Building on the reduced-form results in Table 3, the decomposition in Table 5 calculates and compares the magnitudes of the three main motives from the theoretical framework: persuasion, altruism, and self-interest. I use the estimated parameters from Column 1 of Table 3 and the sample averages of the senders' beliefs for the Direct Pressure Message. For example, the persuasive motive in Equation 4 is represented by the term $\gamma\Delta_S x_R$; the point estimate for γ from Column 1 of Table 3 is 0.05, and the sample average of the expected effectiveness $\Delta_S x_R$ for a Direct Pressure Message is 13.9 pp. Consequently, the average value of pressuring purely from persuading the recipient to register to vote is $0.05 \times 13.9 = +\$0.7$ (Delta method s.e.=0.1). Under the same approach, the average value of the altruism motive is -\$0.2 (s.e.=0.1), which is negative since senders expect recipients to dislike being pressured on average ($\overline{\Delta_S \tilde{u}_R} = -0.6$). In this case, altruism for the recipient deters senders from pressuring by a monetary equivalent of -\$0.2. Lastly, the self-interested motives have both positive and negative components. Exposing oneself to the risk of retaliation dissuades pressuring by -\$0.4 (s.e.=0.1), and initiating an awkward, unsolicited interaction by a further -\$0.9 (s.e.=0.2). Meanwhile, vindictive "cold glow" utility outweighs any effort costs, and the two combined are worth +\$0.3 (s.e.=0.2).

The last column shows the weight of each motive in the decision to give pressure, calculated as the absolute value of each motive divided by the sum of the absolute values for all the motives (\$2.4). While the average WTP to send a Direct Pressure Message seems small in magnitude (though still statistically significant) at -\$0.4, the average in fact conceals strong positive and negative motives that sway senders both to pressure and not to pressure.

The most powerful motive in the psychology of socially pressuring others is self-interest, which comprises almost two-thirds (64%, s.e.=6) of the senders' (un)willingness to pressure. Enforcing the norm and persuading others to comply drives a sizable share of the decision (29%, s.e.=6), but is still less than half of the self-interested motives. Comparatively, altruistic concerns matter very little (7%, s.e.=3).

Beyond the overall average, the decomposition can be further applied within each response on the Likert scale of how much the sender expects the recipient to (dis)like being pressured. To do so, I relax the assumption in Equation 5 that restricts the relationship between $WTP_{S,R}$ and $\Delta_S \tilde{u}_R$ to be linear, and instead estimate it nonparametrically in the following equation:

$$WTP_{S,R} = \gamma\Delta_S x_R + \sum_{k=-2}^{2} \alpha^k \mathbf{1}\{\Delta_S \tilde{u}_R = k\} + (\sum_{k=-2}^{2} \theta^k \mathbf{1}\{\Delta_S \tilde{u}_R = k\} - w) d_{S,R} + \nu + \varepsilon_{S,R} \quad (6)$$

Table 5: Decomposition

Motive	Parameter	Estimate (SE)	Absolute weight (SE)
Persuasion	$\gamma\Delta_S x_R$	\$0.70 (0.14)	29% (6)
Altruism	$\alpha\Delta_S\tilde{u}_R$	-\$0.17 (0.06)	7% (3)
Self-interest:			
Conflict-avoiding	$\theta\Delta_S\tilde{u}_R$	-\$0.38 (0.09)	16% (4)
Privacy & social costs	w	-\$0.85 (0.17)	36% (5)
Effort costs + cold glow	$-c+v$	\$0.28 (0.19)	12% (7)
Total		-\$0.42 (0.18)	

This table shows the estimates of the dollar equivalent of each motive in the average willingness to send a Direct Pressure Message. Parameter estimates are taken from from Column 1 of Table 3 (Equation 4). Standard errors are clustered by sender and calculated using the Delta method.

Equation 6 now includes separate pure altruism (α^k) and conflict-avoidance (θ^k) coefficients for each response on the 5-point Likert scale ($\mathbf{1}\{\Delta_S\tilde{u}_R = k\}$). Since one of the coefficients must be omitted, I assume that there are no altruistic concerns when the sender thinks the recipient will neither like nor dislike being pressured (i.e., $\alpha^0 = 0$). One downside of this more flexible specification is that w (the coefficient on the dummy for Direct Pressure Messages) cannot be separately identified from the θ^k coefficients.

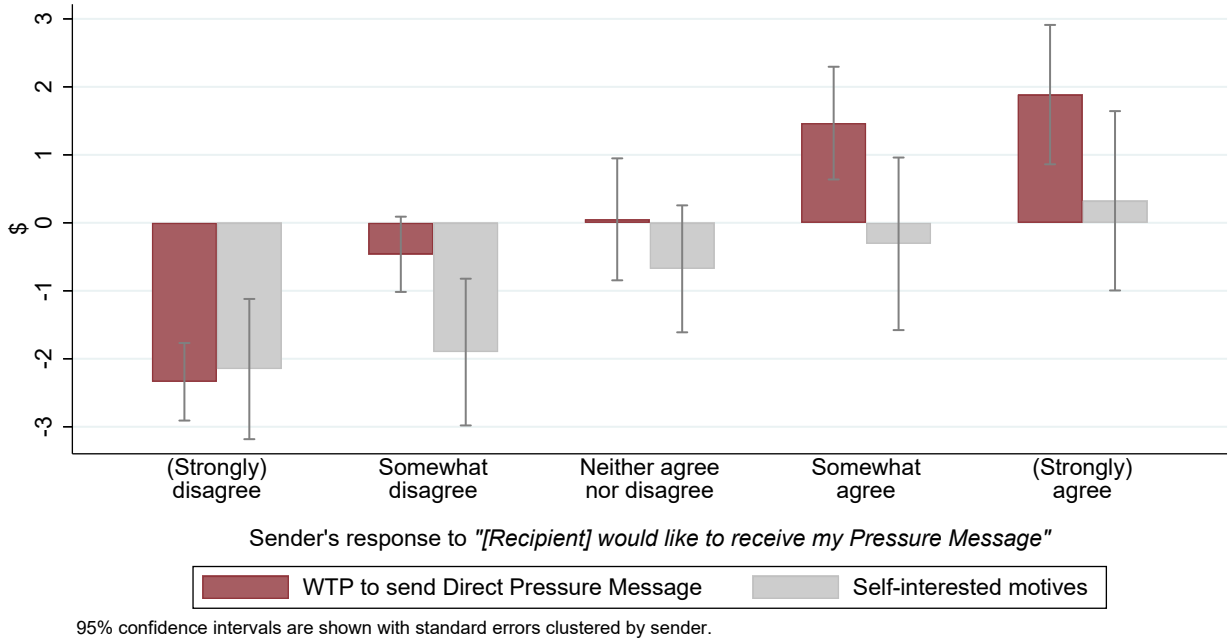
Figure 8 takes the results from this nonparametric specification and shows how the magnitude of the self-interested motives depends on how much the sender anticipates the recipient to (dis)like the message. The self-interested motives for each Likert response k is equal to $\theta^k - w + \nu$, which is roughly the difference in the average WTP within k between Direct and Anonymous Pressure Messages that is not explained by the predicted effectiveness of the message. In line with intuition, self-interested motives are strong and negative when senders expect the recipient to dislike being pressured, which likely entails greater risks of retaliation or social costs. However, when senders forecast that the recipient will either be neutral toward the message or even like it, their self-interested concerns disappear.

The nonparametric specification in Equation 6 also serves as a robustness check for the linearity assumption in Equation 5. Table A.4 repeats the decomposition in Table 5 but with parameter estimates from Equation 6. The results of the parametric and nonparametric decompositions are almost identical, endorsing the linear approximation in Equation 5.

4.6 Effectiveness of social pressure

Although the large majority (79%) of senders expect the Direct Pressure Message to have a positive effect on the recipient's likelihood of registering to vote (median=+10 pp.), the actual effect was in fact marginally negative. Figure 9 shows the intend-to-treat (ITT) effects

Figure 8: Nonparametric decomposition of self-interested motives



on electoral participation and compares the voter registration and turnout rates between the Non-registrant treatment and control groups. The treatment group, which was assigned to receive a Direct Pressure Message from a Registrant participant, had a voter registration rate of 77% and a turnout rate of 69% in the election. In the control group, which was assigned to not receive any messages, the rates were higher by 10 pp. for registration ($p < 0.1$) and by 6 pp. for turnout ($p = 0.5$).

To calculate the treatment-on-treated (TOT) effects, I can verify whether the intended recipients successfully received the Direct Message Pressure using a question on the Follow-up Survey. The Follow-up Survey asked Non-registrants whether they had received an email from another participant, and if so, what the email was about. There were 48 Non-registrants all in the treatment group who answered that they had received a message about registering to vote (see Table A.5 for a detailed breakdown). I use the subset of 130 Non-registrants (87% of the sample) who completed the Follow-up Survey in the TOT analysis.

Table 6 presents both the ITT and TOT effects on the Non-registrants' electoral participation. Columns 1 and 3 correspond to the findings in Figure 9. Columns 3 and 4 show the TOT effects, which are substantively the same as the ITT results. There is a marginally negative effect on the registration rate, but a null effect on the voter turnout rate. Given the small sample of Non-registrants, it is difficult to draw any robust conclusions on the effect of receiving social pressure. Furthermore, the direction of the negative effect runs counter to the studies to date on social pressure.

