Measuring image concern

Emeric Henry[*]

Sciences Po and CEPR

emeric.henry@sciencespo.fr

28, rue des Saints-Pères

75007 Paris, France

Jan Sonntag

Sciences Po

jan.sonntag@sciencespo.fr

28, rue des Saints-Pères

75007 Paris, France

August 10, 2017

Abstract

It is now well documented that individuals tend to change their behavior when their actions are observed by others. Yet there is no systematic way of measuring this dimension of preferences at the individual level. In this paper, we propose a novel experimental game to measure individual sensitivity to social image. We document substantial heterogeneity in the level of image concern and assess its drivers and consequences. While few socioeconomic characteristics seem to have explanatory power, members of ethnic minorities appear less sensitive to being observed by a another member of a minority group. Finally, we show that more image concerned individuals tend to be more selfish and find evidence suggesting that they try to avoid situations where their actions might be observed.

JEL Classification: D03, D64

Keywords: image concern, experimental measurement, repeated prisoner's dilemma

---

[*]Corresponding author

1 Introduction

Individuals behave differently when their choices and actions can be observed by others. This fact is now well documented empirically (Ariely et al. (2009), Andreoni and Petrie (2004), Bursztyn and Jensen (2015)) and some important theoretical implications have been drawn (see for instance Bénabou and Tirole (2006, 2012), Ellingsen and Johannesson (2008) or Andreoni and Bernheim (2009)). Yet little is known about the drivers or the consequences of image concern. One of the main reasons for this gap in the literature is that there is currently no systematic way of measuring individual sensitivity to perceptions by strangers.[1]

The first goal of this paper is to propose an experimental game designed to measure image concern at the individual level. A key feature of its design is its ability to identify image concern separately from other social preferences. The second goal of the paper is to use this game to study the socioeconomic determinants of sensitivity to image concern, to assess whether the observer's characteristics matter and to examine whether image concern is linked to other social preferences.

The *image concern game* we propose involves three players: a dictator (he), a recipient and an observer (she). The dictator determines how much money to transfer to a lottery with two possible outcomes: success, in which case the recipient receives a given amount of money, or failure, in which case the recipient receives nothing. The more money the dictator transfers, the higher the chances of success. The dictator takes his decision knowing that the observer will be informed of the outcome of the lottery. Before the lottery is actually run, the dictator has to reveal his willingness to pay to remain anonymous (in an incentive compatible way), i.e. for his picture not to be revealed to the observer in case the lottery is a failure. The recipient never sees any pictures. The observer sees only the outcome of the lottery, not the amount the dictator actually transferred.

There are two main aspects that drive the structure of this game. First, image concern is easily measured by the willingness to pay to remain anonymous in case the recipient remains empty-handed. Second, and most importantly, if some reasonable properties of the utility function are satisfied, this measurement proves independent of other social preferences including altruism. In case the dictator does not remain anonymous, the observer does not find out how much was contributed to the lottery, only that the lottery was a failure. Thus, the inference the observer makes when she sees the picture is an updated belief on the characteristics of the dictator conditional on the fact that the lottery was a failure, and this belief cannot be conditioned on the actual amount transferred. Separating our measure from other social preferences is essential to understand the specific drivers of image concern and to show how it correlates with these other dimensions of preferences.

The game is sufficiently portable to be used in future lab or lab-in-the-field experiments to yield a measure of image concern that can be correlated with other experimental outcomes. We made sure that the game did not require complicated repeated interactions and could even be

---

[1]Heterogeneity in image concern needs to be measured to understand how it affects behavior. It is also an important element in theoretical models such as Ali and Bénabou (2016).

run without the different parties being present at the same time, as long as the authenticity of the participants' photos could be ensured. It is, however, less portable than other games aimed at measuring social preferences such as the trust game or the dictator game. As any setup aimed at measuring image concern will require a mechanism to vary the degree of anonymity, for instance by using pictures, this seems inevitable. We will discuss this aspect further in our conclusion.

Running this game in the lab, we find substantial heterogeneity: about one third of the participants chooses not to pay anything, while one third gives even large amounts to remain anonymous. We show that no socioeconomic variable has explanatory power. Similarly, few characteristics of the observer significantly impact the willingness to pay to remain anonymous. This is encouraging evidence of the portability of the setup.[2] Nationality seems to be the exception: *Non-French individuals pay significantly less for anonymity when facing other non-French observers and slightly more when observed by French observers*, a fact linking nicely to the literature on discrimination.[3] One possible interpretation is that non-French participants fear that due to prejudice, French observers will interpret a failed outcome of the lottery more adversely than non-French observers.

We validate our measure of image concern in three different ways. First, we show that it significantly correlates with a survey question administered at the end of the experiment.[4] Second, we use a simple model to derive an implication of image concern and show that it stands in the data: more image concerned individuals transfer more in the first lottery to avoid situations where they would have to pay for anonymity.[5] Finally we show that in an infinitely repeated prisoner's dilemma game, more image concerned individuals adapt their behavior to the social norm much more than others when playing in presence of observers.

This last element of validation leverages an infinitely repeated prisoner's dilemma game, which participants played in the second phase of the experiment. In half of the sessions the game was played with observers, in the others without. We asked observers to rate the behavior of those they observe after each round so as to document what actions are judged positively by the community and identify the prevalent social norm. Using the repeated games, we first show that more image concerned individuals, when not observed, tend to cooperate less than others. We argue that this is evidence in favor of the fact that *more image concerned individuals tend to be more selfish*. Second, as mentioned above, comparing treatments run with observers to those without, we can show that more image concerned individuals correct their behavior in the direction of the social norm more than others – at least when they are observed by others.

---

[2]Indeed, if the experiment is run in different settings, different observers will be used. This evidence suggests that the measurements are not sensitive to this fact.

[3]We ask a survey question about nationality and not race, since questions on race are not allowed in France. Most non-French participants are from former French colonies in North Africa.

[4]There is unfortunately no well established question aimed at measuring image concern, contrary to the case of trust where the "Interpersonal Trust" question ("Generally speaking, would you say that most people can be trusted or that you need to be very careful in dealing with people?") is systematically used. We thus constructed our own question: "It is important for me not to be perceived as selfish."

[5]This suggests that image concerned individuals avoid situations where they risk being exposed. This is coherent with the results of Dellavigna et al. (2012), who show that when the date of a door to door fundraising visit is announced, people try to avoid being present. Our individual measurement of image concern allows us to show direct evidence of such mechanism whereby image concerned people avoid situations where their image is at risk.

In the last section of the paper we discuss different properties of the image concern game. We also show that our main results are robust to adjusting for multiple hypothesis testing, following an approach proposed by Benjamini and Hochberg (1995).

Our paper is closely connected to the empirical literature on the influence of being observed, using both field and lab experiments. We differ in both our goal and approach. The goal of most of these papers is to document the average influence of being observed by comparing average differences between treatments. We, on the other hand, are interested in individual measurements and individual consequences. Ariely et al. (2009), for instance, compare effort levels in treatments that varied in three dimensions: Subjects were either observed or unobserved, received monetary incentives or not and contributed either to a "good cause" (Red Cross) or a "bad one" (NRA). They find that being observed increased effort levels only when subjects did not receive monetary incentives and only when they volunteered for a good cause. Ekström (2012) finds that norm compliance can be increased even by subtle clues of observation, such as pictures of human eyes. Riyanto and Zhang (2015) deviate from this tradition by eliciting the willingness-to-pay of participants in a dictator game to avoid seeing other people's opinion on their action. However, their design can only capture the dictator's self-image concern and not their concern for others' opinion of them.

There is also a strand of the literature documenting consequences of being image concerned. Mellström and Johannesson (2008) show that subjects were less likely to become blood donors when offered a small monetary reward and that offering the possibility to donate the reward to a charity immediately restores blood donor supply. This suggest that image concern can be an important driver of unselfish actions (see also Lacetera and Macis (2010)). Dellavigna et al. (2012) show that notifying residents in advance of a door to door fundraiser significantly decreases the share of households opening doors, one possible interpretation being that image concerned individuals attempt to avoid the pressure. Bursztyn and Jensen (2015) show that image concern can affect educational choices and show the importance of who the observer's identity.

Further evidence on the consequences of image concern has been gathered using laboratory experiments, where typically behavior is compared across treatments with and without observers. Andreoni and Petrie (2004) find that contributions in a public goods game increased when the players were not anonymous. Dana et al. (2006) offer participants a costly possibility to opt out of a dictator game and show that giving in the dictator game is in part motivated by participants not wanting to appear selfish. In the same spirit, other contributions find that providing options for the participants to overcome their moral dilemmas significantly lowers transfers (Rege and Telle (2004), Samek and Sheremeta (2014), Dana et al. (2007)).

We point out one branch of the literature that tries to find individual proxies for image concern. Carpenter and Myers (2010) use data on the purchase of vanity plates by firefighters that make them identifiable as such at all times. They show that this proxy can predict higher responses to emergency calls but has no effect on less visible activities such as training. In a study of Wikipedia, Algan et al. (2013) use the size of the contributor's page and the extent to which they choose to

display awards as a proxy for image concern.[6] We share the goal of these papers to find individual proxies for image concern but try to determine a less context specific measure exploitable in a wide variety of settings.

Our approach is similar to some extent to the approach in the literature on trust. Analogously to Glaeser et al. (2000), Fehr et al. (2003) and Sapienza et al. (2013), we compare measurements of preferences obtained by survey questions to those resulting from laboratory experiments. Our results also link us to the literature on racial discrimination and bring a new twist by documenting that non-French subjects are only image concerned when facing French individuals. There is a growing literature experimentally studying issues of discrimination and prejudice (for instance Fershtman and Gneezy (2001)). Here what seems to play a role is the fear of prejudiced reactions.

