

Pan, Jingjing; Li, Jianbiao; Zhu, Chengkang

Working Paper

The Role of Reward in Cooperation-Enhancing and Welfare-Improving Under Imperfect information: Theory and Evidence

Suggested Citation: Pan, Jingjing; Li, Jianbiao; Zhu, Chengkang (2022) : The Role of Reward in Cooperation-Enhancing and Welfare-Improving Under Imperfect information: Theory and Evidence, ZBW - Leibniz Information Centre for Economics, Kiel, Hamburg

This Version is available at:

<https://hdl.handle.net/10419/269216>

Standard-Nutzungsbedingungen:

Die Dokumente auf EconStor dürfen zu eigenen wissenschaftlichen Zwecken und zum Privatgebrauch gespeichert und kopiert werden.

Sie dürfen die Dokumente nicht für öffentliche oder kommerzielle Zwecke vervielfältigen, öffentlich ausstellen, öffentlich zugänglich machen, vertreiben oder anderweitig nutzen.

Sofern die Verfasser die Dokumente unter Open-Content-Lizenzen (insbesondere CC-Lizenzen) zur Verfügung gestellt haben sollten, gelten abweichend von diesen Nutzungsbedingungen die in der dort genannten Lizenz gewährten Nutzungsrechte.

Terms of use:

Documents in EconStor may be saved and copied for your personal and scholarly purposes.

You are not to copy documents for public or commercial purposes, to exhibit the documents publicly, to make them publicly available on the internet, or to distribute or otherwise use the documents in public.

If the documents have been made available under an Open Content Licence (especially Creative Commons Licences), you may exercise further usage rights as specified in the indicated licence.

The Role of Reward in Cooperation-Enhancing and Welfare-Improving Under Imperfect information: Theory and Evidence

Jingjing Pan^{b†}

Jianbiao Li^{a,c†}

Chengkang Zhu^{a*}

June 30, 2022

^a School of Economics, Institute for Study of Brain-Like Economics, Shandong University, Jinan, China

^b Business School, Jinan University, Jinan, China

^c Nankai University Binhai College, Tianjin, China

*Correspondence:

Chengkang Zhu

zhuchengkang@sdu.edu.cn

[†]These authors contributed equally to this work.

Abstract

Although previous literature demonstrates that punishment is more efficient and stable than reward, in our daily life, numerous kinds of rewards permeate. One possible explanation for widely use of reward institution in practice is that it's an efficient and satisfactory way to enhance cooperation and welfare in a social dilemma situation even the contribution is hardly evaluated accurately. Nevertheless, this explanation lacks support from empirical evidence. Our study aims to examine whether the institution with reward option is an efficient and satisfactory way to solve social dilemma problems under imperfect information conditions. We show that reward institutions sustain higher cooperation levels and let participants get more welfare under imperfect information conditions. Furthermore, we find most participants to have a tendency to favor reward institutions, even when the information is highly noisy. Our study sheds light on the superiority of reward institutions over punishment institutions in a realistic world.

Keywords: Public goods games; Reward; imperfect information; Cooperation; Welfare

JEL Classification : C91 Laboratory, Individual Behavior; C92 Laboratory, Group Behavior; D01 Microeconomic Behavior: Underlying Principles; D91 Role and Effects of Psychological, Emotional, Social, and Cognitive Factors on Decision Making.

1 Introduction

The evolutionary success of humans has often been attributed to our cooperation in non-relative groups. However, a large amount of theoretical and empirical studies has documented that individually costly cooperation tends to break down in the long run, which is called as the social dilemma problem (Dietz et al., 2003; Nowak, 2006; Fischbacher & Gächter, 2010; Gächter et al., 2017). Punishment and reward are considered key mechanisms proposed solving the social dilemma, nevertheless, punishment is demonstrated to be more efficient and stable than reward (Yamagishi, 1986; Ostrom, 1992; Fehr & Gächter, 2000, 2002; Walker, 2004; Sefton, 2007; Gächter et al., 2008; Rand et al., 2009; Boyd et al., 2010; Sigmund et al., 2010; Faillo et al., 2012; Andreoni & Gee, 2012; Kamijo et al., 2014; Yang et al., 2019). Based on these findings of previous literature, it seems that reward should not have been used in fostering cooperation when punishment is available. In our daily life, however, numerous kinds of rewards permeate. For example, the partners of Alibaba are rewarded by each other from the bonus pool at the end of each year based upon their contributions to the corporation. Does Alibaba use peer reward institution because the punishment for their partner is not available? In addition, non-monetary rewards such as honor and esteem are often used by individuals for their partners who contribute more. Motivated by these observations, the basic question we address in this paper is why do groups sometimes choose the reward other than punishment in enforcing cooperation?

Social welfare is a different important consideration for groups choosing incentive institutions (i.e., reward and punishment) besides cooperation since groups with higher welfare can survive in the evolutionary process (Dong, 2019). The reward is chosen probably because it is better in improving welfare than punishment in a realistic environment. In a more realistic environment, one fundamental characteristic of group cooperation is that the amount of each member's contribution is hardly observed nor evaluated. That is, the observed information on other group member's contribution is likely to be inaccurate. Both decentralized and centralized punishments are demonstrated to be a poor tool to enforce cooperation when subjects observe other's decisions with a small amount of noise because they are not able to increase the welfare of group members (Ambrus & Greiner, 2012; Nicklisch et al., 2016; Ambrus et al., 2019). As the level of noise increases to large, both decentralized and centralized punishment institutions even reduce social welfare (Nicklisch et al., 2016). Therefore, one possible explanation for widely use of reward institution in practice is that it's an efficient and satisfactory way to improve cooperation and welfare in a social dilemma situation even the contribution is hardly evaluated accurately. Whether do reward institutions sustain higher cooperation and welfare in the imperfect information environment? And whether do subjects prefer reward institutions even if the noise level is large? We aim to examine these crucial issues.

In this paper, we first investigate how the option of costly reward affects cooperation and welfare in the imperfect information environment, in which the observed information on other group member's contribution is liable to be wrong. In particular, we investigate a voluntary contribution mechanism with reward (VCMR) game in which, after each contribution decision, the private signal of a player's contribution to another subject—that is, the information on the player's contribution announced to another subject—might differ from the true contribution of the player. Similar to Nicklisch et al. (2016), we employed two imperfect information treatments: (i) *high-noise* treatment (represents high noisy environment); and (ii) *low-noise* treatment (represents low noisy

environment). Specifically, the possibility that the private signal differs from the true contribution equals roughly 10% in the *low-noise* treatment and equals roughly 50% in the *high-noise* treatment¹. This design corresponds to collective production situations in which even if a group member contributes to a joint project, the others do not recognize the contribution. In our design, such mistakes in private signal only influence the subjects' information, but not their payoffs². As a control design, we also employed the *perfect-information* treatment, under which the private signal always matches the true contribution level.

The other dimension of our design is the reward option. We employed three institutions: (i) a reward-free institution (hereafter referred as *NR*); (ii) a decentralized reward institution (hereafter referred as *DR*); and (iii) a centralized reward institution (hereafter referred as *CR*)³. In *CR*, an additional (randomly drawn) subject (the authority) can reward the players in this institution while players are not permitted to reward each other. The authority's payoff is increasing in the players' contributions, and the authority is not required to bear the costs of rewards. In *DR* players can reward other players in the same institution, at their own expense. Hence, our experiment facilitates investigating the effects of *DR* and *CR* on cooperation and welfare in both perfect and imperfect information environments.