Figure 9: Effect of receiving a Direct Pressure Message on electoral participation rates

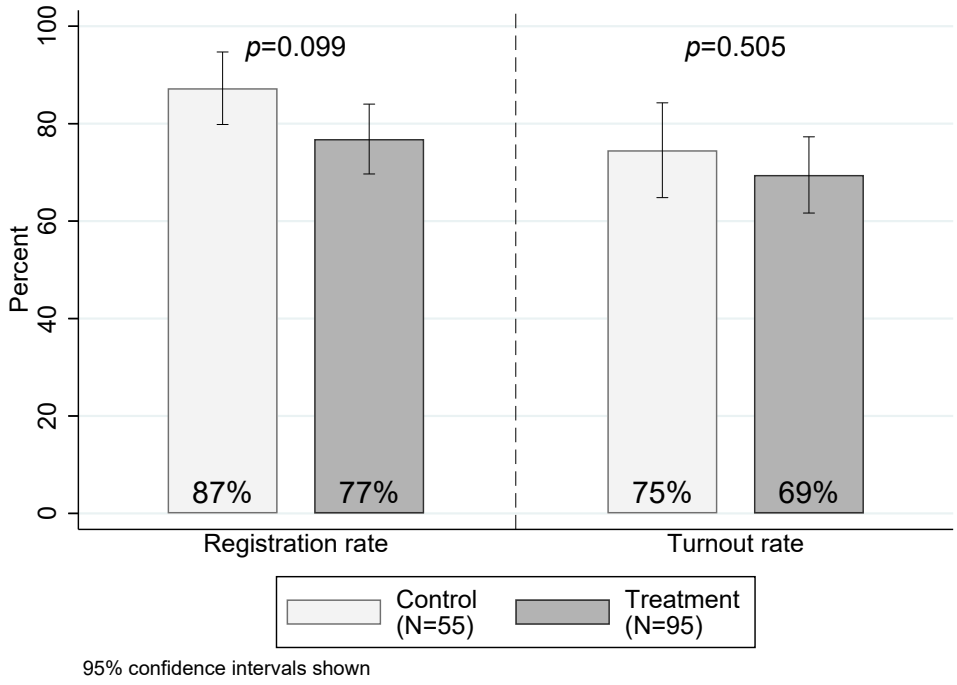


Table 6: Effect of receiving a Direct Pressure Message on rates of electoral participation

	Voter registration		Voter turnout	
	ITT	TOT	ITT	TOT
	(1)	(2)	(3)	(4)
Receive Direct Pressure Message	-0.10	-0.19	-0.05	-0.11
	(0.06)	(0.11)	(0.08)	(0.14)
No message	0.87	0.88	0.75	0.76
	(0.05)	(0.05)	(0.06)	(0.06)
Non-registrants	150	130	150	130
K-P F-stat		116.01		116.01

This table compares the voter registration and turnout rates between Non-registrant participants in the treatment group (assigned to receive a Direct Pressure Message) and in the control group (assigned to receive no message). Columns 1 and 3 show the intend-to-treat (ITT) results, and Columns 3 and 4 the treatment-on-treated (TOT) results. Robust standard errors are displayed in parentheses.

To discern whether the marginally negative result on voter registration is simply noise or a hint of a true effect, I conducted a Post-election Survey on the Non-registrants in March 2021, a few months after the main experiment and the election. The survey was offered to the 130 Non-registrants who completed the main study as a means to better understand the results from the main experiment and gathered 100 responses. The survey explained that

participants who had not registered before the study either did not receive any email from another participant (Group A), or received an email from another participant telling them to register to vote before the deadline (Group B). The survey also displayed a template of the Pressure Message (without mentioning "pressure") that the participants in Group B were assigned to receive from another participant. The survey disclosed that in Group A (who did not receive an email), 87% registered to vote and 75% voted in the election. The survey asked the Non-registrants to predict the corresponding rates for Group B. As in the main experiment, the survey incentivized the predictions by awarding a $1 bonus if either forecast was within 5 percentage points of the actual rates. The survey also asked them to briefly explain the reasoning behind their predictions.

Only 2 out of the 100 responses predict a negative effect, which suggests that the marginally negative effect on voter registration observed in Figure 9 and Table 6 is likely due to noise. The median expected effect is +3 pp. for both registration and turnout. Compared to the overoptimistic Registrants who forecast a median effect of +10 pp., the Non-registrants have a subdued perspective on the effectiveness of peer persuasion, perhaps since they are more familiar with the barriers to electoral participation. Eight of the respondents reflected on the experience of receiving the message themselves; 5 of them stated that the message had no bearing on their electoral participation, and the remaining 3 said that it had a positive effect on their own decision to vote. Interestingly, although "pressure" was never mentioned in the survey, 8 responses used the word to refer to the message.

The lack of a positive effect of the pressure message may be due to several reasons. The 2020 General Election was an extremely charged election. With expanded mail-in voting in the midst of the coronavirus pandemic, voter turnout was the highest since 1900.[15] Those who did not participate even in this election may be facing barriers too high for a peer "nudge" to overcome. In addition, before Non-registrants began the study, they did not know that their lack of compliance with the norm (i.e., their unregistered status) would be shared with others. Their actions were made public *before* they had time to adjust their status, whereas experiments such as the seminal Gerber, Green, and Larimer (2008) study revealed subjects' voting records to their neighbors *after* they had been given prior notice. The effect in this experiment may have been similar to these previous studies if Non-registrants who received a Pressure Message were told that their sender would be updated after the election on whether they registered and voted. This anticipation of punishment, as Fudenberg and Pathak (2010) find in a public-goods game, is a crucial component in the peer enforcement of norms. Lastly, email as a medium for persuasion tends to be less effective than others such as in-person communication (Bronchetti, Huffman, and Magenheim, 2015; DellaVigna

[15]Source: United States Election Project.

and Linos, 2021).

5 Application

This section discusses a simple application of the results to the literature on the community enforcement of norms. Section 5.1 presents a highly stylized model to make a theoretically straightforward point on the importance and realistic appeal of decentralized social pressure.

5.1 "Costly" punishment in the community enforcement of norms

Since Kandori (1992), Ellison (1994), and Harrington (1995), the repeated Prisoner's Dilemma game under random anonymous matching has been a classic setting to study strategies of community enforcement to sustain cooperative social norms. This stylized setting represents social interactions between strangers typically in large populations, when players cannot communicate or identify past norm violators.

In the model, a large, even number of N players are matched into pairs every round. Each player's partner is selected at random from the $N-1$ other players with uniform probability. All pairs of players do not know the identities and histories of their partner, and can only observe the outcome of their game in each round. The players have a discount factor δ and engage in an infinite number of rounds of the stage game shown in Table 7 where $l, g > 0$.

Table 7: Prisoner's Dilemma stage game

	Cooperate	Defect
Cooperate	$(1, 1)$	$(-l, 1+g)$
Defect	$(1+g, -l)$	$(0, 0)$

Despite these stringent conditions, Kandori (1992) and Harrington (1995) show that community enforcement can sustain cooperation in a Nash equilibrium through a grim trigger strategy. In this strategy, players begin by cooperating, but switch to defecting when they encounter a defector. The threat of community punishment—the contagious breakdown of cooperation and the consequent loss in future payoffs—deters defection. Ellison (1994) refines this strategy to form a sequential equilibrium in two ways: by introducing a public randomization device that resets to cooperation periodically, or by spreading retaliatory defections across rounds.

While these papers establish seminal theoretical contributions, even the authors themselves doubt the realism of such grim trigger strategies. Kandori (1992) writes that "we do

not observe such a norm very often." In addition, Ellison (1994) notes that if there are "crazy" players who always defect, then the contagious strategy does not sustain cooperation; he considers the crucial assumption of common rationality "fairly implausible" in large populations.

Sugaya and Wolitzky (2020) generalize Ellison's observation, and prove that for *any* strategy, cooperation is impossible in a population that contains a proportion of irrational "bad apples" who never cooperate. The rough intuition is that the payoff of the rational players cannot be much lower than that of the irrational players against the distribution of actions in equilibrium. As defecting is the dominant strategy in the stage game, it follows that the rational players too must almost always be defecting. This small yet reasonable addition—that there are "anti-social" types who refuse to cooperate with the norm—poses a problem for realistic models of community enforcement.

Yet in real life, we also often have an opportunity to punish those who violate norms. This type of decentralized peer enforcement has long been a feature in theories of social norms. Early influential papers such as Axelrod (1986), Hirschleifer and Rasmusen (1989), Boyd and Richerson (1992), Sethi and Somanathan (1996), and Henrich and Boyd (2001) represent this social interaction by adding a second stage in each round of the model when cooperators can sanction defectors. More recently, Calvó-Armengol and Jackson (2010) build a model of peer pressure in which players can alter each other's payoffs in the stage game at a cost to themselves, Levine and Modica (2016) derive conditions for cooperation in a setting where peers in a network randomly audit each other's contributions, and Acemoglu and Wolitzky (2020) develop a model where specialized enforcers can punish defectors after each round of cooperative play.

Drawing from these aforementioned papers, I enrich the classic setup discussed thus far with two realistic extensions. First, as in Sugaya and Wolitzky (2020), I assume that a proportion of $\varepsilon \in (0,1)$ in the population are irrational types who always defect. Next, I add a second stage to the model to allow for peer punishment. If a player cooperates but her partner defects in the first stage, then at some net cost or benefit x to herself, she can punish her partner and lower his payoff by $p > g$ in the second stage. A defector who is punished has his payoff reduced to $1 + g - p$, and a cooperator who punishes a defector has a payoff of $-l + x$ for the round, as summarized in Table 8.

Table 8: Second stage choice for a cooperator encountering a defector

	Cooperator	Defector
Don't punish	$-l$	$1+g$
Punish	$-l+x$	$1+g-p$

The outcome of this augmented model hinges on x, the net cost or benefit of punishing. If punishing is costless ($x = 0$) as in Hirschleifer and Rasmusen (1989), then defection will be prevented by the threat of punishment in the second stage. Most other models, however, assume that punishment is costly ($x < 0$). Intuitively, expressing disapproval seems like a negative experience; it also consumes time and energy, and introduces the potential of retaliation or conflict (Elster, 1989; Falk, Fehr, and Fischbacher, 2002).