Finally, the second phase of the experiment relates to the literature on infinitely repeated games in the lab (Dal Bó and Fréchette (2011); Dal Bó (2005) among others). To the best of our knowledge, it is the first time an infinitely repeated prisoner's dilemma is played with observers, a side contribution of our paper.[7] In addition to our analysis on image concern, our study also allows for a better understanding of the social norms governing those games, using the ratings by observers of the behavior of participants.

The remainder of this article is organized as follows. Section 2 introduces the image concern game and a conceptual model to analyze it. Section 3 presents its results. Section 4 discusses correlates of our image concern measure. Section 5 contains a detailed analysis of the infinitely repeated prisoners dilemma game. Section 6 discusses potential concerns and section 7 concludes.

2   Measuring image concern: concept and procedure

Although being observed by strangers has been shown to impact behavior and many models include a term corresponding to image in the utility function, there is no unified concept, model or even terminology. We therefore start in section 2.1 by defining the concept we want to measure, then introduce in section 2.2 the image concern game we propose to perform this measurement and finally analyze the game in section 2.3.

2.1   Conceptual framework

We define the image concern of an individual as the degree to which anonymous strangers' opinion of him affects his utility. Specifically, the utility of individual i, when he consumes c_i and others consume c_{-i} and when the image others have of him is denoted R_i (defined formally below), is

---

[6]Algan et al. (2014) in an analysis of open source software programmers use the answer to a survey question to identify image concern.

[7]Charness et al. (2007) do examine a prisoner's dilemma with observers, but without repeat interaction and where the observers are group members who have a stake in the game. Other papers have focused mainly on punishment by third-parties for norm enforcement, rather than observation alone (see for instance Fehr and Fischbacher (2004). In our setup observers have no possibility to punish players. Sutter et al. (2009) use a much weaker form of observation, where observers only know about decisions and payoffs but do not see players' pictures or anything that could identify them.

given by:

$$U_i = V_i(c_i, c_{-i}) + \mu R_i \tag{1}$$

The term μ measures image concern, the dimension of preferences we want to measure. This has been sometimes called "concern for social image" (Andreoni and Bernheim (2009)) or "image motivation" (Ariely et al. (2009)). Note that the utility function can include other social preferences such as altruism, since the consumption of others directly enters the utility function.

Suppose the characteristics or type of the individual is multidimensional $v^i = (v_1^i, .., v_K^i)$. The characteristics could be altruism, reciprocity or other individual characteristics that could potentially influence the shape of the function V_i. The characteristic k for individual i is drawn from the distribution Φ_k^i. The image term R_i corresponds to the beliefs others hold on i's characteristics. Specifically, we assume that for a given characteristic k, R_i is the difference between the expected value of characteristic k for individual i and the average value of the characteristic in the population v_k^0. Finally we assume that individual i might not care in the same way about the image he conveys on the different characteristics: we therefore introduce weights γ_k^i on characteristic k, so that R_i is given by:

$$R_i = \sum_{k=1}^{K} \gamma_k^i \left(E\left[v_k^i\right] - v_k^0 \right) \tag{2}$$

The expectation $E\left[v_k^i\right]$ is affected by observable actions taken by individual i as in the case of the image concern game we introduce in the next section, where the observer updates her beliefs about the sender based on the outcome of the lottery if she observes it.

Finally we insist on the fact that, in our definition, the image concern term μR_i is the pure effect of image and excludes both material benefits that the individual could derive from a good reputation with individuals he might interact with later on as well as self-image issues.[8]

2.2 The image concern game

The image concern game we propose is played between three players: the dictator (he), the recipient and the observer (she). The game is played as follows:

1. The dictator sees the photo of the observer but the observer does not see any pictures yet. The recipient never sees any pictures (neither of the observer nor of the dictator) throughout the game.

2. The dictator receives 100 tokens. He decides how much to allocate to a lottery. The lottery has two possible outcomes: success, in which case the recipient receives 50 tokens or failure, in which case the recipient receives nothing. For each token paid by the dictator, the chances of success increase by one percent, i.e. if the dictator gave an amount $X \in (0, 100)$, the probability that the lottery is a success is $X/100$.

---

[8]In the literature, for instance Bénabou and Tirole (2006), the reputation term is allowed to cover all the different dimensions.

3. Before the lottery is run, the dictator chooses the maximum amount b he is willing to pay to remain anonymous in case the lottery results in a failure. To ensure truthful answers, we use a Becker-DeGroot-Marschak (BDM) type mechanism.[9]

4. The lottery is carried out:

   (a) If it is a success, the recipient receives 50 tokens and the picture of the dictator appears on the screen of the observer.

   (b) If it is a failure, the recipient receives nothing and the BDM mechanism comes into play. A random number $r \sim U(0, 100)$ is drawn. If $r \leq b$, the dictator pays r and remains anonymous (the observer does not see the dictator's picture). If $r > b$, the dictator pays nothing and the observer sees the picture of the dictator. In both cases, the observer learns that the lottery outcome was a failure.

   No matter the result of the lottery, neither the observer nor the recipient ever learn the amount actually chosen by the dictator in either step. They are only informed about the outcome of the lottery.

As suggested in the introduction, there a several key ideas that underly the setup of this game.

First, the individual image concern can be measured by the willingness to pay b to remain anonymous, chosen in step 3. This is formally shown in proposition 1.

Second, we chose to have three players, rather than just a dictator and a recipient, in order to withhold the identities of dictators and recipients from each other. This allows us to separate image concern from fear of retaliation. If the picture was shown directly to the recipient instead of an outside observer, it would not be possible to separate these two motives. This design also guarantees that dictators are not influenced by the recipients' characteristics visible on the picture (such as the perceived needs).

Third, the decision to pay for anonymity is separated from the amount actually transferred in the lottery by the dictator. Regardless of how much the dictator gave in step 1, the inference that an observer makes about the dictator's generosity when she sees a failure is the same since she does not observe the actual amount transferred. This is also clarified below in proposition 1. Intuitively, the fact that the measurement is not confounded by altruism becomes clear when considering two dictators with the same image concern but different levels of generosity. In our setting, the two dictators will give different amounts in the lottery but will bid the same way for anonymity. We could have chosen a setting where dictator and recipient play a classical dictator game and the dictator first has to bid for anonymity given that the amount transferred will be revealed to the observer. In such a case, the more generous dictator would still transfer more than the other in the lottery, but would then bid less for anonymity since he would have less to be ashamed of. We

---

[9]This method of incentive compatible WTP elicitation was introduced by Becker et al. (1964) and is very common in the literature. To avoid concerns that it is not well understood by participants, the instructions clearly stated that the best strategy is to honestly report WTP.

would thus mistakenly conclude that the first dictator was less image concerned. Our game, at a slight cost of complexity, is designed to overcome this potential issue.

Fourth, we are measuring pure image concern. It cannot be self image since the bid determines whether the picture is shown to an outside observer. Moreover, image cannot entail future material payoffs as most of the lab participants do not know each other and we control for this factor when they do.

2.3 Analysis of the game

We clarify the claims made above by deriving theoretically the equilibrium choices in the experiment if the utility of participants is given by equation 1.

The dictator has two choices to make: the amount X he transfers to the lottery and the amount b he bids to remain anonymous (both variables are normalized by 100, so that for instance the probability of winning the lottery is X). We put more structure on the function V_i introduced in equation 1. We denote $v^i(1)$ the utility of dictator i net of costs and reputation if the lottery is a success (and the recipient receives the 50 tokens), while $v^i(0)$ is the corresponding value in case of a failure. Thus $v^i = v^i(1) - v^i(0)$ measures the altruism of individual i. This will be the only characteristic defining the type (i.e. the number of characteristics is $k = 1$). We denote $c_1(X)$ the strictly increasing and convex cost function of giving to the lottery and $c_2(b)$ the strictly increasing and convex cost function of bidding for anonymity. We assume these two costs are separable, corresponding to the idea that these two decisions are made sequentially.

Individual i thus chooses X and b to maximize

$$X \underbrace{\left[v^i(1) + \mu R_1^*\right]}_{\text{payoff when lottery succeeds}} + (1-X)\underbrace{\left[v^i(0) + (1-b)\mu R_0^* - bc_2(b)\right]}_{\text{payoff when lottery fails}} - c_1(X)$$

where $R_l^* = E[v|l] - v_0$, where v_0 is the average altruism under the ex ante distribution, and $E[v|l]$ is the expectation of v conditional on the outcome $l \in \{0, 1\}$ of the lottery. A successful outcome brings a positive image since it signals that the dictator has likely transferred more money, $R_1^* > 0$, while a failure is a bad signal $R_0^* < 0$, as we show in the proof of proposition 1. Note that compared with equation 2, we implicitly suppose that $\gamma_1^i = 1$. In any case in the data, μ and γ_1^i cannot be identified separately, since there is a single characteristic $k = 1$. In this environment, we obtain the following results:

Proposition 1: In an interior equilibrium:

1. The bid for anonymity b^* of an individual is strictly increasing in μ

2. The bid for anonymity b^* of an individual is independent of the altruism v

3. The transfer to the lottery X^* is increasing in the altruism v and in image concern μ

Proof: see appendix

The first result clarifies the fact that the bid for anonymity measures the image concern. More image concerned individuals suffer more from the revelation of their picture when the lottery is a failure.[10]

The second result shows that, under the assumption of our model, our measurement of image concern is independent of altruism v. This result relies on the assumption that the cost functions c_1 and c_2 are separable. The idea is that the bid matters only in the case the lottery is a failure, and conditional on a failure, the amount X actually invested no longer plays a role. This is a key feature that drove the design of our experiment.

