Peer Approval Concerns Promote Cooperation in Public Goods Experiments

Marco Faillo[a] Daniela Grieco[b] Luca Zarri[c]

Pre-print version

Abstract

In this study, we experimentally analyze the effectiveness of peer-to-peer ratings as a cooperation enforcement device in a finitely repeated public goods game setting. In particular, we wonder whether players approve of others' contribution decisions in the expected direction and whether, even under anonymity and in the absence of monetary consequences, mechanisms based on peer approval encourage cooperation. We run two treatments that differ for the amount of information on rating points provided to the players, whereas, in a third treatment, we analyze peer approval when assigning ratings to others is costly. Our findings reveal that, in the two core treatments, peer approval concerns lead to higher contributions and efficiency, compared to our control. Introducing a small fixed cost for assigning rating points results in a very high percentage of subjects deciding not to rate others' behavior.

JEL Classification: D02; D03 ; D63.

Keywords: Peer Approval; Economic Experiments; Cooperation; Public Goods Games.

[a] Department of Economics and Management, University of Trento.

[b] Department of Law Cesare Beccaria, University of Milano.

[c] Department of Economics, University of Verona.

We thank James Andreoni, Stefano Barbieri, Ted Bergstrom, Werner Güth, Maria Vittoria Levati, Ginevra Marandola, Joseph Tao-yi Wang, Adam Zylbersztjn and participants at the 2017 NYUAD-APET Workshop on "At the Forefront of Public Economics" and at the 2018 Behavioral and Experimental Economics Network (BEEN) Workshop in Bologna for helpful suggestions and comments. The usual caveats apply. We gratefully acknowledge the University of Verona (2011 Joint Project on "Punishment and Decision-making: Neuroeconomic Foundations, Behavioural Experiments and Implications for Law and Economics"), the Department of Economics and Management of the University of Trento and the Department of Economics of the University of Verona for financial support.

1. Introduction

In social dilemmas, a tension exists between individual payoff maximization and maximization of social welfare. Daily life abounds in examples of collective action problems where everyone would be better-off when universal cooperation is achieved, compared to a zero cooperation outcome, but a strong temptation to free ride on others' efforts arises. These span a series of seemingly unrelated interaction problems including reducing CO2 emissions by driving fuel-efficient cars, volunteering on school boards or town councils, helping to maintain the roads and fire departments in our cities and collaborating to teamwork in the office or within a firm.

Recent research on voluntarily provided public goods has been increasingly shedding light on the link between individuals' decisions over whether and how much to contribute and their *peers' attitude and behavior towards giving* (see e.g., Kumru and Vesterlund, 2010; Castillo et al., 2014). As pointed out by Castillo et al. (2014), the advent of social media generates increased opportunities for peer-to-peer fundraising that might benefit charitable organizations. As they note, online giving has been significantly growing over time, together with the number of platforms and tools that promote donations, and peer-to-peer components that encourage solicitation through individual networks (such as friends asking friends to give) seem to play a crucial role in driving their success. In light of this, in our study, we sought to contribute to the rapidly expanding stream of experimental literature (reviewed in Section 2) examining the effectiveness of various forms of sanction-based and reward-based institutional arrangements on cooperation by focusing on the role of decentralized, *non-pecuniary* mechanisms based on peer *approval*.

The human desire to get social approval by other people has been extensively studied within the social sciences. In economics, prominent scholars such as Bernard de Mandeville (1714) and Adam Smith (1759) have shed light on the importance for human beings of being approved of by others. In the last years, with the advent of social media, our search for peer approval has taken new forms and we cannot even rule out that new technologies have been magnifying our "need" to be socially recognized by our peers for who we are (including what we pretend to be in the eyes of others) as well as for what we do[1].

Therefore, apart from purely strategic and purely altruistic considerations (see on this Andreoni's (1989; 1990) classical studies), a plausible reason why people might decide to voluntarily engage in prosocial activities such as contributing to public goods is that they care about others' views. In this study, we test the idea that one's willingness to contribute may positively and significantly depend on receiving non-monetary benefits associated with *knowing that other individuals in the group (i.e., other peers) like what she is doing*. For example, in teamwork or volunteering on local parks maintenance, it might be the case that, for a single employee or volunteer, having her effort towards a common goal being praised by

[1] In this regard, a few weeks ago Facebook co-founder Sean Parker interestingly declared that Facebook exploits human weakness and creates a *social-validation feedback loop*: "we need to sort of give you a little dopamine hit every once in a while, because someone liked or commented on a photo".

her peers works in motivating her to contribute more. As a consequence, this may turn into higher cooperation and welfare, compared to groups in which peer approval plays no role. Therefore, the core question we address in this paper is the following: does asking individuals to express their approval of their peers' behavior in a public goods game environment significantly affect contribution levels and efficiency as time unfolds?

Our broad conjecture is that people might contribute to public goods also due to internalized social norms and, in particular, that the search for social approval by one's peers can be a powerful motivator. As noted by Rege and Telle (2004), classical theories of social exchange (Homans, 1961; Blau, 1964) posit that the possibility to receive social approval in return for pecuniary rewards may lead individuals to give more, in social dilemma situations: "Such exchanges take place because people are anxious to receive social approval from others. Moreover, a person receives social approval from another person if his actions imply a pecuniary reward to that person" (Rege and Telle, 2004; p. 1630).