Under the common assumption that punishing someone is costly, cooperation among the $1-\varepsilon$ rational types cannot be sustained, since the threat of being punished for defecting is not credible. When facing a defector, each rational player has an incentive to shirk the responsibility to enforce and to avoid the cost x—punishing itself is a public good. Consequently, peer discipline fails among the rational types, and cooperation disintegrates across the community. To circumvent this issue, some models have added "metanorms" (Axelrod, 1986) under which the failure to punish defectors is itself punished (as in Levine and Modica, 2016), or haven taken an evolutionary game theoretic approach (Boyd and Richerson, 1992; Henrich and Boyd, 2001).

Yet the common assumption that $x < 0$ seems at odds with numerous laboratory experiments that find a substantial fraction of people choose to lower their own payoff to punish a defector, even when there is no instrumental value—they behave as if the act of punishing yields a net *positive* payoff. In these studies, people are willing to pay to punish non-cooperative behavior (Yamagishi, 1986; Ostrom, Walker, and Gardner, 1992; Fehr and Gachter, 2000), even in one-shot anonymous environments with new, random partners every round (Fehr and Gachter, 2002), and even when the punishments are only revealed at the end of all the rounds (Fudenberg and Pathak, 2010). Informed by these findings, Falk, Fehr, and Fischbacher (2005) conclude that hedonic, retaliatory motives drive the peer enforcement of norms, and Ouss and Peysakhovich (2015) consider such sanctions not a public good, but rather a private good that provides "cold glow" utility.

In addition, the option to punish defectors prevents the breakdown of cooperation in these studies. Interestingly in Fudenberg and Pathak (2010), punishments are still effective even when they are administered only after the entire game. The fear of punishment, not necessarily the act, is enough to sustain cooperation. That is, participants seem to accurately predict that there are those in the population who, for their own satisfaction, are willing to punish defectors.

Based on this experimental evidence, what if x varies across players such that a proportion of rational players are "confrontational" with $x > 0$ and gain utility from punishing a defector? Then cooperation can be a Nash equilibrium among the $1-\varepsilon$ rational types, if the expected threat of being confronted for defecting outweighs the expected gain from defect-

ing. Denote F_x as the commonly known cdf of x and suppose $1 - F_x(0) \geq 0$. The following inequality is the incentive compatibility condition for the rational non-confrontational types to cooperate (which is also sufficient for the rational confrontational types):

$$(1 - \varepsilon)p(1 - F_x(0)) \geq (1 - \varepsilon)g + \varepsilon l$$

This stylized model of peer norm enforcement underscores the importance of understanding the motives for punishing someone who fails to comply with a norm. Depending on the distribution of x, cooperation endures or unravels. Although this result is theoretically simple, it provides a method to sustain the cooperative norm among most of the community even when there are types who refuse to comply. Furthermore, this mechanism of community enforcement has perhaps more realistic appeal than contagious grim trigger strategies or convoluted communication structures.

In real life, *how* do we enforce the norm? A common way we socially pressure others is through verbal disapproval and confrontation (Kandel and Lazear, 1992; Falk, Fehr, and Fischbacher, 2002; Bowles and Gintis, 2002). Some lab studies have introduced this natural form of norm enforcement and allowed subjects to communicate "verbal sanctions" face-to-face or in a chat room (Ostrom, Walker, and Gardner, 1992; Sally, 1995) or to assign nonmonetary disapproval points to one another instead (Masclet et al., 2003). They find that the usage and effectiveness of these pressuring communications are comparable to those of the monetary punishments.

In the experiment in this paper, I find that 23% of the senders exhibit a strictly positive willingness to pressure when communicating directly, and an even higher percentage of 32% when communicating anonymously. Roughly a quarter are "confrontational" types who seem to gain utility from asserting the norm to a noncomplier. Although the experiment was not explicitly designed to measure the distribution of x (and is certainly not a Prisoner's Dilemma game), it still illustrates that people who are willing to confront do exist, and more so when they can be anonymous and immune to the risk of retaliation. Theoretically, this outspoken minority who have preferences for confronting could sustain norms that would otherwise unravel.

This paper also helps to bridge one gap in the literature between the high rate of peer punishments in the lab (e.g., 84% in Fehr and Gachter, 2002) but the relatively low rate of confrontation in the field. In a couple interesting field studies intended to mimic the lab setting, actors violated a social norm next to a single bystander in a public place, such as littering in a train station (Balafoutas and Nikiforakis, 2012; Balafoutas, Nikiforakis, and Rockenbach, 2014). These field experiments find that less than 20% of bystanders are willing

to verbally confront the norm violator, which the authors use to measure the rate of "norm enforcement" or "direct punishment". Their survey evidence suggests that the bystanders are unwilling to do so from fear of retribution. The key feature of these field studies, however, is that the interactions are direct; in the lab, communication tends to be anonymous. As the experiment in this paper demonstrates, anonymous communication negates self-interested costs such as the risk of retaliation, and people are more willing to confront others.

6 Conclusion

In this experiment, Californian university students encounter their peers who have yet to register to vote for an imminent election, and must decide whether they want to socially pressure those peers by sending an email telling them to register. The context of this experiment is intended to reflect similar situations that we face in real life. Do we confront a shirking colleague, tell a neighbor to recycle, or ask a stranger in the grocery store to wear a mask during the COVID-19 pandemic? Why or why not?

To investigate the motives behind such decisions, the experiment elicits the participants' preferences and beliefs about socially pressuring their peers to register. This paper finds that self-interest dominates, comprising almost two-thirds of all the motives. Participants are generally unwilling to risk retaliation and to initiate an awkward, unsolicited conversation. Although this may seem like a negative finding for community norm enforcement, persuading their peers to follow the norm is worth roughly half as much as the senders' own self-interest, even in California where the election outcome was practically predetermined.

The sample of young Americans in this study is also relevant for policies in civic engagement and social pressure. Young adults are a pivotal group in U.S. elections, but have historically registered and voted in lower rates than the general population. Although they tend to be more sensitive to peer pressure (Pasupathi, 1999; Gardner and Steinberg, 2005), the results in this study suggest that one-time peer messaging without accountability is ineffective for mobilizing young voters. This age group is also enmeshed with the "call-out culture" that usually involves broadcasting bad behavior on social media, instigating shaming from a wide online audience.[16] This paper shows that the options to post anonymously or to indirectly criticize the perpetrator who has no real means of retaliating are important catalysts for the proliferation of this culture.

Despite the flourishing work on social pressure in the economics literature, there

[16] "The destructiveness of call-out culture on campus" (*The Atlantic*, 2017); "Tales from the teenage cancel cultures" (*The New York Times*, 2019); "Social medias call-out culture continues to improve customer service" (*Forbes*, 2020)

still remain numerous complexities to be explored. This paper is just one extension that focuses on direct confrontation between peers. In real life, nonverbal communication such as disapproving glances can be a powerful means of enforcement that is not observed in this experiment. While past literature has made significant progress comparing public versus private behavior, future studies can peel down the finer layers and disclose the rich ways through which we communicate and sustain norms.

References

Acemoglu, D. and A. Wolitzky. 2020. "Sustaining cooperation: Community enforcement versus specialized enforcement." *Journal of the European Economic Association* 18 (2): 1078-1122.

Ali, S. and C. Lin. 2013. "Why people vote: Ethical motives and social incentives." *American Economic Journal: Microeconomics* 5 (2): 73-98.

Allcott, H. 2011. "Social norms and energy conservation." *Journal of Public Economics* 95: 1082-1095.

Andreoni, J. 1989. "Giving with impure altruism. Applications to charity and Ricardian equivalence." *Journal of Political Economy* 97 (6): 1447-1458.

Andreoni, J., J. Rao, and H. Trachtman. 2017. "Avoiding the ask: A field experiment on altruism, empathy, and charitable giving." *Journal of Political Economy* 125 (3): 625-653.

Ashraf, N., O. Bandiera, & B. Jack. 2014. "No margin, no mission? A field experiment on incentives for public service delivery." *Journal of Public Economics* 120: 1-17.

Axelrod, R. 1986. "An evolutionary approach to norms." *American Political Science Review* 80 (4): 1095-1111.

Balafoutas, L. and N. Nikiforakis. 2012. "Norm enforcement in the city: A natural field experiment." *European Economic Review* 56: 1773-1785.

Balafoutas, L., N. Nikiforakis, and B. Rockenbach. 2014. "Direct and indirect punishment among strangers in the field." *Proceedings of the National Academy of Sciences* 111 (45): 15924-15927.

Bandiera, O., I. Barankay, and I. Rasul. 2005. "Social preferences and response to incentives: Evidence from personnel data." *The Quarterly Journal of Economics* 120 (3): 917-962.

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

Bhaskar, V. and C. Thomas. Forthcoming. "Community enforcement of trust with bounded memory." *Review of Economic Studies.*

Bowles, S. and H. Gintis. 2002. "Social capital and community governance." *Economic Journal* 112: 419-436.

Boyd, R. and P. Richerson. 1992. "Punishment allows the evolution of cooperation (or anything else) in sizable groups." *Ethology and Sociobiology* 13: 171-195.

Bradler, C., R. Dur, S. Neckermann, and A. Non. 2016. "Employee recognition and performance: A field experiment." *Management Science* 62 (1): 3085-3099.

Bronchetti, E., D. Huffman, and E. Magenheim. 2015. "Attention, intentions, and follow-through in preventive health behavior: Field experimental evidence on flu vaccination." *Journal of Economic Behavior and Organization* 116: 270-291.

Bursztyn, L., T. Fujiwara, and A. Pallais. 2017. "'Acting wife': Marriage market incentives and labor market investments." *American Economic Review* 107 (11): 3288-3319.

Bursztyn, L., A. González, and D. . 2020. "Misperceived social norms: Women working outside the home in Saudi Arabia." *American Economic Review*, 110 (10): 2997-3029.

Burszytn, L. and R. Jensen. 2015. "How does peer pressure affect educational investments?" *The Quarterly Journal of Economics* 130 (3): 1329-1367.

Butera, L., R. Metcalfe, W. Morrison, and D. Taubinsky. Forthcoming. "Measuring the welfare effects of shame and pride." *American Economic Review.*

Calvó-Armengol, A. and M. Jackson. 2010. "Peer pressure." *Journal of the European Economic Association* 8 (1): 62-89.