The last point indicates that, under the assumption that v and μ are not correlated, more image concerned individuals will tend to invest more in the lottery to avoid having to pay for anonymity. Indeed, more image concerned individuals know that a failed outcome will be more costly for them as they will have to pay more for anonymity. This is consistent with the interpretation given in Dellavigna et al. (2012), who show that notifying residents in advance when a door to door fundraiser takes place significantly decreases the share of households opening doors. One possible interpretation for this finding is that image concerned individuals attempt to avoid the pressure. Our individual level measure of image concern allows us to show even more precise evidence of such a mechanism. This last prediction of proposition 1 is used in section 3.2 as a method to validate our measure of image concern. Note that it does rely on the assumption that the level of image concern is uncorrelated with altruism. A sufficiently large negative correlation could reverse the result.

Two assumptions are necessary for some of the results. First the cost is assumed to be separable between the amount transferred to the lottery and the bid for anonymity. This assumption could be relaxed but it would introduce an interaction between image concern and altruism. Second, risk neutrality is assumed. Note that in all our empirical results we control for attitudes towards risk.

2.4 Experimental setup

2.4.1 Organization of the sessions

The experiment was computer-based and there was no communication between subjects. All participants were seated in the same room, separated by screens, and briefed together. Before the experiment started, a picture was taken of each participant and fed into the experimental software, so that subject anonymity could be removed in a controlled manner. The participants were informed that the photo would be destroyed immediately after the end of the session. Each session was organized in three phases:

1. Participants played four successive and independent rounds of the *image concern game*. They

---

[10]However, the relation between the bid and μ is not linear, unless the cost function c_2 is quadratic.

were randomly assigned to be either a dictator, a recipient or an observer and informed of their assignment. They kept this role for the four rounds. At the beginning of each round, a photo of the observer was shown to the dictator (in the right panel of his screen as shown in Figure 6). At the end of each round, dictators were rematched with different observers and recipients. In each round an observer was assigned two dictators. No dictator encountered the same observer or recipient twice.

The players were informed that they would play four rounds but that only one of them would be selected at random to determine their payoffs. Nevertheless, at the end of each round, they observed the outcome. The payoff of the dictator and the recipient depended on the dictator's choice and the outcome of the lotteries, as described in section 2.2. The observers, on the other hand, received a fixed payment of 40 tokens per round, independent of other players' actions.

2. Subjects played a repeated prisoner's dilemma game described below. In half of the sessions, the prisoner's dilemma games were run with third party observers, in the others without.

3. A survey was conducted containing the main question we use as validation for our image concern game, as well as questions on socioeconomic information. We also included a general question on risk-taking that has been shown to be strongly correlated with incentive compatible measures of risk-preferences and to predict risky behavior (Dohmen et al. (2011)).

At the beginning of each phase, participants received a copy of the instructions, which were then read out loud by the experimenters. Participants filled out a brief questionnaire to check their own understanding and could ask questions in private. The experimenter then read out the correct answers to the questionnaire, making us confident that subjects accurately understood the instructions.

2.4.2 Infinitely repeated prisoner's dilemma

Players were organized in pairs and played the following prisoner's dilemma with payoffs presented in table 1.

TABLE 1: PAYOFFS OF PRISONER'S DILEMMA

	C	D
C	8, 8	0, 10
D	10, 0	4, 4

The infinitely repeated game was implemented using a random continuation rule where at the

end of each round there was a probability of 7/8 that another round was played in the game.[11] After the last round of a game had ended, the participants were rematched so that no group of subjects encountered each other more than once. The participants were informed that they would play exactly three games. In practice, as in Peysakhovich and Rand (2015) and Fudenberg et al. (2012), we did not randomize the number of rounds within a game *during* the session but *before* since we wanted to compare behavior across treatments and thus wanted games of identical length. We followed the randomization chosen by Peysakhovich and Rand (2015), who also used a continuation probability of 7/8, and we chose exactly the same length as in the first three games described in their paper. Given this approach, each participant played three games, the first with 12 rounds, the second with one round and the third with three rounds.[12]

In half of the sessions, the actions of both players in the prisoner's dilemma were visible to an observer who did not have any stakes in the game. Each observer was assigned to two pairs of players. The observers saw the players' photos and computer names, as well as the decisions they made in the game. A picture of their observer and his or her computer name was visible on the players' screens while they took their decisions. Observers had to indicate whether they had met the other participants before and were asked in each round, after having observed the choice of the players, how they rated their behavior in the game.

The payoff in this phase of the experiment was the sum of payoffs in all rounds. The observers received a flat payment of 5 tokens per round that was independent of the players' actions.

2.4.3 The sample

The experiment was conducted in May and September 2014 at the Laboratoire d'Economie Experimentale de Paris. The lab has access to a subject pool that comprises individuals not affiliated to any university as well as students and staff. Table 2 provides descriptive statistics both for the full sample and for the sample of participants for whom we have an image concern measure, i.e. who played the image concern game as dictators. The sample is fairly balanced in terms of gender and marital status and is not exclusively composed of students (61 percent of students in the dictators sample). The majority of participants is French and most of those who report not being French are from North African countries.

While the experiment was ongoing, we documented whether participants smiled or frowned on their pictures, and whether the picture was blurry or participants had their eyes closed.[13] To accommodate privacy concerns, we deleted all pictures immediately after each session.

Each of the 13 sessions involved exactly 20 participants. In the image concern game, eight of those participants were assigned to the role of dictator, eight to the role of recipient and four to

---

[11]With this continuation probability, cooperation can be sustained in a subgame perfect and risk dominant equilibrium.

[12]In the instructions it was not made clear whether the randomization at the end of each round was done on the spot or had been done before (as was in fact the case). It was only stated that at the end of each round, there were 7 chances out of 8 to have another round in the game.

[13]The experimenter was unable to see any of the decisions participants made while he was coding these variables.

the role of observer (observers were in charge of two dictators in each round). For the prisoner's dilemma, half of the subjects played with observers and half without. Overall this gives 260 subjects, out of which 104 played the image concern game in the position of dictator. On average, subjects received €16.74 for participating in the experiment, including a fixed €4 show-up fee.[14]

3   Measuring image concern: the results

3.1   Heterogeneity in image concern

The image concern game is designed to measure image concern in a straightforward way using the willingness to pay for anonymity. In principle, only one round of the game is needed to obtain a measurement. However, we ran four successive rounds with rematching in order to assess the impact of observers' characteristics on the measure. For each individual, we thus have four individual measures that we could potentially combine in different ways. However, we find little variability in aggregate bidding behavior: the average value of μ does not vary much across rounds, except for the last round where it is slightly but not significantly lower. For the remainder of the paper, we will therefore use the bid for anonymity in the first round as an individual's measure for image concern.[15] To facilitate interpretation we divide it by its standard deviation wherever it is used as an explanatory variable.

The results of the experiment reveal significant heterogeneity in terms of image concern. The distribution of the bids for anonymity chosen in the first round is given in figure 1.[16] 40 percent of the sample is completely insensitive to image, i.e. does not pay to stay anonymous. On the contrary, more than 22 percent seem quite sensitive and give more than 20 tokens to stay anonymous in the case of an adverse lottery outcome. Since this is the first study to measure individual sensitivity to image using an experimental game, it is difficult to compare the distribution to existing results. As a reference point, Carpenter and Myers (2010) find that 23 percent of firefighters in their sample purchase a vanity plate for their car, which is the proxy the authors interpret to identify image concerned individuals.

For most of the paper we exclude three outliers who bid more than 90 for anonymity whereas the highest bid among the rest of the population is 75. The results are robust to the inclusion of these individuals. This leaves us with 101 individuals for whom we have a measure of image concern.

3.2   Validation of the game

One key question immediately arises: Are we indeed measuring image concern? There is not, as in the case of trust, a widely accepted survey question convincingly capturing the degree of image concern. We therefore constructed a question that reflects this construct: "It is important for me not
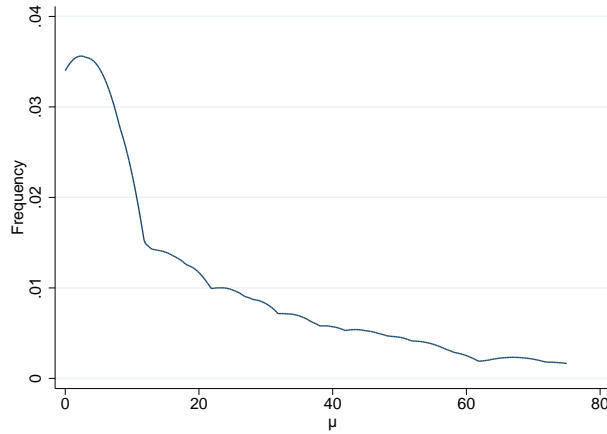
---

[14]The conversion rate of tokens to euros was 10 to 1.

[15]Our results are mostly robust to using alternative measures such as the average bids.

[16]Its shape is very similar for the average value of the bids (see online appendix).

NOTES. This figure plots a kernel density estimate of participants' image concern as measured by their first bid for anonymity (Epanechnikov, bandwidth = 5.3)

to be perceived as selfish" on a 0-5 scale. We show in figure 2 the average image concern measure for each of the answers to this question. The more people agree with the statement, the more they were willing to pay for their anonymity in case the lottery outcome in the experiment reflected badly on them. This suggests that our measure does capture the sensitivity to the perception by others.

We confirm these graphical results in table 3. In column 1 we present the results of an ordered logit where the dependent variable is the answer to the survey question. There is a positive and significant correlation between the bid for anonymity and the answers.[17] It is important to note that it is the only variable that can explain variations in the answer to that question. In particular, none of our socioeconomic variables turns out to significantly impact the answers.

