American Economic Review 2022, 112(1): 122–168
https://doi.org/10.1257/aer.20190433

Measuring the Welfare Effects of Shame and Pride†

By Luigi Butera, Robert Metcalfe,
William Morrison, and Dmitry Taubinsky*

Public recognition is frequently used to motivate desirable behavior, yet its welfare effects—such as costs of shame or gains from pride—are rarely measured. We develop a portable empirical methodology for measuring and monetizing social image utility, and we deploy it in experiments on exercise and charitable behavior. In all experiments, public recognition motivates desirable behavior but creates highly unequal image payoffs. High-performing individuals enjoy significant utility gains, while low-performing individuals incur significant utility losses. We estimate structural models of social signaling, and we use the models to explore the social efficiency of public recognition policies. (*JEL* C93, D64, D82, D91)

> *"What do you regard as most humane? To spare someone shame."*
> – Friedrich Nietzsche, The Joyful Wisdom

> *"A soldier will fight long and hard for a bit of colored ribbon."*
> – Napoleon Bonaparte[1]

The human desire to avoid negative social image and appear "good" is a powerful motivator (Loewenstein, Sunstein, and Golman 2014; Bursztyn and Jensen 2017). For instance, 89 percent of businesses use some form of public recognition programs

* Butera: Copenhagen Business School (email: lbu.eco@cbs.dk); Metcalfe: University of Southern California and NBER (email: robert.metcalfe@usc.edu); Morrison: UC Berkeley (email: wmorrison@berkeley.edu); Taubinsky: UC Berkeley and NBER (email: dmitry.taubinsky@berkeley.edu). Roland Bénabou was the coeditor for this article. We thank Ryan Johnson and Erik Daubert from the YMCA of the USA for their invaluable guidance and expertise in setting up our research partnership, as well as Maria-Alicia Serrano for her continuing support. We thank the YMCA of the Triangle Area, particularly Tony Campione, Mark Julian, Brian Spanner, and Janet Sprague, for their support in implementing our field experiment. We are indebted to Afras Sial for outstanding research assistance. We are grateful to David Birke, Stefano DellaVigna, and Devin Pope for generously sharing their experimental software. We thank Leonardo Bursztyn, Gary Charness, Alain Cohn, Jonathan Davis, Keith Ericson, Ray Fisman, Jana Gallus, David Gill, Alex Imas, John List, Fatemeh Momeni, Ricardo Perez-Truglia, Na'ama Shenhav, Daniel Sznycer, Daniel Tannenbaum, Jessica Tracy, as well as conference and seminar participants for helpful comments. The views expressed herein do not necessarily reflect the views of the YMCA of the USA or any other YMCA member association. Morrison thanks the National Science Foundation Graduate Research Fellowship for support. Taubinsky thanks the Alfred P. Sloan Foundation for research funding. The authors declare that they have no other relevant or material financial interests that relate to the research described in this paper. The YMCA experiment was approved by University of Chicago IRB IRB15-1647; the charitable contribution experiments were approved by Boston University IRB 5473X (BU) and University of California, Berkeley IRB 2020-01-1288. AEA RCT ID: AEARCTR-0004004 and AEARCTR-0005737. This paper was previously circulated under the title "The Deadweight Loss of Social Recognition."

† Go to https://doi.org/10.1257/aer.20190433 to visit the article page for additional materials and author disclosure statements.

[1] We thank an anonymous referee for this quote.

(WorldatWork 2017), including examples like "employee of the month" (Kosfeld and Neckermann 2011). Bloom and Van Reenen (2007) find that 60 percent of manufacturing companies publicly reveal and compare employees' performance data. Governments use public recognition programs to motivate citizens to pay their taxes (Bø, Slemrod, and Thoresen 2015; Perez-Truglia and Troiano 2018), to motivate bureaucrats to do a better job (Gauri et al. 2018), and to encourage teachers, doctors, and managers in schools and hospitals to improve their performance.

Recent field studies confirm that public recognition of individuals' behavior has substantial effects in a number of economically important domains. Examples include charitable and political donations (Soetevent 2005, 2011; Perez-Truglia and Cruces 2017), tax compliance (Perez-Truglia and Troiano 2018), education and career choices (Bursztyn and Jensen 2015; Bursztyn, Fujiwara, and Pallais 2017; Bursztyn, Egorov, and Jensen 2019), employee productivity (Ashraf, Bandiera, and Jack 2014; Neckermann, Cueni, and Frey 2014; Bradler et al. 2016; Kosfeld, Neckermann, and Yang 2017; Neckermann and Yang 2017), voter turnout (Gerber, Green, and Larimer 2008), blood donation (Lacetera and Macis 2010), childhood immunization (Karing 2019), energy conservation (Yoeli et al. 2013), and credit card take-up (Bursztyn et al. 2017).[2]

The *financial* costs of utilizing public recognition to motivate behavior are typically low, but the *image* costs—such as the emotional costs of shame—may not be. Although behavioral scientists sometimes refer to social-influence-based interventions as light-touch, innocuous "nudges" (Halpern 2015, Benartzi et al. 2017), it is well understood that such a label would not be appropriate for a policy that leads to a significant number of individuals experiencing shame (see, e.g., Bernheim and Taubinsky 2018, for a review). Indeed, there is a vigorous debate about the appropriateness of public policies that generate feelings of shame, with some political and legal theorists arguing that such policies are an unjustifiable offense to human dignity and a form of mob justice (Massaro 1991, Nussbaum 2009; see also Bénabou and Tirole 2011 for formal analysis).[3] On the other hand, public recognition policies that mostly generate warm feelings of pride are arguably a "win-win." Developing quantitative methods for measuring the welfare effects of public recognition is therefore crucial for both positive and normative progress.

In this paper, we develop a portable approach for directly quantifying the image utility effects of public recognition. We deploy our approach in two different experimental designs conducted with four different subject pools. In each experiment, we address three research questions. First, do people have a significant willingness to pay to seek out or avoid public recognition of their behavior, implying that public recognition has a direct image utility effect? Second, how does utility from public recognition depend on people's realized behavior? In particular, are individuals choosing high levels of socially desirable behavior made better off (e.g., from experiencing pride), and are individuals choosing low levels of the desirable behavior made worse off (e.g., from experiencing shame)? Third, are the net image payoffs

[2] Laboratory experiments also show that public recognition can enhance prosocial behavior; e.g., Andreoni and Petrie (2004); Rege and Telle (2004); Andreoni and Bernheim (2009); Ariely, Bracha, and Meier (2009); Jones and Linardi (2014); Bernheim and Exley (2015); Exley (2018); and Birke (2020).

[3] Others promote such policies as instruments for the internalization of community norms (Etzioni 1999, Kahan and Posner 1999).

negative or positive? As we show, this third question relates to both the curvature of the public recognition utility function (PRU), and to the reference standard at which image payoffs transition from negative to positive.

Our first experiment was conducted in the field, in partnership with the YMCA of the USA and the YMCA of the Triangle Area (YOTA) in Raleigh, North Carolina.[4] We invited all members of YOTA to participate in a newly designed one-month program called "Grow and Thrive." This program encouraged members to attend their local YMCA more often by having an anonymous donor give $2 to the local YMCA for each day that an individual attended the YMCA. While this charity incentive was provided to everyone, participants could also be assigned to a public recognition program, which would reveal each participant's attendance and donation raised to all other participants in the program.

Our second set of experiments was conducted online and builds on the Ariely, Bracha, and Meier (2009) and DellaVigna and Pope (2018) "Click for Charity" task. The online experiments complement our field experiment by utilizing a design that gives us greater flexibility and control over the decision environment. In this real-effort task, participants raise money for the American Red Cross by repeatedly pressing two keys on a computer keyboard. The design was within subjects, and participants took part in three rounds. In the anonymous effort round, participants' scores were not shared with anyone. In the anonymous and paid effort round, participants additionally received pay for their effort. In the publicly shared effort round, participants' contributions to the Red Cross were publicly shared with others in the experiment through a web page that posted individuals' photos, amount raised, rank relative to other participants, and, for two of the subject pools, names.[5]

We administered the online protocol simultaneously to three different subject pools that differ in individuals' familiarity with each other: (i) the online panel Prolific Academic, where participants almost surely do not know each other (henceforth, *Prolific sample*); (ii) the University of California, Berkeley's pool of subjects for economics and psychology experiments, where some participants might know each other (henceforth, *Berkeley sample*); and (iii) a section of Boston University's statistics class for second- and third-year undergraduate business majors, where students are likely to know each other (henceforth, *BU sample*).

Our revealed-preferences approach to estimating the effects of shame and pride utilizes the incentive-compatible Becker-DeGroot-Marschak (BDM) mechanism to elicit participants' (possibly negative) willingness to pay (WTP) for public recognition at various possible realizations of their performance. An advantage of this "strategy method" approach is that it is robust to possible misforecasting of one's future behavior. In the YMCA experiment, participants' WTP to be publicly recognized was elicited in an initial online survey before the start of the monthlong period during which incentives for attendance were provided. Participants were asked to state their WTP to be publicly recognized for all levels of attendance ranging from 0 to 30 days. To generate random assignment, as well as to minimize any negative

[4] The YMCA of the USA is a national, nonprofit, charitable organization that supports local communities with a focus on youth development, healthy living, and social responsibility. YOTA primarily serves the Raleigh-Durham, North Carolina, and surrounding communities. It is one of 850 member association YMCAs.

[5] Birke (2020) utilizes a similar approach to public recognition of online participants. We thank him for his advice and for kindly sharing his code.

inferences that could be drawn about participants who are not publicly recognized, the BDM responses were used to determine assignment to public recognition with only 10 percent chance. With 90 percent chance, assignment was random.

In the charitable contribution experiments, we again used the BDM mechanism to elicit participants' WTP to have their contribution to the Red Cross publicly recognized, for different possible levels of performance. As before, participants' elicited preferences were implemented with 10 percent chance, while with 90 percent chance, participants were randomly assigned to have their outcome based on one of the three rounds. In the 10 percent of cases where participants' preferences were implemented, participants' contribution was based on a randomly chosen score from one of the three rounds, and participants with a preference to be recognized were listed alongside the participants randomly assigned to the publicly shared effort round.

We present six sets of results. First, we find that public recognition substantially increased desirable behavior. In the YMCA experiment, it significantly increased attendance by 17 percent, and in the charitable contribution experiments, it significantly increased contributions by 13 percent, 14 percent, and 13 percent in the Prolific, Berkeley, and BU samples, respectively.

Second, we find that a majority of participants have a nonzero WTP for public recognition. The fraction of participants with positive WTP to either opt in or opt out of public recognition at some level of performance is 93 percent, 73 percent, 78 percent, and 89 percent in the YMCA, Prolific, Berkeley, and BU samples, respectively. Participants' eagerness to pay for social image is consistent with a long intellectual tradition of incorporating "psychic" or emotional effects into otherwise standard economic models using money metrics (starting with, e.g., Becker 1968, Ehrlich 1973).

Third, the WTP data allows us to examine how participants' image payoffs vary with performance. In all experiments, image payoffs are strictly increasing in performance, participants in the bottom quartile of performance receive negative payoffs, while participants in the top quartile of performance receive positive payoffs, on average. The robust presence of negative payoffs from public recognition is consistent with leading economics models of social signaling (e.g., Bénabou and Tirole 2006, Andreoni and Bernheim 2009), but it is not a robust implication of psychological theories of shame (Tangney et al. 1996; Tangney, Stuewig, and Mashek 2007). Psychologically, raising *any* amount of money for the Red Cross could have been perceived as commendable prosocial behavior.[6]

Fourth, we estimate structural models of social signaling. We consider "action-signaling" models in which individuals directly care about how their action compares to the population behavior (e.g., Becker 1991; Besley and Coate 1992; Blomquist 1993; Lindbeck, Nyberg, and Weibull 1999), and "characteristics-signaling" models in which individuals care about what their action reveals about their characteristics (e.g., Bénabou and Tirole 2006, Andreoni and Bernheim 2009, Ali and Bénabou 2020). We provide a key out-of-sample test of the validity of our methodology and modeling framework by showing that data on (i) the treatment effect of public recognition and (ii) people's WTP for public recognition can be used to predict (iii) the

[6]From a psychological perspective, shame is an emotion that accompanies moral transgressions (Tangney et al. 1996; Tangney, Stuewig, and Mashek 2007), and ex ante it was unclear that any action in our experiments could be labeled as such.

effect of financial incentives on behavior. In the charitable contribution experiments, the financial incentive was randomized, and thus we estimate its effects directly. In the YMCA experiment, we compare our models' predictions to individuals' forecasts of how they would respond to a financial incentive. Across all four subject pools, we find that the models' predictions only slightly overestimate the effects of the financial incentives, and that the difference is not statistically significant at conventional levels. This suggests that our monetization of image payoffs is accurately capturing the (presumably nuanced) psychological effects of public recognition.

Fifth, we study the shape of the PRU. In our models, whether the net image payoffs are negative or positive depends on the degree of concavity and the reference standard for positive image. Intuitively, more concavity leads individuals to be more sensitive to negative image, while a higher standard increases the fraction of individuals who experience negative effects. For example, if people derive positive image if and only if they are "better than average," then, by Jensen's inequality, a concave PRU makes public recognition negative-sum while a convex PRU would make public recognition positive-sum.

Both the reduced-form analyses and the structural estimates imply significant concavity in the YMCA and Prolific samples. We cannot reject linearity in the Berkeley and BU samples, although we also cannot reject that those samples feature as much concavity as the YMCA and Prolific samples. We also find that the standard for positive image payoffs is higher than the population average behavior in the YMCA and BU samples, is equal to the average in the Berkeley sample, and is lower than the average in the Prolific sample. Collectively, these results imply that public recognition is negative-sum in the YMCA and BU samples, is approximately zero-sum in the Berkeley sample, and is positive-sum in the Prolific sample.

Sixth, we use our structural estimates to generate out-of-sample predictions about the welfare and behavior effects of scaling up the public recognition intervention in the YMCA experiment to all of YOTA. We find that at the parameters estimated for the YMCA sample, public recognition is substantially negative-sum. However, if the PRU more closely resembled our estimate in the Prolific sample, then public recognition would be positive-sum.

Collectively, our results illustrate the importance of directly measuring the welfare effects of public recognition, and the potential benefits of our methodology. Our findings about the prevalence of negative image utility imply that the appropriateness of public recognition in settings such as ours could be legitimately debated. From a pure economic efficiency perspective, we find that public recognition could be a socially inefficient tool for behavior change in the YMCA field setting despite the low financial cost of the intervention and initial enthusiasm of our field partners. On the other hand, our results from the Prolific sample also illustrate that public recognition could be an efficient tool in other settings. This illustrates that it is inappropriate to judge the success of a public recognition policy solely by its effect on behavior, and how our methodology could help enrich the applied work on social signaling by helping researchers study both behavior and welfare.

The remainder of the paper is organized as follows. Section I further reviews the related literature. Section II introduces our theoretical framework. Section III describes the YMCA experiment and Section IV reports the reduced-form results. Section V describes the charitable contribution experiments and Section VI reports the reduced-form results. Section VII presents our estimates of structural models of social signaling and welfare implications. Section VIII concludes.

I. Discussion of Related Literature

Our research is related to several literatures. The most closely related is the large and growing experimental literature studying the effects of public recognition on individual behavior, summarized above. However, this literature studies behavior, and does not assess welfare effects of positive or negative image. We build on this literature by developing a portable approach for measuring image utility, which can be productively incorporated into future experiments on public recognition.

Our work also relates to a recent literature that evaluates the welfare effects of scalable, nonfinancial policy instruments such as reminders (Damgaard and Gravert 2018), energy-use social comparisons (Allcott and Kessler 2019), calorie labeling (Thunstrom 2019), and defaults (Carroll et al. 2009; Bernheim, Fradkin, and Popov 2015).[7] Our paper contributes to this literature by analyzing a different and highly popular nonfinancial policy instrument, and by providing new methods for testing and estimating models of social signaling. Unlike this prior work, our experiments utilize a new "strategy method" design technique that eliminates the need to rely on the assumption that individuals can correctly forecast their future behavior.[8]

Finally, our model-based design allows us to produce the first structural estimates of leading models of social signaling such as those of Bénabou and Tirole (2006).[9] We therefore also contribute to a recent and growing literature in structural behavioral economics (see DellaVigna 2018 for a review). The work by DellaVigna, List, and Malmendier (2012) and DellaVigna et al. (2017) is closest in spirit to our paper in this literature, although they do not study the scalable lever of revealing people's behavior to others, nor do they estimate the leading social signaling models. These two papers quantify the social pressure effects of face-to-face interaction in charitable contributions and voting, respectively. They do this by using structural methods to infer the cost of social pressure from the degree to which individuals avoid interaction with others. In contrast, we use conceptually different and more direct experimental techniques that leverage the richness of our action space and allow us to directly observe the shape of utility from the social motives. The richer data provided by our approach enables the estimation of structural models of social signaling.

II. Theoretical Framework for Analysis

A. *The Models*

We consider individuals who choose the level of intensity $a \in \mathcal{A} \subset \mathbb{R}^+$ to engage in some activity. Choosing a generates *material utility* $u(a;\theta) + y$, where

[7] Additionally, our work relates to the theoretical work of Kaplow and Shavell (2007), who derive conditions for when and how much to use policies that invoke shame or pride when the objective is to maximize social welfare.

[8] See Bernheim and Taubinsky (2018) for a more detailed discussion of the literature and potential confounds.

[9] Ariely, Bracha, and Meier (2009); Exley (2018); Bursztyn, Egorov, and Jensen (2019); and Karing (2019) test comparative statics of the Bénabou and Tirole (2006) model, and Karing (2019) quantifies the value of sending a positive (but not fully revealing) signal. These papers do not estimate the underlying public recognition utility function.

y is the individual's income and $\theta \in \mathbb{R}$ is the type of the individual, which we typically interpret as the individual's intrinsic motivation to engage in socially desirable behavior.[10] We assume that $u(a;\theta)$ is single-peaked in a and that $\frac{d}{da}u(a;\theta)$ is increasing in θ and is bounded. Thus, each individual has some optimal intensity level $a^*(\theta)$, and higher types θ derive more benefit from choosing higher levels of a. In addition to material utility, individuals also derive public recognition utility S, which we define below.