Danz, D., L. Vesterlund, and A. Wilson. 2020. "Belief Elicitation: Limiting Truth Telling with Information on Incentives." Working paper.

Deb, J., T. Sugaya, and A. Wolitzky. 2020. "The Folk Theorem in repeated games with anonymous random matching." *Econometrica* 88 (3): 917-964.

DellaVigna, S. and M. Gentzkow. 2010. "Persuasion: Empirical evidence." *Annual Review of Economics* 2: 643-669.

DellaVigna, S. and E. Linos. Forthcoming. "RCTs to scale: Comprehensive evidence from two nudge units." *Econometrica.*

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

DellaVigna, S., J. List, U. Malmendier, and G. Rao. 2017. "Voting to tell others." *Review of Economic Studies* 84: 143-181.

DellaVigna, S., J. List, U. Malmendier, and G. Rao. 2021. "Estimating social preferences and gift exchange at work." Working paper.

Denant-Boemont, L., D. Masclet, and C. Noussair. 2007. "Punishment, counterpunishment and sanction enforcement in a social dilemma experiment." *Economic Theory* 33: 145-167.

Ellison, G. 1994. "Cooperation in the Prisoner's Dilemma with anonymous random matching." *Review of Economic Studies* 61: 567-588.

Elster, J. 1989. "Social norms and economic theory." *The Journal of Economic Perspectives* 3 (4): 99-117.

Falk, A., E. Fehr, and U. Fischbacher. 2002. "Appropriating the commons: A theoretical explanation." In Ostrom, E. et al. (Eds), *The Drama of the Commons* (pp. 157-191). The National Academies Press: Washington D.C.

Falk, A., E. Fehr, and U. Fischbacher. 2005. "Driving forces behind informal sanctions." *Econometrica* 73 (6): 2017-2030.

Fehr, E. and S. Gachter. 2000. "Cooperation and punishment in public goods experiments." *American Economic Review* 90 (4): 982-993.

Fehr, E. and S. Gachter. 2002. "Altruistic punishment in humans." *Nature* 415: 137-140.

Fudenberg, D. and P. Pathak. 2010. "Unobserved punishment supports cooperation." *Journal of Public Economics* 94: 78-86.

Funk, P. 2010. "Social incentives and voter turnout: Evidence from the Swiss mail ballot system." *Journal of the European Economic Association* 8 (5): 1077-1103.

Gardner, M. and L. Steinberg. 2005. "Peer influence on risk taking, risk preference, and risky decision making in adolescence and adulthood: An experimental study." *Developmental Psychology* 41 (4): 625-635.

Gerber, A., D. Green, and C. Larimer. 2008. "Social pressure and voter turnout: Evidence from a large-scale field experiment." *American Political Science Review* 102 (1): 33-48.

Harrington, J. 1995. "Cooperation in a one-shot Prisoners' Dilemma." *Games and Economic Behavior* 8 (2): 364-377.

Henrich, J. and R. Boyd. 2001. "Why people punish defectors: Weak conformist transmission can stabilize costly enforcement of norms in cooperative dilemmas." *Journal of Theoretical Biology* 208: 79-89.

Hirshleifer, D. and E. Rasmusen. 1989. "Cooperation in a repeated Prisoners' Dilemma with ostracism." *Journal of Economic Behavior and Organization* 12: 87-106.

Hossain, T. and R. Okui. 2013. "The binarized scoring rule." *Review of Economic Studies* 80 (3): 984-1001.

Allcott, H. 2011. "Social norms and energy conservation." *Journal of Public Economics* 95: 1082-1095.

Johnson, M. 2020. "Regulation by shaming: Deterrence effects of publicizing violations of workplace safety and health laws." *American Economic Review* 110 (6): 1866-1904.

Kandel, E. and E. Lazear. 1992. "Peer pressure and partnerships." *Journal of Political Economy* 100 (4): 801-817.

Kandori, M. 1992. "Social norms and community enforcement." *Review of Economic Studies* 59 (1): 63-80.

Karing, A. 2021. "Social signaling and childhood immunization: A field experiment in Sierra Leone." Working paper.

Lagomarsino, B. et al. 2017. "Peer pressure: Experimental evidence from restroom behavior." *Economic Inquiry* 55 (3): 1579-1584.

Levine, D. and A. Mattozzi. 2020. "Voter turnout with peer punishment." *American Economic Review* 110 (10): 3298-3314.

Levine, D. and S. Modica. 2016. "Peer discipline and incentives within groups." *Journal of Economic Behavior and Organization* 123: 19-30.

Mas, A. and E. Moretti. 2009. "Peers at work." *American Economic Review* 99 (1): 112-145.

Masclet. D., C. Noussair, S. Tucker, and M.-C. Villeval. 2003. "Monetary and nonmonetary punishment in voluntary contributions mechanism." *American Economic Review* 93 (1): 367-380.

Ostrom, E., J. Walker, and R. Gardner. 1992. "Covenants with and without a sword: Self-governance is possible." *American Political Science Review* 86 (2): 404-417.

Ouss, A. and A. Peysakhovich. 2015. "When punishment doesn't pay: Cold glow and decisions to punish." *Journal of Law and Economics* 58: 625-654.

Pasupathi, M. 1999. "Age differences in response to conformity pressure for emotional and nonemotional material." *Psychology and Aging* 14 (1): 170-174.

Perez-Truglia, R. and U. Troiano. 2017. "Shaming tax delinquents: Theory and evidence from a field experiment in the United States." *Journal of Public Economics*, 167: 120-137.

Sally, D. 1995. "Conversation and cooperation in social dilemmas." *Rationality and Society* 7 (1): 58-92.

Sethi, R. and E. Somanathan. 1996. "The evolution of social norms in common property resource use." *American Economic Review* 86 (4): 766-788.

Sugaya, T. and A. Wolitzky. 2020. "A few bad apples spoil the barrel: An Anti-folk Theorem for anonymous repeated games with incomplete information." *American Economic Review* 110 (12): 3817-3835.

Yamagishi, T. 1986. "The provision of a sanctioning system as a public good." *Journal of Personality and Social Psychology* 51 (1): 110-116.

Yoeli, E., M. Homan, D. Rand, and M. Nowak. 2013. "Powering up with indirect reciprocity in a large-scale field experiment." *Proceedings of the National Academy of Sciences* 110: 10424-10429.

Zizzo, D. 2010. "Experimenter demand effects in economic experiments." *Experimental Economics* 13: 75-98.

A Additional survey details

A.1 Initial Survey: Non-registrants' WTP to share unregistered status

Before the end of the Initial Survey, Non-registrants took an additional section where they provided their willingness to pay to (not) share their registration status with other participants. Their WTP responses capture the loss of privacy and the welfare cost of revealing to others that they have not registered to vote, which could be damaging to their social reputation. They are informed that if their status is shared, they may also receive an email from another participant about the election. To reduce experimenter demand effects, the topic of the email is merely described as being "about the election" without specifying voter registration. The WTP is elicited through the same binary search procedure as for the Registrants in Section 3.3.1. The Non-registrants, however, made only one WTP choice for sharing their status, as opposed to the three WTP choices made by Registrants who were each assigned three potential recipients.

Figure A.8 shows the distribution of the WTP responses from the Non-registrants who passed the attention check preceding the WTP elicitation. The average WTP is -$1.3 (s.e.=0.3), and 36% of the Non-registrants strictly prefer not to share their information. Over half (53%) are indifferent, and 11% state a positive WTP. After a response of $0, the most common WTP is the negative extreme of $7, which holds 16% of the responses.

A.2 Verification, randomization, and matching

Verification. Before the Follow-up Survey, I verified all the Registrants against the California state voter registration records using a combination of their name, date of birth, and email address. For those who I could not find in the database, I sent them an email asking them to forward a screenshot verifying their voter registration online within 24 hours. The 24-hour rule disqualified Registrants who registered after starting the study, as voter registration applications took 48 hours to process. Only verified Registrants were allowed to continue the study.

For Non-registrants, there was no clear method to verify their self-reported status. Those who were missing from the voter database were clearly valid cases, but others could be registered under an old address and still technically be Non-registrants. There was no apparent reason for dishonesty, but it is likely that some Non-registrants misunderstood the question or were unaware that they were already registered. For instance, California's Motor Voter program automatically registers those who are eligible and in the DMV system.

While this is not an issue for the main results of giving social pressure from the Registrant side, it may confound the effects of receiving it on the Non-registrant side. Dropping the "pre-registered" Non-registrants who were in the voter file before the study yet voted without updating their address, Figure A.9 reproduces Figure 9 and shows the effects of the pressuring email on this subsample. The rates in the treatment group becomes even lower relative to the control group. Nevertheless, the evidence from the Post-election Survey described in Section 4.6 still casts doubt on whether these differences reflect true effects or merely noise.

Randomization and matching. The randomization and matching process in this study was involved, since there were separate treatments that needed to be randomized on both Non-registrant and Registrant sides, and the treatment for the Non-registrants required interaction from the Registrants. The process was conducted in the following steps after the Initial Survey.

For 10% of the Non-registrants, their WTP choices on the Initial Survey for sharing their name and voter registration status were exercised (see Section A.1); one of the rows on the multiple price list was randomly selected, and their preferred option on that row was implemented. These Non-registrants with endogenous sharing or withholding of their information with other participants are excluded from the analysis in Section 4.6.

The remaining 90% of the Non-registrants shared their full name and registration status with Registrants on the Follow-up Survey, regardless of their WTP choices. Roughly two-thirds were assigned to either receive a Pressure Message from a Registrant (treatment group), and the remaining one-third was assigned to not receive any message (control group) using stratified randomization with bins based on their WTP, gender, race/ethnicity, and age. More Non-registrants were assigned to the treatment group to counteract attrition among the Registrant senders.