The association with the natural question in our questionnaire offers a strong initial validation of our measure of image concern. Our second method to validate the measure is to test the last prediction of proposition 1 that stated that more image concerned individual should bid more in the initial lottery phase of the experiment to avoid situations where they will have to pay to preserve their anonymity. We find in the second column of table 3 that more image concerned individuals transfer significantly higher amounts to the lottery: A standard deviation increase in the image concern measure is associated with an increase in transfers by eight tokens. An alternative interpretation of this finding could be that image concerned individuals are just more generous. We in fact show in the next section that the correlation goes in the other direction.

Our third and final method to validate the measure is presented in section 5.4 where we compare sessions where the repeated game was played with observers to those where it was not. We show that more image concerned individuals, as measured by our game, react more to being

---

[17]It is unlikely that participants answered the question in this way purely to appear consistent with their behavior in the game, since between the survey and the image concern game, other games had been played.

FIGURE 2: IMAGE CONCERN LEVELS BY ANSWERS TO THE STATEMENT "IT IS IMPORTANT FOR ME NOT TO BE PERCEIVED AS SELFISH" (SCALE 0-5)
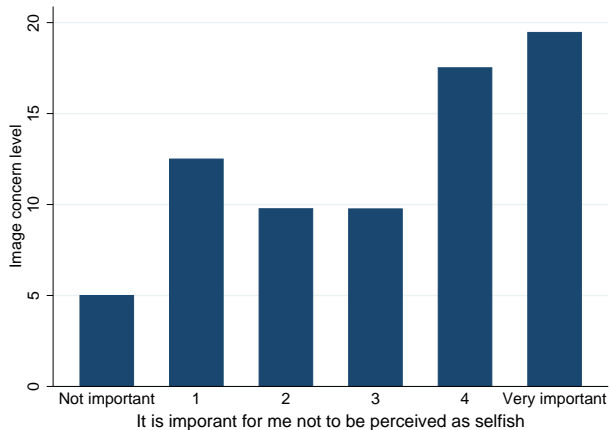
NOTES. This figure reports the average image concern levels as measured by the first bid for anonymity. The survey question was "It is important for me not to be perceived as selfish" and participants could answer on a 0-5 scale ranging from "not important" to "very important."

observed and in particular are more likely to choose the action that observers judged positively.

4   Drivers and consequences of image concern

The first goal of this paper is to propose a game to measure individual sensitivity to image concern and to validate the measure we obtain as was presented in the previous sections. The second goal, the object of this section, is to explore what socioeconomic characteristics are correlated with image concern and start examining whether more image concerned individuals behave differently than others.

4.1   Drivers of image concern

Sensitivity to image concern does not seem to correlate with socioeconomic characteristics as shown in table 4.[18] In column 1 we consider the full sample, in column 2 we restrict the sample to those having bid a positive amount for anonymity, while in column 3, we use an indicator variable of whether the bid for anonymity is positive as dependent variable. The only fact that seems to emerge is that women might be more prone to bid *something* to remain anonymous, even if the amount is low. Age, level of education and profession appear to play no role.

At the end of the study, a detailed survey was conducted that allows us to determine other dimensions of behavior. We find that more image concerned individuals report having significantly fewer friends and are less involved in political parties, as reported in table 5.[19] Although

---

[18]Larger samples could potentially uncover correlations that we are not able to find.

[19]Other dimensions such as being involved in a nonprofit or the amount of donations made seem unrelated to image

the sample size warrants prudence, what seems to emerge is that the sensitivity to image concern cannot be explained by socioeconomic characteristics, but has close connections with behavioral patterns, such as the formation of friendships.

4.2   Does the observer matter?

Examining the impact of observers' characteristics on dictators' willingness to pay for image, we find that few of them alter dictators' behavior, thus confirming the portability of our experiment. We investigate the four features that participants could easily infer from the pictures: gender, approximate age, nationality (French vs. non-French) and facial expression. Table 6 presents the results. Column 1 shows that none of these features have any influence on average. The fact that the observer smiles has a slight positive impact that tends not to be significant. Interestingly, the fact that the observer knows the dictator has no effect.[20] Because the game may have very different dynamics when players already know each other, table 6 also contains models where such observations are excluded altogether.

Columns 3 to 8 introduce interaction terms to understand the role of observers in more detail. Two main results emerge. First, we find weakly significant evidence that *women give slightly more when observed by men*. Second, *non-French dictators appear to be significantly less concerned about their image when they are observed by another non-French participant and slightly but not significantly more when observed by a French*.[21] This fact appears clearly in figure 3, which plots the marginals of the last regression presented it table 6.

There are two plausible explanations. The first is that non-French subjects are just intrinsically sensitive only to the perceptions French people might have of them. The second is that they expect the French to be prejudiced against them and thus interpret a bad outcome more negatively, pushing the non-French to initially give more. This could be driven by the experience of discriminatory behavior. Note that this prejudice, if it were indeed present, would not be justified since non-French participants give more in the lottery, even though the difference is not significant.[22] These interaction terms are significant even though none of the socioeconomic variables have explanatory power.

5   Impact of image concern in an infinitely repeated prisoner's dilemma

We now turn to the analysis of the behavior of participants in phase two of the experiment, where they played the prisoner's dilemma. We will exploit both the differences across sessions (sessions
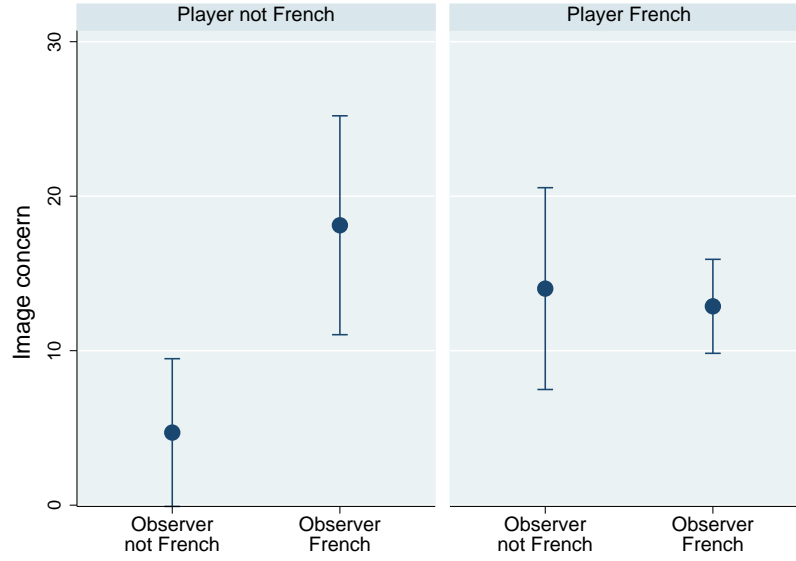
concern.

[20]Note that there is no clear expectation regarding the direction of this effect. We might suppose image concern should be higher for individuals the dictator knows since he might be brought to interact with them outside the lab. At the same time, the individuals he knows might have already formed an opinion and infer less based on the outcome of the experiment.

[21]As a reminder, we use the variable coding nationality as French or not, as a proxy for the race of individuals, which we cannot ask directly. Most of the non-French are nationals from North African countries.

[22]Non-French: 22.9, French: 18.3, t=1.21 (adjusted for participant clusters).

NOTES. This figure plots the predictive margins corresponding to column 8 of table 6 along with 95% confidence intervals.

with and without observers), as well as individual heterogeneity in image concern within sessions.

The purpose of this analysis is threefold. First, to establish what subjects see as the behavioral norms that prevail in the repeated prisoner's dilemma. Second, to determine whether image concerned individuals behave differently than others with respect to these social norms in the absence of observers, for instance in terms of cooperation rates. Finally, we examine whether image concerned individuals react more than others to the fact of being observed. This will provide further validation of our measure of image concern and inform us on the average effect of being observed in an infinitely repeated prisoner's dilemma game.[23]

## 5.1 Prevailing social norms in the repeated prisoner's dilemma

We hypothesize that more image concerned individuals adhere more closely to social norms when they are observed compared to when they are not. Showing that this is the case requires as a first step to determine what actions in the repeated prisoner's dilemma are considered appropriate by the community, i.e. what is the social norm. We are able to address this interesting and novel question by studying observers' ratings of players' actions. After each round of the game, observers were asked to rate the actions of both players. This allows us to document two social norms:

- Cooperation is judged favorably by observers – independently of partners' previous actions
- Consistent, reliable cooperation garners additional approval by observers

---

[23]Measures of image concern obtained from the survey questions perform less well at explaining behavior in this game compared to the measure obtained by the experiment and presented in this section.
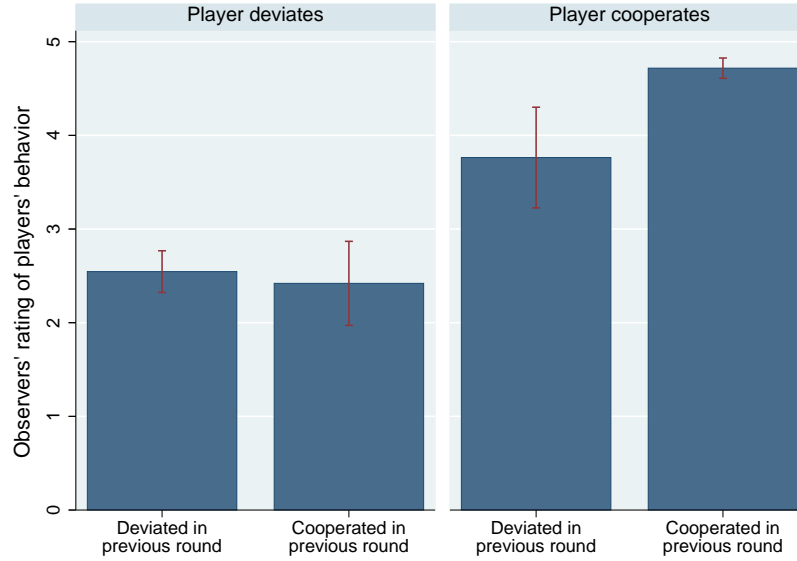
NOTES. This figure reports average ratings of players' decisions in the infinitely repeated prisoner's dilemma game by outside observers. 95% confidence intervals in red.