Furthermore, we investigate whether the reward option is preferred by group members in the imperfect information environment. In particular, we investigate which institution is chosen by the players after they have experienced three institutions in turn. For this purpose, we divided the main experiment into two parts. In the first part, we used a within-group design, in which a group of players participate in all three institutions in random order.⁴ Specifically, in each institution a player interacts with 3 other fixed group members for 5 periods. At the end of this part, players observed the accurate information about average contribution and welfare (that is, payoff that related to contributions, imposed rewards and received rewards) for each period of each institution. In the second part, players decide to participate in one of the three institutions and then complete the game with other players who choose the same institution. Since the payoff is related to the chosen institution and subsequent decisions, our experiment elicited the group member's preference for *CR*, *DR* and *NR*.

We have determined that in the benchmark *perfect-information* treatment, both *CR* and *DR* increase the amounts of contribution and the welfare. This strengthens the findings in Rand et al. (2009), a key literature demonstrating that the decentralized reward institution improves both cooperation and welfare. In *NR*, players contribute about half endowments at the beginning, and this contribution level quickly decays. *DR* not only provides a higher contribution level at the

¹ These two imperfect information treatments represent two typical information environments. We employed the *low-noise* treatment to examine whether the positive effects of reward on cooperation and welfare and subjects' preference for reward institutions are undermined by a slight noise. In addition, we employed the *high noise* treatment to examine whether reward institutions are still efficient and satisfactory to improve cooperation and welfare under a relatively large noisy environment.

² The payoffs of players are determined by their true actions and are revealed at the end of the second stage.

³ These two reward institutions are most typical institutions for reward used in previous literature (Walker,2004; Sefton,2007; Rand et al., 2009).

⁴ In experimental economics, within-group design might incur "demanding effect" and hence reduces the validity of treatment effect found in the experiments. Our design mitigates the possible concerns about "demanding effect". First, we provide a substantial incentive for players' decisions, which mitigates the player's motivation to meet the experimenter's demand in the first stage. Second, after experienced three institutions, players are asked to choose one institution and interact with other group members in the same institution. This task catches players' attention and prevents them from guessing the experimenter's demand for their decisions in the first stage, which mitigates the demanding effect on players' contributions and reward actions in the first stage.

beginning, but also diminishes the decaying trend occurring in *NR*. Build on our theoretical model and experimental data, we simulate the decisions of players in three institutions with a long finite period (50 periods). The simulated result shows that *DR* sustains a higher level of contribution and welfare than *NR* even in the last period, which was similar to the finding of Rand et al. (2009). We also found the same effect of *CR* with *DR* on cooperation tendency. More specifically, *CR* even sustains a higher contribution and welfare level than *DR*, which makes most players to prefer *CR*.

In the imperfect information environment, the observed patterns are similar to perfect-information. Both *CR* and *DR* significantly increase the contribution and welfare level, and sustain higher contribution and welfare than *NR* in the last period. Simulation results demonstrate that players maintain a moderate cooperation even in the 50th period. In a low noisy environment, *CR* sustains higher contribution and welfare than *DR*. In contrast, in a high noisy environment, *CR* has a similar effect on contribution with *DR*, but sustains higher welfare than *DR*. Moreover, we also find most players to prefer *CR* than other two institutions notwithstanding in the imperfect information environment. To summarize, whether the information is perfect or not, clearly *DR* and *CR* are beneficial for society, notwithstanding in the long run. Moreover, *CR* is the most preferred institution.

Compared to the perfect information environment, we find a smaller effect of *DR* and *CR* on improving cooperation and welfare in the imperfect information environment. A closer look at the data provides hints for why the costly reward option is less effective under the imperfect information environment than the perfect information environment in improving cooperation and welfare. First, in the case of *DR* and *CR*, the linkage between contribution and reward is reduced by the high level of noise. This suggests that players contributing high are less likely to believe to be rewarded in the high noisy environment because others might receive a private signal of low contribution, or others might suspect they contribute at a low level even if the private signal of contribution is at a high level. Hence, players are more likely to contribute less in the high noisy environment because they might get similar rewards to higher contribution level, even if they contribute at a low level. Second, players learn the other's cooperation level more slowly with noisy information. That is, in the imperfect information environment, players hardly expect that others increase their contribution level with the reward option, leading conditional cooperators to contribute less under imperfect information than perfect information in *DR* and *CR*. Third, only in the case of *DR*, the amount of reward is reduced by the high level of noise. This results in a lower welfare level in *DR* than *CR* under high noisy environment, even if the contribution levels of two reward institutions are comparable.

Our paper contributes to the existing literature in several important ways. First, we complement and expand the recent discussion on how to enforce cooperation in a realistic noisy environment. Bereby-Meyer and Roth (2006) find that players hardly to learn to cooperate in a repeated social dilemma situation when payoffs are noisy. It suggests that the reputation mechanism does not sustain cooperation in the imperfect information environment. Grechenig et al. (2010) investigate whether the option to punish others costly improves cooperation in a short repeated social dilemma situation with uncertainty about the contributions of others. They point out that punishment is inefficient to improve cooperation and welfare if there is some nontrivial degree noise. That is, perfect-information is crucial to the efficiency of sanction mechanism. Ambrus and Greiner (2012) focus on the effects of increasing the severity of punishment on cooperation and welfare in a long repeated

noisy environment⁵. They find a U-shape relationship between severity and welfare with small noise. More specifically, a standard punishment significantly decreases welfare while a severe punishment leads to roughly the same welfare as with no punishment option even in the long run. Ambrus and Greiner (2019) compare the effect of the democratic punishment institution with decentralized punishment and centralized punishment institution on cooperation and welfare in a medium repeated noisy environment (20 periods). Although they reveal that democratic punishment leads to a higher level of contribution and welfare with the small level of noise, a clear decaying trend of contribution and welfare level is not checked by democratic punishment. Hence democratic punishment is likely to be uneconomical in the long run. Moreover, neither decentralized punishment nor centralized punishment institution is found to improve cooperation or welfare even in a medium run. To summarize, these results suggest that punishment option does not increase welfare in a long run even if there exists a small noise. In contrast to the previous literature, we find that even with a large level of noise, both *DR* and *CR* sustain a higher contribution and welfare level in the long run. These findings shed light on the superiority of reward institutions over punishment institutions in promoting cooperation and welfare in a realistic world where cooperation information is likely observed inaccurate.

Second, we also contribute to the growing literature about the preference for incentive institutions. Most literature focuses on the institution choice under perfect information environment. Sutter et al. (2010) investigate the endogenous choice of reward and punishment institutions. They find a preference on the reward option in perfect-information environment, although punishment is much more effective in sustaining a high level of cooperation. Dickinson et al. (2015) replicate this finding and validate the external validity by using police commissioners' sample. Markussen et al. (2014) focus on the endogenous choice of institutions in sanction. They reveal that centralized punishment institution (automatically executed) is more popular than decentralized punishment in perfect-information environment. To our knowledge, only one study concentrates on the institution choice under imperfect information environment. Nicklisch et al. (2016) show that decentralized punishment is the most popular sanctioning institution in perfect-information environment. However, this preference is significantly affected by the level of noise. Subjects do not show a clear preference for decentralized punishment, centralized punishment and sanction-free institutions in a low noisy environment while sanction-free institution is chosen with the highest shares in a high noisy environment. In contrast to Nicklisch et al. (2016), our paper concentrates on the endogenous choice of institutions in reward. We find a stable preference for centralized reward institution, whether there exists a small or large level of noise in contribution information. Combining the findings of Nicklisch et al. (2016), our paper suggests that individuals are most likely to prefer reward institutions in the imperfect information environment and hence explains the prevalence of centralized reward institution in the real world.