The Non-registrants in the treatment group who were assigned to receive a Pressure Message were randomly matched to a Registrant selected with uniform probability to be either (a) same gender and race/ethnicity, (b) same gender and different race/ethnicity, (c) different gender and same race/ethnicity, or (d) different gender and race/ethnicity. These randomly selected Registrant senders were automatically assigned to the Pressure Message condition, saw their matched Non-registrant recipient under the Direct Message condition, and had to send the Direct Pressure Message to their matched recipient to complete the study, regardless of their WTP choices.

Non-registrants in the 10% who endogenously shared their information with others were uniformly assigned to either the Direct Info, Anonymous Info, Direct Pressure, or Anonymous Pressure condition, and were matched with any remaining Registrant who had

not been assigned to send a message. For these matched Registrant senders, their WTP choices were binding on the Follow-up Survey. Therefore, for Registrants in any experimental arm, there was a small chance that their WTP choices would be exercised.

The remaining Registrants who had not been selected as senders were randomly assigned to the Info Message arm (roughly 40%) or to the Pressure Message arm (roughly 60%). The proportion of the two treatments was chosen to balance the number of times each Non-registrant's information was shared with other participants, and to favor more power in the Pressure Message treatment.

All Registrants randomly drew three potential recipients (two more if they had already been assigned a recipient) to be shown on the Follow-up Survey. The drawing procedure ensured that each Non-registrant would be shown under both the Direct and Anonymous Pressure Message conditions to two different Registrants, and at least once under the Info Message condition (which could be Direct or Anonymous). All Registrants (regardless of whether they were assigned to the Pressure or Info Message condition) saw at least one recipient under the Direct Message condition, and at least one recipient under the Anonymous Message condition. Registrants assigned to the Pressure Message condition only saw Non-registrants as potential recipients, while Registrants assigned to the Info Message condition could see both Non-registrants and other Registrants as potential recipients.

Ultimately, the randomization ensured that a Non-registrant could receive only one message (if any), and a Registrant could send only one message (if any). For the 90% of Non-registrants who shared their information with no choice, the only message they could have received was a Direct Pressure Message. The rationale was to maximize the power (which was already low) to detect an effect of receiving a pressure message on voter registration rates.

Figure A.1: Within-sender assignment of Direct and Anonymous Messages

If you're selected to message a participant, it could be 1 of 2 types:

1. **Direct Message**: You'll have to email your message to the participant directly from your lionelmessi@ucla.edu account.

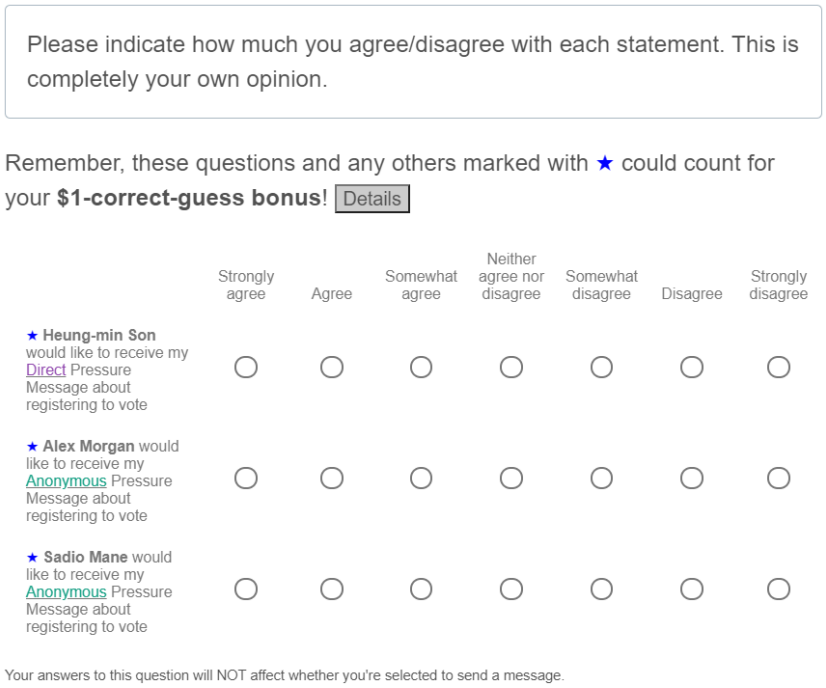
2. **Anonymous Message**: We'll email your message on your behalf, say it's "from another participant", and won't mention your name.

The type of message you'll have to send is different for each participant:

Participant	Message type
Heung-min Son	Direct
Alex Morgan	Anonymous
Sadio Mane	Anonymous

Figure A.2: Eliciting the sender's beliefs on the recipient's (dis)like of the message

(a) Pressure Message



(b) Info Message

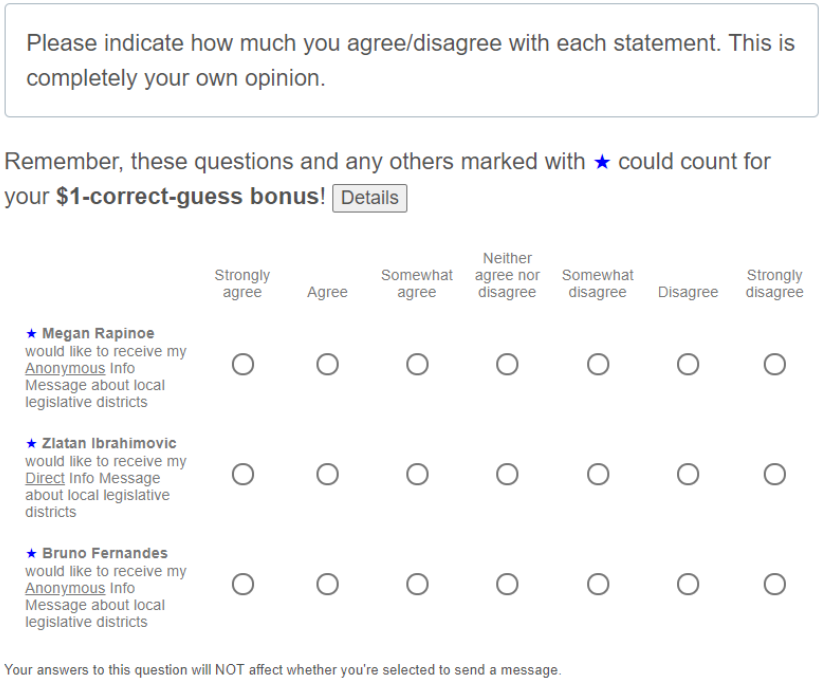


Figure A.3: Eliciting the sender's beliefs on the effectiveness of the message

Figure A.4: Eliciting the sender's willingness to pay to send the message

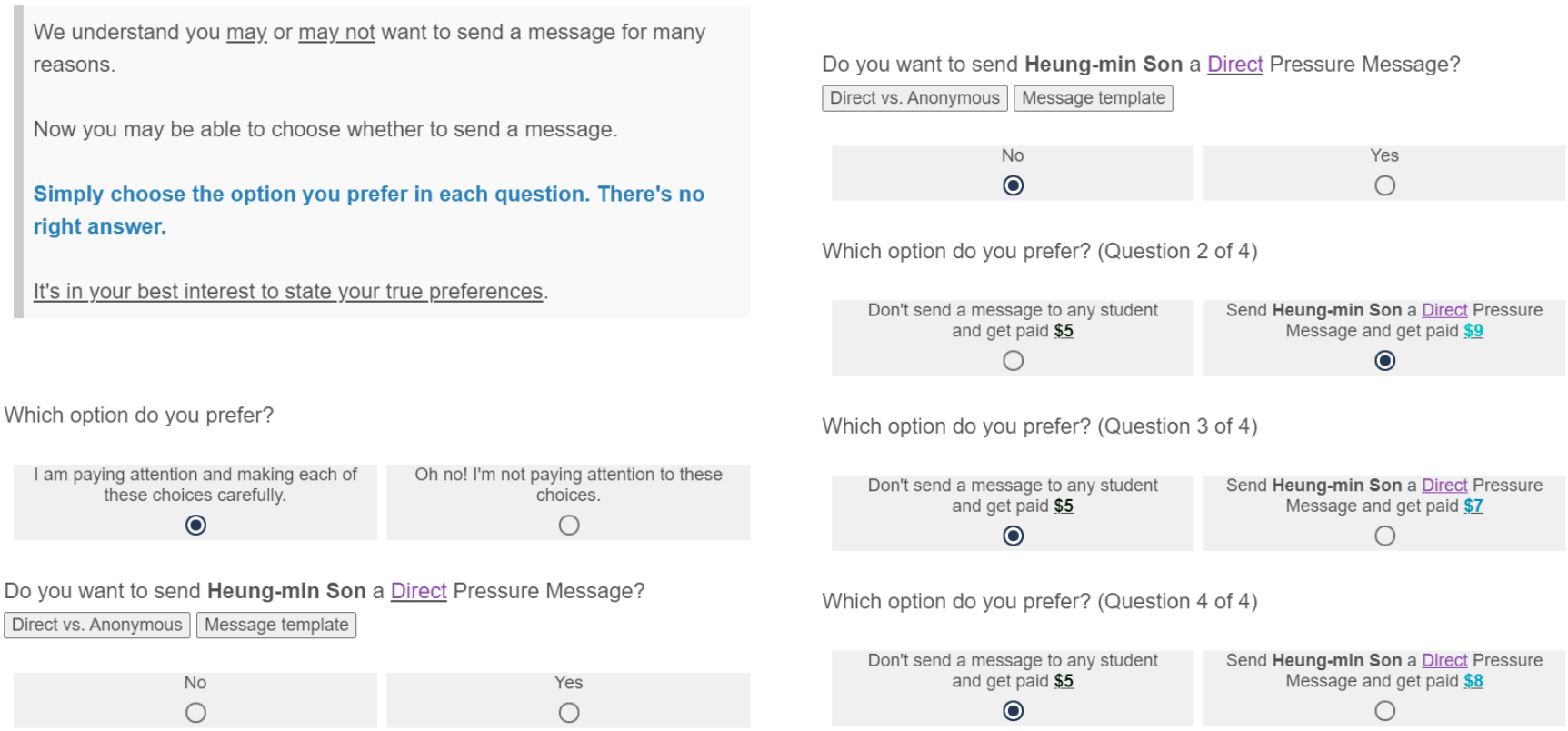


Figure A.4: Eliciting the sender's willingness to pay to send the message (continued)

We've filled out the 15 rows of choices below to be logically consistent with your previous responses.