Through our experiment, based on a finitely repeated public goods game setting, we analyze different mechanisms based on peer approval and seek to understand not only whether peer-to-peer evaluations affect contribution behavior, but also whether the public vs. private nature of information over received approval points and the costs to be incurred for assigning approval points matter. To this aim, our experimental design includes a Baseline and three treatments. In all the treatments, receiving approval points has no monetary consequences and, in two of them ("Private" and "Public"), subjects are asked to costlessly assign rating points to their co-players.

In the "Private" treatment, in each period, all players in a group know their co-players' contributions, but the information on the peer approval points received in one period is *privately provided to the player who receives them only*. In the "Public" treatment, the information on the approval points received in one period is provided to the player together with information over others' contribution levels and amount of approval points received by them in that period. A reason why knowing the amount of approval points received by others might enhance cooperation is that this information may provide relevant reference points to individuals, insofar as they are sensitive not only to receiving social approval, but also to comparisons with their peers with regard to received social approval. Finally, we analyze a third treatment that has the same information characteristics as the Private treatment, but in which subjects have to incur a *monetary cost* if they wish to rate their peers: we call it "Private with Cost". The reason why we also introduce this treatment is that it should help making the decision of rating peers more salient to the players: potential recipients of peer approval are expected to focus and care more about the rating decisions insofar as they know that their peers have incurred a cost to express approval of their behavior.

The remainder of the paper is organized as follows. Section 2 discusses prior research related to our experiment. Section 3 illustrates our experimental design and Section 4 reports our key results. Section 5 concludes.

2. Background

In the last years, a rapidly expanding stream of experimental economics literature has been exploring the effectiveness of various decentralized mechanisms based on rewards and/or sanctions as cooperation enforcement devices. Several papers focused on the role of *monetary incentives* in fostering cooperation. Fehr and Gächter's (2000; 2003) pathbreaking studies convincingly document that, whereas in non-punishment treatments contribution levels tend to decline as time unfolds, monetary punishment is an extremely powerful tool to increase cooperation in public goods games, thanks to many subjects' willingness to incur costs to decrease others' monetary payoffs. The authors show that sanctioning acts effectively both under anonymous random matching and with fixed groups playing a finite number of times. Andreoni et al. (2003) analyze demands for costly reward and punishment and their effects on cooperation in a series of simple two-person proposer-responder games. They offer evidence of substantial demands for both rewards and punishments and show that the two tools perform complementary roles: while rewards alone are relatively ineffective, combining rewards and punishments has a very strong effect on cooperation. Sefton et al. (2007) investigate both monetary punishment and monetary rewards in a public goods game experiment in which reward and punishment institutions are compared with regard to their effects on cooperation and efficiency. They find that while the presence of a reward option per se is insufficient to sustain cooperation, sanctioning is an effective mechanism for sustaining contributions. Next, when both punishment and reward are allowed to occur, a synergistic relationship between the two arises, as this treatment leads to high earnings and contribution levels[2]. Rand et al. (2009) investigate the effectiveness of monetary punishment and reward options (both separately and jointly) in promoting cooperation within a finitely repeated public goods game setting in which individual identities persist from round to round, so that reputational considerations might play a role. Their findings indicate that, when both options are available, reward is associated to higher contributions and payoffs, while sanctioning does not impact contributions and leads to lower payoffs. Faillo et al.'s (2013) VCM experiment sheds light on the role of "legitimate punishment" (i.e. in each round subjects are allowed to punish only co-players who contribute to the public good less than they do) in leading to higher contributions and welfare over time.

While the aforementioned experimental papers examined the role of monetary punishments and rewards in promoting cooperation, other studies have been shedding light on the power of different *non-pecuniary factors* in increasing contribution levels and efficiency in social dilemmas. In Masclet et al.'s (2003) VCM experiment, subjects can assign from 0 to 10 punishment points to their co-players in the

[2] Sefton et al. (2007) also document that high contributors (i.e., players who contribute more than the group average) are more willing to use rewards and sanctions and more likely to receive rewards, whereas low contributors are more likely to receive sanctions. In later rounds, rewards are used less frequently than sanctions.

group, so having the opportunity to express *disapproval* of others' contribution decisions. Their work documents that, even though punishments points are costless to assign and have no direct effect on final earnings, they succeed in significantly raising contribution levels. Gächter and Fehr (1999) is an early experiment on social approval and public goods provision and shows that while approval incentives alone do not effectively deter free riding, they significantly enhance cooperation when combined with some minimal social familiarity. Noussair and Tucker (2005) examine the impact of both formal and informal sanctions on public goods provision, finding that contributions and welfare are higher under a sanctioning system combining formal and informal penalties, compared to institutional arrangements based on one of the two types of sanctions only. López-Pérez and Vorsatz (2010) run a Prisoner's Dilemma experiment aimed at investigating the role of approval and disapproval in affecting players' behavior in the absence of monetary consequences. Their results suggests that disapproval aversion is an important force in leading players to cooperate. They also offer evidence that actual cooperators approve (resp., disapprove) of potential cooperators (defectors) more than actual defectors do. Dugar (2010) focuses on the impact of nonmonetary sanctions and rewards (proxied by performance-based disapproval and approval ratings, respectively) in an experimental game with Pareto-ranked equilibria. He finds that even though these ratings are costless and payoff neutral, induced disapproval ratings are more effective than approval ratings in facilitating players' coordination on the most efficient equilibrium. Dugar (2013) analyzes (non-pecuniary) approval and disapproval ratings in a linear public goods game with three treatments in which subjects can assign approval or disapproval ratings to other participants. As he notes, his findings seem to suggest that incentives induced via the social disapproval scheme are more effective in deterring free riding than the ones induced via the social approval device. In a recent laboratory experiment, Simpson et al. (2017) examine the links between interpersonal expression of (positive and negative) moral judgments and cooperation and show that groups whose members are allowed to morally judge others turn out to be better able to foster cooperation compared to groups with no capacity to sanction. Kumakawa (2013) allows players in a voluntary contribution mechanism to send a free-form written message to their co-players after observing their contribution decision. However, his results reveal that the opportunity to praise or blame per se is not enough to sustain cooperation, while, when subjects actually experience being blamed, they substantially increase cooperation in the following round. By contrast, the experience of being praised seems to be ineffective in raising cooperation[3]. This finding is in line with Masclet et al. (2003), showing that, when disapproval points can be assigned, cooperation increases (see on this also Carpenter and Seki, 2011).