Finally, our paper is at the forefront of papers in the economics literature studying how antisocial rewards influence subsequent contribution behaviors in the repeated public goods game⁶. The literature on the harmful effect of antisocial incentive behaviors most focus on antisocial

⁵ Ambrus and Greiner (2012) employ two punishment treatments: a standard punishment technology and a strong punishment technology. In the standard punishment technology, every token spent on punishment incurs a damage of three tokens. In the strong punishment technology, very token spent on punishment incurs a damage of six tokens. Therefore, the strong punishment technology is more severe than the standard punishment technology.

⁶ According to the definition of antisocial punishment in previous literature, we define the reward followed a low contribution by the subject as antisocial reward.

punishments. Herrmann et al. (2008) is the first paper to study antisocial punishments across societies, demonstrating that the antisocial punishment of contributors lowers their subsequent contributions in the perfect information environment⁷. Ambrus & Greiner (2012) and Ambrus & Greiner (2019) then replicate this finding using a binary public goods game, documenting that subjects who were punished “unfairly” by other group members are less likely to contribute in the next period⁸. In addition, Ambrus & Greiner (2019) reveal that this adverse effect is mitigated under democratic and dictator punishment⁹. To the best of our knowledge, our study takes the lead in examine how participants make cooperation decisions after received antisocial rewards from other group members or an authority. In contrast to antisocial punishments, we find that antisocial rewards do not affect players’ subsequent contributions in the perfect information and low noisy environment but decrease players’ subsequent contributions in the imperfect information environment. This finding suggests that when the information is perfect or the noise is small, players are able to expect the rewards assigned to them accurately, and that the received antisocial rewards per se do not affect players’ subsequent contributions. This has significant implications for understanding the mechanisms underlying subjects’ cooperation choices in reward institutions, and for helping us better understand how reward differences from punishment in enhancing cooperation under noisy environment.

This paper is structured as follows. Section 2 describes our basic game and derives some important propositions for the effect of reward institutions. In Section 3 we introduce the experimental setting. Section 4 presents the results and Section 5 concludes the paper with a discussion.

2 The Model

We use an overarching model that combines (i) social norms of conditional cooperation, and (ii) sequential reciprocity, using psychological game theory.

2.1 Game structure

Consider a standard VCM, let $I = \{1,2,3,4\}$ denote a group of 4 subjects who interact in T periods. The IDs of players changed from period to period, meaning that the reputation effect is ruled out. In each period $t \in \{1,2, \dots, T\}$, individual $i \in I$ receives an endowment $m \in \mathbb{N}^+$, which can be allocated to either a private good or a public good. We use index j for other group members, $j \in I \setminus \{i\}$. The voluntary contribution of player i in period t must satisfy $g_{it} \in G$, $G = \{0,1,2, \dots, m\}$. Let \bar{g}_t denote the average contribution towards public goods (i.e., $\bar{g}_t = \frac{1}{4} \sum_{i=1}^4 g_{it}$). Individual member i ’s payoff from her contribution in period t , denoted by π_{it}^1 , is

⁷ Herrmann et al. (2008) uses a standard linear public goods game with/without punishment. In their no-punishment condition, the endowment is set to 20 tokens and the marginal per capital return is set to 0.4. In their punishment condition, the punishment point is set to 10 and the efficiency of punishment is set to 3.

⁸ In Ambrus & Greiner (2012), subjects can only choose between contributing all or none of their endowments.

⁹ In Ambrus & Greiner (2019), there are three treatments: Individual punishment, democratic punishment and dictator punishment. In their individual punishment treatment, each group member could directly reduce the earnings of another group member. In their democratic punishment treatment, group members simultaneously cast votes for each group member whether to punish the member or not. If a majority is reached, that participant will be punished. In their dictator punishment treatment, one of the group members is randomly chosen to decide whether to punish other group members.

given by

$$\pi_{it}^1(m, g_{it}, \bar{g}_t) = m - g_{it} + 4\gamma \cdot \bar{g}_t \quad (2.1)$$

The marginal *per capita* return (MPCR) from investing into the public good is denoted as γ , which satisfies $\frac{1}{4} < \gamma < 1$, meaning that the self-interested choice and the socially optimal one are at odds.

We set up three variants of this basic game representing different institutions: (i) the standard VCM without reward option (representing *NR*), (ii) the VCM with centralized reward option (representing *CR*) and (iii) the VCM with decentralized reward option (representing *DR*). We denote by C , D and N the set of players in the three institutions. In period $t \in \{1, 2, \dots, T\}$ of each institution, there are two stages: a contribution stage and a reward stage. The contribution stage for all institutions is the same as the standard VCM described above. All group members make decisions about their contributions simultaneously. The only difference is that in *CR*, an authority a joins the group but takes no actions in this stage.

In the reward stage, all players including the authority receive a signal s valued at $s_{kjt} \in S$ about the contribution of player j , $S = \{0, 1, 2, \dots, m\}$, $k \in I \cup \{a\}$ and $j \in I \setminus \{i\}$ ¹⁰. For each k , a signal is produced, such that

$$s = \begin{cases} g_{jt} & \text{with prob} = \lambda \\ \tilde{g}_{jt} & \text{with prob} = 1 - \lambda \end{cases} \quad (2.2)$$

Where \tilde{g}_{jt} is randomly drawn from the set $G = \{0, 1, 2, \dots, m\}$ with uniform probabilities. Thus, for each player, there is an independent random draw determining whether the signal corresponds to the true contribution or not. Parameter λ characterizes the observability of contribution information¹¹. We define perfect observability as $\lambda = 1$ and imperfect observability as $\lambda \in [0, 1)$ ¹². The player j doesn't know whether the subject k receives a true or false signal of her contribution. Let $P(X = x)$ denote the probability of a variable X equaling x . We then define the probability of a signal s equaling s_{kjt} conditional on g_{jt} by $P^s(s_{kjt}|g_{jt}) = P(s = s_{kjt}|g_{jt})$ ¹³.

$$P^s(s_{kjt}|g_{jt}) = \begin{cases} \frac{1-\lambda}{m+1} + \lambda, & \text{if } s_{kjt} = g_{jt} \\ \frac{1-\lambda}{m+1}, & \text{if } s_{kjt} \neq g_{jt} \end{cases} \quad (2.3)$$

In addition, all players receive an extra endowment $\bar{r} > 0$. Depending on their institution, players assign reward points (that is, the player i in *DR* and the authority a in *CR*), and then the final payoffs are realized. The three institutions mainly differ in the reward stage. For player i in *NR* the payoff equals the profit after the contribution stage plus the extra endowment:

$$\pi_{it}^N = \pi_{it}^1 + \bar{r} \quad (2.4)$$

In *DR* all player i s decide simultaneously over reward $r_{i \rightarrow j, t}^D \in [0, \bar{r}]$ toward player j . Each reward point assigned to player j leads to an addition of $\kappa > 1$ units from the payoff of player j and reduces the payoff of player i by one unit. Each player can spend up to her extra

¹⁰ s_{kjt} is the value of a signal of player j 's contribution received by subject k .