Please confirm that your choice in each row reflects your true preferences! Each row is a separate choice.

If you would like to change your choices, please use the back ← button.

	You want LEFT	You want RIGHT	
Row 1) Don't send a message to any student and get paid $5	○	●	Send Heung-min Son a Direct Pressure Message and get paid $12
Row 2) Don't send a message to any student and get paid $5	○	●	Send Heung-min Son a Direct Pressure Message and get paid $11
Row 3) Don't send a message to any student and get paid $5	○	●	Send Heung-min Son a Direct Pressure Message and get paid $10
Row 4) Don't send a message to any student and get paid $5	○	●	Send Heung-min Son a Direct Pressure Message and get paid $9
Row 5) Don't send a message to any student and get paid $5	●	○	Send Heung-min Son a Direct Pressure Message and get paid $8
Row 6) Don't send a message to any student and get paid $5	●	○	Send Heung-min Son a Direct Pressure Message and get paid $7
Row 7) Don't send a message to any student and get paid $5	●	○	Send Heung-min Son a Direct Pressure Message and get paid $6
Row 8) Don't send a message to any student and get paid $5	●	○	Send Heung-min Son a Direct Pressure Message and get paid $5
Row 9) Don't send a message to any student and get paid $6	●	○	Send Heung-min Son a Direct Pressure Message and get paid $5
Row 10) Don't send a message to any student and get paid $7	●	○	Send Heung-min Son a Direct Pressure Message and get paid $5
Row 11) Don't send a message to any student and get paid $8	●	○	Send Heung-min Son a Direct Pressure Message and get paid $5
Row 12) Don't send a message to any student and get paid $9	●	○	Send Heung-min Son a Direct Pressure Message and get paid $5
Row 13) Don't send a message to any student and get paid $10	●	○	Send Heung-min Son a Direct Pressure Message and get paid $5
Row 14) Don't send a message to any student and get paid $11	●	○	Send Heung-min Son a Direct Pressure Message and get paid $5
Row 15) Don't send a message to any student and get paid $12	●	○	Send Heung-min Son a Direct Pressure Message and get paid $5

Next, we'll explain what will happen with these choices. You'll be able to come back to change your choices if you'd like.

After you answer the same questions for the other two students, you'll end up with 45 rows of choices in total (15 rows × 3 students). Then, either:

A) 1 of the 45 rows will be randomly selected, and you'll get whichever option you chose **in that 1 row**. You'll be compensated the amount you chose.*

- OR -

B) The computer will just randomly assign you to send a message or not. Your choices won't count in this case, so you'll be compensated the max $12.*

It's okay if you don't fully understand this procedure! This has just been set up so that it's best for you to answer the questions truthfully.

*Any bonuses from referrals and correct guesses will be added to this amount.

Figure A.5: Within-sender differences in WTP to send Direct vs. Anonymous Messages