[3] Peeters and Vorsatz (2013) analyze non-monetary rewards and sanctions in a VCM experiment and show that both tools raise cooperation only when individuals interact within the same group, but also that the effect is insignificant.

In our study we sought to contribute to the existing experimental literature on public goods provision by focusing on the role of decentralized mechanisms based on peer *approval* only[4]. Importantly, as we make clearer in the next section, a key feature of our institutional arrangement is that we do this without having recourse to any form of contributors' identification, i.e., we preserve *anonymity* and rule out *reputational* concerns. Next, through our three-treatment design, we examined the role that private vs. public information over received ratings as well as the cost of approving of others' behavior can play in fostering cooperation.

3. The Experiment

3.1. Experimental design

Our experiment consists of a finitely repeated (20 rounds) voluntary contribution mechanism As we explain below, overall the experiment consists of three treatments (Private, Public, and Private with Cost) with two stages (a contribution stage and a rating stage), plus a control (Baseline), where only the contribution stage occurs and there is no assessment of peers' behavior through the assignment of bullet points. In the three treatments, assigning rating points to one's peers in the rating stage has no monetary consequences for the recipients.

3.2. Treatments

In all treatments subjects are informed about the vector of individual contributions in their group in each period. The key distinguishing features of the treatments are as follows: (i) in a first treatment (Private), subjects are asked to *costlessly* assign rating points to their peers in each period and the information on the rating points received in one period is *privately provided only to the subject* who receives them in that period; (ii) in a second treatment (Public), subjects are asked to *costlessly* assign rating points to their peers in each period and the information on the rating points received in one period is provided to the subject who receives them in that period *and to all the other members of the group*; (iii) in a third treatment (Private with Cost), subjects have to *pay a cost* if they wish to assign rating points to others and the information on the rating points received in one period is *privately provided only to the subject* who receives them in that period. Additionally, we run a control where only the contribution stage is played and participants are prevented from assigning rating points to others (Baseline).

[4] Holländer (1990) proposed an early formal model of social exchange in which cooperative behavior is driven by the expectation of (emotionally prompted) social approval.

For all treatments, in each session there were 20 periods of interaction, which proceeded under identical rules. The participants in a session were randomly assigned to groups of size four, so that they did not know the identities of the other members of their group. We used a partner protocol that kept the composition of each group constant over rounds, so that, at the end of each period, individuals remained in the same group. However, to prevent the formation of individual reputation histories, individuals' labels were reassigned on a random basis in each period. For example, the same player could be designated as player 13 in period t, as player 5 in period t + 1, and as player 24 in period t + 2. Therefore, our partner protocol was also characterized by anonymity of the components of the group and change of participants' labels across rounds.

At stage 1, at the beginning of each period all participants receive a fixed amount $e = 20$ of tokens. Each participant i has to decide whether she wants to invest into a public project or not an amount $g_i \leq e$. Decisions are made simultaneously and with no information about peers' choices. At the end of stage 1, each participant is informed about the contribution of other group members and of her current earnings, which consist of two components:

a. The proportion of her initial 20 tokens that she has kept for herself (i.e. 20 tokens – Her contribution to the project);
b. Her income from the project. The income to her is equal to 40% of the total of the four individual contributions to the project.

Therefore, her earnings at the end of stage 1 are calculated by the computer in the following way:

Each participant's earnings after stage = (20 – her contribution to the project) + 40%(total group contribution to the project)*

$$= (20 - g_i) + 0.4\sum_{j=1}^{4} g_j$$

In the Baseline, each round is made of this stage only, whereas, as we anticipated above, the three treatments (Private, Public, and Private with Cost) are also characterized by a second stage in which rating decisions take place. In all three treatments, receiving one or more rating points has no impact on the recipient's earnings. The other main features of these treatments are summarized here below.

Private Treatment. At stage 2, subjects are asked to simultaneously rate the behavior at the stage 1 (contribution stage) of each of their peers by assigning points from zero ("Not positive at all") to ten

("Totally positive")[5]. Individuals may differ in the personal weight they put on rating points. However, we follow the abundant literature on punishment in assuming that subjective interpretations will on average be consistent with the zero-ten scale subjects receive, and whose meaning is clearly stated in the instructions. Assigning rating points is costless. Each subject is provided with information on the total number of rating points she received in the current stage, but has no information on the number of rating points received by her three group co-players.

Public Treatment. At stage 2, subjects are asked to simultaneously rate the behavior at the stage 1 (contribution stage) of each of their peers by assigning points from zero ("Not positive at all") to ten ("Totally positive"). Each subject is provided with information on the total number of rating points she received in the current stage, plus information on the total number of rating points received by her three group co-players.