Consistent with psychological theories, we recognize that people can derive image payoffs either directly from their behavior a or from their characteristics θ (see, e.g., Leary 2007). We thus consider models of both of these mechanisms.

To simplify exposition, in the body of the paper we consider fully revealing equilibria in which each individual's choice of action a is perfectly observed, and in which there is a one-to-one mapping between types θ and actions a. We present the models and solution concepts in full generality in online Appendix A.

Formally, let S be an increasing function that satisfies $S(0) = 0$, and let $\nu \in \mathbb{R}^+$ be the "visibility parameter" (Ali and Bénabou 2020), which might depend on the number of observers, or the extent to which the observers are paying attention to an individual's behavior. The *action-signaling model* posits that when an individual's action is made public, the individual cares about how his action compares to a weighted average of behavior in the population (Becker 1991; Besley and Coate 1992; Blomquist 1993; Lindbeck, Nyberg, and Weibull 1999; 2003):

$$u(a;\theta) + y + \nu S(a - \rho\bar{a}), \tag{1}$$

where $\bar{a}$ is the average action in the population, and $\rho\bar{a}$ is the standard for what constitutes a positive versus negative image. The *characteristics-signaling model* posits that individuals derive utility from what their action reveals about their characteristics to the audience (e.g., Andreoni and Bernheim 2009, Bénabou and Tirole 2006, Ali and Bénabou 2020):

$$u(a;\theta) + y + \nu S(E[\theta|a] - \rho\bar{\theta}), \tag{2}$$

where $E[\theta|a]$ is the inference about a person's type given their behavior, $\bar{\theta}$ is the average type in the population, and $\rho\bar{\theta}$ is the standard for what constitutes positive versus negative image.[11]

The parameter ρ determines how many individuals experience positive versus negative image. When $\rho = 0$, all individuals choosing $a > 0$ receive positive image payoffs from public recognition. When $\rho > 1$, the standard is particularly demanding, as individuals must perform better than average to receive positive image payoffs.

As the general model in online Appendix A clarifies, the parameter ρ is a reduced-form parameter that is endogenous to the information structure. In our

[10] Assuming that utility is linear in income is a simplifying assumption that is not crucial for our theoretical exposition, but that is realistic given the relatively small financial stakes of our experimental setting.

[11] Note that there always exists a separating equilibrium in the characteristics-signaling model when u is smooth and $\mathcal{A}$ is convex and compact (Mailath 1987).

empirical estimates, the parameter should be regarded as a rough, not definitive, measure of whether individuals generally have high or low standard for positive image payoffs. In particular, in the case where (almost) nothing is revealed about individuals' behavior and characteristics, the general model makes the sensible prediction that individuals incur no image payoffs. Roughly speaking, the parameter ρ tends to one as the information structure coarsens. Additional parametric assumptions are necessary to use our estimates of ρ to make out-of-sample predictions about the impacts of other types of public recognition schemes.

B. *Net Image Payoffs*

Although theoretical work often makes the simplifying assumption that the net image payoff is zero by assuming that S is linear and that $\rho = 1$, it is well understood that both assumptions are not without loss of generality (e.g., Bénabou and Tirole 2006, 2011). From a psychological perspective, because shame and pride are separate emotions of different valences (Tangney, Stuewig, and Mashek 2007), people's well-being may not be equally sensitive to these two emotions, implying nonlinearity in S. And to the extent that shame is an emotion that accompanies moral transgressions (Tangney et al. 1996; Tangney, Stuewig, and Mashek 2007), it is also not clear that ρ might even be strictly positive for all behaviors. For example, raising *any* amount of money for charity might always lead to pride.

Both the curvature of S and the value of ρ determine the net image payoff. In particular, let $a^*(\theta)$ denote individuals' equilibrium strategies. Then the image payoffs in the two models are, respectively, given by

$$E\left[S\left(a^*(\theta) - \rho\bar{a}\right)\right], \tag{3}$$

$$E\left[S\left(E\left[\theta | a^*(\theta)\right] - \rho\bar{\theta}\right)\right]. \tag{4}$$

If S is concave and $\rho \geq 1$, then Jensen's inequality implies that the net image payoffs in the two models are given by

$$E\left[S\left(a^*(\theta) - \rho\bar{a}\right)\right] \leq S\left(E\left[a^*(\theta) - \rho\bar{a}\right]\right) \leq 0,$$

$$E\left[S\left(E\left[\theta | a^*(\theta)\right] - \rho\bar{\theta}\right)\right] \leq S\left(E\left[E\left[\theta | a^*(\theta)\right] - \rho\bar{\theta}\right]\right) \leq 0.$$

Thus, net image payoffs are negative when the function is concave and the standard for behavior/characteristics is at least as demanding as the average. Conversely, net image payoffs are positive when $\rho \leq 1$ and S is convex.[12] In general, the net image payoff decreases in ρ, decreases in the slope of $S(x)$ in the region $x < 0$, and increases in the slope of S in the region $x \geq 0$.

As we show in online Appendix A, the relationship between $E[S]$ and the shape of S holds more generally for any kind of public recognition scheme, such as two-tier public recognition schemes that publicize only the behavior of the top performers.

[12] In a similar vein, Corneo (1997) models trade union membership as a signaling game between workers, and shows that the reputation effect of trade union membership increases with union density if and only if S is concave.

Thus, if, for example, S is concave and people compare themselves to the average ($\rho = 1$), then the two-tier scheme will lead to a net negative image payoff as well. Intuitively, *not* being recognized as a top performer is worse than not having *any* information revealed about oneself, and thus the two-tier scheme cannot avoid inducing some amount of negative image payoff among those in the lower tier. Thus, our findings about the shape of S have implications beyond the fully revealing public recognition schemes that we study in this paper.

In online Appendix B, we show that the net image payoff $E[S]$ connects to a key economic question: whether public recognition is an efficient tool for behavior change relative to standard financial incentives. In addition to $E[S]$, the other three key inputs to this question are (i) the cost of implementing the public recognition scheme (e.g., due to the need to set up monitoring and distribution of information), (ii) the shadow cost of public funds, and (iii) the extent to which public recognition or financial incentives are best targeted toward people with the highest social marginal value of behavior change.

C. *Structural versus Reduced-Form Estimates of the PRU*

Often, the economic questions of interest are about the effects of utilizing public recognition on a whole population, not just the experimental sample. Answering this question requires an additional step of analysis, because scaling up public recognition to more people can change the equilibrium.

To formalize, call $R : \mathcal{A} \to \mathbb{R}$ the *reduced-form public recognition function* that assigns, for each value a, a public recognition payoff $R(a)$. Let R_{exp} denote the function elicited for the experimental population during the experiment, and let R_{pop} denote the reduced-form public recognition function that would result if public recognition was applied to the whole population of interest. These two objects can be meaningfully different: when the public recognition lever is applied to the whole population, population behavior changes, and thus the benchmark for what is considered relatively good behavior may change as well.

As a simple example, suppose that $\rho = 1$ and suppose that in our YMCA setting, an individual is observed to have attended the YMCA four times during the month of the experiment, and that average population attendance is 3.5 attendances. In the context of the experiment, an individual attending four times would thus receive positive public recognition payoffs in the action-signaling model. However, suppose that after applying the public recognition intervention to the whole population, average attendance would increase to 4.5 attendances. Then an attendance of four would actually generate negative public recognition utility. Our reading of existing literature studying social comparisons and social pressure is that it often stops at R_{exp}.[13]

[13] For example, suppose that individuals' utility in Allcott and Kessler (2019) is a decreasing function of the difference between their energy use and the energy use of the neighbors they are shown. Then the utility that they receive from the information mailer depends on whether the mailer goes out to their neighbors as well. However, since not everyone received the mailer in the experiment, the reduced-form effects that they estimate cannot be used to directly evaluate the policy of sending out mailers to all households. To perform such an evaluation, it would be necessary to take a stand on the structural utility function for social comparisons, to estimate it using the experimental results, and to estimate the counterfactual equilibrium of sending the mailers to everyone in the population. As another example, consider evaluating individuals' utility from encountering a surveyor who asks about voting behavior. DellaVigna et al. (2017) estimate the utility of doing so after votes have already been cast. But to evaluate

III. YMCA Field Experiment

A. *Recruitment*

The field experiment was conducted in collaboration with the YMCA of the USA and YOTA in North Carolina, and was publicly called "Grow and Thrive." YMCA members of two large YMCA facilities from YOTA were invited via email to sign up for this program by completing a survey. They were informed that for every day that they attended the YMCA during the program month, an anonymous donor would make a \$2 donation to their YMCA branch.

The Grow and Thrive program ran from June 15, 2017 to July 15, 2017. On June 1, 2017, the 15,382 members of the two YOTA branches received an email from their local YMCA announcing the launch of a new pilot program aimed at helping YMCA members to stay active and support their community at the same time. The initial email informed participants about the Grow and Thrive program and included a link to an online survey. YMCA members were told that they could sign up for the program by completing the survey and agreeing to participate.[14]

B. *Experimental Protocol*

The survey began by explaining the nature of the incentives during the program.[15] Participants were told that an anonymous benefactor with an interest in promoting healthy living and supporting the broader community provided funds to incentivize YOTA members to attend their local YMCA more frequently. During the month of the Grow and Thrive program, a \$2 donation was made on each participant's behalf for each day they visited the YMCA, up to a total donation of \$60 per person (i.e., 30 visits).

Participants were then told that they might also be randomly selected to participate in the public recognition program. We explained that if a participant was selected into this program, they would receive an email at the end of Grow and Thrive, which would (i) list the names of everyone in the program, (ii) list their attendance during Grow and Thrive, and (iii) list the total donations generated by them during Grow and Thrive. We explained that only participants in the public recognition program would receive and be listed in the email. Figure 1 provides a screenshot of what this public recognition email entailed.

We then utilized an incentive-compatible BDM mechanism to elicit participants' (possibly negative) WTP for public recognition for various possible realizations of

---

the equilibrium impact of increasing the visibility of one's voting behavior, it is necessary to account for the fact that visibility also changes voting behavior, which changes the payoffs one receives from telling a surveyor if one has voted or not. Evaluating the equilibrium outcomes would thus require one to estimate the structural microfoundations of why individuals like to tell others that they voted.

[14] The "pilot" language was important for our field partner, but we recognize that in principle it could have affected people's perceptions about the longer-run consequences of their choices. However, recent work by and de Quidt, Haushofer, and Roth (2018) and DellaVigna and Pope (2019) suggests that framing effects of this sort seem to have muted effects on behavior. DellaVigna and Pope (2019) also suggest that academics seem to overestimate the extent to which such framing matters.

[15] The experimental instructions Appendix contains text and screenshots of the instructions and questions used in the experiment.

Thank you for joining Grow & Thrive from your friends at YMCA!		
	# of visits	Dollars Raised
1. John Doe	25	$50
2. Mary Adams	24	$48
..		
49. Jack Black	10	$20
..		

Figure 1. Illustration of Public Recognition Information

Note: This figure shows an illustration of how individuals' attendance was publicized in the YMCA experiment.

their performance. The incentive-compatible method elicited WTP for public recognition the following possible contingencies of a person's performance: 0 visits, 1 visit, 2 visits, 3 visits, 4 visits, 5 or 6 visits, 7 or 8 visits, 9 to 12 visits, 13 to 17 visits, 18 to 22 visits, and 23 or more visits. For each of the 11 intervals, participants were first asked whether they would want to be publicly recognized if their attendance during Grow and Thrive fell in that interval. Participants were then asked how much they were willing to pay to guarantee that their choice was implemented.

Each of the 11 questions had the following structure: "*If you go to the YMCA* [*X times*] *during Grow and Thrive*, *do you want to participate in the personal recognition program?*" Participants were then asked to state, for each of the 11 levels of possible attendance, how much of an experimental budget of $8 they would be willing to give up to guarantee that their decision about public recognition was implemented. The question asked, "*You said you would rather* [*participate*] [*NOT participate*] *in the personal recognition program if you go* [*X times*] *to the Y. How much of the $8 reward would you give up to guarantee that you will indeed* [*participate*] [*NOT participate*] *in the personal recognition program?"*[16] The details were then explained in simple and plain language, and participants were told, in bold font, that *"it is in your interest to be honest about whether you want to participate in the personal recognition program*, *and how much of the $8 reward you would give up to ensure that you will or will not participate in the personal recognition program.*" Figure 2 provides a screenshot from the survey of one of the pairs of questions.

To preserve random assignment, as well as to minimize any negative inferences that could be drawn about those not in the public recognition group, we informed participants that their responses would be used to determine assignment with 10 percent chance, and that with 90 percent chance their assignment would be determined randomly. For participants in the 10 percent, a computer would check their attendance during Grow and Thrive and match it with their answers. With 50 percent chance, they would receive an $8 Amazon gift card and they would be assigned to the public recognition group if and only if they indicated a preference to be in that group. Otherwise, with 50 percent chance, the BDM mechanism was used to

[16] Each of these 11 questions was presented to subjects on a separate screen. To make it clear which attendance level was relevant to their WTP elicitation, we highlighted it.

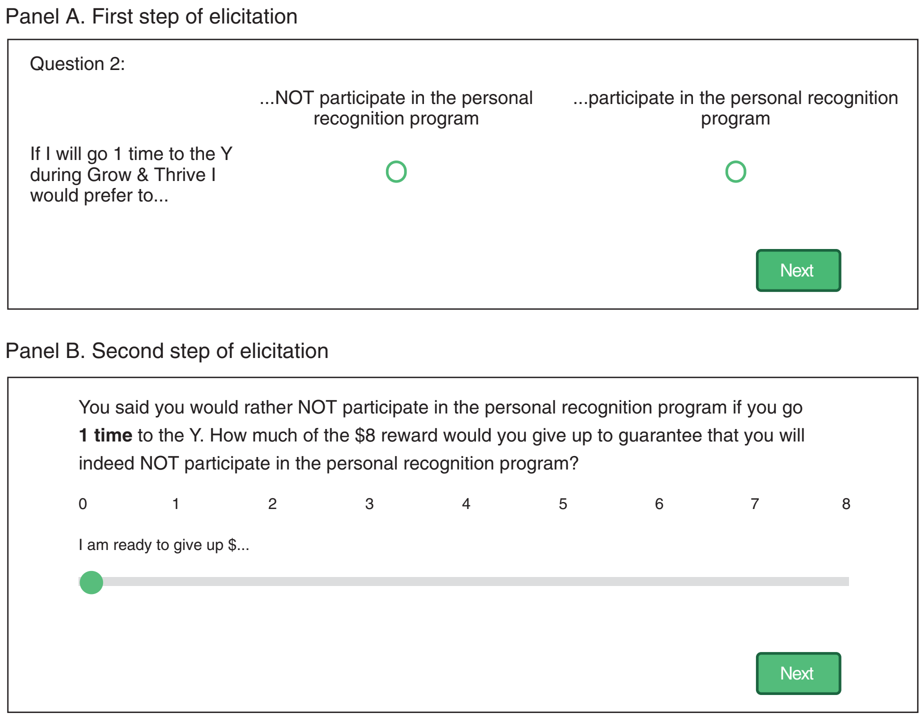

Figure 2. An Example of WTP for Public Recognition in the YMCA Experiment

Notes: These figures present screenshots of the procedure for elicitation of WTP for public recognition. The example above shows the elicitation of WTP for attending the YMCA once during Grow and Thrive. The top panel presents the first step of the elicitation, where participants are asked whether they want to be publicly recognized. The bottom panel presents the second step, where participants are asked how much they are willing to pay (from $0 to $8) to guarantee that their preference from the first step is implemented. Participants choose the amount by moving the slider bar.

determine the participant's extra reward and assignment to the public recognition group.[17]

To obtain intuition for why truth telling is incentive compatible with our mechanism, first note that a participant's chance of receiving public recognition is always higher if they indicate a preference for it in the first part of the elicitation. Second, after a participant commits their answer of whether or not they want public recognition, note that the bidding component of the elicitation is just a standard second-price sealed-bid auction against the computer. In summary, the procedure allowed participants to indicate a WTP for public recognition between −$8 and $8. For the 10 percent of participants whose decisions would be used to determine assignment, a bid of $8 guaranteed that the participant would be in the public recognition group, a bid of $0 generated a 50 percent chance of being in the public recognition group, and

[17] Specifically, the computer generated a random number between zero and eight, and a participant's preference for being in the public recognition program would be implemented if and only if the participant's WTP was higher than the random number. In this case, the computer's random number was subtracted from the participant's budget. If the computer chose a value greater than the participant's WTP to implement their choice, then the participant's preferred choice for being part of the public recognition program would NOT be implemented, and the participant would receive the $8.

a bid of $-\$8$ guaranteed that the participant would not be in the public recognition group.[18]

Because others' behavior plays a role in the models summarized in Section II, it was important to help participants have accurate beliefs about others' behavior. Prior to making their decisions about being part of the public recognition program, participants were provided an estimate of the average YOTA monthly attendance in the past year.

In the last component of the survey, we elicited participants' beliefs about their future attendance during Grow and Thrive with and without public recognition and under different levels of financial incentives. In this part, we also elicited participants' preferences over different financial incentives, which we describe later in the analysis. Finally, we reminded participants that a computer would randomly determine whether they would be part of the public recognition group, and we asked them to explicitly agree to participate in Grow and Thrive.

All participants were notified via email about their treatment assignment on the morning of the first day of Grow and Thrive. Participants assigned to the public recognition treatment received a reminder summary of the public recognition treatment when they were notified of their assignment.