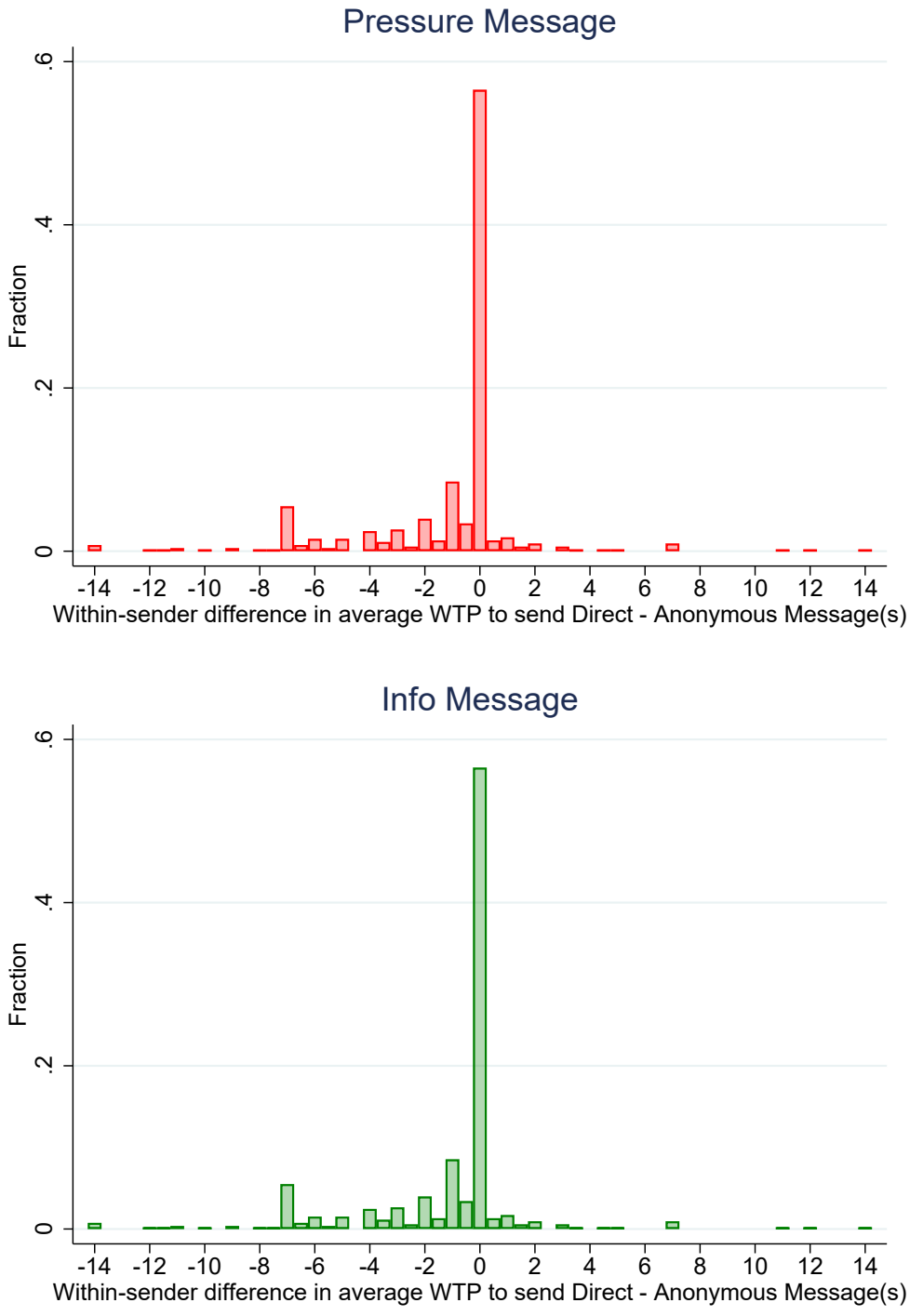


Figure A.6: WTP to pressure histograms by predicted recipient's (dis)like

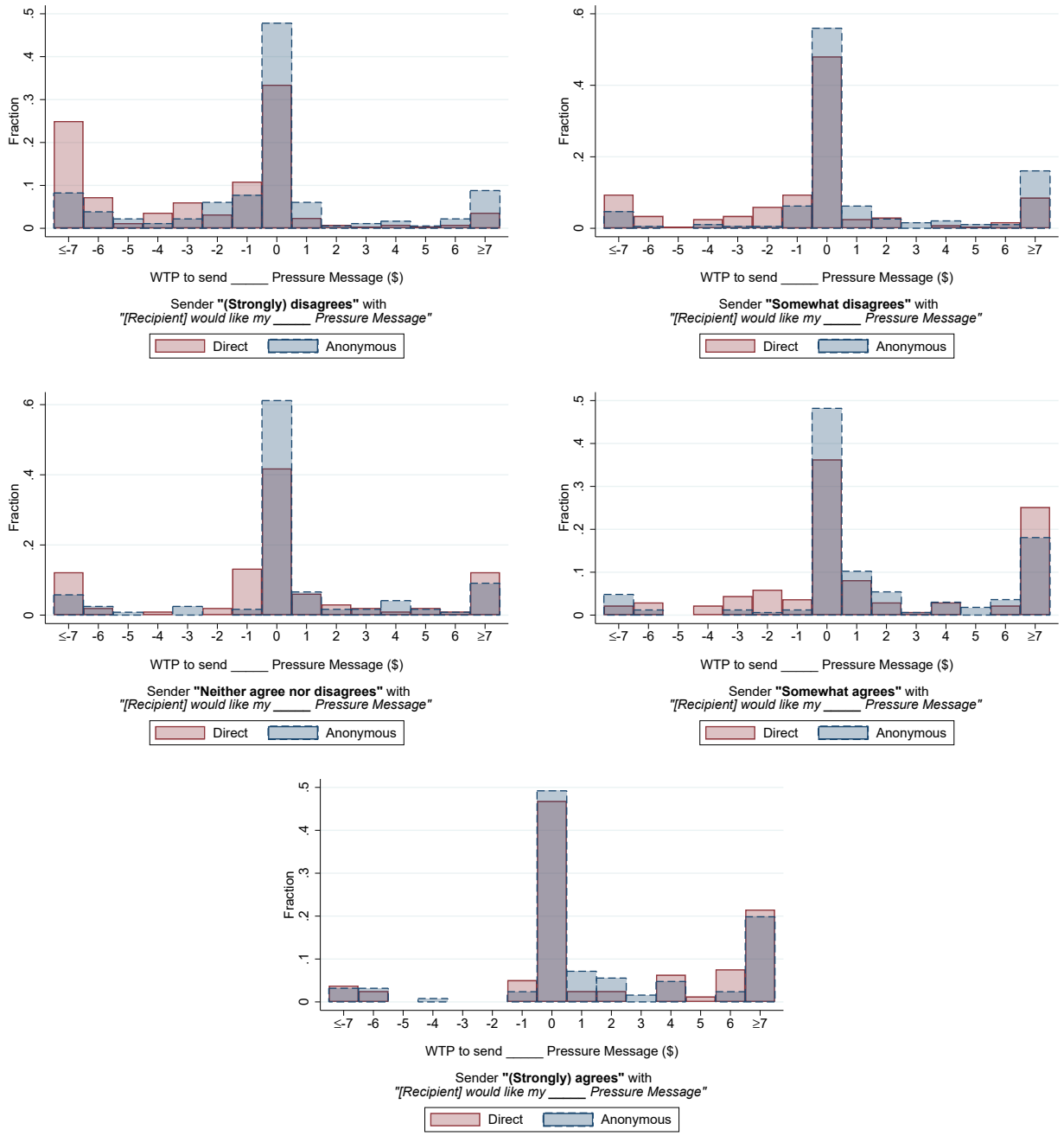
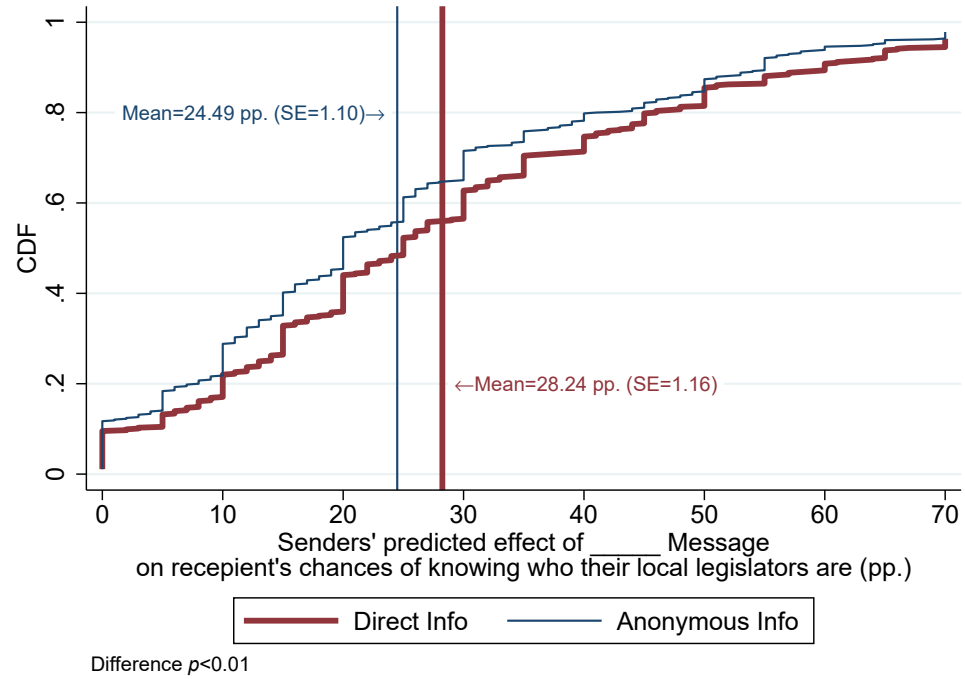
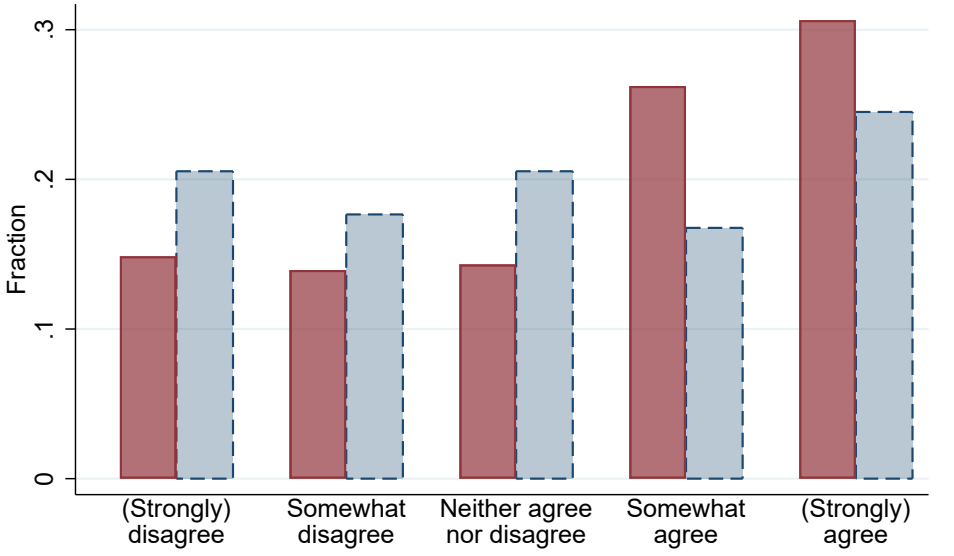


Figure A.7: Senders' predictions on the effects of the Info Message



Difference $p<0.01$



χ^2=28.79 (p=0.00)

Figure A.8: Non-registrants' willingness to share unregistered status

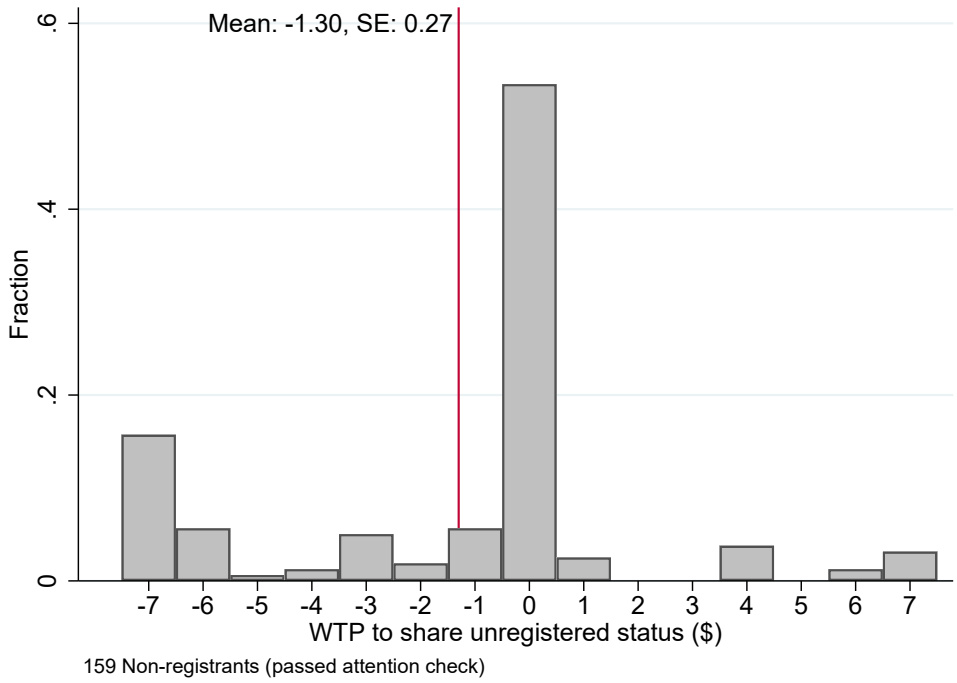


Figure A.9: Effect of receiving a Direct Pressure Message excluding "pre-registered" Non-registrants

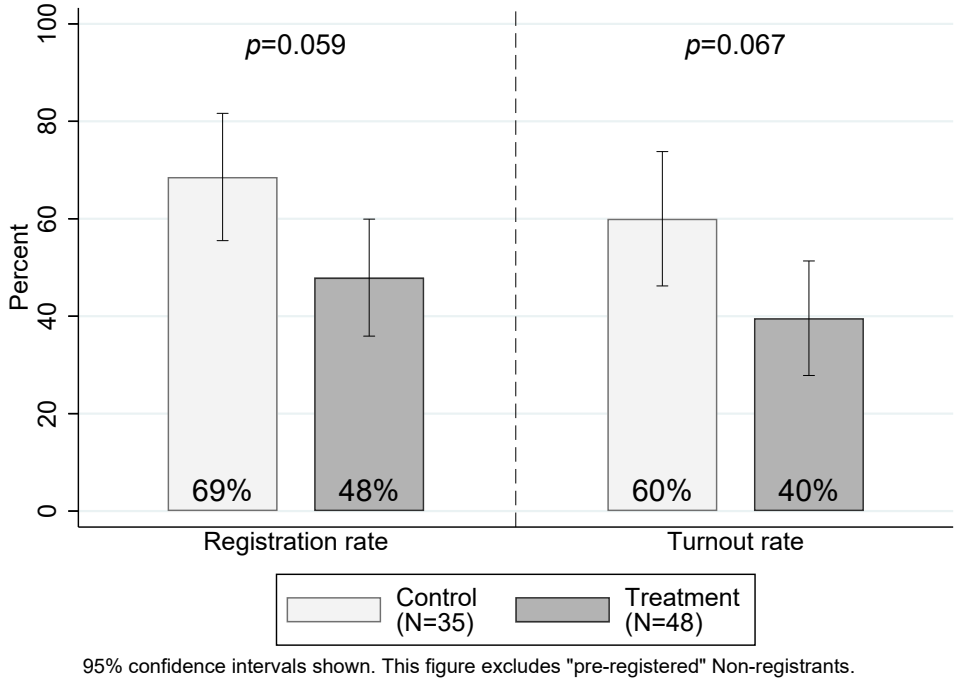


Table A.1a: Balance: Registrants

	Pressure Message	Info Message	Difference (SE)	Diff. p-value
Gender				
Male	0.31 [0.46]	0.31 [0.46]	0.00 (0.03)	0.94
Female/other	0.69 [0.46]	0.69 [0.46]	-0.00 (0.03)	0.94
Age (years)	21.62 [3.27]	21.46 [2.81]	0.16 (0.20)	0.44
Race/ethnicity				
White	0.29 [0.46]	0.32 [0.47]	-0.03 (0.03)	0.40
Asian	0.47 [0.50]	0.42 [0.49]	0.05 (0.03)	0.11
Other	0.24 [0.43]	0.26 [0.44]	-0.03 (0.03)	0.36
Degree				
Masters/doctoral	0.14 [0.35]	0.12 [0.33]	0.02 (0.02)	0.49
Undergraduate	0.86 [0.35]	0.88 [0.33]	-0.02 (0.02)	0.49
STEM major/field	0.71 [0.46]	0.68 [0.47]	0.02 (0.03)	0.43
Year degree started	2018.09 [1.23]	2018.08 [1.28]	0.01 (0.09)	0.88
Predictions				
Current reg. rate (%)	60.53 [16.81]	61.63 [17.03]	-1.10 (1.15)	0.34
Final reg. rate (%)	74.20 [14.65]	75.13 [14.33]	-0.93 (0.98)	0.35
Recipients characteristics				
Mixed gender comp.	0.67 [0.47]	0.66 [0.48]	0.02 (0.03)	0.60
Mixed racial comp.	0.85 [0.36]	0.87 [0.33]	-0.03 (0.02)	0.28
	527	367		0.80 (joint F-test)