The first point unambiguously appears in figure 4: Cooperating is highly rated by the observers. However, the perception of observers is also based on a more subtle reaction to the history of play: the positive rating of cooperation is particularly strong when the player also cooperated in the previous round. The observers rate consistency in cooperation very highly. One interpretation could be that they value unconditional cooperators, who consistently avoid deviating.

The regression analysis presented in table 7 confirms these findings. We first note that a certain number of facts not linked to behavior affect the ratings. Students and those who smile in their picture are better rated than others. One interesting fact linking back to the issue of perception of discrimination is that French observers tend to give significantly lower ratings to their fellow citizens and higher ratings to non-French participants. A natural interpretation is that they fear being perceived as prejudiced.

In terms of observed behavior, cooperation indeed significantly increases the rating (table 7). We also confirm that cooperation following cooperation in the previous round has an additional positive effect on ratings – stability is valued (column 3). On the other hand, there is no significant dependency of the rating on the partner's action in the previous round (column 2).[24] These social norms will be used in the next sections to determine how image concerned individuals react to observation.

---

[24]It would be natural to think that ratings would also depend on what the other player did in the past. However, both table 7 and figure 7 in the appendix show that there is no extra negative rating coming from a deviation that follows cooperation by the partner, i.e. there does not seem to be a judgment on betrayal of the partner.

5.2 Cooperation rates of image concerned individuals when not observed

In environments where they are not observed, more image concerned individuals are more likely to defect – they fail to comply with the social norm. To show this, we exploit the variation within sessions of the image concern parameter to determine whether more image concerned individuals behave differently than others, particularly in terms of cooperation rates.

In table 8 we report the results of a probit regression of the cooperation variable explained by socioeconomic characteristics and by the image concern parameter, restricting attention to treatments where the prisoner's dilemma was played without observers. Regardless of whether we focus on all rounds (column 1), only the first round of each game (column 2) or exclusively on the last game (column 3), we find that the more image concerned individuals cooperate significantly less. This is particularly striking since no socioeconomic characteristic has consistently significant explanatory power.

Cooperation rates in infinitely repeated games can reflect both an intrinsic level of cooperativeness (due to altruism, for instance), but also beliefs about how likely others are to cooperate.[25] We favor the interpretation based on altruism since the result seems even stronger when we restrict the sample to the last game (column 3) where beliefs should be less heterogeneous since learning will already have occurred during earlier games. This suggests an initial picture where image concerned individuals appear less altruistic: they are more concerned with themselves and the impression they give to others – a result that should be explored in later work.

5.3 The effect of observers in infinitely repeated prisoner's dilemma

On average, individuals are more likely to follow the social norm in the repeated prisoner's dilemma when they are observed. Comparing the average behavior in treatments with observers to those without, we find that cooperation is more likely in the presence of observers. In treatment without observers, the average cooperation rate is 0.3 while it is 0.53 with observers. The difference is significant ($t = 3.04$, adjusting for participant clusters).

This is also true for the second dimension that appears to influence ratings, the consistency in cooperation. When we restrict the analysis to rounds where the player cooperated in the previous round, the rate of cooperation is 0.66 without observers and 0.77 with observers and this difference is (weakly) significant ($t = 1.70$, adjusting for participant clusters). When on the contrary, we focus on the case where the player deviated in the previous round, the rate of cooperation falls to 0.11 with observers and 0.15 without observers, a difference no longer statistically significant ($t = 1.02$, adjusting for participant clusters). Being observed pushes players to follow the social norm more closely, even along subtle dimensions of the norm. This is coherent with the literature that analyzes behavior in public good games when observed (Andreoni and Petrie (2004)) and other experimental games played with observers.

---

[25]Whether an individual cooperated in the very first round can be seen as a measure for his base level of cooperativeness. In most experiments on infinitely repeated games, this measure significantly increases cooperation in the remainder of the games (see for instance Dal Bó and Fréchette (2011)).

5.4 Further validation of our image concern measure

Finally, we use the effect of observers in the repeated prisoner's dilemma to provide further validation of our image concern measure: More image concerned individuals become more prone to behave according to the social norm when observed. As reported in table 9, image concerned participants react more than others to the presence of observers. The interaction term between image concern and being observed in column 3 positive and significant. However, this is not the case immediately and only becomes apparent in later games when participants have learned the dynamics of the game (column 4).

FIGURE 5: COOPERATION RATES DEPENDING ON PAST BEHAVIOR AND PRESENCE OF OBSERVERS
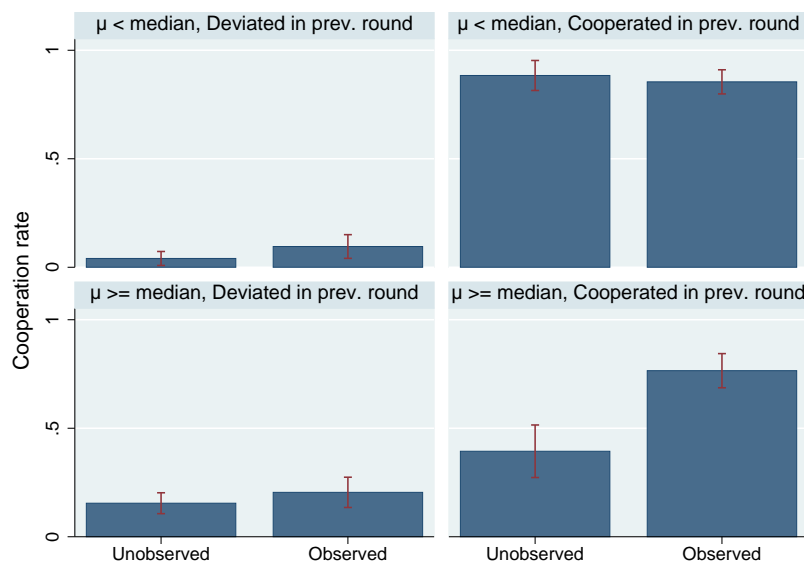
NOTES. This figure contains average cooperation rates in the infinitely repeated prisoner's dilemma game along with 95% confidence intervals. The top and bottom row report averages for participants with image concern levels below and above the median of 10, respectively. The left column contains cases in which the player deviated in the previous round, the right column corresponds to previous cooperation.

Similarly, under observation, image concerned individuals also react more strongly to the social norm of consistency in cooperation, which was highly rated by observers. This can be seen graphically in figure 5, where the first line is for individuals who are not very image concerned (below median of 10) and the second line is for those who are very concerned (above median): For individuals who cooperated in the previous round, the effect on cooperation of being observed is much stronger when the individual is image concerned than when not.

These results are confirmed in table 10. In column 1 we restrict the sample to non-observed sessions and we see that more image concerned individuals tend to cooperate less after a round where they cooperated than less image concerned individuals. However we see in column 2 that for observed sessions, this effect disappears. Image concerned individuals seem to take into

account the fact that this lack of consistency in cooperation is badly perceived by observers. This is confirmed in column 3 where the interaction term between the level of image concern, cooperation in previous round and the fact of being observed is positive and significant.

6 Discussion

We have presented in this paper a novel experimental game to measure image concern, validated the measure and presented initial facts about image concerned individuals. We now discuss some of the possible issues that might be raised about this game.

Experimenter as observer

One might worry that the dictator is under the impression that he is being observed, not only by the observer in the game, but also by the experimenter. If the dictator believes that the experimenter can see how much he pays to be anonymous and if he thinks the experimenter will adversely interpret payments for anonymity, it could imply that image concerned individuals could be less inclined to give than if the experimenter was not present.

We took several precautions to limit this potential problem. First, we clearly told the participants that they would remain anonymous from the point of view of the experimenter: they were told that the photos would be deleted at the end of the session, and that we would of course preserve their anonymity while we conduct the analysis. Furthermore, while the picture of the observer was always visible on the screen of the dictators, the experimenter was not visible during the experiment.

However, even if despite these precautions subjects were still influenced by the experimenter, this would have little impact on the results presented here. It would only decrease the variance in the answers but not change the ranking of individuals in terms of μ.[26] The fact that we find a high degree of heterogeneity in the population, suggests that even if this effect were present, it would not be large.

Shy individuals

Another potential worry could be that our experiment just measures how shy an individual is: Some people might be ready to pay for their picture not to be shown, even if no actions are revealed. We believe that the three different validation methods of our measurement that we presented in section 3.2 and 5.4 reject this alternative story. There is no reason for shy people to systematically answer our survey question differently or to follow the social norm when they are observed in the infinitely repeated game. It might nevertheless be interesting to run an alternative experiment where the decision would be to pay not to have a picture revealed without any actions being attached to it.

---

[26]Unless of course there are two dimensions of image concern that can both vary across the population: being concerned about the experimenter's perception of the level of generosity and being concerned about the experimenter's perception about trying to hide one's true type.