¹¹ The parameter λ in our study is different from that in Nicklisch et al. (2016). In their study, λ is the possibility of subject k receiving a true signal of other's contribution. In our study, however, this possibility is $\frac{1-\lambda}{m+1} + \lambda$ instead of λ .

¹² Based on equation (2.3), when λ equals 1, the possibility of subject k receiving a true signal is 1; when λ is less than 1, this possibility is less 1 as well.

¹³ $P^s(s_{kjt}|g_{jt})$ is the possibility of subject k receiving a signal s equaling s_{kjt} if player j 's contribution is g_{jt} . For example, if player j 's contribution is m , the possibility of subject k receiving a signal equaling $\frac{m}{2}$ is $\frac{1-\lambda}{m+1}$ and that of subject k receiving a signal equaling m is $\frac{1-\lambda}{m+1} + \lambda$.

endowment \bar{r} for reward, that is, $\sum_j r_{i \rightarrow j, t}^D \leq \bar{r}$. Units not spent on reward are credited to the payoff of player i . For player i in DR , the payoff equals

$$\pi_{it}^D = \pi_{it}^1 + \left(\bar{r} - \sum_j r_{i \rightarrow j, t}^D \right) + \kappa \cdot \sum_j r_{j \rightarrow i, t}^D \quad (2.5)$$

In CR all reward decisions are delegated to the authority a . The authority decides over reward $r_{a \rightarrow j, t}^C \in [0, 3\bar{r}]$ toward player $j \in I$. Like in DR each reward point assigned to a player assigned to player j leads to an addition of $\kappa \geq 1$ units from the payoff of player j and costs one unit. In CR these costs have to be borne by all players except j in the institution. In sum, the authority can spend up to the extra endowment of all player j s for reward, i.e., $\sum_{j \in I} r_{a \rightarrow j, t}^C \leq 4\bar{r}$.

In addition, maximum reward targeted at a single player is restricted to $3\bar{r}$. Units not spent on reward are credited to the particular player's account. Hence, DR and CR are identical with regard to the feasible set as well as the financial consequences of reward. The only difference is that reward decisions are made by the authority. For player i in CR , the payoff equals

$$\pi_{it}^C = \pi_{it}^1 + \left(\bar{r} - \frac{\sum_j r_{a \rightarrow j, t}^C}{3} \right) + \kappa \cdot r_{a \rightarrow i, t}^C \quad (2.6)$$

The authority's payoff equals the average profit after contribution stage of all players in CR

$$\pi_{at}^C = \frac{\sum_i \pi_{it}^1}{4} \quad (2.7)$$

2.2 Belief updating

In order to construct appropriate *belief hierarchies*, we make some basic assumption on the belief based on our experimental data. The belief of subject k about player j 's contribution in period t known as subject k 's *first-order belief* toward player j is defined by b_{kjt} . Without any correlated signals, *first-order belief* is referred to as the *first-order positive belief*. Denote the probability of *first-order positive belief* by $B_{kj}^1(\cdot): G \rightarrow [0, 1]$. Thus $B_{kj}^1(g_{jt}) = P(b_{kjt} = g_{jt})$ is the probability assigned by subject k that the contribution of the player j in period t equals g_{jt} before the contribution decision is made. Our first assumption is on the *first-order positive belief*.

Assumption 1. *The positive belief of subject k about player j 's contribution equals \tilde{g}_{jt} , where \tilde{g}_{jt} is randomly drawn from the set $G = \{0, 1, 2, \dots, m\}$ with uniform probabilities. Thus, for each $g_{jt} \in G$, the probability of first-order positive belief $B_{kj}^1(g_{jt})$ equals $\frac{1}{m+1}$.*

From Assumption 1, k may unconditionally expect player j averagely contribute $\frac{m}{2}$ to public goods. Let $E[\cdot]$ denote the expectation of a variable. We define k 's first-order expectation as $\mathbb{E}_{kjt}^1[\cdot] = E[b_{kjt} | \cdot]$, which means the expectation of subject k 's *first-order belief* toward player j . Without any signals, k 's unconditional first-order expectation $\mathbb{E}_{kjt}^1[\emptyset]$ equals $\frac{m}{2}$.

Remark 1. *Assumption 1 is easy to understand. When subject k has no information about player j , she may expect player j contribute randomly. Our data showing the average beliefs about other group members in the first period is nearly 10 somewhat supports this assumption¹⁴. Of course,*

¹⁴ In our experiment, the parameter m is set to 20. Hence i 's first order expectation of each j consistently

player i may also expect player j will do the same action with herself. This phenomenon is famous as the false consensus effect and supported by a great deal of empirical evidence. This effect is also backed by our data since the average beliefs about other group members are significantly positive correlated with player i 's contribution in the first period. However, this mechanism does not apply to authority a in CR¹⁵. For simplicity and consistency, we use Assumption 1 in our model to characterize the first-order positive belief of subject k .

Denote player i 's second order expectation as $\mathbb{E}_{jit}^2[\cdot] = \mathbb{E}[b_{jit} | \cdot]$, which equals the expectation of player j 's first-order belief toward player i .

Lemma 1. *The unconditional second order expectation of player i is given by*

$$\mathbb{E}_{jit}^2[\emptyset] = \mathbb{E}[b_{jit}] = \frac{m}{2} \quad (2.8)$$

We denote $(S_{kjt})_{k \in I \cup a, j \in I \setminus \{i\}} = \{s_{kj1}, s_{kj2}, \dots, s_{kjt}\}$ as the signal set of subject k in the reward stage of period t about player j 's contributions¹⁶, and $S_{kt} = \times_{j \in I \setminus \{i\}} S_{kjt}$ as the information set of k in the reward stage of period t . The possibility of player k 's first-order belief toward player j conditional on a signal s_{kjt} in period t is denoted by $B_{kj}^1(g_{jt} | s_{kjt}) = P(b_{kjt} = g_{jt} | s_{kjt})$.

Assumption 2. *Subject k updates her belief about player j 's contribution with the latest signal of j 's contribution according to Bayes rule. That is, $B_{kj}^1(g_{jt} | S_{kt})$ is given by¹⁷,*

$$B_{kj}^1(g_{jt} | S_{kt}) = B_{kj}^1(g_{jt} | s_{kjt}) = \begin{cases} \frac{1 + m\lambda}{1 + m}, & \text{if } g_{jt} = s_{kjt} \\ \frac{1 - \lambda}{1 + m}, & \text{if } g_{jt} \neq s_{kjt} \end{cases} \quad (2.9)$$

From Assumption 2, the possibility that player j 's contribution is believed to be same as her signal s_{kjt} is increasing with information observability λ . When information is perfectly observed (i.e., $\lambda = 1$), k 's first-order belief toward player j is completely consistent with her received signal s_{kjt} .

Remark 2. *In our experiment, players in a group change their IDs from period to period. Therefore, they cannot assign the signals shown to them in previous periods to a certain group member in the current period. For this reason, we assume players update their beliefs about a certain group member in the reward stage only with the signal shown in the current period.*

Lemma 2. *The first-order expectation of subject k toward player j conditional on signals is given by,*

$$\mathbb{E}_{kjt}^1[S_{kt}] = \mathbb{E}[b_{kjt} | s_{kjt}] = \frac{m(1 - \lambda)}{2} + \lambda s_{kjt} \quad (2.10)$$

Lemma 3. *The second order expectation conditional on i 's contribution is given by,*

equals 10 in the first period. As a result, the average beliefs about other group members equal 10 as well.