This table compares the average characteristics of the Registrants assigned to the Pressure Message condition versus the Info Message condition. Standard deviations are shown in brackets, and robust standard errors in parentheses.

Table A.1b: Balance: Non-registrants

	Control (No Msg.)	Treatment (Direct Pressure Msg.)	Difference (SE)	Diff. p-value
Gender				
Male	0.39 [0.49]	0.40 [0.49]	-0.01 (0.08)	0.90
Female/other	0.61 [0.49]	0.60 [0.49]	0.01 (0.08)	0.90
Age (years)	21.06 [2.50]	20.67 [2.00]	0.40 (0.37)	0.29
Race/ethnicity				
White	0.34 [0.48]	0.33 [0.47]	0.01 (0.08)	0.91
Asian	0.38 [0.49]	0.47 [0.50]	-0.09 (0.08)	0.27
Other	0.28 [0.45]	0.20 [0.40]	0.08 (0.07)	0.24
Degree				
Masters/doctoral	0.06 [0.24]	0.05 [0.23]	0.01 (0.04)	0.83
Undergraduate	0.94 [0.24]	0.95 [0.23]	-0.01 (0.04)	0.83
STEM major/field	0.73 [0.45]	0.73 [0.45]	-0.00 (0.08)	0.99
Year degree started	2018.26 [1.19]	2018.42 [1.05]	-0.16 (0.19)	0.41
Preference for sharing status				
Indifferent	0.54 [0.50]	0.49 [0.50]	0.05 (0.09)	0.59
Prefer not to share	0.37 [0.48]	0.38 [0.49]	-0.01 (0.08)	0.87
Prefer to share	0.09 [0.29]	0.13 [0.34]	-0.03 (0.05)	0.55
WTP to share status ($)	-1.23 [3.22]	-1.38 [3.70]	0.15 (0.60)	0.80
Predictions				
Current reg. rate (%)	56.54 [15.75]	53.29 [18.20]	3.25 (2.93)	0.27
Final reg. rate (%)	70.58 [15.07]	67.67 [17.84]	2.91 (2.86)	0.31
	95	55		0.91 (joint F-test)

This table compares the average characteristics of the Non-registrants assigned to the treatment and control groups. Standard deviations are shown in brackets, and robust standard errors in parentheses. This table exlcudes 15 Non-registrants who received or did not receive a message endogenously depending on their or their sender's WTP choices.

Table A.1c: Balance: Non-registrant relation to assigned sender

	Same gender & race	Same gender, diff. race	Diff. gender, same race	Diff. gender & race	Referral
White male	5	3	2	1	2
White female	3	3	3	5	5
Asian male	2	3	4	2	2
Asian female	5	6	5	5	2
Other race male	2	3	3	2	1
Other race female	4	2	2	3	5

$\chi^2 = 11.09$ (p=0.94)

Table A.2: Sample attrition

	Non-registrants	Registrants	
		Info Msg.	Pressure Msg.
Finished Initial Survey	165	461	657
Finished Follow-up Survey	144	386	549
Sent required confirmation email	NA	382 out of 385	471 out of 474
Sent required Direct Message	NA	1 out of 1	70 out of 75[A]
Passed attention check	138	367	529
Completed study	**138**	**367**	**527**[B]

This table shows the number of participants remaining after each requirement of the study.
[A]As pre-registered, 1 participant who finished the Follow-up survey, stated a WTP to send the Direct Message of less than $12, and did not send the Direct Message is considered to have fulfilled the "Sent required Direct Message" condition.
[B]2 participants stated that they did not know how to fill out the WTP responses in the open-ended feedback box. Consequently, although they completed all the study requirements, their WTP responses are not included in the analysis.

Table A.3: Robustness: Sender motives

	Pressure	Info				
Dep. Var.: WTP to send message ($)	(1)	(2)	(3)	(1)-(2)	(1)-(3)	(2)-(3)
Sender's predictions on:						
Effectiveness of message (10 pp.)	0.50	0.07	-0.00	0.44	0.51	0.07
	(0.10)	(0.17)	(0.06)	(0.20)	(0.12)	(0.17)
Recipient's (dis)like of message [-2, 2]	0.31	0.21	0.27	0.10	0.04	-0.06
	(0.11)	(0.28)	(0.10)	(0.29)	(0.15)	(0.29)
Direct Message	-0.85	-2.10	-1.05	1.24	0.20	-1.04
	(0.17)	(0.63)	(0.17)	(0.64)	(0.24)	(0.66)
Direct×Predicted recipient's (dis)like	0.68	0.17	0.28	0.51	0.40	-0.11
	(0.14)	(0.32)	(0.15)	(0.34)	(0.20)	(0.35)
Constant	0.28	0.83	0.09	0.83	0.09	0.09
	(0.19)	(0.53)	(0.21)	(0.52)	(0.21)	(0.21)
Recipients						
Non-registrants only	✓	✓				
Registrants only			✓			
Responses	1581	97	1003	1678	2584	1100
Senders	527	80	366	607	893	367
R^2	0.14	0.10	0.07	0.14	0.13	0.07

Coefficient estimates are from linear regressions with standard errors clustered by sender shown below in parentheses. These regressions estimate the specification in Equation 4 and include two predictions made by the senders. First, senders predicted the effect of the message on the recipient's likelihood of registering to vote by the election (Pressure Message) or knowing the names of their local legislators (Info Message). The coefficient on this prediction corresponds to a 10 percentage point (pp.) increase in the predicted effectiveness. Second, senders indicated on a Likert scale how much they agreed with the statement that the recipient would like to receive their message. The coefficient on this prediction corresponds to a 1-point increase on the Likert scale ranging from -2 (strongly disagree) to 2 (strongly agree).

Table A.4: Decomposition (non-parametric specification)

Motive	Parameter	Estimate (SE)	Absolute weight (SE)
Persuasion	$\gamma\Delta_S x_R$	\$0.68 (0.14)	30% (10)
Altruism	$\sum_k \alpha^k \mathbf{1}\{\Delta_S \tilde{u}_R = k\}$	\$0.23 (0.36)	10% (13)
Self-interest:			
Social + conflict costs	$w + \sum_k \theta^k \mathbf{1}\{\Delta_S \tilde{u}_R = k\}$	-\$1.24 (0.16)	55% (19)
Effort costs + cold glow	$-c + v$	-\$0.10 (0.40)	4% (16)
Total		-\$0.42 (0.18)	

This table shows the estimates of the dollar equivalent of each motive in the average willingness to send a Direct Pressure Message. Parameters are estimated under the specification of Equation 6. Standard errors are clustered by sender and calculated using the Delta method.

Table A.5: Non-registrant recipients: Experimental outcomes

	Response in Follow-up Survey			
	Received email	Did not receive email	No response	Total
Panel A. Control group				
Did not register	0 [0%(↓) 0%(→)]	6 [12%(↓) 86%(→)]	1 [17%(↓) 14%(→)]	7 [13%(↓) 100%(→)]
Registered but did not vote	0 [0%(↓) 0%(→)]	6 [12%(↓) 86%(→)]	1 [17%(↓) 14%(→)]	7 [13%(↓) 100%(→)]
Registered and voted	0 [0%(↓) 0%(→)]	37 [76%(↓) 90%(→)]	4 [67%(↓) 10%(→)]	41 [75%(↓) 100%(→)]
Total	0 [0%(↓) 0%(→)]	49 [100%(↓) 89%(→)]	6 [100%(↓) 11%(→)]	55 [100%(↓) 100%(→)]
Panel B. Treatment group				
Did not register	10 [21%(↓) 45%(→)]	9 [27%(↓) 41%(→)]	3 [21%(↓) 14%(→)]	22 [23%(↓) 100%(→)]
Registered but did not vote	4 [8%(↓) 57%(→)]	2 [6%(↓) 29%(→)]	1 [7%(↓) 14%(→)]	7 [7%(↓) 100%(→)]
Registered and voted	34 [71%(↓) 52%(→)]	22 [67%(↓) 33%(→)]	10 [71%(↓) 15%(→)]	66 [69%(↓) 100%(→)]
Total	48 [100%(↓) 51%(→)]	33 [100%(↓) 35%(→)]	14 [100%(↓) 15%(→)]	95 [100%(↓) 100%(→)]

Format: Count [column frequency%(↓) row frequency%(→)]

This table tabulates Non-registrant participants by the post-election experimental outcome (did not register, registered but did not vote, registered and voted) and whether they indicated in the Follow-up Survey that they had received an email from another participant in the study about registering to vote. This table excludes 15 Non-registrants who received or did not receive a message endogenously depending on their or their sender's WTP choices.