Private with Cost Treatment. Each subject is provided with information on the total number of rating points she received in the current stage, but has no information on the number of rating points received by her three group co-players. Assigning rating points in this case is costly: subjects have to pay one token for each peer they wish to rate (using the same 0-10 range as in Private and Public, for each co-player), regardless of the number of rating points they assign to each peer.

A synthetic description of the treatments is reported in Table 1.

[5] As to rating decisions in treatments Private and Public, the experimental instructions were as follows: "You will be asked to assign a certain number of points (from 0 to 10) to each group member. The number of points will be positively associated to your rating of her behavior in the previous stage of the game (0 points: not positive at all; 10 points: totally positive)". Our Private treatment is similar to one of Dugar's (2013) treatments, but a relevant difference is that, since we focus on reaction to approval points rather than on points assignment, we ask subjects to rate others, while in Dugar's study subjects can decide whether to approve of others' behavior or not.

Table 1. Treatments

Treatment	**Information**	**Cost of Assigning Rating Points**	**Impact on Recipients' Earnings**
Private	Other members' contribution and own rating	Zero	No
Public	Other members' contribution and all ratings	Zero	No
Private with Cost	Other members' contribution and own rating	1 token for each peer they wish to rate	No
Baseline	Other members' contribution	N.A.	N.A.

3.3. Experimental procedure

A total of 224 subjects took part voluntarily in the experiment, that has been conducted at the CEEL Lab of the University of Trento between June 2015 and November 2017. We ran three sessions (56 participants per treatment) for the Baseline, the Private, the Public and the Private with Cost treatment. The experiment was programmed by using the zTree platform (Fischbacher, 2007). The subjects were university students (55% from Economics and Business majors, 45% females, 94% Italian; average age: 21.4). We implemented a between-subjects design, so that no individual participated in more than one session. In each session, participants received a show-up fee of 3 euros, plus their earnings from the experiment. The average final payment per participant was 12.90 euros (show-up fee included) and the sessions averaged approximately 1 hour and 30 minutes. Once all of them were seated, the instructions were handed to them in written form before being read aloud by the experimenter. We took great care to ensure that the participants understood the rules of the game: they had to answer several control questions and we did not proceed with the actual experiment until all participants had answered all questions correctly.

4. Results

In this section we illustrate our main experimental results: first, we provide an overview of the dynamics of contributions, earnings and ratings across treatments; then we focus on the effectiveness of rating behavior and its role in explaining this dynamics.

4.1. Contribution

Figure 1 displays the average levels of groups' contribution in the twenty periods of interaction in the four experimental conditions. Table 2 reports the average contributions of the groups across treatments.

Figure 1. Average group contribution

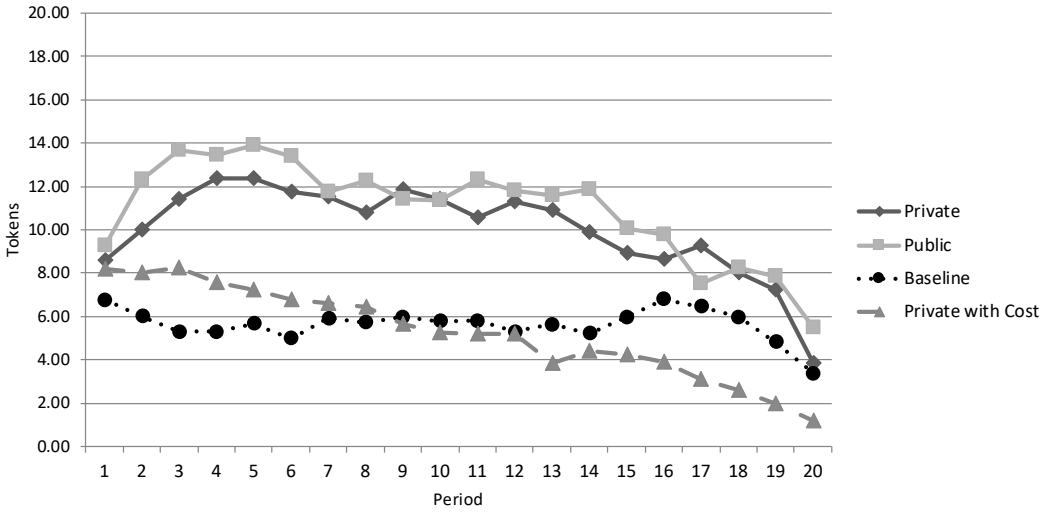


Table 2. Mean contributions

Group	Baseline	Private	Public	Private with Cost
1	7.18	2.68	7.26	10.39
	(2.06)	(3.84)	(2.71)	(7.32)
2	12.41	19.75	14.89	9.80
	(6.67)	(1.12)	(3.20)	(5.52)
3	3.24	4.30	12.89	10.01
	(3.84)	(2.61)	(5.38)	(5.41)
4	5.39	11.86	4.91	1.36
	(2.80)	(6.75)	(5.10)	(2.50)
5	16.29	8.71	15.58	4.25
	(5.56)	(3.72)	(4.46)	(1.69)
6	5.24	18.94	18.05	1.78
	(5.00)	(3.47)	(2.45)	(3.39)
7	2.90	3.10	8.01	1.58
	(1.35)	(1.45)	(4.87)	(1.34)
8	8.49	11.50	7.00	2.33
	(3.96)	(4.25)	(5.01)	(1.36)
9	6.08	5.04	2.20	5.05
	(2.89)	(3.85)	(3.05)	(4.29)
10	3.45	14.25	19.04	6.43
	(1.98)	(4.42)	(2.72)	(3.68)
11	1.31	17.55	14.68	2.89
	(1.38)	(4.61)	(4.56)	(1.78)
12	1.26	12.75	13.75	2.74
	(1.56)	(5.68)	(6.29)	(1.81)
13	4.04	5.13	7.39	8.51
	(2.48)	(1.95)	(1.59)	(2.09)
14	1.64	4.83	7.81	6.83
	(2.53)	(2.40)	(3.18)	(1.76)
Mean	5.63	10.03	10.96	5.28

Stardard deviations in parentheses.