¹⁵ This is because the authority a doesn't need to contribute in our experiment.

¹⁶ S_{kjt} contains signals about player j 's contribution received by subject k in previous periods as well as that received by subject k in the current period. Therefore, the signal set of subject k in the contribution stage of period $t \geq 2$ about player j 's contributions is $S_{kj,t-1}$. When $t = 1$, we denote $S_{kj,t-1} = \emptyset$.

¹⁷ In the contribution stage of period t , as the signal set of subject k in the contribution stage of period t about player j 's contributions is $S_{kj,t-1}$, the possibility of k 's first order belief is defined by $B_{kj}^1(g_{jt} | S_{k,t-1}) = B_{kj}^1(g_{jt} | s_{kj,t-1})$.

$$\mathbb{E}_{jit}^2[g_{it}] = \mathbb{E}[b_{jit}|g_{it}] = \frac{m - \lambda^2 m}{2} + \lambda^2 g_{it} \quad (2.11)$$

2.3 Psychological utility function

2.3.1 Psychological utility function in reward stage

In *DR*, player i may infer the *kindness intentions* of player j based i 's conditional expectation. Following the spirit of Dufwenberg and Kirchsteiger (2004), we define the reciprocity of player i towards player j as follows¹⁸

$$R_{ijt} = (\mathbb{E}_{ij}^1[S_{it}] - \mu \cdot m) \cdot r_{i \rightarrow j, t}^D \quad (2.12)$$

Where μm is player i 's reciprocal reference point. The right-hand side (hereafter, RHS) of equation (2.12) is the product of two terms. The first term $(\mathbb{E}_{ij}^1[S_{it}] - \mu \cdot m)$ captures the perception of player i for the kindness of player j in the contribution stage. Thus, player i perceives player j to be kind, if player i believes that player j contributes more than a fraction of μ of her endowment in the contribution stage. The second term $r_{i \rightarrow j, t}^D$ captures the kindness of player i towards player j in the reward contribution. Thus, player i perceives being kind to player j if she chooses a reward $r_{i \rightarrow j, t}^D > 0$. The psychological utility, U_{it}^{2D} , of player i in the reward stage of *DR* is given by

$$U_{it}^{2D} = \mathbb{E}[\pi_{it}^D] + \sum_j \delta_i \cdot R_{ijt} \quad (2.13)$$

In (2.13), the first term on the RHS is the expectation of player i in the reward stage on her total payoff of period t defined in (2.5). The second term captures conditional reciprocity, the reciprocity of player i towards player j in reward stage was formally defined in (2.12); $\delta_i > 0$ is a parameter related to player i 's reciprocity sensitivity.

In *CR*, the reciprocity of authority a towards player j as follows

$$R_{ajt} = (\mathbb{E}_{aj}^1[S_{at}] - \mu \cdot m) \cdot r_{a \rightarrow j, t}^C \quad (2.14)$$

The psychological utility, U_{at}^{2C} , of authority a in the reward stage of *CR* is given by

$$U_{at}^{2C} = \pi_{at}^C + \sum_j \delta_a \cdot R_{ajt} \quad (2.15)$$

In (2.15), the first term on the RHS gives the utility from authority a 's payoff, defined in (2.7). The second term captures conditional reciprocity, the reciprocity of authority a towards player j in reward stage was formally defined in (2.14); $\delta_a > 0$ is a parameter related to authority a 's reciprocity sensitivity.

2.3.2 Psychological utility function in contribution stage

Following Fehr and Schurtenberger (2018), we define the social norm of conditional cooperation that player i should contribute no less than the average contribution of other group members. In *NR*, the psychology utility function of player i in the contribution stage is denoted

¹⁸ In Dufwenberg and Kirchsteiger (2004), the kindness of player j to player i in the contribution stage is defined by $\pi_i^1(g_j, g_{-j}) - \frac{1}{2}[\pi_i^1(m, g_{-j}) + \pi_i^1(0, g_{-j})]$, which can be simplified by $\gamma(g_j - \frac{1}{2}m)$. Moreover, the kindness of player i to player j in the reward stage can be simplified by $\kappa(r_{i \rightarrow j, t}^D - \frac{1}{2}\bar{r})$. Hence the reciprocity of player i to player j is defined by $R_{ij} = \kappa\gamma(g_j - \frac{1}{2}m)(r_{i \rightarrow j, t}^D - \frac{1}{2}\bar{r})$. For simplicity, we define the reciprocity of player i to player j as equation (2.12). However, we should have derived the same propositions if we had directly used the definition of reciprocity in Dufwenberg and Kirchsteiger (2004) instead.

by U_{it}^{1N} ,

$$U_{it}^{1N} = E[\pi_{it}^N] - \alpha_i \cdot \{(\bar{E}_{it}^1[S_{i,t-1}] - g_{it})^+\}^2 \quad (2.16)$$

Where $\bar{E}_{it}^1[S_{i,t-1}] = \sum_{j \neq i} \frac{E_{ijt}^1[S_{i,t-1}]}{3}$ is the average first-order expectation toward other group members, which equals player i 's expectation about other group members' average contribution. The notion $z = x^+$ means that $z = x$ if $x > 0$ and $z = 0$ if $x \leq 0$, $\alpha_i > 0$ captures the sensitivity to norm incompliance. In (2.16), the first term on the RHS is the expectation of player i in the contribution stage on her total payoff of period t . The second term captures the utility loss that player i may suffer if her contribution is less than her expectation about other group members' average contribution at the end of the contribution stage, i.e., $g_{it} < \bar{E}_{it}^1[S_{i,t-1}]$.

In *DR*, the psychology utility function of player i in the contribution stage is denoted by U_{it}^{1D} ,

$$U_{it}^{1D} = E[\pi_{it}^D] - \alpha_i \cdot \{(\bar{E}_{it}^1[S_{i,t-1}] - g_{it})^+\}^2 + E[\sum_j \delta_i \cdot R_{ijt}] \quad (2.17)$$

Similarly, in *CR*, the psychology utility function of player i in the contribution stage is denoted by U_{it}^{1C} ,

$$U_{it}^{1C} = E[\pi_{it}^C] - \alpha_i \cdot \{(\bar{E}_{it}^1[S_{i,t-1}] - g_{it})^+\}^2 \quad (2.18)$$

In (2.17) and (2.18), the first term on the RHS is the expectation of player i in the contribution stage on her total payoff in period t . The second term captures the utility loss from incompliance with conditional cooperation norm. The third term in (2.17) captures the expectation of player i on her psychological utility of reciprocity¹⁹.

2.4 Psychological best responses

There should be no automatic supposition that a *subgame perfect Nash equilibrium* (SPNE) is the empirically appropriate solution concept for finite repeated experimental games, when the machinery of psychological game theory is used (Battigalli and Dufwenberg, 2022). It is arguably more compelling that players perform best response conditional on their beliefs about other player's actions, and such beliefs have not always been true. Our data shows players generally form false beliefs about other's contributions even in the last period, especially when other group members' contributions cannot be accurately observed (i.e., $\lambda < 1$). Hence, we use player's best response instead of SPNE as the solution of our model. In the contribution stage, the beliefs of players are expectation for the average contribution of other group members, and in reward stage, these are expectations about reward chosen by other group members, conditional on her contribution. Players play a psychological best response if they maximize their psychological utility in each stage, conditional on their beliefs.