Different sensitivities depending on the characteristic

In our experiment, the inferences that the observer can make about the characteristics of an individual is restricted to the level of altruism. It is conceivable that individuals might care differently about their image depending on the characteristic at stake. Using the terminology of the model, for different characteristics k, the γ_k^i might be different (see equation (2)). We cannot test this directly, but we found that those individuals identified as being image concerned in our game also react to observers in situations where their actions reveal different things than altruism: For instance, they respect the second norm of behavior in the repeated prisoner's dilemma more, a norm that has less to do with altruism. This question nevertheless remains one that should be examined in future research.[27]

Multiple hypothesis testing

A growing body of literature highlights the risk of false discoveries when testing multiple hypotheses within the same experiment (see for instance List et al. (2016)). Indeed, when analyzing several subgroups and correlating multiple survey questions and experimental outcomes, one may worry that rejecting null hypotheses at traditional significance levels leads to an elevated risk of type I error. We thus conduct additional robustness checks to minimize the risk that any of the results presented above are false discoveries. Following an approach proposed by Benjamini and Hochberg (1995), we collect all the hypotheses tested in the regressions and apply their method to control the False Discovery Rate (FDR), defined as the share of false rejections among all rejections. The resulting corrected p-values (q-values) associated with all the hypotheses investigated in this paper are reported in table 11. They should be interpreted as the minimum FDR for rejection of the associated null hypothesis of no effect.

Despite the fact that this method is much more conservative than regular hypothesis testing, our main results still hold – even if some of them require lower thresholds to rejection. In particular, all three validations of our image concern measure presented above are confirmed.

7   Conclusion

This paper proposes an experimental procedure to measure individual sensitivity to image concern, validates the measure and starts exploring determinants and consequences of this underexplored dimension of preferences. It opens the way for future research on the topic.

As a first step in this direction, we are able to document two patterns: First, the extent to which individuals are concerned with their social image is highly heterogeneous and does not depend on the observer or on demographic characteristics. As a notable exception, we provide evidence that minority members might be more concerned with the image that they project to members of the majority group. Second, we show that image concerned individuals are less cooperative in

---

[27]Bracha and Vesterlund (2013), for instance, describe an experiment separating generosity signaling from income signaling.

classic infinitely repeated prisoner's dilemma games, but that cooperation rates increase up to the average when the same game is played under the scrutiny of a third party observer. This suggests that for these individuals, norm compliance is essentially linked to the salience of their actions.

Due to the nature of the concept to be measured, the game is of course less portable than other games aimed at measuring social preferences, such as the trust game or the dictator game and might be more difficult to run in a remote field environment. Variations in the setup can, however, be used. For instance the process of taking pictures could be eliminated by asking the dictator to stand up if he loses anonymity. There are also ways to implement the BDM without computer terminal. Overall, we attempted to ensure simplicity of the setup while preserving robustness of the measure, in particular avoiding confounding factors such as other social preferences.

Appendix A1: Proof of Proposition 1

Individual i chooses X and b to maximize

$$X \left[v^i(1) + \mu R_1^* \right] + (1 - X) \left[v^i(0) + (1 - b)\mu R_0^* - bc_2(b) \right] - c_1(X)$$

R_1^* is the image, in equilibrium, when the lottery is a success while R_0^* is the image in case of a failure.

We have

$$
\begin{aligned}
X R_1^* + (1 - X) R_0^* &= X \left(E[v|1] - v_0 \right) + (1 - X) \left(E[v|0] - v_0 \right) \\
&= v_0 - v_0
\end{aligned}
$$

Where the second equality is derived from the law of iterated expectations. Thus if $R_1^* > 0$ then $R_0^* < 0$. $R_1^* > 0$, since the fact that the lottery is a success makes the recipient update positively her beliefs on the amount transferred by the individual, under the assumption that v and μ are not correlated.

In an interior equilibrium, since $v = u^i(1) - u^i(0)$ we have:

$$
\begin{aligned}
c_1'(X^*) &= v + \mu R_1^* - [(1 - b^*)\mu R_0^* - bc_2(b^*)] \\
bc_2'(b^*) + c_2(b) &= -\mu R_0^* > 0
\end{aligned}
$$

The second equation yields immediately the result that b^* is increasing in μ and independent of v. Applying the envelope theorem to the first equation, we have

$$\frac{\partial X^*}{\partial \mu} = -\frac{1}{c_1''(X^*)}(1 - b^*)R_0^* > 0$$

So this establishes the third result.

Appendix A2

TABLE 2: DESCRIPTIVE STATISTICS FOR THE SAMPLE OF PARTICIPANTS

Variable	All Participants Share/Mean	Dictators Share/Mean
Demographics		
Female	.56	.59
In a relation	.49	.58
Age	28.5	30.20
French national	.87	.85
Attitudes to Risk	5.75	5.76
Professional status		
Student	.66	.61
Employed	.21	.22
Unemployed	.09	.11
Retired	.04	.06
Highest degree achieved		
High school or less	.29	.31
College diploma	.44	.45
Master's degree	.24	.22
PhD	.04	.02
Field of study		
Economics and finance	.28	.29
Other social sciences	.15	.16
Law	.15	.11
Picture controls		
Smiles on photo	.07	.09
Frowns on photo	.02	.02
Photo blurry or eyes shut	.03	.03
Observations	260	104

	(1) Survey question *Ordered logit*	(2) Transfer to lottery *OLS*
Image concern (μ)	0.50[a]	7.98[a]
	(0.18)	(2.32)
Economist	-0.93[b]	-0.98
	(0.43)	(3.64)
Female	-0.11	3.53
	(0.32)	(3.61)
Age	0.01	0.10
	(0.02)	(0.16)
In a relationship	-0.41	2.31
	(0.44)	(2.78)
Student	-0.32	-0.14
	(0.54)	(5.68)
French	-0.97	-1.79
	(0.70)	(5.74)
Response time	-0.02[b]	-0.00
	(0.01)	(0.06)
Knows the observer	-0.04	-2.98
	(0.47)	(4.71)
Observations	101	101
Pseudo R^2	0.065	0.244

Session-clustered standard errors in parentheses
[c] $p<0.1$, [b] $p<0.05$, [a] $p<0.01$

NOTES. This table reports in the first column, the estimations of an ordered logit regression of the answers to the question "It is important for me not to be perceived as selfish" (0-5 scale), and in the second, a regression of the amount transferred by the dictator in the image concern game, regressed on image concern measures and socioeconomic characteristics. The image concern measure has been divided by its standard deviation for easier interpretation of the magnitudes. The sample contains all individuals who played as dictators in the image concern game, excluding 3 outliers.

	(1) Image concern *OLS*	(2) Image concern *OLS*	(3) Non-zero image con. *Probit*
Female	-0.45 (3.51)	-6.97 (6.18)	0.43[a] (0.16)
Age	0.12 (0.25)	0.00 (0.23)	0.01 (0.02)
In a relationship	2.22 (3.49)	7.59 (4.94)	-0.29 (0.28)
Student	-8.01 (7.63)	-7.51 (9.56)	-0.65 (0.44)
French	-3.59 (6.91)	-3.15 (5.62)	-0.71 (0.46)
Knows the observer	5.39 (5.88)	3.19 (8.87)	0.01 (0.41)
Observations	101	61	101
Pseudo R^2	0.091	0.130	0.110

Session-clustered standard errors in parentheses

[c] $p<0.1$, [b] $p<0.05$, [a] $p<0.01$

NOTES. Column 1 presents a regression using the bid for anonymity the first time the image concern game is played as dependent variable and socioeconomic characteristics for the full sample as explanatory variables. Column 2 restricts the sample to individuals who bid a positive amount. Column 3 is a probit regression of the indicator variable taking value 1 if the individual bid a positive amount for anonymity. The sample contains all individuals who played as dictators in the image concern game, excluding 3 outliers.

	(1) # Friends *OLS*	(2) Party member *Probit*	(3) Active in NPO *Probit*
Image concern	-0.83[a]	-0.46[a]	0.11
	(0.23)	(0.17)	(0.15)
Female	-1.59[b]	-0.53	-0.04
	(0.54)	(0.44)	(0.40)
Age	0.01	0.01	0.01
	(0.03)	(0.02)	(0.01)
In a relationship	-0.18	0.08	-0.14
	(0.87)	(0.37)	(0.21)
Student	0.45	-0.33	0.45
	(0.95)	(0.71)	(0.43)
French	1.60		
	(1.06)		
Knows the observer	1.73[c]	0.76	-0.43
	(0.94)	(0.65)	(0.55)
N	101	85	85
Pseudo R^2	0.135	0.111	0.019

Session-clustered standard errors in parentheses
[c] $p<0.1$, [b] $p<0.05$, [a] $p<0.01$

NOTES. Column 1 presents the estimations of a regression of the variable measuring the self-reported number of friends. Column 2 presents the estimation of a probit regression of an indicator variable taking the value 1 if the participant is a member of a political party. Column 3 contains analogous results for a binary variable indicating whether the subject is actively engaged in a non-profit organization. The sample contains all individuals who played as dictators in the image concern game, excluding 3 outliers.