We find that our two key treatments (Private and Public), in which (costless) rating decisions follow contribution decisions in each round, succeed in significantly raising cooperation, compared to our Baseline (Wilcoxon-Mann-Whitney two-tailed test taking average contributions of the groups as

independent observations[6]: Private vs. Baseline: z=1.88, p=0.059; Public vs. Baseline: z=2.71, p=0.006). We detect no significant difference in contributions between Private and Public (z=-0.64, p=0.52). Finally, we see that contribution levels are significantly higher in Private than in Private with Cost (z=2.30, p=0.02).

Result 1a*: The introduction of peer rating with public information (Public) has a positive effect on contributions, as compared with the Baseline in which no rating occurs.*

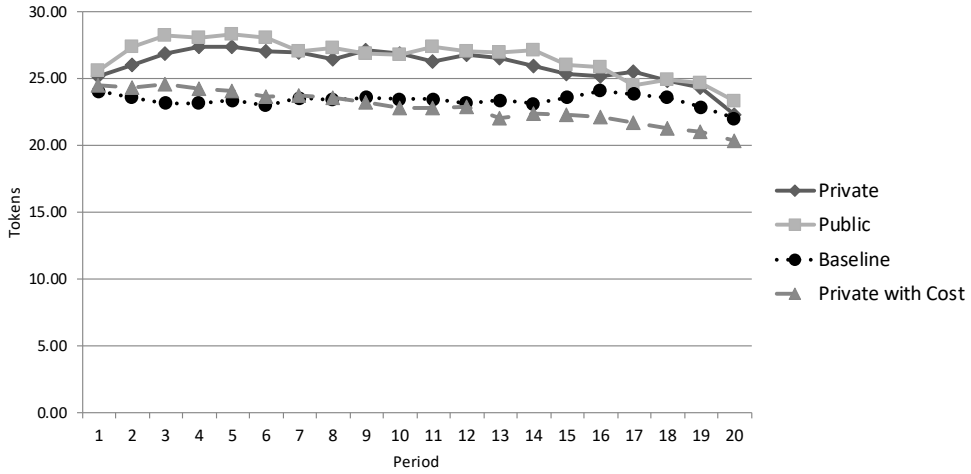
Result 1b*: Under private information, the introduction of peer rating (Private) has a positive effect on contributions, compared to the Baseline. However, the difference in contributions between Private and Baseline is smaller than the one detected comparing Public and Baseline.*

Result 1c*: The introduction of a small fixed cost for assigning rating points (Private with Cost) leads to a significant decrease in the level of contribution, compared to Private.*

4.2. Earnings

Figure 2 reports average earnings in the twenty periods of interaction.

Figure 2. Average group payoffs



The dynamics of earnings over time mirrors the dynamics of contributions: there is no significant difference between Private and Public (z=-0.64, p=0.52), whereas the difference is significant between Private and Baseline (z=1.88, p=0.06) and between Public and Baseline (z=2.71, p=0.006). Next, earnings in Private are significantly higher than in Private with Cost (z=2.52, p=0.01).

[6] In the remainder of the section we will always report the results of this type of test.

Result 2a*: The introduction of peer rating with public information (Public) leads to higher earnings, as compared with the Baseline in which no rating occurs.*

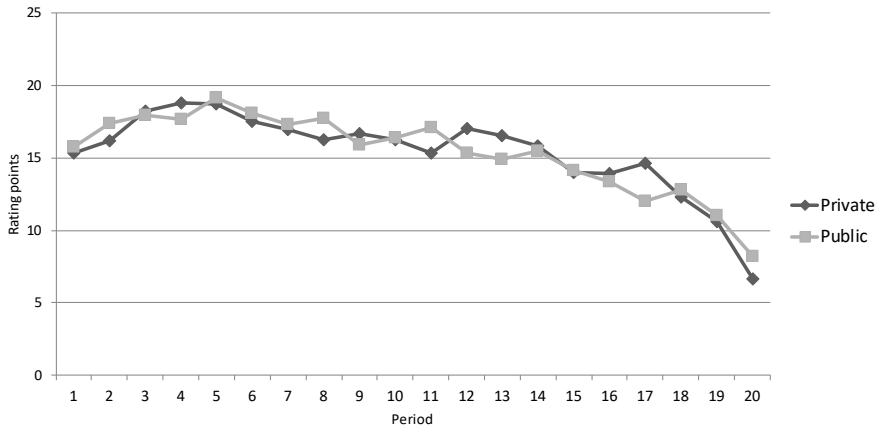
Result 2b*: Under private information, the introduction of peer rating (Private) leads to higher earnings, compared to the Baseline. However, the difference in earnings between Private and Baseline is smaller than the one detected comparing Public and Baseline.*

Result 2c*: The lowest level of earning is observed in the Private with Cost treatment.*

4.3. Rating behavior

Finally, Figure 3 illustrates the way subjects distribute rating points over time in the two treatments in which assigning ratings is costless.