Definition 1. *A psychological best response of player i in period t is a series of contributions and reward levels $(g_{it}^*, r_{i \rightarrow j, t}^*)$ with the following properties:*

1. $g_i^* \in G, \sum_j r_{i \rightarrow j, t}^* \in [0, \bar{r}]$.

¹⁹ Among three institutions, only player i in *DR* is able to reward player j . Therefore, the psychological utility of player i in *DR* but not in *NR* nor *CR* contains the motivation of reciprocity relating to her future reward behaviors.

2. In the reward stage of DR, $r_{i \rightarrow j, t}^* = r_{i \rightarrow j, t}^*(S_{it})$ maximizes U_{it}^{2D} in (2.13) with the information set S_{it} .
3. In the contribution stage of NR, g_i^* maximizes U_{it}^{1N} in (2.16) with the information set $S_{i, t-1}$.
In the contribution stage of DR, g_i^* maximizes U_{it}^{1D} in (2.17) with the information set $S_{i, t-1}$.
In the contribution stage of CR, g_i^* maximizes U_{it}^{1C} in (2.18) with the information set $S_{i, t-1}$.

Definition 2. A psychological best response of authority a is a series of reward levels $r_{a \rightarrow j, t}^*$ with the following properties:

1. $r_{a \rightarrow j, t}^* \in [0, 3\bar{r}]$, $\sum_j r_{a \rightarrow j, t}^* \in [0, 4\bar{r}]$
2. In the reward stage of CR, $r_{a \rightarrow j, t}^* = r_{a \rightarrow j, t}^*(S_{at})$ maximizes U_{at}^{2C} in (2.15) with the information set S_{at} .

Assumption 3. To ensure budget constraint, we assume (a) player i rewards each player j no more than $\frac{\bar{r}}{3}$ in DR, that is, $r_{i \rightarrow j, t}^* \in [0, \frac{\bar{r}}{3}]$; (b) authority a rewards each player j no more than \bar{r} in CR, that is $r_{a \rightarrow j, t}^* \in [0, \bar{r}]$.

Remark 3. In DR of our experiment, player i is able to assign all her reward points \bar{r} to player j . However, this extreme assignment is seldomly found in our data. In order to ensure the sum of rewards no more than \bar{r} , player i forms three independent mental accounts, each containing $\frac{\bar{r}}{3}$ reward points for each player j . In CR, in the same way, authority a forms four independent mental accounts, each containing \bar{r} reward points for each player j . Our data showing that about 95% rewards of player i towards player j are no more than $\frac{\bar{r}}{3}$ in DR, and about 80% rewards of authority a towards player j is no more than \bar{r} in CR supports Assumption 3 to some extent.

We solve the psychological best response of players by backward induction.

2.4.1 Optimal choice in reward stage

For a non-integer x , we denote the function $Ceil(x)$ as the smallest integer larger than x . If x is an integer, $Ceil(x)$ is defined as x .

Lemma 4. If Assumption 3 holds, the best response for player i in reward stage of DR is given by,

$$r_{i \rightarrow j, t}^* = \begin{cases} \frac{\bar{r}}{3}, & \text{if } s_{ijt} \geq \tilde{s}^D \\ 0, & \text{if } s_{ijt} < \tilde{s}^D \end{cases} \quad (2.19)$$

Where \tilde{s}^D equaling $Ceil(\frac{\mu m}{\lambda} + \frac{1}{\delta_i \lambda} - \frac{m(1-\lambda)}{2\lambda})$, is denoted as reward reference point in DR, which represents the minimum amount of signal for player i to reward player j .

Lemma 5. If Assumption 3 holds, the best response of authority a in reward stage of CR is given by,

$$r_{a \rightarrow j, t}^* = \begin{cases} \bar{r}, & \text{if } s_{ajt} \geq \tilde{s}^C \\ 0, & \text{if } s_{ajt} < \tilde{s}^C \end{cases} \quad (2.20)$$

Where \tilde{s}^C equals $Ceil(\frac{\mu m}{\lambda} - \frac{m(1-\lambda)}{2\lambda})$, is denoted as reward reference point in CR , which represents the minimum amount of signal for authority a to reward player j .

2.4.2 Optimal choice in contribution stage

To solve the psychological best response of player i in the contribution stage, we first make a basic assumption about the conditional expectation of player i on the received reward.

Assumption 4. *Subjects are able to expect how rewards are chosen from player j and authority a in DR and CR , respectively. That is, $E[r_{j \rightarrow i,t} | \cdot] = E[r_{j \rightarrow i,t}^* | \cdot]$ and $E[r_{a \rightarrow i,t} | \cdot] = E[r_{a \rightarrow i,t}^* | \cdot]$.*

Lemma 6. *If Assumption 4 holds, the expectation of player i on the reward chosen from player j conditional on player i 's contribution is given by,*

$$E[r_{j \rightarrow i,t} | g_{it}] = \begin{cases} \frac{\bar{r}}{3} \cdot [1 - \frac{(1-\lambda)Ceil(\tilde{s}^D)}{m+1}] & , \text{if } g_{it} \geq \tilde{s}^D \\ \frac{\bar{r}}{3} \cdot [1 - \lambda - \frac{(1-\lambda)Ceil(\tilde{s}^D)}{m+1}] & , \text{if } g_{it} < \tilde{s}^D \end{cases} \quad (2.21)$$

Lemma 7. *If Assumption 4 holds, the expectation of player i on the reward chosen from authority a conditional on player i 's contribution is given by,*

$$E[r_{a \rightarrow i,t} | g_{it}] = \begin{cases} \bar{r} \cdot [1 - \frac{(1-\lambda)Ceil(\tilde{s}^C)}{m+1}] & , \text{if } g_{it} \geq \tilde{s}^C \\ \bar{r} \cdot [1 - \lambda - \frac{(1-\lambda)Ceil(\tilde{s}^C)}{m+1}] & , \text{if } g_{it} < \tilde{s}^C \end{cases} \quad (2.22)$$

Assumption 5. *The player i 's average first-order expectation in the contribution stage is no less than $\frac{1-\gamma}{2\alpha_i}$, which is no less than 1. That is, $\bar{E}_{it}^1[S_{i,t-1}] \geq \frac{1-\gamma}{2\alpha_i} \geq 1$.²⁰*

We denote the function $Round(x)$ as the integer nearest to x . If there are two integers with the same distance to x , $Round(x)$ is defined as the larger one.

Lemma 8. *In NR , the best response of player i in contribution stage is given by,*

$$g_{it}^* = Round(\bar{E}_{it}^1[S_{i,t-1}] - \frac{1-\gamma}{2\alpha_i}) \quad (2.23)$$

For simplicity, we denote a symbol \tilde{s}_u as follows:

$$\tilde{s}_u = \begin{cases} Round(\bar{E}_{it}^1[S_{i,t-1}] - \frac{1-\gamma}{2\alpha_i}) + \frac{\lambda\kappa\bar{r}}{(1-\gamma)} + \frac{\alpha_i c^2}{(1-\gamma)} & , \text{if } \lambda\kappa\bar{r} \geq (1-\gamma) \cdot c - \alpha_i c^2 \\ \bar{E}_{it}^1[S_{i,t-1}] - \frac{1-\gamma}{2\alpha_i} - \frac{|c - \frac{1-\gamma}{2\alpha_i}|}{c - \frac{1-\gamma}{2\alpha_i}} \cdot \sqrt{\frac{\lambda\kappa\bar{r}}{\alpha_i} + (c - \frac{1-\gamma}{2\alpha_i})^2} & , \text{if } \lambda\kappa\bar{r} \in (0, (1-\gamma) \cdot c - \alpha_i c^2) \end{cases} \quad (2.24)$$

Where c represents the difference between the average first-order expectation $\bar{E}_{it}^1[S_{i,t-1}]$ and the best response g_{it}^* in NR , that is $c = \bar{E}_{it}^1[S_{i,t-1}] - Round(\bar{E}_{it}^1[S_{i,t-1}] - \frac{1-\gamma}{2\alpha_i})$. \tilde{s}_u represents an indifference continuous contribution amount which means if players increased their contributions

²⁰ Assumption 5 is used for simplicity of our model. Without this assumption, our proposition 1-5 still keep unchanged.

to this amount, the earned reward should have been just right offset by the cost of the increasing contribution. However, since the contribution is set to an integer in our model, indifference contribution amount can hardly be chosen by player i .