All communications with YMCA members took place via email. We prepared an FAQ document covering common questions YMCA members might have about the program. To guarantee the consistency of the responses, and to minimize the burden on YMCA employees, we instructed employees working at the front desk to encourage members to address their questions via email to a specific contact person at the YMCA; the contact person would then use the answers provided in the FAQ to respond.[19]

C. *Attendance Data*

We received administrative attendance records from May 1, 2016 to July 15, 2017 for YMCA members in the branches where we conducted the experiment, including those not in Grow and Thrive. Attendances were recorded whenever a member accessed the YMCA facilities by swiping their personal YMCA access card on a turnstile. Before a member could swipe in, a front desk employee verified that the access card belong to the member.[20] We utilize attendance data for nonexperimental participants in the out-of-sample predictions in Section VII.

[18] To formally see that this procedure is incentive compatible, let v denote a participant's preferences to be publicly recognized at a particular attendance level. Then if a participant indicated a preference for public recognition and bid a value b, their expected payoff would be $\pi_1(b) = \$8 + 0.5v + 0.5(v - b/2)(b/8)$. Conversely, if the participant indicated a preference for no public recognition and bid b to not get it, then the expected payoff is $\pi_2(b) = \$8 + 0.5v + 0.5(-v - b/2)(b/8)$. Clearly, $\pi_1 = \pi_2$ if and only if $b = 0$, with $\pi_1 \geq \pi_2$ if and only if $v \geq 0$. Conditional on $v \geq 0$, the bid that maximizes π_1 is $b = v$. Conditional on $v < 0$, the bid b that maximizes π_2 is $b = -v$.

[19] The YMCA contact reported that only one participant contacted him, asking if he could be added to the public recognition group. After the (negative) response, there were no further questions from the participant.

[20] While YMCA members have to swipe in to access the YMCA, they do not have to swipe out to leave. Therefore we do not have information about how much time participants spent at the YMCA. YMCA employees were told to track any unusual activities among YMCA members. YMCA employees did not report any unusual pattern of access to the facilities during the experiment.

D. *Discussion of the Design*

What Are Individuals Signaling?—Due to the nature of our setting and the wishes of the YMCA, we were not able to implement a treatment in which participants received public recognition without the Grow and Thrive incentive of raising \$2 per attendance for YOTA. As such, we cannot fully differentiate between whether YMCA members were motivated by the desire to be recognized for being health conscious, or for being charitable. However, like charitable giving, pursuing good health through exercise is also perceived by many as a social and moral obligation (Conrad 1994, Whorton 2014, Cederström and Spicer 2015), and thus it is plausible that both motivations give rise to PRUs of similar structure.

Preferences for Signaling versus Preferences for Information.—Although all participants were given the average YOTA monthly attendance from the past year, only the public recognition group received information about others' behavior. To the extent that there was demand for this additional information, our WTP data is an upper bound on demand for public recognition alone. We chose to give any information to individuals only in the public recognition group to better capture the reality of how such interventions are usually implemented. In practice, the counterfactual to a public recognition scheme is not anonymized information provision—it is nothing at all.

Anticipated versus Realized Image Payoffs.—Although our approach does not require people to correctly forecast their future attendance, it does rely on the assumption that people can anticipate the image payoffs of public recognition. Testing this assumption would require a design that elicits people's WTP for public recognition after their attendance is realized. This design is significantly less well powered as it elicits only one data point per person, and thus is left for future work where larger samples can be acquired. However, because people experience shame and pride often, it is likely that they can accurately anticipate the intensity of these feelings, as is consistent with psychological evidence (Sznycer et al. 2016, 2017; Cohen, Chun, and Sznycer 2020).

IV. Reduced-Form Results from the YMCA Experiment

A. *The Experimental Sample*

A total of 428 YOTA members completed the survey and agreed to participate in Grow and Thrive; 192 participants were randomly assigned to participate in Grow and Thrive but not in the public recognition program and 193 participants were randomly assigned to participate in both Grow and Thrive and the public recognition program.[21] Forty-three participants were randomly assigned to receive the extra \$8

[21] We randomized our 428 participants into the public recognition group by blocking and balancing over WTP survey responses and attendance in the twelve months preceding the experiment. All participants were notified by the YMCA of the Triangle via email about their treatment assignment the morning of the first day of Grow and Thrive.

Table 1—Balance Table for YMCA Experiment

	No PR treatment	PR treatment	p-value
Average WTP (over all possible number of visits)	1.10	1.09	0.98
	(5.13)	(5.03)	
Average monthly past attendance	5.75	5.64	0.86
	(5.64)	(5.67)	
Beliefs about attendance assuming public recognition	13.90	13.41	0.44
	(5.88)	(6.18)	
Beliefs about attendance assuming no public recognition	12.51	11.83	0.28
	(5.94)	(6.09)	
Gender (0 = Male; 1 = Female)	0.74	0.76	0.63
	(0.44)	(0.43)	
Age	44.24	43.70	0.65
	(11.19)	(11.60)	
Number of subjects	185	185	

Notes: This table reports summary statistics across all coherent participants, by assignment to the public recognition (PR) group. Variable "average WTP (over all possible number of visits)" is the average participant WTP across all possible intervals of future attendance. Variables "beliefs about attendance assuming (no) public recognition" report the average forecast of future attendance conditional on (not) being part of the public recognition treatment. The last column reports two-sided p-values to test for balance across our experimental treatment. The analysis excludes 15 participants with "incoherent" preferences for public recognition. Standard deviations are reported in parentheses.

reward for themselves, which they were able to use to affect their likelihood of being publicly recognized. These 43 participants for whom participation in the public recognition program is endogenous are excluded from our empirical analysis.

Unless otherwise noted, from the remaining 385 participants we also exclude 15 participants who indicate a demand for public recognition that has no discernible relation to the number of attendances, and are thus likely confused or disengaged from the study.[22] The remaining *coherent sample* includes individuals whose WTP for public recognition is monotonically increasing in attendance, as well as individuals with preferences that are monotonically decreasing in attendance (i.e., a desire to be recognized as not wanting to attend the YMCA), or individuals with preferences that peak at intermediate levels of attendance (i.e., wanting to look "average"). In online Appendix C.8, we also analyze the slightly smaller group of participants whose preferences for public recognition are monotonically increasing in YMCA visits.

Table 1 shows that all pre-experiment outcomes, as well as preferences elicited through our online component, are balanced by whether participants were randomly assigned to be in the public recognition group. One noteworthy property of our sample is the high average past attendance of 5.69, which is approximately twice as high as the past attendance of 3.02 of all YOTA members. However, we show below that past attendance does not vary meaningfully with people's preferences over public recognition.

[22] The results are qualitatively and quantitatively similar when using all participants.

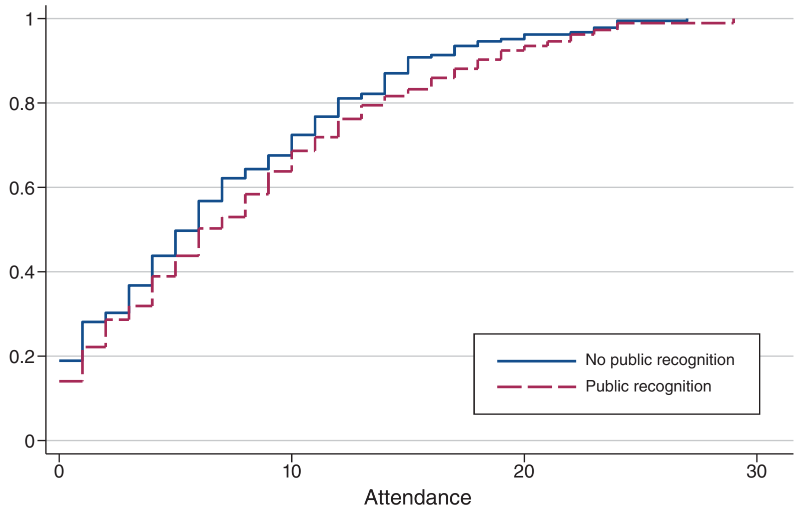

FIGURE 3. CUMULATIVE DISTRIBUTIONS OF ATTENDANCE DURING THE YMCA EXPERIMENT, BY TREATMENT

Notes: This figure plots the CDFs of attendance during the experiment, by whether participants were in the public recognition group. The analysis excludes 15 participants with "incoherent" preferences for public recognition.

B. *The Effect of Public Recognition on Behavior*

Figure 3 displays the cumulative distribution functions (CDFs) of attendance by treatment, showing that the impact of public recognition is positive across all levels of attendance. We quantify these results in Table 2. The table shows that public recognition increased attendance by approximately 1.2 visits. Given an average attendance of approximately seven visits in the control group, this corresponds to an approximately 17 percent increase in attendance. This estimate is just outside the range of marginal statistical significance without controlling for participants' past attendance, but becomes highly statistically significant when controlling for participants' past attendance.

C. *WTP for Public Recognition*

The significant effect of public recognition on behavior suggests that it constitutes a meaningful incentive to participants. Consistent with this, we find that 93 percent of participants have a strict preference to opt in or opt out of public recognition for at least one level of attendance.

Figure 4 plots the average WTP by the attendance level that would be publicized to other participants. These WTP profiles constitute model-free measures of the reduced-form PRU R_{exp} introduced in Section IIC. We identify each set of possible visits from our elicitation with its midpoint, meaning that the first five sets $\{0\}, \{1\}, \ldots, \{4\}$ are identified with $0, 1, \ldots, 4$; the "5 or 6 visits" set is identified with 5.5; the "9 to 12 visits" set is identified with 10.5; and so forth. Panel A presents data for participants with monotonic preferences for public recognition, panel B

Table 2—The Impact of Public Recognition on YMCA Attendance

	(1)	(2)	(3)
Public recognition	1.10	1.19	1.27
	(0.69)	(0.46)	(0.45)
Average past attendance		0.88	0.77
		(0.04)	(0.05)
Beliefs			0.19
			(0.05)
Control mean	6.91	6.91	6.91
	(0.47)	(0.47)	(0.47)
Number of subjects	370	370	370

Notes: This table reports regression estimates of the effects of public recognition on attendance during the experiment. "Beliefs" reports the expectations YMCA members had about their attendance assuming that they would be part of the public recognition treatment. The analysis excludes 15 participants with "incoherent" preferences for public recognition. The control mean is the average attendance for participants in the experiment who are not in the public recognition program. Standard errors are clustered at the participant level and reported in parentheses.

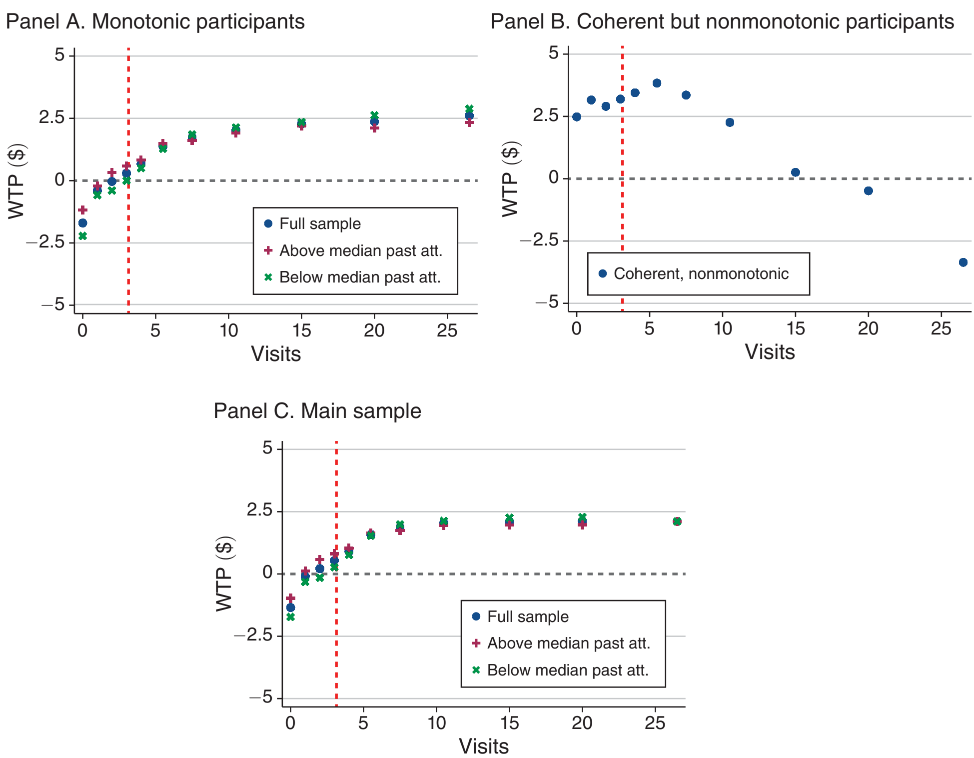

Figure 4. WTP for Public Recognition, by YMCA Attendance

Notes: These figures plot the average WTP for public recognition by each of the 11 intervals of possible future attendance. For intervals including more than one value of visits (e.g., "5 or 6 visits"), the WTP is plotted at the midpoint of the interval. Panel A reports the average WTP for participants with monotonic preferences for public recognition, as well as for this sample split by median past attendance. Panel B reports the average WTP for participants included in the main sample, but with nonmonotonic preferences for public recognition. Panel C reports the average WTP for the full sample, as well as for this sample split by median past attendance. The average YOTA attendance during Grow and Thrive is indicated by the dashed red line.

presents data from participants with coherent but nonmonotonic preferences, and panel C presents data from the full coherent sample (the combination of panels A and B). In panels A and C, we also plot the WTP of participants with above versus below-median past attendance. The vertical dashed line in each panel corresponds to the average YOTA attendance of 3.14, which is a potential reference standard for positive versus negative image payoffs. As discussed in Section II, the net effect image payoff is decreasing in the magnitude of the reference standard.

On average, as shown in panel C, the WTP for public recognition is strictly increasing in the number of visits. It is negative at low numbers of visits and positive at high numbers of visits. This pattern is more pronounced in the monotonic panel, as shown in panel A. Panel B shows that the remaining participants with nonmonotonic preferences have a distinct WTP profile that peaks at approximately seven attendances and declines steeply afterward. Consistent with this nonmonotonic profile, we find an essentially null (but noisy) effect of public recognition on the attendance of these 31 participants (0.39; 95 percent CI $[-2.59, 3.38]$).

Figure 4 also shows that participants' PRUs do not vary with their past attendance. We verify this formally in regression analysis in online Appendix Table A.1. This is important for two reasons: First, because participants in our study had a higher-than-average attendance, and thus a strong interaction between past attendance and WTP for public recognition could limit the external validity of our results. Second, this suggests that participants in our study did not self-select based on sensitivity to public recognition. If low attenders self-selected on being relatively insensitive to public recognition, while high attenders self-selected on being relatively sensitive to public recognition, then the WTP profiles for the above and below-median groups in Figure 4 would diverge.

Table 3 quantifies the descriptive results in Figure 4 by presenting regressions of WTP for public recognition on the midpoint of the visit intervals. We present results both from ordinary least squares (OLS) and Tobit regressions. Because some participants' WTPs were at the maximum possible amount of \$8 or the minimum possible amount of $-$\$8 for some of the elicitation intervals, some preferences were likely to be censored by our elicitation, and thus the Tobit models may give a more accurate assessment of how WTP for public recognition varies with the number of visits. We present linear regressions in odd-numbered columns, and we include a quadratic term for visits in even-numbered columns to study the curvature of the PRU. In this and all subsequent analyses of the WTP data, we cluster standard errors by participant.

All specifications in Table 3 generate two robust results, which are visually apparent in Figure 4. First, the WTP for public recognition is significantly increasing in the number of visits. Second, this relationship is significantly concave, as implied by the negative coefficient on visits squared.

The quadratic regression models allow us to quantify the curvature of the reduced-form PRU, R_{exp}. One measure of curvature is $-R''_{exp}/R'_{exp}(\bar{a}_{pop})$, where $\bar{a}_{pop}$ is the average attendance of YOTA participants, which is analogous to the coefficient of absolute risk aversion. Another measure of curvature is $-R''_{exp}/R'_{exp}(\bar{a}_{pop})$ multiplied by the standard deviation of attendance of YOTA participants. This second measure quantifies the percent decrease in R'_{exp} from a one standard deviation change in behavior, and is a unitless measure akin to the coefficient

Table 3—WTP for Public Recognition by YMCA Attendance

Model	OLS	OLS	Tobit	Tobit
Dependent variable	WTP	WTP	WTP	WTP
	(1)	(2)	(3)	(4)
Number of visits	0.10	0.36	0.19	0.62
	(0.01)	(0.04)	(0.03)	(0.07)
Number of visits squared		−0.01		−0.02
		(0.00)		(0.00)
Constant	0.20	−0.57	−0.03	−1.35
	(0.30)	(0.32)	(0.59)	(0.63)
$-R''/R'\ (\bar{a}_{pop})$	—	0.069	—	0.068
95 percent CI	—	[0.064, 0.075]	—	[0.062, 0.074]
$-R''/R'\ (\bar{a}_{pop}) \times$ SD	—	0.337	—	0.329
95 percent CI	—	[0.310, 0.364]	—	[0.299, 0.358]
Observations	4,070	4,070	4,070	4,070
Number of subjects	370	370	370	370

Notes: This table reports regression estimates from linear and quadratic models of willingness to pay for public recognition by attendance. Measures of the curvature of the estimated reduced-form public recognition function are $-R''_{exp}/R'_{exp}(\bar{a}_{pop})$ and $-R''_{exp}/R'_{exp}(\bar{a}_{pop}) \times$ SD, where $\bar{a}_{pop}$ and SD $=$ 4.86 are the average attendance and standard deviation of attendance for the general YOTA population, respectively. The analysis excludes 15 participants with "incoherent" preferences for public recognition. Standard errors are clustered at the participant level and reported in parentheses. Ninety-five percent confidence intervals for the curvature statistics are computed using the delta method.

of relative risk aversion. The unitless property allows us to compare our estimates of curvature across both the YMCA and the charitable contribution experiments.