TABLE 6: ROLE OF OBSERVERS

	(1) Image concern	(2) Image concern	(3) Image concern	(4) Image concern	(5) Image concern	(6) Image concern	(7) Image concern	(8) Image concern
Observer female	0.62 (2.01)	0.16 (2.01)				-0.16 (1.99)		-0.14 (2.01)
Observer age	-0.01 (0.07)	-0.04 (0.07)		-0.03 (0.07)				-0.04 (0.07)
Observer french	-0.24 (2.83)	0.98 (2.98)		1.31 (2.89)		0.64 (3.13)		
Observer smiles	5.18 (7.42)	5.30 (7.62)		5.08 (7.47)		5.01 (7.69)		4.46 (7.63)
Observer frowns	2.95 (6.35)	6.04 (7.71)		6.37 (7.57)		6.26 (7.93)		6.25 (7.28)
Picture blurry	-5.18 (6.44)	-4.35 (6.62)		-4.07 (6.43)		-4.06 (6.85)		-3.88 (6.58)
Male × observer female			3.68 (2.93)	3.19 (3.38)				
Female × observer male			5.83^c (3.35)	5.47 (3.44)				
Female × observer female			3.17 (2.76)	2.96 (3.02)				
>24y old × observer <=24y old					-2.71 (2.47)	-2.10 (2.60)		
<=24y old × observer >24y old					-5.72 (3.57)	-5.50 (3.41)		
<=24y old × observer <=24y old					-4.09 (3.41)	-3.34 (3.44)		
Non-french × observer french							9.45^b (4.03)	13.42^a (4.33)
French × observer non-french							7.20 (4.38)	9.32^b (3.71)
French × observer french							4.88 (3.31)	8.17^a (2.88)
Knows the observer	0.88 (3.10)	obs. excluded	0.99 (2.93)	obs. excluded	0.81 (2.92)	obs. excluded	1.13 (3.00)	obs. excluded
Observations	404	367	404	367	404	367	404	367
R^2	0.141	0.160	0.147	0.168	0.146	0.166	0.146	0.172

Participant-clustered standard errors in parentheses
^{c} $p<0.1$, ^{b} $p<0.05$, ^{a} $p<0.01$

NOTES. This table presents regressions of the bid for anonymity on dictators' and observers' characteristics. All regressions control for round fixed effects and for outcomes of the previous round, though the results are qualitatively unchanged when these controls are omitted. The full sample contains all individuals who played as dictators in the image concern game, excluding 3 outliers. Columns 2, 4, 6 and 8 exclude players that knew each other.

	(1) Rating	(2) Rating	(3) Rating
Player cooperates	2.94^a	2.47^a	1.78^a
	(0.54)	(0.66)	(0.66)
Player French	1.11	1.06	1.19
	(0.81)	(0.73)	(0.74)
Observer French	0.66	0.53	0.62
	(0.51)	(0.53)	(0.50)
Observer French × player French	-1.40^c	-1.28^c	-1.40^c
	(0.80)	(0.73)	(0.73)
Player is student	1.21^c	1.24^c	1.22^c
	(0.66)	(0.65)	(0.63)
Player smiles	2.37^a	2.16^a	2.00^b
	(0.84)	(0.82)	(0.80)
Subjects know each other	-2.16^a	-2.13^a	-2.16^a
	(0.69)	(0.70)	(0.65)
Deviate after partner cooperated		0.29	
		(0.40)	
Cooperate after partner cooperated		0.75	
		(0.56)	
Deviate after player cooperated			-0.15
			(0.35)
Cooperate after player cooperated			1.60^b
			(0.69)
Observations	520	520	520
Pseudo R^2	0.242	0.246	0.253

Standard errors reported in parentheses are clustered at the participant level
^{c} p<0.1, ^{b} p<0.05, ^{a} p<0.01

NOTES. This table reports the estimations of an ordered logit regression of the ratings given by the observer (1-5 rating) on the player's characteristics and behaviors. We control for players' and observers' age and gender as well as attitudes towards risk. The online appendix contains alternative specifications as robustness checks.

	(1) Cooperate *All rounds*	(2) Cooperate *First rounds*	(3) Cooperate *Last game*
Image concern (μ)	-0.36[b]	-0.73[a]	-0.78[a]
	(0.17)	(0.24)	(0.29)
Female	-0.13	-0.20	-0.63
	(0.35)	(0.38)	(0.39)
Age	0.01	-0.01	-0.02
	(0.01)	(0.02)	(0.02)
Student	-0.14	-0.53	-0.41
	(0.54)	(0.57)	(0.63)
French	-0.81[c]	-0.49	-0.68[c]
	(0.46)	(0.56)	(0.40)
Risk aversion	0.03	0.09	0.08
	(0.09)	(0.08)	(0.09)
Response time	-0.00	0.00	-0.08
	(0.01)	(0.01)	(0.05)
Round fixed effects	Yes	Yes	Yes
Observations	640	120	120
Pseudo R^2	0.07	0.12	0.18

Standard errors reported in parentheses are clustered at the participant level

[c] $p<0.1$, [b] $p<0.05$, [a] $p<0.01$

NOTES. This table reports the estimations of a probit regression of the indicator variable taking the value 1 if the player cooperated, restricting the sample to treatments with no observers. All regressions control for round fixed effects. Column 1 includes all rounds of the prisoner's dilemma game. Column 2 is restricted to the first round of each game. Column 3 restricts the sample to the third game.

	(1) Cooperate *Unobs. sample*	(2) Cooperate *Observed sample*	(3) Cooperate *Full sample*	(4) Cooperate *Full sample*
Image concern (μ)	-0.78^a	0.11	-0.62^b	-0.57^b
	(0.29)	(0.16)	(0.27)	(0.23)
Observed			1.06^a	0.89^a
			(0.27)	(0.26)
Observed $\times\ \mu$			0.64^b	0.31
			(0.29)	(0.27)
Control variables	Yes	Yes	Yes	Yes
Round fixed effects	Yes	Yes	Yes	Yes
Games/Rounds	Last game	Last game	Last game	First rounds
Observations	120	120	240	240
Pseudo R^2	0.18	0.26	0.25	0.17

Standard errors reported in parentheses are clustered at the participant level

^{c} p<0.1, ^{b} p<0.05, ^{a} p<0.01

NOTES. This table reports the estimations of a probit regression of the indicator variable taking the value 1 if the player cooperated. The image concern measure has been divided by its standard deviation for easier interpretation of the magnitudes. Except for column 4, only observations from the last game are included. Column 1 includes only the observations from sessions where the game was played without observers and column 2 those where the players are observed. We control for socioeconomic characteristics, attitudes towards risk, response time and whether the player knew the observer.

	(1) Cooperate *Unobs. sample*	(2) Cooperate *Obs. sample*	(3) Cooperate *Full sample*
Image concern (μ)	0.19 (0.37)	-0.89[c] (0.49)	-0.11 (0.20)
Player cooperated in previous round	1.00[c] (0.53)	2.94[a] (0.76)	1.61[a] (0.31)
Observed			1.00[a] (0.32)
Image con. $\times$ cooperated in prev. round	-2.78[a] (0.95)	2.01[a] (0.75)	-1.48[a] (0.45)
Image con. $\times$ coop. in prev. round $\times$ obs.			2.45[a] (0.61)
Control variables	Yes	Yes	Yes
Game $\times$ round fixed effects	Yes	Yes	Yes
Observations	80	80	160
Pseudo R^2	0.52	0.50	0.48

Standard errors reported in parentheses are clustered at the participant level

[c] $p<0.1$, [b] $p<0.05$, [a] $p<0.01$

NOTES. This table reports the estimations of a probit regression of the indicator variable taking the value 1 if the player cooperated. The image concern measure has been divided by its standard deviation for easier interpretation of the magnitudes. Only observations from the last game are included. Column 1 includes only the observations from sessions where the game was played without observers and column 2 those where the players are observed. We control for socioeconomic characteristics, attitudes towards risk, response time and whether the player knew the observer.

TABLE 11: RESULT OF BENJAMINI-HOCHBERG (1995) MULTIPLE TESTING CORRECTION

Table	Coefficient	Point estimate	p-value	q-value
3 (1)	Image concern (μ)	0.50	0.01*	0.04*
3 (2)	Image concern (μ)	7.98	<0.01*	0.04*
4 (1)	Female	-0.45	0.91	0.94
	Age	0.12	0.56	0.73
	In a relationship	2.22	0.57	0.74
	Student	-8.01	0.20	0.45
	French	-3.59	0.55	0.73
	Knows the observer	5.39	0.43	0.65
5 (1)	Image concern (μ)	-0.83	<0.01*	0.03*
5 (2)	Image concern (μ)	-0.46	0.01*	0.04*
5 (3)	Image concern (μ)	0.11	0.47	0.67
6 (11)	Non-French × observer French	19.61	<0.01*	<0.01*
	French × observer non-French	13.50	<0.01*	<0.01*
	French × observer French	13.12	<0.01*	<0.01*
6 (2)	Observer female	0.05	0.98	0.98
	Observer age	-0.07	0.31	0.57
	Observer french	2.49	0.37	0.61
6 (5)	Male × observer female	2.97	0.30	0.57
	Female × observer male	4.93	0.21	0.45
	Female × observer female	2.51	0.45	0.67
6 (8)	>24y old × observer <=24y old	-1.19	0.70	0.83
	<=24y old × observer >24y old	-6.83	0.11	0.30
	<=24y old × observer <=24y old	-4.00	0.34	0.61
7 (1)	Cooperate	2.94	<0.01*	<0.01*
7 (3)	Cooperate	1.78	0.01*	0.04*
	Cooperate after player cooperated	1.60	0.02*	0.09*
8 (3)	Image concern (μ)	-0.77	0.01*	0.04*
9 (3)	Observed	1.06	<0.01*	<0.01*
	Image concern (μ)	-0.62	0.02*	0.09*
	Image concern× observed	0.64	0.03*	0.11
10 (3)	Image concern (μ)	-0.11	0.59	0.75
	Cooperated in previous period	1.61	<0.01*	<0.01*
	Image con. × coop. in prev. round	-1.48	<0.01*	0.01*
	Image × coop. prev. × observed	2.45	<0.01*	<0.01*
	Observed	1.00	<0.01*	0.02*

Control variables not reported here but included in the correction procedure. p/q-values ≤ 0.1 are starred.