Lemma 9. (a) In *DR*, the best response of player i in contribution stage is given by,

$$g_{it}^* = \begin{cases} \text{Round}\left(\bar{\mathbb{E}}_{it}^1[S_{i,t-1}] - \frac{1-\gamma}{2\alpha_i}\right), & \text{if } \tilde{s}^D \in [0, \text{Round}\left(\bar{\mathbb{E}}_{it}^1[S_{i,t-1}] - \frac{1-\gamma}{2\alpha_i}\right)] \cup (\text{Ceil}(\tilde{s}_u) - 1, m] \\ \tilde{s}^D & , \text{if } \tilde{s}^D \in (\text{Round}\left(\bar{\mathbb{E}}_{it}^1[S_{i,t-1}] - \frac{1-\gamma}{2\alpha_i}\right), \text{Ceil}(\tilde{s}_u) - 1] \end{cases} \quad (2.25)$$

(b) In *CR*, the best response of player i in contribution stage is given by,

$$g_{it}^* = \begin{cases} \text{Round}\left(\bar{\mathbb{E}}_{it}^1[S_{i,t-1}] - \frac{1-\gamma}{2\alpha_i}\right), & \text{if } \tilde{s}^C \in [0, \text{Round}\left(\bar{\mathbb{E}}_{it}^1[S_{i,t-1}] - \frac{1-\gamma}{2\alpha_i}\right)] \cup [\text{Ceil}(\tilde{s}_u) - 1, m] \\ \tilde{s}^C & , \text{if } \tilde{s}^C \in (\text{Round}\left(\bar{\mathbb{E}}_{it}^1[S_{i,t-1}] - \frac{1-\gamma}{2\alpha_i}\right), \text{Ceil}(\tilde{s}_u) - 1) \end{cases} \quad (2.26)$$

2.5 Comparison between institutions with and without reward

Proposition 1. (a) For period 1, the contribution of player i in *DR* is no less than in *NR*. If

$\tilde{s}^D \in (\text{Round}(\frac{m}{2} - \frac{1-\gamma}{2\alpha_i}), \text{Ceil}(\tilde{s}_u) - 1]$, the contribution of player i in *DR* is strictly more than

that in *NR*. For any period $t > 1$, given the same first-order expectation of the average contribution of other group members, the contribution of player i in *DR* is no less than in *NR*.

If $\tilde{s}^D \in (\text{Round}(\bar{\mathbb{E}}_{it}^1[S_{i,t-1}] - \frac{1-\gamma}{2\alpha_i}), \text{Ceil}(\tilde{s}_u) - 1]$, the contribution of player i in *DR* is strictly more than that in *NR*.

(b) For period 1, the contribution of player i in *CR* is no less than in *NR*. If $\tilde{s}^C \in (\text{Round}(\frac{m}{2} - \frac{1-\gamma}{2\alpha_i}), \text{Ceil}(\tilde{s}_u) - 1]$, the contribution of player i in *CR* is strictly more than that in *NR*. For any

period $t > 1$, given the same first-order expectation of the average contribution of other group members, the contribution of player i in *CR* is no less than in *NR*. If $\tilde{s}^C \in$

$(\text{Round}(\bar{\mathbb{E}}_{it}^1[S_{i,t-1}] - \frac{1-\gamma}{2\alpha_i}), \text{Ceil}(\tilde{s}_u) - 1]$, the contribution of player i in *CR* is strictly more than that in *NR*.

Discussion of proposition 1: In *DR* and *CR*, without considering the reward, players' best contribution is smaller than their expectation about the average contribution of other group members. To earn the reward, players may choose their best contribution amount according to the reward reference point. When the reward reference point is larger than their best contribution without reward and less than their indifference contribution amount, players will increase their contributions for the reward. Otherwise, they will keep their contributions unchanged. The intuition behind this result is that, only if more contributions lead to more earning, players have motivation to increase their contributions. If the reward reference point is lower than the best contribution without reward, players have earned the reward without changing their contributions. If the reward reference point is more than the indifference contribution amount, the reward cannot cover the cost of increasing

contribution. That is, contributing more to get reward makes them earn less. Hence, in these two conditions, players still keep their contributions though they considering the reward.

Proposition 2. (a) *For any period t , the expectation of player i 's average first-order expectation in DR is no less than that in NR .* (b) *For any period t , the expectation of player i 's average first-order expectation in CR is no less than that in NR .*

Discussion of proposition 2: At the beginning of the game, players don't have any information about other players' contribution amounts. Hence, players in NR, DR and CR expect other players in the same institution to contribute $\frac{m}{2}$ to the public good. Based on the equivalent first-order expectation in period 1, players contribute no less both in DR or CR than in NR . In perfect-information environment, players update their beliefs entirely depending on the other player's contribution information in the previous period. That is, the average first-order expectation in period 2 equals the average contribution level of other players in period 1. In imperfect information environment, as signals of other players' contribution are likely incorrect, players update their beliefs partly depending on their received signals in the previous period and partly depending on their positive beliefs. That is, their average first-order expectations in period 2 are positively correlated with their received signals in period 1. As the expectation of average first-order expectation is positive correlated to the average of received signals which is also increased by other players' contribution amount, we know that the expectation of average first-order expectation in period 2 is no less in DR or CR than in NR . By parity of reasoning, in any period, the expectation of average first-order expectation is no less in DR or CR than in NR . Therefore, we expect the average first-order expectation is no less in DR or CR than in NR .

Proposition 3. (a) *For any period $t > 1$, the expectation of player i 's contribution in DR is no less than in NR .* (b) *For any period $t > 1$, the expectation of player i 's contribution in CR is no less than in NR .*

Discussion of proposition 3: In DR and CR , players expect others contribute no less than in NR . If players' first-order expectations about the average contribution of other group members are same in DR and NR , they contribute no less in DR than NR . Moreover, if players' first-order expectations about the average contribution of other group members are more in DR than NR , their contribution levels are strictly larger in DR . Similarly, if players' first-order expectations about the average contribution of other group members are same in CR and NR , they contribute no less in CR than NR . Moreover, if players' first-order expectations about the average contribution of other group members are more in CR than NR , their contribution levels are strictly larger in CR . According to proposition 3, we expect the average contribution level in DR and CR is no less than that in NR .