Table 3 shows that while the coefficients in the Tobit models are almost twice as large as the corresponding coefficients in the OLS models, our measure of curvature is very stable. This suggests that while the censoring likely led to a linear rescaling of the PRU, it did not affect the *shape*.

In addition to censoring, another potential concern is that participants may have been less serious about the WTP elicitation when asked to evaluate public recognition for an attendance level that was outside the range of what they thought was likely. This could lead participants with low expectations of attendance to be relatively insensitive to variation at the upper range of potential visits, and participants with high expectations of attendance to be relatively insensitive to variation at the lower range of potential visits. We investigate this possibility in Figure 5 and Table 4.

Figure 5 presents the WTP data analogously to Figure 4, but restricts to data points that involve visit intervals whose midpoints are within four visits of individuals' forecasts of attendance in the event that they are randomized into the public recognition group. The standard deviation of the difference between participants' past attendance and their attendance during Grow and Thrive is 4.42; thus visits within four of individuals' forecasted attendance should not seem unlikely. Like Figure 4, Figure 5 shows that WTP for public recognition is strongly increasing and concave in the number of visits, and is close to zero at the YOTA average of 3.14 attendances. The key difference is that the WTP profile in Figure 5 is significantly steeper. While the profile in Figure 4 spans payoffs between approximately −\$2 and \$2, the profile in Figure 5 spans payoffs between approximately −\$4 and \$4. This difference is consistent with the possibility that the data reported in Figure 4

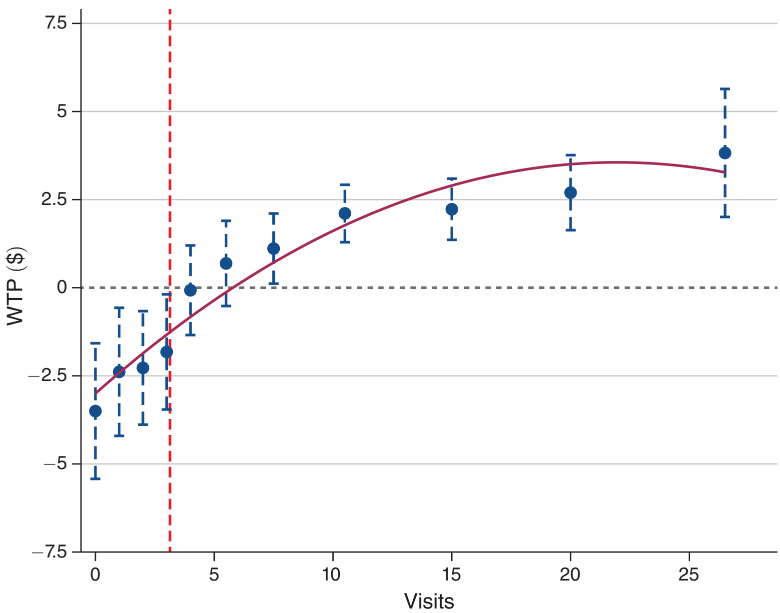

Figure 5. WTP for Public Recognition by YMCA Attendance, Restricting to Questions about Visits Close to Participants' Expectations

Notes: These figures plot the average WTP for public recognition by each of the 11 intervals of possible future attendance. For intervals including more than one value of visits (e.g., "5 or 6 visits"), the WTP is plotted at the midpoint of the interval. The data in these figures are restricted to visit intervals with a midpoint within four of a participant's predicted attendance if assigned to the public recognition group. The analysis excludes 15 participants with "incoherent" preferences for public recognition. The average YOTA attendance is indicated by the dashed red line. Ninety-five percent confidence intervals are constructed from standard errors clustered by participant. The quadratic fit curve is plotted in red.

feature some attenuation due to participants being less sensitive to variation in visits that are outside the range of what they consider plausible.

Table 4 quantifies the results suggested by Figure 5. Columns 1–4 present estimates that restrict to data points where the midpoints of the visit intervals are within four visits of participants' expected attendance if they are assigned to the public recognition group. Columns 5–8 restrict to data points where the visit interval contains participants' expected attendance. Relative to Table 3, the estimated coefficients in Table 4 are on net almost twice as large. The lack of a meaningful difference between the estimates in columns 1–4 versus columns 5–8 suggests that the attenuation is mostly due to considering visits that are very far from one's expectations. However, our estimates of curvature are very similar to the estimates in Table 3, which suggests that participants' reduced sensitivity to variation in unlikely attendance levels is affecting the scale, but not the shape of the WTP profile. Online Appendix C.3 shows that the results in Table 4 do not vary by past attendance, further reinforcing that past attendance is not a correlate of preferences for public recognition.

While a pure linear scaling bias cannot affect qualitative results about the welfare effects of public recognition, it does affect the magnitudes, as well as the out-of-sample predictions of our structural models. For this reason, our structural analysis in Section VII restricts to data where the midpoint of visit intervals is within

Table 4—WTP for Public Recognition by YMCA Attendance, Restricting to Questions about Visits Close to Participants' Expectations

Model	OLS	OLS	Tobit	Tobit	OLS	OLS	Tobit	Tobit
Dependent variable	WTP	WTP	WTP	WTP	WTP	WTP	WTP	WTP
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Number of visits	0.23	0.56	0.40	0.88	0.21	0.59	0.39	1.03
	(0.04)	(0.13)	(0.08)	(0.25)	(0.05)	(0.18)	(0.09)	(0.35)
Number of visits squared		−0.01		−0.02		−0.01		−0.02
		(0.00)		(0.01)		(0.01)		(0.01)
Constant	−1.27	−2.60	−2.47	−4.40	−0.69	−3.02	−1.90	−5.71
	(0.65)	(0.89)	(1.16)	(1.62)	(0.69)	(1.23)	(1.29)	(2.38)
$-R''/R'$ $(\bar{a}_{pop})$	—	0.057	—	0.053	—	0.051	—	0.048
95 percent CI	—	[0.039, 0.076]	—	[0.029, 0.078]	—	[0.031, 0.071]	—	[0.024, 0.073]
$-R''/R'$ $(\bar{a}_{pop}) \times$ SD	—	0.279	—	0.260	—	0.247	—	0.236
95 percent CI	—	[0.188, 0.371]	—	[0.143, 0.377]	—	[0.149, 0.345]	—	[0.114, 0.357]
Restriction	≤4	≤4	≤4	≤4	Exact	Exact	Exact	Exact
Observations	923	923	923	923	370	370	370	370
Number of subjects	370	370	370	370	370	370	370	370

Notes: These tables report regression estimates from linear and quadratic models of willingness to pay for public recognition by attendance. Columns 1–4 restrict to visit intervals with a midpoint within four of a participant's predicted attendance if assigned to the public recognition group. Columns 5–8 restrict to intervals that contain the participant's predicted attendance if assigned to the public recognition group. Measures of the curvature of the estimated reduced-form public recognition function are $-R''_{exp}/R'_{exp}(\bar{a}_{pop})$ and $-R''_{exp}/R'_{exp}(\bar{a}_{pop}) \times$ SD, where $\bar{a}_{pop}$ and SD $= 4.86$ are the average attendance and standard deviation of attendance for the general YOTA population, respectively. The analysis excludes 15 participants with "incoherent" preferences for public recognition. Standard errors are clustered at the participant level and reported in parentheses. Ninety-five percent confidence intervals for the curvature statistics are computed using the delta method.

four of participants' expectations, and utilizes the parametric assumptions of Tobit models to address censoring in the WTP data.

D. *Further Robustness Checks*

Potential Bias from High Visits Questions.—Online Appendix Table A3 shows that excluding high visit intervals slightly increases our estimate of curvature. Thus, our estimates are not biased by WTP for attendance in intervals that might fall outside the range of people's expected attendance.[23]

Potential Bias from Visit Intervals Increasing in Size.—To equalize the number of participants whose attendance falls within each bin and to avoid overburdening participants with too many WTP elicitations, we made the possible visit intervals larger at higher attendance levels. One concern is that this could have created an experimenter demand effect by signaling to participants that we expect differences in WTP for public recognition to be approximately constant across the intervals. This, in turn, could lead us to overestimate concavity. To gauge this concern, in online Appendix C.5 we index the 11 attendance intervals with the integers 0 through 10, and investigate how WTP for public recognition varies across these index values.

[23] Ten percent of participants expected to attend the YMCA as many as 23 times.

We find that WTP for public recognition is significantly concave, and even slightly larger, with respect to this recoding of the intervals.[24]

Demand for Public Recognition as Commitment.—Individuals with perceived self-control problems could in theory try to use our WTP elicitation to motivate their future selves to attend the YMCA more. We argue that this is unlikely for three reasons. First, the method for creating a commitment device using our WTP elicitation is nuanced. This entails individuals lowering expected payoffs for low attendance levels to discourage those low attendance levels. To do so, an individual needs to deviate from "truth telling" by placing a bid that is not equal to the image payoff at that attendance level. Thus, the bias, if it exists, is unsigned, because the individual can place a bid that is either higher or lower than their true expected image payoff. However, we think it is psychologically unrealistic that individuals would try to manipulate their future behavior in such subtle and sophisticated ways. For example, while individuals could in principle use incentivized belief elicitations as a form of a commitment device, Yaouanq and Schwardmann (2019) provide evidence against this.

Second, as shown by Carrera et al. (forthcoming) and others, demand for commitment is unlikely in environments featuring at least moderate uncertainty about future behavior, such as ours. In our sample, the standard deviation of the difference between attendance in two adjacent months is 4.74, which suggests a level of uncertainty that would likely make dominated incentive schemes costly. Third, in online Appendix C.6, we use additional survey questions to analyze whether people's perception of their time inconsistency correlates with their profile of WTP for public recognition, and find no evidence of this. We do this by utilizing the behavior change premium measure developed by Carrera et al. (forthcoming) and Allcott et al. (forthcoming).

E. *Realized Image Payoffs*

We end our reduced-form analysis by estimating the realized image payoffs induced by public recognition. We used the reduced-form PRU obtained from our WTP data, together with participants' actual attendance levels, to compute participants' average payoffs by quartile of attendance. To address the potential scaling bias discussed in Section IVC, we estimate payoffs for each level of attendance using the specification in column 4 of the two panels in Table 4: we use the Tobit model, and we restrict to WTP data that involves attendance intervals with midpoints within four visits of participants' expected attendance. To compute a participant's realized image payoff, we use the estimated regression to estimate the payoff associated with the participant's realized attendance during the month of the experiment. We present results using the raw WTP data in online Appendix C.7.

[24] To see why the estimate of curvature could increase, recall that quadratic functions are *locally* linear. A quadratic function that has a moderately smaller derivative at say 20 visits than at say 0 visits should in fact have similar derivatives at 0 visits and 10 visits. The fact that we find moderately smaller derivatives at an index value of 10 than at an index value of 0 thus implies substantial curvature with respect to the rescaled interval values.

Figure 6 presents the results, both for the monotonic and the coherent sample. On average, participants who were publicly recognized received a net-zero image payoff. Participants in the lowest quartile of attendance receive significantly negative payoffs, participants in the second quartile receive somewhat negative payoffs, and participants in the top two quartiles receive significantly positive payoffs.

Importantly, because participants in our experiment have significantly higher YMCA attendance than the average YOTA member, these reduced-form calculations constitute an upper bound on the net image payoff that would result from scaling up our public recognition intervention to the whole YOTA population. This suggests that scaling up the public recognition program to all of YOTA would generate a significantly negative average image payoff, consistent with our findings in Section VII.

F. *Over-Optimism and the Benefits of the Strategy Method*

A key feature of our design is that our elicitation of people's WTP for public recognition does not require them to form beliefs about their future attendance. In online Appendix C.2 we assess the accuracy of individuals' beliefs, and find significant overestimation of attendance, consistent with other work (e.g., DellaVigna and Malmendier 2006, Acland and Levy 2015, Carrera et al. forthcoming).

Because the PRU is (on average) monotonically increasing in attendance, this misprediction implies that simply eliciting WTP for being in the public recognition program, without conditioning on attendance, would create upward bias in conclusions about the welfare effects of public recognition. Related considerations apply to other social-influence-based interventions, such as the social comparisons studied in Allcott and Kessler (2019).

V. Charitable Contribution Experiment

A. *Recruitment*

The charitable contribution experiments were administered online to three separate subject pools: (i) members of the online platform Prolific Academic; (ii) participants from the University of California, Berkeley's Experimental Social Science Laboratory (Xlab), who are primarily undergraduate students; and (iii) undergraduate students from a mandatory statistics class, QM222, at Boston University's Questrom School of Business. We refer to these pools as the Prolific, Berkeley, and BU samples, respectively.

For all samples, the experiment ran for one week from April 18, 2020 to April 24, 2020.[25] For the Prolific sample, we recruited only participants who (i) reside in the United States, (ii) had a 95 percent or higher approval rating, and (iii) had completed at least 15 prior studies on Prolific. For the Berkeley sample, we restricted

[25] Before the experiment started, we preregistered our design and analysis plan on the AEA RCT Registry (AEARCTR-0005737). We had originally planned to also recruit from the QM221 statistics class for first-year students (who know each other less well than the QM222 students), but the response rate was too low to make use of this data. The results for the limited QM221 data ($N = 52$) are in online Appendix D.3.

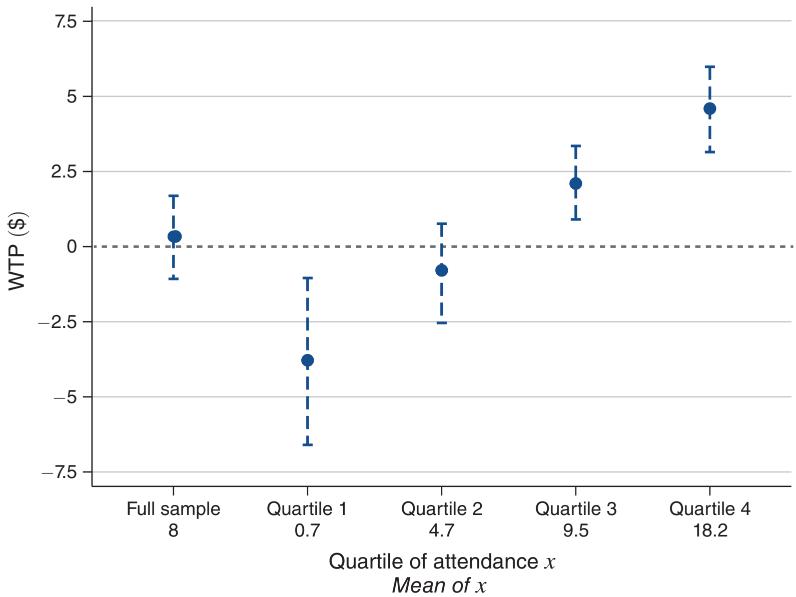

Figure 6. The Net Image Payoffs in the YMCA Experiment

Notes: These figures plot the average realized public recognition payoff of participants assigned public recognition, for both the full sample and each quartile of actual attendance. The average attendance is reported below each subsample label. A participant's payoff is defined as the WTP predicted by the regression in column 4 of Table 4, given the participant's realized attendance. The analysis excludes 15 participants with "incoherent" preferences for public recognition. Bootstrapped percentile-based confidence intervals, sampled by participant with 1,000 iterations, are displayed.

participants to those who had not taken any studies involving deception through Xlab. For the BU sample, all 350 students enrolled in QM222 received an email from their professor inviting them to participate in the experiment.[26] Participants from all subject pools were informed they could only complete the experiment on a laptop or personal computer with a working webcam.

B. *Experimental Protocol*

Except where noted below, the experimental protocol was identical for each of the three samples.[27] Perhaps the biggest implementation difference was the difference in incentive levels. Relative to the Prolific sample, we scaled up all incentives by a factor of 2.5 in the Berkeley and BU samples. This was done to reflect differences in payment norms across the samples. Prolific requires researchers to pay all participants at least \$6.50 per hour, Berkeley Xlab requires researchers to pay at least \$20 per hour, and BU requires researchers to pay at least \$15 per hour.

[26] The course was broken up into nine classes taught by five professors. Coauthor Robert Metcalfe taught three of the classes.

[27] The experimental instructions Appendix contains text and screenshots of the instructions and questions used in the experiment. An online example of the experiment is available here: https://wharton.qualtrics.com/jfe/form/SV_2mImcVP4XP3Pmf3.

In the experiment, participants could raise money for the Red Cross by successively pressing the "a" and "b" keys on the computer. Each pair of button presses earned a point, which translated to money donated to the Red Cross by the experimenters, and in some cases also to additional payments to the participants.

After consenting to participate in the experiment, participants first reviewed instructions about the button-pressing task. Participants then practiced the task for up to 30 seconds.

Participants were then presented with an overview of the structure of the experiment. Participants were told that they would complete three rounds of the button-pressing task (presented in random order), and that each round would last up to 5 minutes. We gave participants the option to finish each round early, since this "extensive margin" option appears to lead to more elastic labor supply, as suggested by DellaVigna et al. (2019), DellaVigna and Pope (2019), and our own pilots.

In all rounds, participants in the Berkeley and BU samples raised 5 cents for the Red Cross for every 10 points that they scored, while participants in the Prolific sample raised 2 cents for every 10 points. In the anonymous effort round, this was the only incentive, and participants' performance remained anonymous. In the anonymous and paid effort round, participants also earned financial compensation for themselves, which was identical to their Red Cross contribution (5 cents per 10 points in the Berkeley and BU samples, and 2 cents per 10 points in the Prolific sample). Participants' performance in this round also remained anonymous.