Figure 3. Average number of rating points assigned in the Private and Public treatments



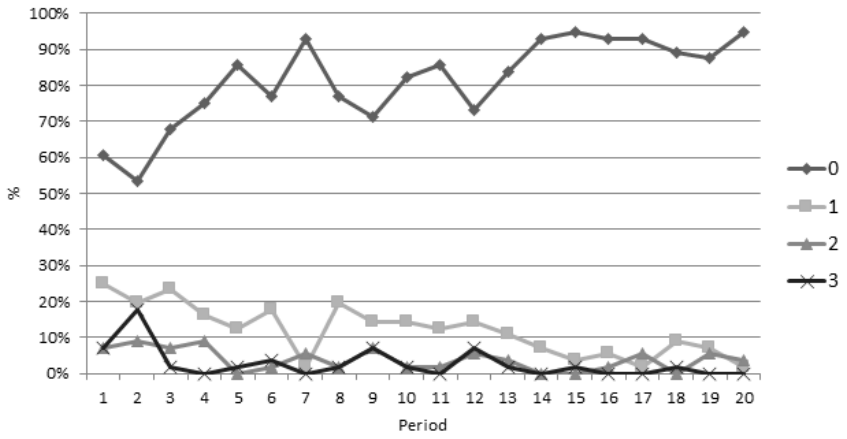
In the case of treatments in which rating assignment is costless (Private and Public), we observe a general decline in the average rating that mirrors the decline in the level of cooperation, with no significant differences in the use of rating points (Private vs. Public: z=0.13, p=0.89).

Result 3a*: The dynamics of assigned rating points are similar in Private and Public and display a decreasing pattern over time, in line with the dynamics of contributions in the two treatments.*

As for the treatment with costly rating, the most relevant evidence regards the high percentage of subjects deciding *not to rate* their peers. Since subjects pay a fix sum (equal to one token) for each peer they decide to rate, the rating costs they face range from 0 (no peers evaluated) to 3 (all the three peers are evaluated. Figure 4 illustrates the dynamics of the distribution of rating costs over time.

***Result 3b**: The introduction of a small fixed cost for ratings results in a very high and substantially increasing percentage of subjects deciding not to use the rating option.*

Figure 4. Distribution of ratings costs in the Private with Cost treatment



4.4. Regression analysis

Table 3 investigates the determinants of contribution levels when considering treatment effects and adding demographic controls.

Table 3. Determinants of contribution levels

VARIABLES	
Baseline (β_{Base})	-3.151**
	[1.597]
Public (β_{Public})	0.652
	[1.832]
Private with Cost	-4.514***
	[1.453]
Average_first	0.543***
	[0.082]
Constant	8.497***
	[2.636]
β_{Public}-β_{Base}	3.802**
	[1.623]
Observations	4,460
Number of subjects	223

Random Effects GLS (clustered standard errors in parentheses).
The dependent variable takes values from 0 to 20. Baseline, Public and Private with Cost are dummy variables assuming value equal to 1 in the corresponding treatment and 0 elsewhere. Average_first represents the average contribution level of the group in the first period. Controls: age, nationality, major, gender and number of experiments in which the subject has been involved in the past.
Estimation of a Tobit model with clustered standard errors and pooled data provides the same results.
*** significant at 1%; ** significant at 5%; * significant at 10%.

The findings reported in Table 3 confirm the results about contribution of section 4.3: there is no significant difference in contribution levels between the Private treatment (taken as reference) and the Public. In contrast, the Baseline treatment displays a significantly lower level of contributions either when compared to the Private or to the Public treatment. There is also a significant difference between the Private and the Private with Cost treatments: average contribution levels in Private with Cost are significantly lower.

Table 4 focuses on subjects' rating behavior and its impact on contribution levels. Our conjecture is that rating is effective if (a) high contributors assign a low amount of rating points to low contributors and (b) low contributors assign a high amount of rating points to high contributors. In the set of regressions reported above, we classify subjects' behavior as assigning a "high amount of rating points" if the subject assigns at least 15 rating points to her peers (over a maximum amount of 30, i.e. a maximum of ten rating points to each peer). We consider the whole sample of subjects (columns 1), high contributors only (column 2) and low contributors only (column 3), respectively.

Table 4. Probability of assigning a high amount of rating points

VARIABLES	(1) all	(2) high contributors	(3) low contributors
Public	-0.198	-0.298	-0.042
	[0.212]	[0.219]	[0.247]
Distance_from_average	0.008	-0.056***	0.070***
	[0.008]	[0.017]	[0.020]
Average_contribution_others	0.096***	0.112***	0.102***
	[0.016]	[0.018]	[0.021]
Rating points given_span			
Constant	0.951	1.322	0.860
	[0.907]	[0.943]	[1.034]
Observations	4,460	2,513	1,947

Probit (clustered standard errors in parentheses).
The dependent variable is a dummy variable assuming value equal to 1 if the subject assigns a high (≥15) amount of rating points, and 0 elsewhere. Public *and* Private with Cost *are dummy variables assuming value equal to 1 in the corresponding treatment and 0 elsewhere.* Average_contribution_others *is the average contribution of the subject's peers in the group.* Distance_from_average_contr_others *is calculated as the difference between the subject's contribution level and* Average_contribution_others.
High contributors *are subjects whose contribution in period* t *is higher than or equal to the average contribution of other peers in* t. Low contributors *are subjects whose contribution in* t *is lower than the average contribution of other peers in* t.
Controls: age, nationality, major, gender and number of experiments in which the subject has been involved in the past.
Estimation of a Tobit model with clustered standard errors and pooled data provides the same results.
**** significant at 1%; ** significant at 5%; * significant at 10%.*

The probability of assigning a high amount of rating points is generally not significantly different when comparing the Public treatment to the Private. In this table and in the next one, we do not consider the Private with Cost because only a very small percentage of subjects assigns rating points.