NOTES. This table reports the coefficients of interest from previous regressions along with their associated p-values and the q-values obtained when adjusting for multiple testing using the Benjamini-Hochberg (1995) procedure. The first column indicates the number and column of the table from which the coefficient is taken.
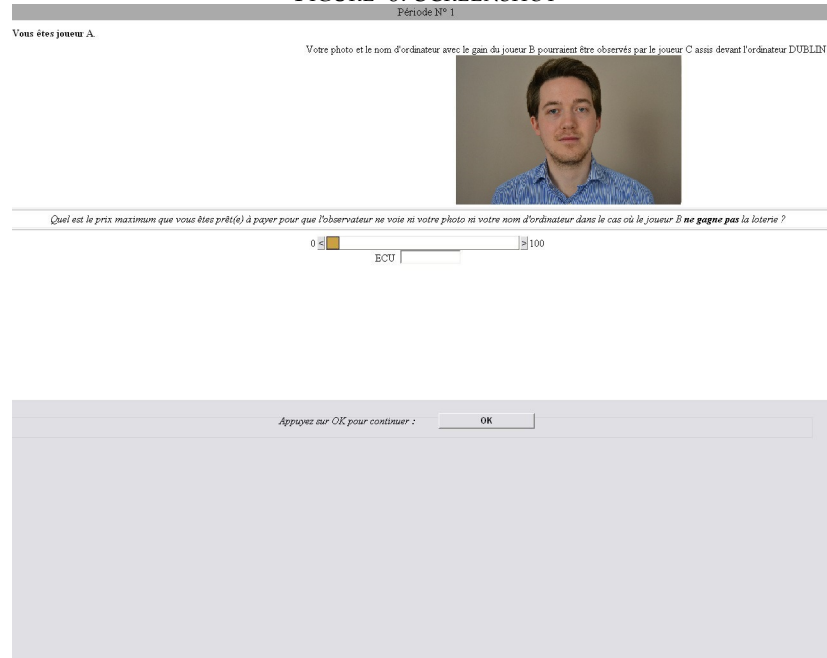
NOTES. Screenshot of the software used in the experiment. This screen reads "You are player A. Your photo and computer name could become visible to the player C sitting at computer 'Dublin' [photo of player C]. What is the maximum price you are willing to pay so that this observer does not see your picture and computer name in case player B does not win the lottery? Click 'OK to continue."

FIGURE 7: RATING OF BEHAVIOR BY OBSERVER DEPENDING ON PAST CHOICES OF PARTNER
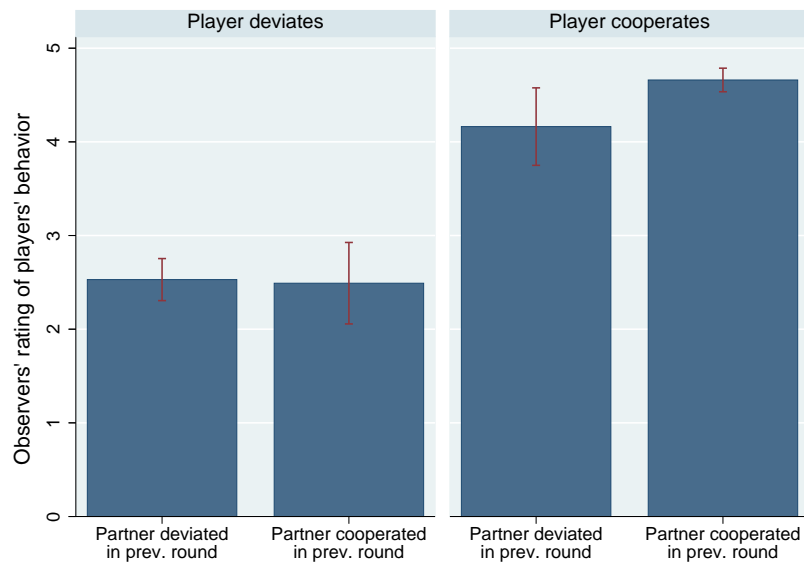
NOTES. This figure reports average ratings of players' decisions in the infinitely repeated prisoner's dilemma game by outside observers. 95% confidence intervals in red.

References

Algan, Y., Benkler, Y., Henry, E., and Hergueux, J. (2014). Social Motives and the Organization of Production : Experimental Evidence from Open Source Software. *Working paper*.

Algan, Y., Benkler, Y., Morell, M. F., and Hergueux, J. (2013). Cooperation in a peer production economy - experimental evidence from Wikipedia. *Working Paper*.

Ali, N. S. and Bénabou, R. (2016). Image Versus Information: Changing Societal Norms and Optimal Privacy. *NBER Working Paper*.

Andreoni, J. and Bernheim, B. D. (2009). Social Image and the 50-50 Norm: A Theoretical and Experimental Analysis of Audience Effects. *Econometrica*, 77(5):1607–1636.

Andreoni, J. and Petrie, R. (2004). Public goods experiments without confidentiality: A glimpse into fund-raising. *Journal of Public Economics*, 88(7-8):1605–1623.

Ariely, D., Bracha, A., and Meier, S. (2009). Doing good or doing well? Image motivation and monetary incentives in behaving prosocially. *American Economic Review*, 99(1):544–555.

Becker, G. M., DeGroot, M. H., and Marschak, J. (1964). Measuring utility by a Single Response Sequential Method. *Systems Research and Behavioural Science*, 9(3):226–232.

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

Bénabou, R. and Tirole, J. (2012). Laws and Norms. *IZA Discussion Paper*, 6290:1–44.

Benjamini, Y. and Hochberg, Y. (1995). Controlling the false discovery rate: a practical and powerful approach to multiple testing. *Journal of the Royal Statistical Society B*, 57(1):289–300.

Bracha, A. and Vesterlund, L. (2013). How Low Can You Go ? Charity Reporting When Donations Signal Income and Generosity. *Working paper*.

Bursztyn, L. and Jensen, R. (2015). How does peer pressure afect educational investments? *Quarterly Journal of Economics*, 130(3):1329–1367.

Carpenter, J. and Myers, C. K. (2010). Why volunteer? Evidence on the role of altruism, image, and incentives. *Journal of Public Economics*, 94(11-12):911–920.

Charness, G., Rigotti, L., and Rustichini, A. (2007). Individual behavior and group membership. *American Economic Review*, 97(4):1340–1352.

Dal Bó, P. (2005). Cooperation under the Shadow of the Future: Experimental Evidence from Infinitely Repeated Games. *American Economic Review*, 95(5):1591–1604.

Dal Bó, P. and Fréchette, G. R. (2011). The Evolution of Cooperation in Infinitely Repeated Games. *American Economic Review*, 101(1):411–429.

Dana, J., Cain, D. M., and Dawes, R. M. (2006). What you don't know won't hurt me: Costly (but quiet) exit in dictator games. *Organizational Behavior and Human Decision Processes*, 100(2):193–201.

Dana, J., Weber, R. A., and Kuang, J. X. (2007). Exploiting moral wiggle room: Experiments demonstrating an illusory preference for fairness. In *Economic Theory*, volume 33, pages 67–80.

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

Dohmen, T., Falk, A., Huffman, D., Sunde, U., Schupp, J., and Wagner, G. G. (2011). Individual risk attitudes: Measurement, determinants, and behavioral consequences. *Journal of the European Economic Association*, 9(3):522–550.

Ekström, M. (2012). Do watching eyes affect charitable giving? Evidence from a field experiment. *Experimental Economics*, 15(3):530–546.

Ellingsen, T. and Johannesson, M. (2008). Pride and Prejudice: The Human Side of lncentive Theory. *American Economic Review*, 98(3):990–1008.

Fehr, E. and Fischbacher, U. (2004). Third-party punishment and social norms. *Evolution and Human Behavior*, 25(2):63–87.

Fehr, E., Rosenbladt, B. V., Wagner, G. G., Fischbacher, U., and Schupp, J. (2003). A Nation-Wide Laboratory : Examining Trust and Trustworthiness by Integrating Behavioral Experiments into Representative Surveys. *IZA Discussion Paper*, (715).

Fershtman, C. and Gneezy, U. (2001). Discrimination in a segmented society: An experimental approach. *Quarterly Journal of Economics*, 116(1):351–377.

Fudenberg, D., Rand, D. G., and Dreber, A. (2012). Slow to anger and fast to forgive: Cooperation in an uncertain world. *American Economic Review*, 102(2):720–749.

Glaeser, E. L., Laibson, D. I., Scheinkman, J. A., and Soutter, C. L. (2000). Measuring Trust. *Quarterly Journal of Economics*, 115(3):811–846.

Lacetera, N. and Macis, M. (2010). Social image concerns and prosocial behavior: Field evidence from a nonlinear incentive scheme. *Journal of Economic Behavior and Organization*, 76(2):225–237.

List, J. A., Shaikh, A. M., and Xu, Y. (2016). Multiple Hypothesis Testing in Experimental Economics. *NBER Working Paper*.

Mellström, C. and Johannesson, M. (2008). Crowding Out in Blood Donation: Was Titmuss Right? *Journal of the European Economic Association*, 6(4):845–863.

Peysakhovich, A. and Rand, D. G. (2015). Habits of virtue: creating norms of cooperation and defection in the laboratory. *Management Science*, 62(3):631 – 647.

Rege, M. and Telle, K. (2004). The impact of social approval and framing on cooperation in public good situations. *Journal of Public Economics*, 88(7-8):1625–1644.

Riyanto, Y. E. and Zhang, J. (2015). Putting a price tag on others' perceptions of us. *Experimental Economics*, 19(2):480–499.

Samek, A. S. and Sheremeta, R. M. (2014). Recognizing contributors: an experiment on public goods. *Experimental Economics*, 17(4):673–690.

Sapienza, P., Toldra-Simats, A., and Zingales, L. (2013). Understanding trust. *Economic Journal*, 123(573):1313–1332.

Sutter, M., Lindner, P., and Platsch, D. (2009). Social norms, third-party observation and thirdparty reward. *Working paper*.