In the publicly shared effort round, participants' performance would be revealed to all participants in their experimental group after the conclusion of the study. In this round, participants' effort only translated to Red Cross donations, not to their own earnings. Specifically, after the end of the study, all participants would receive a link to view the pictures and contributions raised for the Red Cross of all participants in their group who were assigned to have their effort publicly shared with others. The information shared would include participants' photos, their scores and donations in the button-pressing task, their ranks relative to other publicly recognized participants and, for the Berkeley and BU samples, their names.[28] All participants were required to take a picture of themselves using their webcam, and they were given the option to upload an alternative picture or retake their picture. In summary, we included one baseline round where participation remained anonymous, one round where participants earned performance-based financial compensation, and one round where effort was publicly recognized.

Each round had a 30 percent chance of being randomly chosen to determine a participant's outcome. With 10 percent chance, participants' preferences for public recognition would be used to determine whether their performance would be publicly recognized or remain anonymous—we called this the choose your visibility option.

The choose your visibility option involved an incentive-compatible elicitation procedure that was analogous to that of the YMCA experiment. We asked 18 pairs of questions about WTP for public recognition, corresponding to 18 possible

[28] We did not collect and reveal participants' names in the Prolific sample because this would violate the platform's privacy requirements.

intervals of performance. The 18 intervals were 0–99 points, 100–199 points, . . . , 1,600–1,699 points, and 1,700 or more points. For each interval, we first asked participants if they wanted their effort to be publicly shared if it fell in one of those intervals, and we then asked them to state their WTP to have their preference implemented. Participants were given a $10 budget for this elicitation in the Prolific sample, and a $25 budget in the Berkeley and BU samples. As in the YMCA experiment, we told participants, in bold font, that "carefully and honestly answering the questions is in your best interest."

If the choose your visibility option was randomly chosen to determine a participant's outcome, then the score from one of the three rounds was randomly chosen to determine the participant's contribution to the Red Cross. However, the web page identifying participants' contributions did not differentiate between participants who were randomly assigned to be in the publicly shared effort round and participants assigned to the choose your visibility option—all recognized participants and their contributions were presented identically. Thus, the proper inference about any publicized participant is that their score was probably based on the publicly shared effort round, and that the reason their contribution was publicized was likely due to random chance rather than because of the preferences elicited in the choose your visibility option. This procedure also ensured that participants' performance in all three rounds carried equal importance and, by creating some uncertainty about the score used, broadened the range of scores that participants could consider relevant for the choose your visibility elicitation.

Because others' behavior can play a role in social image utility, we first collected an initial round of data to provide participants with signals of others' performance in the publicly shared effort round. Participants in the Prolific sample were presented with information from a 79-person pilot, and participants in the Berkeley and BU samples were given information from a 52-person pilot. Participants were informed of the average performance and the twenty-fifth, fiftieth, and seventy-fifth percentiles of performance from these samples. Participants were also informed of the sample size of the data, and were also provided a link to view a full CDF of past performance.

For the Berkeley and Prolific samples, participants were also informed about the size of their experimental group. In the Berkeley sample, participants were randomly divided into groups of approximately 75 participants, and they were told that approximately 25 participants in their group would have their effort publicly shared with all others in the group. In the Prolific sample, participants were randomly assigned to be in a group of 300, 75, or 15 participants, and were told that approximately 100, 25, or 5 participants in their respective group would have their effort publicly shared with all others in the group. We did not include language about group size in the BU sample because we did not have a sufficiently precise prediction about the response rate to provide truthful information. Importantly, the group assignment in the Prolific and Berkeley samples was completely random, which implies that standard errors need only be clustered at the participant level in all analyses.

The timing of the experiment was as follows. First, participants learned about the three rounds and the choose your visibility option. Second, participants received information on past performance and their group size, and answered an attention check question that instructed them to leave the question blank and advance to the next screen. Third, participants indicated their preferences for public recognition in the choose your visibility option. Fourth, participants completed the three button-pressing rounds. The order of the rounds was fully randomized. In each round, participants were reminded of the conditions of the round. In the publicly shared effort round, participants were also shown the image that would be seen by other participants.

Participants were informed of which round was randomly selected to count as soon as they completed the study. Within three days of the end of the study, participants were randomly divided into groups and were sent a link to view the performance information of all participants in their group who were assigned to have their effort publicly shared with others. Participants had 72 hours to view this information, and could only access it by entering the Prolific ID or university email address they had entered when completing the study. If participants clicked to view the additional information, they would receive an additional \$0.50 if in the Prolific sample, or \$1 if in the Berkeley or BU samples. The experimenters did not match the identities and scores of any participant who was not selected to be publicly recognized, and the participants were informed that they would be anonymous even from the experimenters if they were not assigned to be publicly recognized.

C. *Discussion of the Design*

Within-Person Variation.—We chose to have participants complete all three possible rounds for two reasons. First, this ensured that there would not be differential attrition. In a between-subjects design where each participant completed only one of the three rounds, a realistic possibility is that participants might be more likely to attrit from conditions in which they did not receive additional pay for their performance, or conditions in which they might incur negative image payoffs. Second, our design maximizes statistical power for comparisons of performance across the three rounds, and allows for some additional analyses of individual differences.

Relation to the YMCA Experiment.—The charitable contribution experiments complement the YMCA experiments in five key ways.

First, the experiments explore a different domain, and one that is arguably a more common target of public recognition: giving time and effort to charity. This permits an initial investigation of the cross-domain stability of various aspects of people's preferences over public recognition.

Second, by simultaneously running the experiment on three different samples, we are able to explore cross-population stability. One notable difference between our three samples is people's familiarity with each other.

Third, the charitable contribution experimental design more directly eliminates the possibility that participants might use the WTP for public recognition elicitation as a type of commitment device. There is only a 5–15 minute gap between when participants complete the elicitation and when they begin the real-effort rounds, and thus all of these decisions are likely to be regarded as "now." Augenblick's (2018) estimates of discounting in real-effort tasks similar to ours strongly support this interpretation.[29]

Fourth, the large size of the Prolific sample allows us to analyze how group size might affect participants' preferences to be publicly recognized. This analysis is helpful for refining out-of-sample predictions that involve larger groups than those in the experiment. The possible effects of group size can be captured by the ν parameter in the structural models in Section II, but the effects are ambiguous. On the one hand, larger group sizes imply larger audiences. On the other hand, larger group sizes imply that any recognized participant is likely to receive less attention.

Fifth, the charitable contribution experimental design has other features that make analysis and interpretation more straightforward: (i) the design provides subjects not just with the mean of past performance, but with the whole distribution, which could be important if people care about statistics other than average performance; (ii) the design has a significantly larger allowable range in the WTP elicitation, which essentially eliminates all censoring; (iii) the elicitation interface has evenly sized performance intervals, which eliminates potential worries about what participants might infer from variable interval widths; (iv) all participants, not just those publicly recognized, see the performance of the publicly recognized group, which implies that WTP for public recognition cannot be affected by a demand for additional information.

VI. Reduced-Form Results from the Charitable Contribution Experiment

A. *The Experimental Samples*

A total of 1,017, 407, and 121 participants completed the Prolific, Berkeley, and BU experiments. We make two preregistered exclusions for our analysis. We exclude participants failing the attention check, and we exclude participants with "incoherent" preferences for public recognition, where "incoherent" is defined analogously to the YMCA analysis. This yields a final sample of 968, 384, and 118 participants in the Prolific, Berkeley, and BU experiments. Out of the remaining participants, almost all (all but 1.0, 1.8, and 1.7 percent of Prolific, Berkeley, and BU participants, respectively) had monotonically increasing preferences for public recognition, and our results are qualitatively and quantitatively unchanged if we restrict to this monotonic sample. Thus, to simplify the analysis, we present results only for the coherent sample.

In this final sample, Prolific participants were divided into 17 groups of 13–15 participants each, 6 groups of 71–79 participants each, and 1 group of 278 participants. All Berkeley participants were divided into 5 groups of 75–79 participants each, and all BU participants were in the same group.

[29] Augenblick (2018) estimates discount factors for a real-effort task very similar to ours at time horizons varying between a few hours and seven days, using the Berkeley Xlab pool. The estimates imply no plausible discounting for time horizons that are shorter than 15 minutes. For example, while Augenblick (2018) estimates a discount factor of 0.87 for a 7-day horizon, he estimates discount factors of 0.91 and 0.94 for 24-hour and 3-hour horizons, respectively. Extrapolating with any reasonable parametrization of the discount factor to a horizon of 0.15 hours would imply virtually no discounting at that horizon.

There was minimal censoring in the WTP for public recognition elicitation. Prolific, Berkeley, and BU participants chose to use all of their budget in only 6, 4, and 6 percent of all cases, respectively.

Our 100-point intervals in the WTP elicitation generated nearly complete coverage of the distribution of effort. Only 1.1, 2.6, and 2.0 percent of scores in the Prolific, Berkeley, and BU samples, respectively, were 1,700 points or higher.

The average age was 35, 21, and 20 for the Prolific, Berkeley, and BU samples, respectively. The percent of Prolific, Berkeley, and BU participants who identified as female was 50, 69, and 51, respectively.

The averages of the standard deviations of the difference in points scored between any two rounds were 390.9 points, 423.4 points, and 469.7 points in the Prolific, Berkeley, and BU samples, respectively. These scores suggest a fair amount of uncertainty about the score that would be used if selected for the choose your visibility option.

B. *The Effects of Public Recognition on Behavior*

Figure 7 displays the CDFs of points scored by treatment, showing that the impact of public recognition is positive across all levels of points scored in each of the three samples. The figure also suggests that the effect of public recognition is about half of the effect of financial incentives in the Prolific sample, and is only somewhat smaller than the effect of financial incentives in the Berkeley and BU samples.

Table 5 quantifies the effects depicted in Figure 7. The table reports results from OLS regressions of points scored on the experimental round. Column 1 presents results from the Prolific sample, column 2 presents results from the Berkeley sample, and column 3 presents results from the BU sample. Column 4 analyzes whether the effects of public recognition in the Prolific sample vary by group size. In all columns, we control for the order of the round by including dummies for whether the round appeared first, second, or third to a given participant, although the F-tests presented in Table 5 do not detect any fatigue or other order effects. We cluster standard errors at the participant level in this all and subsequent analyses.

As columns 1–3 of Table 5 show, public recognition increases participants' total effort by over 10 percent in all three rounds, which is highly statistically significant. The effects of the financial incentive are substantially larger in the Prolific sample, and modestly larger in the Berkeley and BU samples. Column 4 presents preliminary evidence that the three different group sizes considered in our Prolific experiment do not seem to moderate the effects of public recognition. Thus, the results suggest that the effect of a larger audience is offset by the decrease in attention any recognized individual receives.

Robustness.—We find no evidence that within-subject estimates differ from between-subject estimates. Online Appendix Table A10 analyzes pure between-subject variation by limiting to the first round the participants encountered. The effects of public recognition and financial incentives are virtually identical to the within-subject estimates in the Prolific and Berkeley samples. The effects of both public recognition and financial incentives are substantially smaller in the BU

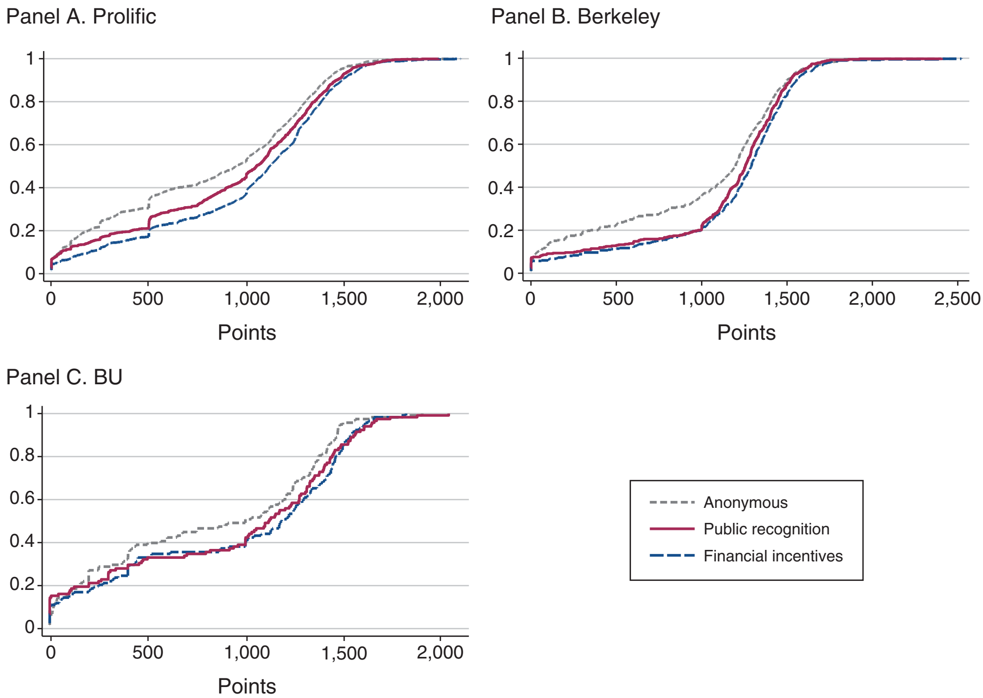

Figure 7. Cumulative Distributions of Points Scored in Each of the Three Rounds of the Charitable Contribution Experiments

Notes: These figures plot the cumulative distribution functions of points scored in the anonymous effort round, the anonymous and paid effort round, and the publicly shared effort round. Panel A presents results for the Prolific sample, panel B presents results for the Berkeley sample, and panel C presents results for the BU sample. The analysis excludes 40 Prolific participants, 11 Berkeley participants, and 2 BU participants with "incoherent" preferences for public recognition.

sample, although they are measured very imprecisely due to the small size of this sample, and the confidence bands include the within-subject estimates.

C. *WTP for Public Recognition*

Consistent with the significant effect of public recognition on behavior in all three samples, we find that 73 percent, 78 percent, and 89 percent of participants in the Prolific, Berkeley, and BU experiments, respectively, have a nonzero WTP for public recognition at one or more levels of performance.

Figure 8 plots the WTP for public recognition by level of publicized effort to raise money for the Red Cross, measured in points. We identify each interval below 1,700 with its midpoint, so that the first interval corresponds to 50 points, the second interval corresponds to 150 points, and so forth. The last point in the figure corresponds to the "1,700 or more" points interval. Panel A presents data from the Prolific sample, panel B presents data from the Berkeley sample, and panel C presents data from the BU sample. In addition to the sample averages, each panel also summarizes the WTP for participants with above- and below-median performance in the anonymous effort round. In all three panels, the vertical dashed line corresponds to the average

Table 5—The Effect of Public Recognition and Financial Incentives on Performance in the Charitable Contribution Experiments

Model	OLS	OLS	OLS	OLS
Dependent variable	Points	Points	Points	Points
	(1)	(2)	(3)	(4)
Public recognition	105.01	134.41	103.61	106.70
	(12.25)	(22.56)	(45.25)	(18.72)
Financial incentives	185.74	177.76	118.33	191.96
	(12.56)	(22.04)	(39.62)	(18.98)
Group of 300				20.61
				(39.85)
Group of 300 × Public recognition				−3.12
				(28.43)
Group of 300 × Financial incentives				−18.85
				(29.05)
Group of 15				17.70
				(41.13)
Group of 15 × Public recognition				−3.21
				(31.13)
Group of 15 × Financial incentives				−3.27
				(31.90)
Control mean	807.9	989.8	815.9	
	(16.7)	(27.2)	(52.8)	
Round order dummies	Yes	Yes	Yes	Yes
Order dummies F-test	0.180	0.497	0.116	0.178
Sample	Prolific	Berkeley	BU	Prolific
Observations	2,904	1,152	354	2,904
Number of subjects	968	384	118	968

Notes: This table reports regression estimates of the effects of public recognition and financial incentives on points scored. Columns 1, 2, and 3 report estimates for the Prolific, Berkeley, and BU samples, respectively. Column 4 includes interactions with group size variables in the Prolific sample, which indicate the approximate number of individuals in the participant's randomly assigned public recognition group. The control mean is the mean points scored in the anonymous effort round. Dummy variables for the order in which the round appeared (first, second, or third) are included, and the p-value from a test of their joint significance is reported. The analysis excludes 40 Prolific participants, 11 Berkeley participants, and 2 BU participants with "incoherent" preferences for public recognition. Standard errors are clustered at the participant level and reported in parentheses.

score in the publicly shared effort round, which is a potential reference standard for positive versus negative image payoffs. As discussed in Section II, the net image payoff is decreasing in the magnitude of the reference standard.

On average, WTP for public recognition is strictly increasing in points scored in all three samples. In all samples, it is negative at low levels of points scored and positive at high levels of points scored. Figure 8 also shows that participants' PRUs do not vary meaningfully with their score in the anonymous effort round, suggesting that preferences for public recognition do not vary meaningfully with their cost of effort or intrinsic motivation to help the Red Cross. Online Appendix Figure A9 presents confidences intervals for the average WTP in each interval.