Furthermore, the probability of assigning a high rating increases significantly in the peers' average contribution level: the more peers are cooperative, the higher the probability that the subject assigns a high amount of rating points. Both these effects are robust to all the specifications we present. Interestingly, the distance between the subject's contribution and her peers' average one plays a role when distinguishing between high (column 2) and low contributors (column 3). For high contributors, the lower the distance between the subject's contribution and the average contribution of her peers in the group, the higher the probability to assign rating points; for low contributors, the higher the distance between the subject's contribution and the average contribution of her peers in the group, the higher the probability to assign rating points. This means that high contributors assign a lower amounts of points to low contributors, and low contributors assign a higher amount of points to high contributors: when using rating points, subjects turn out to do it in a way that makes the rating mechanism effective. This is in line with our conjecture.

Result 4*: For low contributors, the probability of assigning rating points increases significantly in the distance between own contribution and peers' contribution. For high contributors, the probability of assigning rating points decreases significantly in the distance between own contribution and peers' contribution.*

Table 5 investigates the determinants of an increase in contribution level from period *t-1* to period *t*, focusing on subjects who *increased* their contribution. Column 1 focuses on the effectiveness of the amount of rating points the subject received in *t-1*; column 2 considers the role of receiving a low amount of rating points in *t-1*.

Table 5. Determinants of the increase in contribution levels

VARIABLES	(1)	(2)
Rating points received in *t-1*	-0.157***	
	[0.061]	
Rating points received in t-1*Public	-0.069	
	[0.084]	
Contribution	0.462***	0.420***
	[0.037]	[0.044]
Public	1.402	-2.302*
	[0.899]	[1.193]
Low rating received in *t-1*		1.359**
		[0.538]
Low rating received in t-1*Public		2.636*
		[1.368]
Constant	-2.948	-3.974
	[3.078]	[2.704]
Observations	1,809	2,855
Number of subject	174	223

Robust standard errors in brackets
*** p<0.01, ** p<0.05, * p<0.1
Random Effects GLS (clustered standard errors in parentheses).
The dependent variable takes values from 0 to 20, since we only consider the cases where the subject increases her contribution level or keeps it constant. Public is a dummy variable assuming value equal to 1 in the Public treatment and 0 elsewhere. Rating points received in t-1 represents the amount of rating points the subject has received in the previous period. Rating points received in t-1*Public represents the interaction between the amount of rating points the subject has received in the previous period and the dummy Public.
Low rating received in t-1 is a dummy variable assuming value equal to 1 if the subjects has received 15 rating points or less in the previous period, and 0 elsewhere. Low rating received in t-1*Public represents the interaction between the dummy Low rating received in t-1 and the dummy Public.
Controls: age, nationality, major, gender and number of experiments in which the subject has been involved in the past.
Estimation of a Tobit model with clustered standard errors and pooled data provides the same results.
*** significant at 1%; ** significant at 5%; * significant at 10%.

In general, the amount of rating points received in the previous period negatively and significantly affects the increase in contributions: receiving a lower amount of rating points leads to a significantly higher increase in subjects' contribution (Column 1). The Public treatment does not exhibit any difference with the Private treatment that we use as reference.

If we focus on the effectiveness of low ratings (when the subjects receives by her three peers an amount of points lower or equal to 15) as we do in Column 2, the effects of these rating points are still significant, but we observe a positive relation: if the subject has received a low amount of rating points in *t-1*, then the increase in contribution is significantly higher. In this case, however, the effect of receiving a low rating is significantly stronger in the Public treatment, that on average shows a significantly lower increase in contribution. Both regressions (columns 1 and 2) are run by controlling for the subject's contribution level.

Result 5*: The increase in contribution is negatively and significantly related to the amount of rating points a subject receives. In particular, when receiving a low rating, subjects do increase contributions significantly.*

5. Concluding remarks

On the whole, the experiment we conducted led us to conclude that the two peer-based approval mechanisms we examined in our core treatments are effective in raising cooperation and welfare, compared to our control. Insofar as participants are asked to (costlessly) assign rating points to their peers (as it was the case in both Private and Public), individuals do it in the expected direction (with high contributors assigning a lower amounts of points to low contributors, and low contributors assigning a higher amount of points to high contributors) and peer approval effectively fosters cooperation and efficiency. In contrast, insofar as we introduce a small fixed cost for assigning rating points to others, individuals fail to use this option and the mechanism turns out to be ineffective, so that making points assignment costly turns out to be *detrimental* to cooperation and welfare.

Prior research has convincingly documented that various forms of donors' identification can be very effective in raising contributions (Rege and Telle, 2004; Andreoni and Petrie, 2004; Bekkers and Wiepking, 2011, Meer, 2011), so that solicitation methods based on identifiability of contributors (e.g., publishing names of major donors) and peer recognition can be extremely successful (Castillo et al., 2014). In this regard, our experimental findings from the two key treatments offer *complementary* insights, showing that identification of contributors need not be an integral part of effective fundraising strategies that aim to rely on social approval concerns: costless peer-to-peer approval mechanisms may help fostering cooperation and welfare in public goods environments even under *anonymity*. We argue that, especially today, in a period in which web users' privacy concerns are increasingly relevant in many parts of the world, this finding has interesting implications for online fundraising based on peer-to-peer components.