Table 6 quantifies the descriptive results in Figure 8 by presenting results from regressions of WTP for public recognition on effort to raise money for the Red Cross, measured in points. Because very few participants' responses are censored

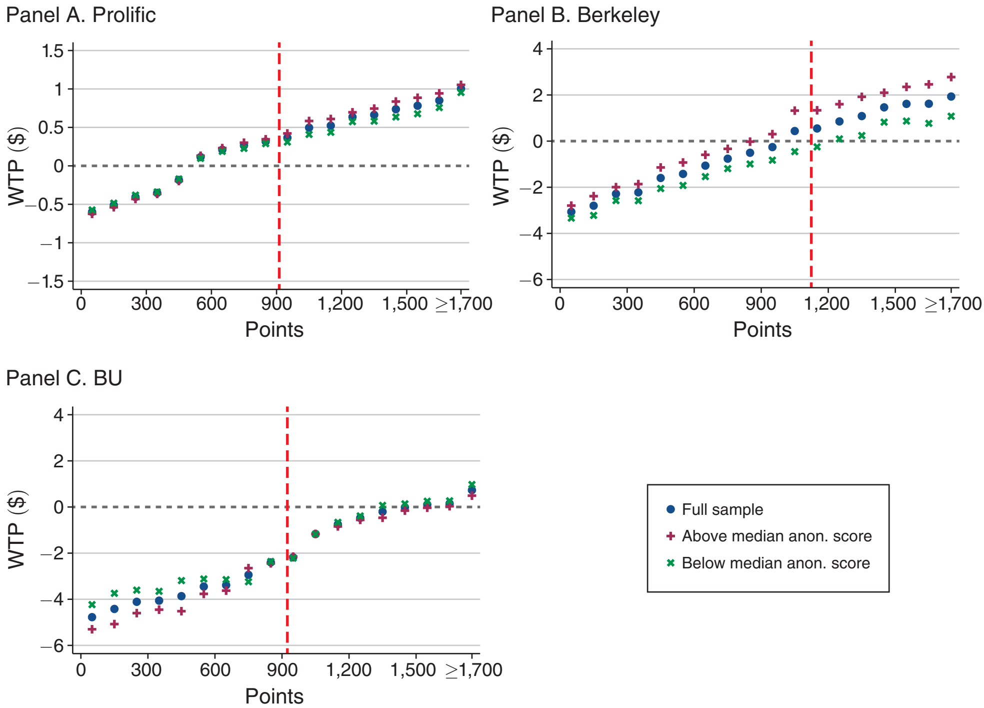

Figure 8. WTP for Public Recognition by Effort in the Charitable Contribution Experiments

Notes: These figures plot the average WTP for public recognition by each of the 18 possible intervals of points scored. The WTP is plotted at the midpoint of each of the first 17 intervals and at ≥1,700 points for the 1,700 or more points interval. Panel A presents results for the Prolific sample, panel B presents results for the Berkeley sample, and panel C presents results for the BU sample. The mean publicly shared effort round scores are indicated by dashed red lines. The analysis excludes 40 Prolific participants, 11 Berkeley participants, and 2 BU participants with "incoherent" preferences for public recognition.

at their full budget, we report results from OLS regressions only. The results are virtually identical in Tobit regressions. Columns 1 and 2 report results from the Prolific sample, columns 3 and 4 report results from the Berkeley sample, and columns 5 and 6 report results from the BU sample. We present linear regressions in odd-numbered columns, and we include a quadratic term for visits in even-numbered columns to study the curvature of the PRU. For this analysis, we exclude the ≥1,700 points interval as it does not represent a narrow band of performance like the other intervals. We make this exclusion in other analyses unless otherwise noted.

Consistent with Figure 8, all regressions imply that the WTP for public recognition is strongly increasing in the level of publicized effort. The implications for curvature are more mixed. The regressions imply significant concavity in the Prolific experiment, and smaller but imprecisely estimated levels of curvature in the Berkeley and BU samples. In the Berkeley and BU samples, we cannot reject linearity, although the 95 percent confidence intervals for curvature, $-R''/R'(\bar{a}_{pop})$, also include the point estimate from the Prolific sample.

Online Appendix D.4 uses the Bayesian information criterion (BIC) to formally show that the linear and quadratic models in Table 6 are the best fit to the data

Table 6—WTP for Public Recognition by Effort in the Charitable Contribution Experiments

Dependent variable	OLS WTP (1)	OLS WTP (2)	OLS WTP (3)	OLS WTP (4)	OLS WTP (5)	OLS WTP (6)
Points (100s)	0.093	0.155	0.310	0.379	0.347	0.309
	(0.007)	(0.018)	(0.033)	(0.070)	(0.060)	(0.116)
Points (100s) sqd.		−0.004		−0.004		0.002
		(0.001)		(0.004)		(0.006)
Constant	−0.557	−0.733	−3.130	−3.325	−5.186	−5.076
	(0.113)	(0.121)	(0.400)	(0.420)	(0.791)	(0.810)
$-R''/R'\ (\bar{a}_{pop})$		0.076		0.027		−0.013
95 percent CI		[0.047, 0.106]		[−0.021, 0.075]		[−0.079, 0.052]
$-R''/R'\ (\bar{a}_{pop}) \times$ SD		0.245		0.114		−0.085
95 percent CI		[0.186, 0.303]		[−0.047, 0.275]		[−0.559, 0.388]
Sample	Prolific	Prolific	Berkeley	Berkeley	BU	BU
Observations	16,456	16,456	6,528	6,528	2,006	2,006
Number of subjects	968	968	384	384	118	118

Notes: This table reports regression estimates from linear and quadratic models of willingness to pay for public recognition by the level of publicized effort. Effort is measured in 100s of points scored. The regressions exclude the $\geq$1,700 points interval. Measures of the curvature of the estimated reduced-form public recognition function are $-R''_{exp}/R'_{exp}(\bar{a}_{pop})$ and $-R''_{exp}/R'_{exp}(\bar{a}_{pop}) \times$ SD, where $\bar{a}_{pop}$ and SD $=$ 4.86 are the average and standard deviation of points scored in the anonymous round (in units of hundreds of points), respectively. The analysis excludes 40 Prolific participants, 11 Berkeley participants, and 2 BU participants with "incoherent" preferences for public recognition. Standard errors are clustered at the participant level and reported in parentheses. 95 percent confidence intervals for the curvature statistics are computed using the delta method.

presented in Figure 8. We show that the slight convexity visible around some multiples of 500 is best explained by moderate "round number bias." When dummies at multiples of 500 are included, higher-order terms beyond the quadratic specification are estimated to be near zero. Second, the round number bias is sufficiently small that the BIC-minimizing models are a quadratic polynomial (without dummies at multiples of 500) in the Prolific sample and a simple linear model (without dummies at multiples of 500) in the Berkeley and BU samples.

The slight uptick in the WTP at the $\geq$1,700 interval is consistent with theory, as individuals should assign a particularly high WTP to that interval if they believe that a score in that interval is perceived to be substantially higher than 1,750. The mean performance conditional on being in that interval is 1,791.6 (SE 28.2), 1,871.6 (SE 94.2), and 1,884.7 (SE 86.4) in the Prolific, Berkeley, and BU samples. Online Appendix Figure A10 plots a variation of Figure 8 where the location of the $\geq$1,700 interval on the x-axis is set equal to the average score in that interval; the figure reveals no trend break around that interval. Consistent with this, online Appendix Table A12 replicates Table 6 on the full data that include the $\geq$1,700 interval, and finds essentially identical regression estimates.

We can compare our unitless measures of curvature, $-R''/R'(\bar{a}_{pop})$ multiplied by the standard deviation of behavior, across the YMCA and charitable contribution experiments. In the charitable contribution experiments, we use the standard deviation of behavior in the anonymous round. Column 2 shows that our estimate of normalized curvature in the Prolific sample is strikingly similar to the estimates in Tables 3 and 4 for the YMCA sample. The analogous estimates for the Berkeley and

BU samples in columns 4 and 6 are smaller in magnitude, although the 95 percent confidence intervals include all point estimates from Tables 3 and 4. Overall, in the Berkeley and BU samples we can neither reject linearity nor the degree of curvature estimated in the YMCA and Prolific samples.

Any potential differences in WTP data among the Prolific, Berkeley, and BU samples are unlikely to be explained by group size. Consistent with our results about the effects on behavior not being affected by group size, online Appendix Table A15 shows that there is no interaction between group size and WTP for public recognition in the Prolific sample. We estimate fairly precise null effects for all interactions, which supports the hypothesis that the effect of a larger audience is offset by the decrease in attention any recognized individual receives.

Robustness and Heterogeneity.—In the YMCA experiment, participants' elicited WTP for public recognition was less sensitive to variation in performance that was outside the range of what they construed as likely behavior for themselves. We investigate this possibility in the charitable contribution experiments in online Appendix Table A11, which presents results from regressions analogous to those in Table 6, but restricting to data where the intervals for which WTP is elicited are within 500 points of participants' average performance in the three rounds. The estimates in online Appendix Table A11 are almost identical to those in Table 6. This is perhaps due to the fact that participants who have experienced economics experiments are better at answering more hypothetical/abstract questions.

We find some evidence for heterogeneity in preferences for public recognition, but consistent with our YMCA results, we find that these preferences do not covary with intrinsic motivation to raise money for the Red Cross, as measured by performance in the anonymous effort round. Online Appendix Table A13 shows that participants with an above-median difference in scores between the public and anonymous rounds also have a steeper PRU—that is, their WTP for public recognition is more steeply increasing in performance. This interaction is significant in the Prolific and BU samples in linear regressions of WTP on performance, but is more noisily estimated in the smaller BU sample, and in regressions that include a quadratic performance term. On net, these results suggest some stable individual differences in preferences for public recognition: some participants have steeper PRUs, and thus their performance is more sensitive to public recognition. Online Appendix E uses random coefficient models to estimate heterogeneity in PRUs more directly, and shows that while there is indeed significant heterogeneity in sensitivity to public recognition, there is little heterogeneity in the curvature of PRUs. Online Appendix Table A14 shows that there is no relationship between the PRU and participants' intrinsic motivation, consistent with the graphical evidence in Figure 8.

D. *Realized Image Payoffs*

Finally, we estimate the net image payoff induced by public recognition. We do this by assigning to each participant the average WTP for public recognition that corresponds to the interval containing the participant's score in the publicly shared effort round. We use the sample average WTP, instead of the participant's own WTP, to maximize statistical power. As discussed above, the PRU does not vary with participants'

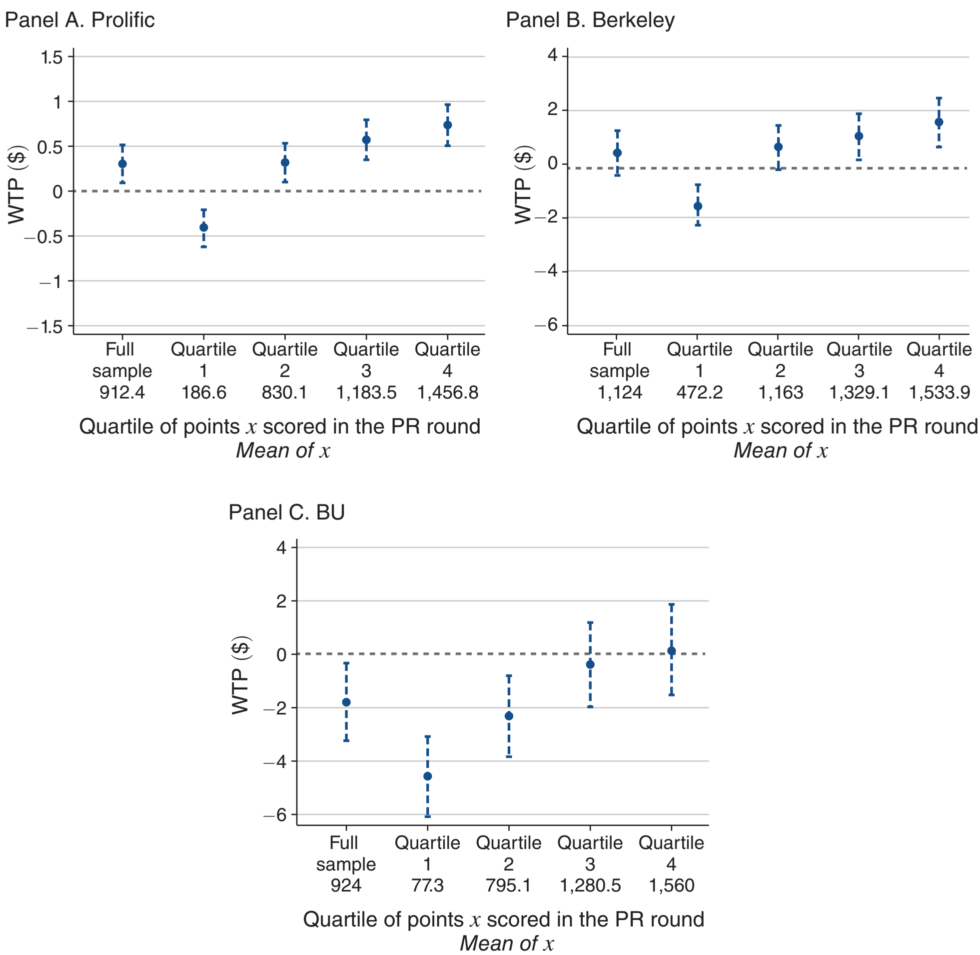

Figure 9. Image Payoffs in the Charitable Contribution Experiments

Notes: These figures plot the average realized image payoff of participants assigned to public recognition, for both the full sample and each quartile of actual attendance. The average points scored in the public recognition round are reported below each subsample label. Panel A presents results for the Prolific sample, panel B presents results for the Berkeley sample, and panel C presents results for the BU sample. The analysis excludes 40 Prolific participants, 11 Berkeley participants, and 2 BU participants with "incoherent" preferences for public recognition. The average realized image payoff is defined as the average WTP reported across all participants for the points interval corresponding to the participant's score in the public recognition round. Bootstrapped percentile-based confidence intervals, sampled by participants with 1,000 iterations, are displayed.

intrinsic motivation or with their score in the public recognition round, and thus using average WTP for a given interval increases statistical power without creating bias.

Figure 9 presents the results. The net image payoff of public recognition is positive in the Prolific sample, is statistically zero in the Berkeley sample, and is negative in the BU sample. The bottom quartile of participants experiences significantly negative payoffs in all three samples. In the Prolific and Berkeley samples, the top three quartiles of participants all experience positive payoffs, while in the BU sample no quartile of performers experiences positive payoffs.

Although there are many differences between the three samples, one key difference is the degree of familiarity among participants. Our results provide suggestive evidence that greater familiarity increases the prevalence of shame, which is consistent with hypotheses and results from psychological research (e.g., Tajfel 1970, Hogg 1992, Bicchieri et al. 2020).

E. *Consistency with Financial Incentive Effects*

Before turning to structural estimation, we provide back-of-the-envelope calculations to validate our money-metric approach to measuring the PRU. The fundamental assumption of our approach is that the effects of public recognition on behavior can be fully captured by the money-metric measures of the PRU in Table 6. For example, column 1 of the table implies that the motivating effects of public recognition are approximately equivalent to a financial incentive of 0.93 cents per 10 points in the Prolific sample. Thus, a key test of our approach is whether a financial incentive of 0.93 cents per 10 points indeed has a similar effect on behavior in the Prolific sample as does public recognition.

Simple calculations suggest remarkable consistency. In the Prolific sample, column 1 of Table 5 shows that public recognition increases performance by 105 points. A linear extrapolation thus implies that a 2 cents per 10 points incentive should increase performance by $105 \times (2/0.93) = 226$ points, which closely matches the 186-point effect estimated in column 1 of Table 5. Analogous arguments in our Table 6 estimates imply that the financial incentive should increase performance by 216 and 150 points in the Berkeley and BU samples, respectively. Empirically, Table 5 reveals only slightly smaller effect sizes of 178 and 118 points, respectively. Our structural estimates in the next section facilitate more formal tests of consistency.

VII. Structural Estimates

Our results thus far provide estimates of the reduced-form public recognition function R_{exp}. In this section, we build on the reduced-form results in three ways. First, we estimate parametric forms of the models presented in Section II. Second, we validate our experimental and structural methodology by more formally implementing the consistency tests from Section VIE. Third, we study the welfare effects of scaling up the public recognition intervention. Our main focus is on scaling up in the YMCA setting because it constitutes an important domain of behavior where there is significant interest in behavior change, and where social influence interventions such as ours are of potential interest. Online Appendix F contains the details of the structural models, their equilibrium predictions, and our approach to identifying these models.

A. *Estimation Methodology*

Functional Form Assumptions.—For tractability, we follow Bénabou and Tirole (2006) in assuming that in the absence of public recognition, people's material utility u is quadratic:

$$u(a;\theta) = \theta a - ca^2/2,$$

where $\theta \in \mathbb{R}^+$ is the intrinsic motivation, and ca is the marginal cost of increasing a. We also assume that the structural PRU in both the action-signaling and characteristics-signaling models in Section II is quadratic. Letting $\bar{a}$ denote the average action, and $\bar{\theta}$ denote the average type, we assume that

$$(5) \qquad \nu S^a(a - \rho\bar{a}) = \gamma_1^a(a - \rho\bar{a}) + \gamma_2^a(a - \rho\bar{a})^2,$$

$$(6) \qquad \nu S^\theta\left(E[\theta|a] - \rho\bar{\theta}\right) = \gamma_1^\theta\left(E[\theta|a] - \rho\bar{\theta}\right) + \gamma_2^\theta\left(E[\theta|a] - \rho\bar{\theta}\right)^2,$$

for the action-signaling and characteristics-signaling models, respectively.[30] As shown in online Appendix F, the resulting reduced-form PRU, $R(a)$, will be quadratic with both microfoundations.

To close the models, it is necessary to take a stand on the comparison sample that generates $\bar{a}$ and $\bar{\theta}$. In the YMCA setting, where participants were members far before the experimental period, and where they have the opportunity to observe and interact with many members outside of Grow and Thrive, the most natural assumption is that individuals care about how they are seen relative to the other YOTA members of their YMCA branch.[31] In our charitable contribution experiments, by contrast, participants did not have a previously established connection to the task, as the task was only introduced to them in the experiment. We thus assume that participants' comparison populations are simply those individuals who also completed the task—our experimental samples.[32]

Estimation.—Let $R_{exp}(a) = r_0 + r_1 a + r_2 a^2$ be the reduced-form PRU that is revealed by our WTP elicitation. We estimated this directly in column 4 of Table 4b for the YMCA sample, and in columns 2, 4, 6 of Table 6 for the Prolific, Berkeley, and BU samples.[33] As shown in online Appendix F, estimates of the structural parameters γ_i^j and ρ from the structural PRUs in (5) and (6) can be obtained as functions of the reduced-form parameters r_0, r_1, r_2.

Given estimates of R_{exp}, the treatment effect of public recognition on behavior identifies the cost parameter c. In the absence of public recognition, the marginal benefit of increasing a is θ, and the marginal cost of increasing a is ca. Thus, individuals choose $a^*(\theta) = \theta/c$, and average performance in the absence of public recognition is

$$(7) \qquad E[a|PR = 0] = E[\theta]/c.$$

[30] To ensure that S is increasing, we further assume that $a \in [0, a_M]$ and that $\gamma_1^j + 2\gamma_2^j a_M \geq 0$.

[31] Moreover, individuals had little reason to expect that participants in Grow and Thrive were different from other YMCA members since we only provided information about the broader base of YOTA members.

[32] An alternative benchmark might be the hypothetical performance of all Prolific, Berkeley Xlab, or BU Section QM222 members. This assumption is equivalent to ours if our experimental participants believed the participants in our experiment were representative of these larger pools.

[33] As discussed in the reduced-form results, the specification in column 4 of Table 4 for the YMCA sample addresses potential attenuation resulting from censoring, and from participants' relative insensitivity to variation of publicized attendance that they consider unlikely.

In the presence of public recognition, the marginal benefit of increasing a is $\theta + r_1 + 2r_2 a$. Thus, individuals choose $a^*(\theta) = (\theta + r_1)/(c - 2r_2)$, and average performance in the presence of public recognition is

$$(8) \qquad E[a|PR = 1] = E[\theta]/(c - 2r_2) + r_1/(c - 2r_2)$$

$$= E[a|PR = 0] \cdot c/(c - 2r_2) + r_1/(c - 2r_2).$$

Given an estimated average treatment effect $\bar{\tau}$ of public recognition on performance, the cost parameter c is identified by setting the difference between (8) and (7) equal to $\bar{\tau}$. We use the treatment effect estimates from column 5 of Table 2 for the YMCA sample, and estimates from columns 1–3 of Table 5 for the Prolific, Berkeley, and BU samples.

Consistency with Financial Incentive Effects.—The calculations above show that the structural models are identified using only data on the treatment effects of public recognition and participants' WTP for public recognition. The estimated models can then be used to make predictions about the effects of financial incentives on behavior, which can be compared to direct estimates from our data. In the presence of a constant marginal incentive of p and no public recognition, the marginal benefit of increasing a is $\theta + p$, and the marginal cost is ca. This implies that individuals choose $a^*(\theta) = (\theta + p)/c$, and thus that the financial incentive increases average performance by p/c.

For the charitable contribution experiments, we benchmark the model predictions against the effects of financial incentives estimated in Table 5. For the YMCA experiment, we were not able to randomize a purely financial incentive, but we did elicit participants' forecasts of how much they would attend the YMCA under three different scenarios: (i) if assigned to the Grow and Thrive control group, (ii) if assigned to the Grow and Thrive public recognition treatment group, (iii) if assigned to the Grow and Thrive control group but given a financial incentive of $1 per attendance. Although forecasted attendance may differ from actual attendance due to overoptimism, Carrera et al. (forthcoming) find that people accurately predict how their attendance will *vary* with incentives for attendance. Consistent with this, participants in our experiment predicted that public recognition would increase their attendance by 1.50 visits, which is similar to, and statistically indistinguishable from, our empirical estimate of 1.19 visits.

Note that the predictions about the effects of financial incentives on behavior in the experiment depend only on the reduced-form PRU R_{exp}, and thus are identical for both the action- and characteristics-signaling models.

Heterogeneity.—In online Appendix F.3, we generalize the model to include heterogeneity in individuals' cost of effort functions and PRUs, and show that our estimation approach is robust to this.

Uncertainty.—Suppose that at the time of the WTP elicitation, individuals are unsure about their type θ or the marginal costs, and that they learn this only after the elicitation when they choose their performance a. For example, individuals might be unsure about how motivated they will feel to work hard in the Click for Charity task, and only accurately learn that when they begin the task. This does not affect our analysis because of the strategy-method nature of our elicitation. All of our computations pertain to the signaling game that is played once individuals learn their type. This signaling game leads to the reduced-form PRU R, and our WTP elicitation exactly elicits $R(a)$ for each a. This robustness rests on the key feature of our design that WTP for public recognition is elicited in a performance-contingent fashion.

B. *Estimation Results*

Table 7 presents the structural estimation results. Panel A presents estimates of the action-signaling model and panel B presents estimates of the characteristics-signaling model. Panel C presents results on consistency with the effects of financial incentives.

Although the model parameters γ_i^j in panels A and B are in different units and thus have different magnitudes, the two panels deliver a similar message, which is consistent with the reduced-form results. First, there is significant concavity of the structural PRU in the YMCA and Prolific samples, although the curvature estimates are more ambiguous in the Berkeley and BU samples. The concavity is particularly pronounced in the characteristics-signaling model in the Prolific sample. Second, the standard at which negative image payoffs transition to positive image payoffs varies across the samples. In the YMCA sample, ρ is above one in both models, although we cannot reject the hypothesis that participants simply care about the average $(\rho = 1)$. In the Berkeley sample, we estimate ρ close to one in both models. In the Prolific sample, we estimate ρ significantly below one in both models, indicating a lower standard for prideworthy behavior. In the BU sample we estimate ρ substantially above one, indicating a high standard for prideworthy behavior.

Panel C shows that in all four samples, the models' predictions about the effects of financial incentives closely match the directly estimated effects. On net, we find slight overestimation, although the last column in panel C shows that this overestimation is not statistically distinguishable from zero at conventional levels. Moreover, the slight overestimation could be explained by a number of realistic features not incorporated into our intentionally parsimonious models.[34]

C. *Welfare Effects of Scaling up Public Recognition*

We now use our structural estimates to assess the average image utility generated by public recognition. Motivated by our results on group size effects in the Prolific sample, we assume that increasing the number of exposed individuals would not change the visibility parameter ν.

Under the assumption that our Prolific, BU, and Berkeley samples are representative of those respective populations, and that individuals in those samples construct the reference point from how the samples performed in the public recognition

[34] For example, our quadratic cost of effort function implies a unit elasticity and thus that behavior is linear in the magnitude of incentives. This assumption would cause us to overestimate the effects of financial incentives if instead behavior were a concave function of financial incentives, as would be the case for isoelastic cost functions with elasticities below one. Various forms of correlated heterogeneity could explain the underestimation as well.

TABLE 7—STRUCTURAL ESTIMATES AND TESTS OF CONSISTENCY

Panel A. Action-signaling model parameter estimates				
Sample	$\hat{\gamma}_1^a$	$\hat{\gamma}_2^a$	$\hat{\rho}^a$	$\hat{c}$
YMCA	0.64	−0.020	1.85	0.46
	[0.35,0.92]	[−0.038,−0.003]	[0.94,2.52]	[0.20,1.95]
Prolific	0.12	−0.004	0.58	0.08
	[0.09,0.14]	[−0.005,−0.002]	[0.40,0.80]	[0.06,0.11]
Berkeley	0.30	−0.004	0.87	0.21
	[0.22,0.38]	[−0.011,0.003]	[0.63,1.15]	[0.14,0.33]
BU	0.38	0.002	1.61	0.34
	[0.19,0.53]	[−0.009,0.013]	[1.14,2.33]	[0.16,1.47]
Panel B. Characteristics-signaling model parameter estimates				
Sample	$\hat{\gamma}_1^\theta$	$\hat{\gamma}_2^\theta$	$\hat{\rho}^\theta$	$\hat{c}$
YMCA	1.28	−0.079	1.40	0.46
	[0.28,2.34]	[−0.292,−0.003]	[0.31,2.19]	[0.20,1.95]
Prolific	1.30	−0.458	0.49	0.08
	[0.98,1.65]	[−0.765,−0.241]	[0.25,0.76]	[0.06,0.11]
Berkeley	1.35	−0.082	0.85	0.21
	[0.90,1.81]	[−0.330,0.046]	[0.55,1.17]	[0.14,0.33]
BU	1.13	0.021	1.68	0.34
	[0.11,2.40]	[−0.101,0.252]	[1.16,2.40]	[0.16,1.47]

Panel C. Predicted and actual effects of financial incentives (on attendance or points (00s))			
Sample	Model prediction	Actual	Pred.− Act.
YMCA	2.16	1.77[a]	0.39
	[0.42,5.00]	[1.30,2.26]	[−1.27,3.19]
Prolific	2.41	1.82	0.60
	[1.81,3.12]	[1.56,2.07]	[0.07,1.25]
Berkeley	2.33	1.78	0.55
	[1.49,3.53]	[1.35,2.24]	[−0.18,1.63]
BU	1.48	1.18	0.29
	[0.21,2.89]	[0.47,1.94]	[−0.89,1.56]

Notes: These tables report parameter estimates of the action-signaling and characteristics-signaling models described in Section VIIA, equations (5) and (6). For panel C, the financial incentive is $1 per attendance for the YMCA sample, 2 cents per 10 points for the Prolific sample, and 5 cents per 10 points for the Berkeley and BU samples. The analysis excludes participants with "incoherent" preferences for public recognition (15 in YMCA participants, 40 Prolific participants, 11 Berkeley participants, and 2 BU participants). Bootstrapped percentile-based confidence intervals from 1,000 replications, clustered at the participant level, are reported in brackets.

[a]Based on individuals' forecasted rather than realized behavior.

round, the welfare effects are immediately given by our reduced-form results in Section VI, and are summarized in Table 8.

For the YMCA sample, however, the natural assumption (discussed above) is that individuals evaluate their performance relative to the performance of all members of YOTA. This implies that our reduced-form estimates of welfare effects are only partial equilibrium, and necessitates the use of our structural model. This need is particularly pronounced because the YMCA sample is not representative of the broader YOTA population.

We present the results in Table 9. Column 1 shows the net image payoffs and column 2 presents the predicted change in behavior. Panel A presents results from the action-signaling model and panel B presents results from the characteristics-signaling model. Except in several special cases, these models have somewhat different

Table 8—Welfare Estimates of Public Recognition in the Charitable Contribution Experiments

Row	Sample	Image payoffs (1)	Change in points scored (2)
1.	Prolific	0.30	13.00%
2.	Berkeley	0.42	13.58%
3.	BU	−1.80	12.70%

Notes: This table reports the average realized image payoff of participants assigned to public recognition. The analysis excludes 40 Prolific participants, 11 Berkeley participants, and 2 BU participants with "incoherent" preferences for public recognition. The average realized image payoff is defined as the average WTP reported across all participants for the points interval corresponding to the participant's score in the public recognition round. The estimates in column 1 match the "full sample" estimates reported in Figure 9. Column 2 reports the change in points scored from public recognition as a percentage of the average points scored in the anonymous round, which are 808, 990, and 816 for the Prolific, Berkeley, and BU samples, respectively.

Table 9—Welfare Estimates of Scaling Up Public Recognition at the YMCA

Panel A. Action-signaling model

Row	Scenario	Parameter estimates			Image payoffs	Change in attendance
		γ_1^a	γ_2^a	ρ^a	(1)	(2)
1.	Baseline (YMCA)	0.64	−0.020	1.85	−3.41	55.77%
2.	ρ from Prolific sample	0.64	−0.020	0.58	0.70	39.31%
3.	ρ from Berkeley sample	0.64	−0.020	0.87	−0.04	42.73%
4.	ρ from BU sample	0.64	−0.020	1.61	−2.46	52.41%
5.	Curvature from Prolific sample	0.64	−0.022	1.85	−3.51	57.06%
6.	Curvature from Berkeley sample	0.64	−0.010	1.85	−2.92	49.53%
7.	Curvature from BU sample	0.64	0.005	1.85	−2.25	41.93%
8.	ρ and curvature from Prolific sample	0.64	−0.022	0.58	0.66	38.81%
9.	ρ and curvature from Berkeley sample	0.64	−0.010	0.87	0.16	43.57%
10.	ρ and curvature from BU sample	0.64	0.005	1.61	−1.61	42.59%

Panel B. Characteristics-signaling model

Row	Scenario	Parameter estimates			Image payoffs	Change in attendance
		γ_1^θ	γ_2^θ	ρ^θ	(1)	(2)
1.	Baseline (YMCA)	1.28	−0.079	1.40	−1.18	47.55%
2.	ρ from Prolific sample	1.28	−0.079	0.49	0.51	40.23%
3.	ρ from Berkeley sample	1.28	−0.079	0.85	−0.12	43.12%
4.	ρ from BU sample	1.28	−0.079	1.68	−1.74	49.75%
5.	Curvature from Prolific sample	1.28	−0.077	1.40	−1.17	47.36%
6.	Curvature from Berkeley sample	1.28	−0.060	1.40	−1.07	45.72%
7.	Curvature from BU sample	1.28	0.022	1.40	−0.63	39.17%
8.	ρ and curvature from Prolific sample	1.28	−0.077	0.49	0.52	40.24%
9.	ρ and curvature from Berkeley sample	1.28	−0.060	0.85	−0.02	42.45%
10.	ρ and curvature from BU sample	1.28	0.022	1.68	−1.13	38.61%

Notes: These tables report welfare estimates based on the structural estimates of the action-signaling and characteristics-signaling models described in Section VIIA. In row 1, γ_1^j, γ_2^j, and ρ^j are set equal to the parameter estimates from the YMCA sample. In rows 2–4, ρ is set equal to the ρ estimated for the Prolific, Berkeley, and BU samples, respectively. In rows 5–7, γ_2^j is set to match the curvature—defined as $\gamma_2^j \cdot SD^j/\gamma_1^j$ for any sample j (where SD denotes standard deviation)—from the Prolific, Berkeley, and BU samples, respectively; that is, it is equal to the product of $\gamma_1^{YMCA}/SD^{YMCA}$ and $\gamma_2^j \cdot SD^s/\gamma_1^j$ from the Prolific, Berkeley, and BU samples, respectively. In panel A, SD^{YMCA} is the standard deviation of attendance for the general YOTA population, and SD^j for the online experiments is the standard deviation of points scored in the anonymous round (in units of hundreds of points). In panel B, SD^{YMCA} and SD^j are the standard deviation of types in the YMCA experiment and online experiments, respectively. Rows 8–10 repeat rows 5–7 with ρ^j set equal to ρ^j from the Prolific, Berkeley, and BU samples, respectively. Column 2 reports the change in attendance from scaling up public recognition as a percentage of the average YOTA attendance, 3.14.

equilibrium implications for behavior and welfare, illustrating the importance of working out the consequences of microfounded models.

We explore the welfare effects across a range of different structural assumptions. Row 1 in both panels considers the baseline estimates for the YMCA sample. Rows 2–4 explore the importance of varying ρ by considering the point estimates from the Prolific, Berkeley, and BU samples. Rows 5–7 consider the importance of varying curvature by using the point estimates from the Prolific, Berkeley and BU samples. Rows 8–10 jointly set ρ and curvature equal to the point estimates from the Prolific, Berkeley, and BU samples.

The table reveals two main insights. First, the average image utility from scaling up public recognition to the full YOTA population is predicted to be substantially negative, particularly in the action-signaling model.

Second, as rows 2–10 illustrate, variation in the reference point parameter ρ has a larger effect on net images payoffs than variation in curvature. Decreasing ρ to the Berkeley sample estimate, while holding curvature fixed at the YMCA estimate, results in a net image payoff near zero. Further reducing ρ to the estimate in the Prolific sample results in a positive image payoff. However, rows 5–7 show that holding ρ constant at the YMCA estimate and varying the curvature to match the estimates in the online samples always results in negative image payoffs. The welfare estimates in rows 8–10 are much closer to those in rows 2–4 than in rows 5–7. This implies that the large variation in social image payoffs between all four of our samples is largely due to variation in the ρ estimate.[35]

In online Appendix B we formalize how the estimates of image utility in this section can be combined with several other statistics to determine whether public recognition or financial incentives are a more efficient means of changing behavior.

VIII. Concluding Remarks

A recent and growing literature establishes that public recognition can meaningfully influence behavior in a number of economically consequential field settings. We build on this literature by developing an empirical methodology for directly quantifying individuals' utility from public recognition. Across two different experimental designs and four different samples, we find that image payoffs from public recognition are significant and highly unequal: some experience significantly negative payoffs, consistent with shame, while other experience significantly positive payoffs, consistent with pride. In the YMCA setting, our results suggest that motivating exercise with public recognition might be less socially efficient than utilizing financial incentives. Our work illustrates how the social costs or benefits of public recognition can be substantial, and provides a framework for measurement and welfare analysis.

[35] Note that the impacts on behavior, in percentage terms, are predicted to be larger in rows 8–10 than in Table 8 because the distribution of individuals' types is different. In particular, the effects will be proportionally larger for more left-skewed distributions.

Of course, our results come with many caveats and leave open many research questions. First, our methods quantify only the direct effects of public recognition on utility, and are not designed to measure other key inputs for a holistic welfare analysis. Online Appendix B provides a formal framework for welfare analysis, and in particular for answering whether another policy lever, such as financial incentives, might be more efficient in creating the same behavior change.[36]

Second, while our methodology is easily imported into many of the domains where researchers have studied the effects of public recognition on behavior, our specific results constitute only an initial set of data points on the welfare effects of public recognition. Consequently, extrapolation to other populations or domains of behavior must be done with caution. Indeed, while our results suggest that the effects of public recognition are invariant to some factors such as group size, our estimates appear to be less stable with respect to other factors such as individuals' familiarity with each other.

Third, even within the specific contexts of our experiments, our *quantitative* welfare estimates cannot be immediately applied to public recognition schemes that produce different information structures such as ones that recognize only the top performers. Although standard economic models imply that coarsening the information structure cannot eliminate feelings of shame if such feelings are prevalent in fully revealing schemes (see online Appendix A), and although our estimates of structural models can be used to generate predictions about these alternative schemes, limited attention or failures of equilibrium thinking could weaken the predictive power of standard economic models. Our flexible online experimental protocol can be augmented to further study how the effects of public recognition vary with the signal structure.