This work might be extended in several directions. First, it would be interesting to see whether the peer-based approval mechanisms we examined in our experiment succeed in raising cooperation and welfare insofar as we introduce some form of donors' identification. Next, future research might seek to rigorously shed light on the interplays between social approval concerns and monetary incentives in affecting contribution levels and efficiency. We leave these extensions as interesting avenues for future experimental research on the determinants of cooperation in social dilemma environments.

References

Andreoni, J., 1989, Giving with Impure Altruism: Applications to Charity and Ricardian Equivalence, Journal of Political Economy, 1447-1458.

Andreoni, J., 1990, Impure Altruism and Donations to Public Goods: A Theory of Warm-Glow Giving?, Economic Journal, 100 (401), 464-477.

Andreoni, J., Harbaugh, W., Vesterlund, L., 2003, The Carrot or the Stick: Rewards, Punishments and Cooperation, American Economic Review, 93, 893-902.

Andreoni, J., Petrie, R., 2004, Public Goods Experiments without Confidentiality: A Glimpse into Fund-Raising, Journal of Public Economics, 88, 1605-1623.

Bekkers, R., Wiepking, P., 2011, A Literature Review of Empirical Studies of Philanthropy: Eight Mechanisms that Drive Charitable Giving, Nonprofit Voluntary Sector Quarterly, 40 (5), 924-973.

Blau, P.M., 1964, Exchange and Power in Social Life, John Wiley, New York.

Carpenter, J., Seki, E., 2011, Do Social Preferences Increase Productivity? Field Experimental Evidence from Fishermen in Toyama Bay, Economic Inquiry, 49, 612-630.

Castillo, M., Petrie, R., Wardell, C., 2014, Fundraising through Online Social Network: A Field Experiment on Peer-to-Peer Solicitation, Journal of Public Economics, 114, 29-35.

Dugar, S., 2010, Nonmonetary Sanctions and Rewards in an Experimental Coordination Game, Journal of Economic Behavior and Organization, 73, 377-386.

Dugar, S., 2013, Non-Monetary Incentives and Opportunistic Behavior: Evidence from a Laboratory Public Good Game, Economic Inquiry, 51, 2, 1374-1388.

Faillo, M., Grieco, D., Zarri, L., 2013, Legitimate Punishment, Feedback, and the Enforcement of Cooperation, Games and Economic Behavior, 77, 1, 271-283.

Fehr, E., Gächter, S., 2000, Cooperation and Punishment in Public Goods Experiments, American Economic Review, 90, 980-994.

Fehr, E., Gächter, S., 2002, Altruistic Punishment in Humans, Nature, 415, 137-140.

Fischbacher, U., 2007, z-Tree: Zurich Toolbox for Ready-made Economic Experiments, Experimental Economics, 10 (2), 171-178.

Gächter, S., Fehr, E., 1999, Collective Action as a Social Exchange, Journal of Economic Behavior and Organization, 39, 341-369.

Holländer, H., 1990, A Social Exchange Approach to Voluntary Cooperation, American Economic Review, 80 (5), 1157-1167.

Homans, G.C., 1961, Social Behavior: Its Elementary Forms. Harcourt Brace Jovanovich, New York.

Kumakawa, T., 2013, Social Rewards and Sanctions: A Voluntary Contribution Mechanism Experiment, Journal of Neuroscience, Psychology, and Economics, 6 (3), 189-200.

Kumru, C.G., Vesterlund, L., 2010, The Effect of Status on Charitable Giving, Journal of Public Economic Theory, 12 (4), 709-735.

López-Pérez, R., Vorsatz, M., 2010, On Approval and Disapproval: Theory and Experiments, Journal of Economic Psychology, 31, 527-541.

de Mandeville, B., 1714, The Fable of the Bees: Or, Private Vices, Publick Benefits, London, J. Roberts.

Masclet, D., Noussair, C., Tucker, S., Villeval, M. C., 2003, Monetary and Nonmonetary Punishment in the Voluntary Contributions Mechanism, American Economic Review, 93, 366-380.

Meer, J., 2011, Brother, Can You Spare a Dime? Peer Pressure in Charitable Solicitation, Journal of Public Economics, 95 (7), 926-94.

Noussair, C., Tucker, S., 2005, Combining Monetary and Social Sanctions, Economic Inquiry, 43 (3), 649-660.

Peeters, R., Vorsatz, M., 2013, Immaterial Rewards and Sanctions in a Voluntary Contribution Experiment, Economic Inquiry, 51 (2), 1442-1456.

Rand, D.G., Dreber, A., Ellingsen, T., Fudenberg, D., Nowak, M.A., 2009, Positive Interactions Promote Public Cooperation, Science, 325, 1272-1275.

Rege, M., Telle, K., 2004, The Impact of Social Approval and Framing on Cooperation in Public Good Situations, Journal of Public Economics, 88, 1625-1644.

Sefton, M., Shupp, R., Walker, J.M., 2007, The Effect of Rewards and Sanctions in Provision of Public Goods, Economic Inquiry, 45 (4), 671-690.

Simpson, B., Willer, R., Harrell, A., 2017, The Enforcement of Moral Boundaries Promotes Cooperation and Prosocial Behavior in Groups, Nature: Scientific Reports, 7, 42844.

Smith, A., 1759, The Theory of Moral Sentiments, London, A. Millar.