More generally, we suggest that our online protocol can be fruitfully extended to facilitate further testing and refinement of social signaling models. Empirical tests of social signaling models typically revolve around comparative statics on behavior, although underlying these comparative statics are predictions about individuals' social image payoffs. By providing a direct estimate of social image payoffs, our methodology can thus enable more direct tests of phenomena such as the overjustification effect and motivation crowding (Gneezy and Rustichini 2000, Bénabou and Tirole 2006), predictions about the effects of social information on prosocial behavior (Bénabou and Tirole 2011), or the evolution of stigma and redistributive norms (Alesina and Angeletos 2005).

With some extension, our approach could also be applied more broadly to study other social influence levers. Although such nonfinancial policy instruments have become popular tools in governments around the world under the banner of "nudge" (OECD 2017), most existing studies focus on how these instruments affect behavior, and have little to say about *welfare* (see, e.g., Bernheim and Taubinsky 2018, for a review).[37] We view this as a limitation of existing research methods, not a reflection of

[36] We note that while financial incentives motivate desirable behavior and have little interaction with public recognition in our domains, there are also important cases where financial incentives could crowd out motivation because they dampen the effects of both shame and pride (e.g., Bénabou and Tirole 2006, Ariely, Bracha, and Meier 2009).

[37] See, e.g., Bénabou and Tirole (2006), Bénabou and Tirole (2011), and Ali and Bénabou (2020) for examples of welfare analysis with social image.

actual social goals. Indeed, in the case of social influence, an honest assessment of the psychological, political, philosophical, and literary studies of human motivation reveals that people's well-being is intensely sensitive to the experience of shame and pride.

REFERENCES

Acland, Dan, and Matthew R. Levy. 2015. "Naiveté, Projection Bias, and Habit Formation in Gym Attendance." *Management Science* 61 (1): 146–60.

Alesina, Alberto, and George-Marios Angeletos. 2005. "Fairness and Redistribution." *American Economic Review* 95 (4): 960–80.

Ali, S. Nageeb, and Roland Bénabou. 2020. "Image versus Information: Changing Societal Norms and Optimal Privacy." *American Economic Journal: Microeconomics* 12 (3): 116–64.

Allcott, Hunt, and Judd B. Kessler. 2019. "The Welfare Effects of Nudges: A Case Study of Energy Use Social Comparisons." *American Economic Journal: Applied Economics* 11 (1): 236–76.

Allcott, Hunt, Joshua Kim, Dmitry Taubinsky, and Jonathan Zinman. Forthcoming. "Are High Interest Loans Predatory? Theory and Evidence from Payday Lending." *Review of Economic Studies*.

Andreoni, James, and B. Douglas Bernheim. 2009. "Social Image and the 50–50 Norm: A Theoretical and Experimental Analysis of Audience Effects." *Econometrica* 77 (5): 1607–36.

Andreoni, James, and Ragan Petrie. 2004. "Public Goods Experiments without Confidentiality: A Glimpse into Fund-Raising." *Journal of Public Economics* 88 (7): 1605–23.

Ariely, Dan, Anat Bracha, and Stephan Meier. 2009. "Doing Good or Doing Well? Image Motivation and Monetary Incentives in Behaving Prosocially." *American Economic Review* 99 (1): 544–55.

Ashraf, Nava, Oriana Bandiera, and B. Kelsey Jack. 2014. "No Margin, No Mission? A Field Experiment on Incentives for Public Service Delivery." *Journal of Public Economics* 120: 1–17.

Augenblick, Ned. 2018. "Short-Term Discounting of Unpleasant Tasks." Unpublished.

Becker, Gary S. 1991. "A Note on Restaurant Pricing and Other Examples of Social Influences on Price." *Journal of Political Economy* 99 (5): 1109–16.

Becker, Gary S. 2000. "Crime and Punishment: An Economic Approach." *Journal of Political Economy* 76 (2): 169–217.

Bénabou, Roland, and Jean Tirole. 2006. "Incentives and Prosocial Behavior." *American Economic Review* 96 (5): 1652–78.

Bénabou, Roland, and Jean Tirole. 2011. "Laws and Norms." NBER Working Paper 17579.

Benartzi, Shlomo, John Beshears, Katherine L. Milkman, Cass R. Sunstein, Richard H. Thaler, Maya Shankar, Will Tucker-Ray, William J. Congdon, and Steven Galing. 2017. "Should Government Invest More in Nudging?" *Psychological Science* 28 (8): 1041–55.

Bernheim, B. Douglas, and Christine L. Exley. 2015. "Understanding Conformity: An Experimental Investigation." Harvard Business School Working Paper 6-070.

Bernheim, B. Douglas, Andrey Fradkin, and Igor Popov. 2015. "The Welfare Economics of Default Options in 401(k) Plans." *American Economic Review* 105 (9): 2798–837.

Bernheim, B. Douglas, and Dmitry Taubinsky. 2018. "Behavioral Public Economics." In *Handbook of Behavioral Economics*, Vol 1, edited by Douglas Bernheim, Stefano DellaVigna, and David Laibson, 381–516. Amsterdam: Elsevier.

Besley, Timothy, and Stephen Coate. 1992. "Workfare versus Welfare: Incentive Arguments for Work Requirements in Poverty-Alleviation Programs." *American Economic Review* 82 (1): 249–61.

Bicchieri, Cristina, Eugen Dimant, Simon Gächter, and Daniele Nosenzo. 2020. "Observability, Social Proximity, and the Erosion of Norm Compliance." Unpublished.

Birke, David J. 2020. "Anti-bunching: A New Test for Signaling Motives in Prosocial Behavior." Unpublished.

Blomquist, N. Soren. 1993. "Interdependent Behavior and the Effect of Taxes." *Journal of Public Economics* 51 (2): 211–18.

Bloom, Nicholas, and John Van Reenen. 2007. "Measuring and Explaining Management Practices across Firms and Countries." *Quarterly Journal of Economics* 122 (4): 1351–408.

Bø, Erlend E., Joel Slemrod, and Thor O. Thoresen. 2015. "Taxes on the Internet: Deterrence Effects of Public Disclosure." *American Economic Journal: Economic Policy* 7 (1): 36–62.

Bradler, Christiane, Robert Dur, Susanne Neckermann, and Arjan Non. 2016. "Employee Recognition and Performance: A Field Experiment." *Management Science* 62 (11): 3085–99.

Bursztyn, Leonardo, Georgy Egorov, and Robert Jensen. 2019. "Cool to Be Smart or Smart to Be Cool? Understanding Peer Pressure in Education." *The Review of Economic Studies* 86 (4): 1487–526.

Bursztyn, Leonardo, Bruno Ferman, Stefano Fiorin, Martin Kanz, and Gautam Rao. 2017. "Status Goods: Experimental Evidence from Platinum Credit Cards." *Quarterly Journal of Economics* 133 (3): 1561–95.

Bursztyn, Leonardo, Thomas Fujiwara, and Amanda Pallais. 2017. "'Acting Wife': Marriage Market Incentives and Labor Market Investments." *American Economic Review* 107 (11): 3288–319.

Bursztyn, Leonardo, and Robert Jensen. 2015. "How Does Peer Pressure Affect Educational Investments?." *Quarterly Journal of Economics* 130 (3): 1329–67.

Bursztyn, Leonardo, and Robert Jensen. 2017. "Social Image and Economic Behavior in the Field: Identifying, Understanding, and Shaping Social Pressure." *Annual Review of Economics* 9 (1): 131–53.

Butera, Luigi, Robert Metcalfe, William Morrison, and Dmitry Taubinsky. 2022. "Replication Data for: Measuring the Welfare Effects of Shame and Pride." American Economic Association [publisher], Inter-university Consortium for Political and Social Research [distributor]. https://doi.org/10.3886/E145141V1.

Carrera, Mariana, Heather Royer, Mark Stehr, Justin Sydnor, and Dmitry Taubinsky. Forthcoming. "Who Chooses Commitment? Evidence and Welfare Implications." *Review of Economic Studies*.

Carroll, Gabriel D., James J. Choi, David Laibson, Brigitte C. Madrian, and Andrew Metrick. 2009. "Optimal Defaults and Active Decisions." *Quarterly Journal of Economics* 124 (4): 1639–74.

Cederström, Carl, and André Spicer. 2015. *The Wellness Syndrome*. Hoboken, NJ: John Wiley & Sons.

Cohen, Adam Scott, Rie Chun, and Daniel Sznycer. 2020. "Do Pride and Shame Track the Evaluative Psychology of Audiences? Preregistered Replications of Sznycer *et al.* (2016, 2017)." *Royal Society Open Science* 7 (5): 191922.

Conrad, Peter. 1994. "Wellness as Virtue: Morality and the Pursuit of Health." *Culture, medicine and psychiatry* 18 (3): 385–401.

Corneo, Giacomo G. 1997. "The Theory of the Open Shop Trade Union Reconsidered." *Labour Economics* 4 (1): 71–84.

Damgaard, Mette Trier, and Christina Gravert. 2018. "The Hidden Costs of Nudging: Experimental Evidence from Reminders in Fundraising." *Journal of Public Economics* 157: 15–26.

de Quidt, Jonathan, Johannes Haushofer, and Christopher Roth. 2018. "Measuring and Bounding Experimenter Demand." *American Economic Review* 108 (11): 3266–302.

DellaVigna, Stefano. 2018. "Structual Behavioral Economics." In *Handbook of Behavioral* Economics, Vol 1, edited by B. Douglas Bernheim, Stefano DellaVigna, and David Laibson, 613–717. Amsterdam: Elsevier.

DellaVigna, Stefano, John A. List, and Ulrike Malmendier. 2012. "Testing for Altruism and Social Pressure in Charitable Giving." *Quarterly Journal of Economics* 127 (1): 1–56.

DellaVigna, Stefano, John A. List, Ulrike Malmendier, and Gautam Rao. 2017. "Voting to Tell Others." *Review of Economic Studies* 84 (1): 143–81.

DellaVigna, Stefano, John A. List, Ulrike Malmendier, and Gautam Rao. 2019. "Estimating Social Preferences and Gift Exchange with a Piece-Rate Design." CEPR Discussion Paper 24931.

DellaVigna, Stefano, and Ulrike Malmendier. 2006. "Paying Not to Go to the Gym." *American Economic Review* 96 (3): 694–719.

DellaVigna, Stefano, and Devin G. Pope. 2018. "What Motivates Worker Effort? Evidence and Expert Forecasts." *Review of Economic Studies* 85(2): 1029–69.

DellaVigna, Stefano, and Devin Pope. 2019. "Stability of Experimental Results." NBER Working Paper 25858.

Ehrlich, Isaac. 1973. "Participation in Illegitimate Activities: A Theoretical and Empirical Investigation." *Journal of Political Economy* 81 (3): 521–65.

Etzioni, Amitai. 1999. "Back to the Pillory?" *American Scholar* 68 (3): 43–50.

Exley, Christine. 2018. "Incentives for Prosocial Behavior: The Role of Reputations." *Management Science* 64 (5): 2460–71.

Gauri, Varun, Julian C. Jamison, Nina Mazar, Owen Ozier, Shomikho Raha, and Karima Saleh. 2018. "Motivating Bureaucrats through Social Recognition: Evidence from Simultaneous Field Experiments." World Bank Policy Research Working Paper 8473.

Gerber, Alan S., Donald P. Green, and Christopher W. Larimer. 2008. "Social Pressure and Voter Turnout: Evidence from a Large-Scale Field Experiment." *American Political Science Review* 102 (1): 33–48.

Gneezy, Uri, and Aldo Rustichini. 2000. "Pay Enough or Don't Pay at All." *Quarterly Journal of Economics* 115 (3): 791–810.

Halpern, David. 2015. *Inside the Nudge Unit: How Small Changes Can Make a Big Difference*. London: Penguin Random House UK.

Hogg, Michael A. 1992. *The Social Psychology of Group Cohesiveness: From Attraction to Social Identity*. New York, London: Harvester Wheastsheaf.

Jones, Daniel, and Sera Linardi. 2014. "Wallflowers: Experimental Evidence of an Aversion to Standing Out." *Management Science* 60 (7): 1757–71.

Kahan, Dan M., and Eric A. Posner. 1999. "Shaming White-Collar Criminals: A Proposal for Reform of the Federal Sentencing Guidelines." *Journal of Law and Economics* 42 (1): 365–91.

Kaplow, Louis, and Steven Shavell. 2007. "Moral Rules, the Moral Sentiments, and Behavior: Toward a Theory of an Optimal Moral System." *Journal of Political Economy* 115 (3): 494–514.

Karing, Anne. 2019. "Social Signaling and Childhood Immunization: A Field Experiment in Sierra Leone." Unpublished.

Kosfeld, Michael, and Susanne Neckermann. 2011. "Getting More Work for Nothing? Symbolic Awards and Worker Performance." *American Economic Journal: Microeconomics* 3 (3): 86–99.

Kosfeld, Michael, Susanne Neckermann, and Xiaolan Yang. 2017. "The Effects of Financial and Recognition Incentives across Work Contexts: The Role of Meaning." *Economic Inquiry* 55 (1): 237–47.

Lacetera, Nicola, and Mario Macis. 2010. "Social Image Concerns and Prosocial Behavior: Field Evidence from a Nonlinear Incentive Scheme." *Journal of Economic Behavior and Organization* 76 (2): 225–37.

Leary, Mark R. 2007. "Motivational and Emotional Aspects of the Self." *Annual Review of Psychology* 58: 317–44.

Lindbeck, Assar, Sten Nyberg, and Jorgen W. Weibull. 1999. "Social Norms and Economic Incentives in the Welfare State." *Quarterly Journal of Economics* 114 (1): 1–35.

Lindbeck, Assar, Sten Nyberg, and Jorgen W. Weibull. 2003. "Social Norms and Welfare State Dynamics." *Journal of the European Economic Association* 1 (2–3): 533–42.

Loewenstein, George, Cass R. Sunstein, and Russell Golman. 2014. "Disclosure: Psychology Changes Everything." *Annual Review of Economics* 6: 391–419.

Mailath, George J. 1987. "Incentive Compatibility in Signaling Games with a Continuum of Types." *Econometrica* 55 (6): 1349–65.

Massaro, Toni. 1991. "Shame, Culture, and American Criminal Law." *Michigan Law Review* 89 (7): 1880–944.

Neckermann, Susanne, Reto Cueni, and Bruno S. Frey. 2014. "Awards at Work." *Labour Economics* 31: 205–17.

Neckermann, Susanne, and Xiaolan Yang. 2017. "Understanding the (Unexpected) Consequences of Unexpected Recognition." *Journal of Economic Behavior and Organization* 135: 131–42.

Nussbaum, Martha C. 2009. *Hiding from Humanity: Disgust, Shame, and the Law*. Princeton: Princeton University Press.

Organisation for Economic Co-operation and Development (OECD). 2017. *Behavioural Insights and Public Policy: Lessons from Around the World*. Paris: OECD Publishing.

Perez-Truglia, Ricardo, and Guillermo Cruces. 2017. "Partisan Interactions: Evidence from a Field Experiment in the United States." *Journal of Political Economy* 125 (4): 1208–43.

Perez-Truglia, Ricardo, and Ugo Troiano. 2018. "Shaming Tax Delinquents." *Journal of Public Economics* 167: 120–37.

Rege, Mari, and Kjetil Telle. 2004. "The Impact of Social Approval and Framing on Cooperation in Public Good Situations." *Journal of Public Economics* 88 (7–8): 1625–44.

Soetevent, Adriaan R. 2005. "Anonymity in Giving in a Natural Context: A Field Experiment in 30 Churches." *Journal of Public Economics* 89 (11): 2301–23.

Soetevent, Adriaan R. 2011. "Payment Choice, Image Motivation and Contributions to Charity: Evidence from a Field Experiment." *American Economic Journal: Economic Policy* 3 (1): 180–205.

Sznycer, Daniel, John Tooby, Leda Cosmides, Roni Porat, Shaul Shalvi, and Eran Halperin. 2016. "Shame Closely Tracks the Threat of Devaluation by Others, Even across Cultures." *Proceedings of the National Academy of Sciences* 113 (10): 2625–30.

Sznycer, Daniel, Laith Al-Shawaf, Yoella Bereby-Meyer, Oliver Scott Curry, Delphine De Smet, Elsa Ermer, Sangin Kim, et al. 2017. "Cross-Cultural Regularities in the Cognitive Architecture of Pride." *Proceedings of the National Academy of Sciences* 114 (8): 1874–79.

Tajfel, Henri. 1970. "Experiments in Intergroup Discrimination." *Scientific American* 223 (5): 96–103.

Tangney, June Price, Rowland S. Miller, Laura Flicker, and Deborah Hill Barlow. 1996. "Are Shame, Guilt, and Embarrassment Distinct Emotions?" *Journal of Personality and Social Psychology* 70 (6): 1256–69.

Tangney, June Price, Jeff Stuewig, and Debra J. Mashek. 2007. "Moral Emotions and Moral Behavior." *Annual Review of Psychology* 58: 345–72.

Thunstrom, Linda. 2019. "Welfare Effects of Nudges: The Emotional Tax of Calorie Menu Labeling." *Judgment and Decision Making* 14 (1): 11–25.

Whorton, James C. 2014. *Crusaders for Fitness: The History of American Health Reformers*. Princeton: Princeton University Press.

WorldatWork. 2017. *Trends in Employee Recognition.* Scottsdale, AZ: WorldatWork.

Yaouanq, Yves Le, and Peter Schwardmann. 2019. "Learning About One's Self." CESifo Working Paper 7455.

Yoeli, Erez, Moshe Hoffman, David G. Rand, and Martin A. Nowak. 2013. "Powering Up with Indirect Reciprocity in a Large-Scale Field Experiment." *Proceedings of the National Academy of Sciences* 110 (2): 10424–29.