Proposition 4. (a) *If the average contribution in DR is strictly more than that in NR during period 1, for any period $t > 1$, the expectation of contribution in DR is strictly more than that in NR as well.* (b) *If the average contribution in CR is strictly more than that in NR during period 1, for any period $t > 1$, the expectation of contribution in CR is strictly more than that in NR as well.*

Discussion of proposition 4: If the average contribution in DR and CR is strictly more than that in NR at the beginning of the game, it means that players contribute more to earn rewards in DR

and *CR*. Based on a higher contribution level of their partner in the previous period, players expect others contribute more with the reward option, leading to a higher contribution level in *DR* and *CR* in the current period, and so on. An implication of proposition 4 is that, if *DR* and *CR* promote contribution at the beginning, these reward institutions sustain a higher contribution level in the long run as well.

Proposition 5. (a) *For any period $t \geq 1$, the expectation of welfare in *DR* is no less than in *NR*. If the average contribution in *DR* is strictly more than that in *NR* in period 1, for any period $t > 1$, the expectation of welfare in *DR* is strictly more than that in *NR* as well.* (b) *For any period $t \geq 1$, the expectation of welfare in *CR* is no less than in *NR*. If the average contribution in *CR* is strictly more than that in *NR* in period 1, for any period $t > 1$, the expectation of welfare in *CR* is strictly more than that in *NR* as well.*

Discussion of proposition 5: In *DR* and *CR*, players contribute no less than in *NR*. In addition, since $\kappa \geq 1$, any reward point assigned to the other group member increases the average earnings of the group ($\kappa > 1$) or at least keeps it constant ($\kappa = 0$). Therefore, players earn no less in *DR* or *CR* than in *NR*. If the average contribution in *DR* and *CR* is strictly more than that in *NR* at the beginning, these reward institutions sustain a higher contribution level in the long run as well. In this condition, players earn more in *DR* and *CR*. An implication of proposition 5 is that, if *DR* and *CR* promote contribution at the beginning, these reward institutions sustain a higher welfare in the long run.

3 Experimental Design

The experiment is played in sessions of thirteen subjects. Prior to the start of the game, we randomly allocate one subject in each session to the role of the authority who always participates in *CR* and twelve subjects to the role of players. Roles remain the same throughout the experiment.

In each session, we divide a repeated game of 20 periods into two parts. The identification numbers of players are randomly assigned between periods. In the first part, we implement a game with three phases consisting of 5 periods each. At the beginning of the first phase, twelve players are randomly and anonymously assigned to 3 different institutions (i.e., *NR* or *DR* or *CR*). Each institution consists of 4 players constant over 5 periods. Then players in each institution participate in another institution together at the beginning of the second and third phases. Our matching mechanism following a within-group design ensures each player to experience three different institutions in a random order. That is, in each session (i) 4 players go through *NR*, then *DR*, and finally *CR*; (ii) other 4 players go through *DR*, then *CR* and finally *NR*; (iii) and the last 4 players go through *CR*, then *NR*, and finally *DR*. In the second part, we implement a game of 5 periods. Based on the previous information of the average contribution and welfare for each institution in the previous 15 periods, all players are asked to choose participating one institution at the beginning of this part and remain there during this part.

Each period consists of 2 stages, a contribution stage and a reward stage²¹. In the contribution stage, each player is endowed with 20 tokens and asked to contribute some of these tokens to a

²¹ In the second part, if there is only one player in an institution, there is no contribution stage nor reward stage. In this case, the player receives a payoff of 20. If there are less than two players in *CR*, the authority receives a payoff of 20.

public good. Each token contributed to the public good is multiplied 1.6 and the resulting amount is shared equally among the players in the respective institution. This payoff function is following Nicklisch et al. (2016) that keeps the marginal social return from the public good constant for different population size²². The rest endowment benefits the player by the same tokens. Then, the player is asked to conjecture about the average amount that other players in the same institution contribute. The more accurate the players' conjectures were, the higher the payoff they could get from their decisions. With the purpose of avoiding hedging, players were paid for either belief elicitation task or for contribution decisions (with the same probability). That is, on average, players were paid for belief elicitation task in 10 periods and paid for contribution decisions in another 10 periods. In contrast to players, the authority was called upon to do nothing in this stage and earned the average profit of players in *CR* for this stage. That is, players in *CR* contribute more, the authority earns more.

In the reward stage, all players and the authority receive the signals from players in their institution. There is a certain probability that the signals are consistent with the true contributions. Taking a between-subjects design, we employed three information treatments: *perfect-information*, *low-noise* and *high-noise*. In the *perfect-information* treatment, the probability is 100%. In the *low-noise* treatment, the probability shifts to 90%. In the *high-noise* treatment, the probability decreases dramatically to 50%. That is, all signals were correct in *perfect-information* treatment, while only 90% and 50% signals were correct in the other two information treatments. Moreover, false signals were randomly drawn from the contribution choice set that excludes the real contribution choice. Players do not receive information about signals for their contributions, that is, they do not know whether other subjects were correctly informed about their exact contribution or not.

In addition, all players receive an extra endowment of three points. In *NR*, players are not allowed to reward. In *DR*, all players decide simultaneously over reward to other players in the same institution. In *CR*, reward decisions are delegated to the authority. The authority decides over reward to players and the rest points are returned equally to players in *CR*. Each point assigned to a player results in an addition of 3 tokens from the rewarded player's payoff. The rest points of each player exchange for tokens in a ratio of 1:1. At the end of each period, players learn their total amount of reward received.

The experimental sessions were computerized and conducted with z-Tree (Fischbacher, 2007). We conducted 36 sessions during September and October 2019 at the Selton Laboratory of Nankai University. Experimental subjects were recruited as volunteers from a university-wide subject pool. None of the subjects participated in more than one session. Overall, 468 subjects took part in 36 sessions, with 13 subjects per session. After entering the laboratory, participants were given instructions to review (sample instructions are provided in Supplementary Material A). The instructions were delivered aloud and explained carefully by the experimenters. The Subjects individually answered a set of control questions to make sure that they understood the task. After all subjects answered the questions correctly, they were randomly allocated to a PC terminal, and the game then started. After the end of the experiment, participants filled out a brief demographic survey. Cash payments from this experiment were calculated by converting the total number of tokens accumulated at a fixed rate. Each session lasted approximately 100 min, and the subjects earned 40 Chinese Yuan (CNY, roughly equal to \$5.64) on average, with a show-up fee of 10 CNY.

²² This setting leads to constant payoff of player i for different population size if the player i 's contribution and the average contribution of other group members are same.

4 Results

The result section is structured as follows: First, we analyze contributions and welfare in the first part to demonstrate that both DR and CR promote contributions and welfare even if signals become noisy. Additionally, we reveal that noise weakens the effects of reward institutions on contributions and welfare (Result 1). Then we show that noise does not substantially influence institutional choice in a systematic way: players opt predominately for CR no matter whether there exists any noise (Result 2). In the next step, we focus on the mechanism by which noise modulates the reward effect. First, we analyze the reward amount and demonstrate that players reward less in DR under *high-noise* than *perfect-information*. We also note that authorities reward less in CR under *low-noise* than *perfect-information* (Result 3). Second, we analyze the linkage between the player's contribution and received reward and show that *high-noise* reduces such linkage both in DR and CR (Result 4). Third, we show that in perfect information environment reward institutions increase players' beliefs about others' average contribution and hence lead to higher contributions, but this effect is impaired by noise (Result 5). Fourth, we analyze the reactions to received reward and reveal a different reaction pattern for *high-noise* compared to other two information environments (Result 6). Finally, using simulation technology based on our model, we show that both DR and CR sustain higher contributions and welfare in a long run (Result 7).

4.1 Contributions and welfare

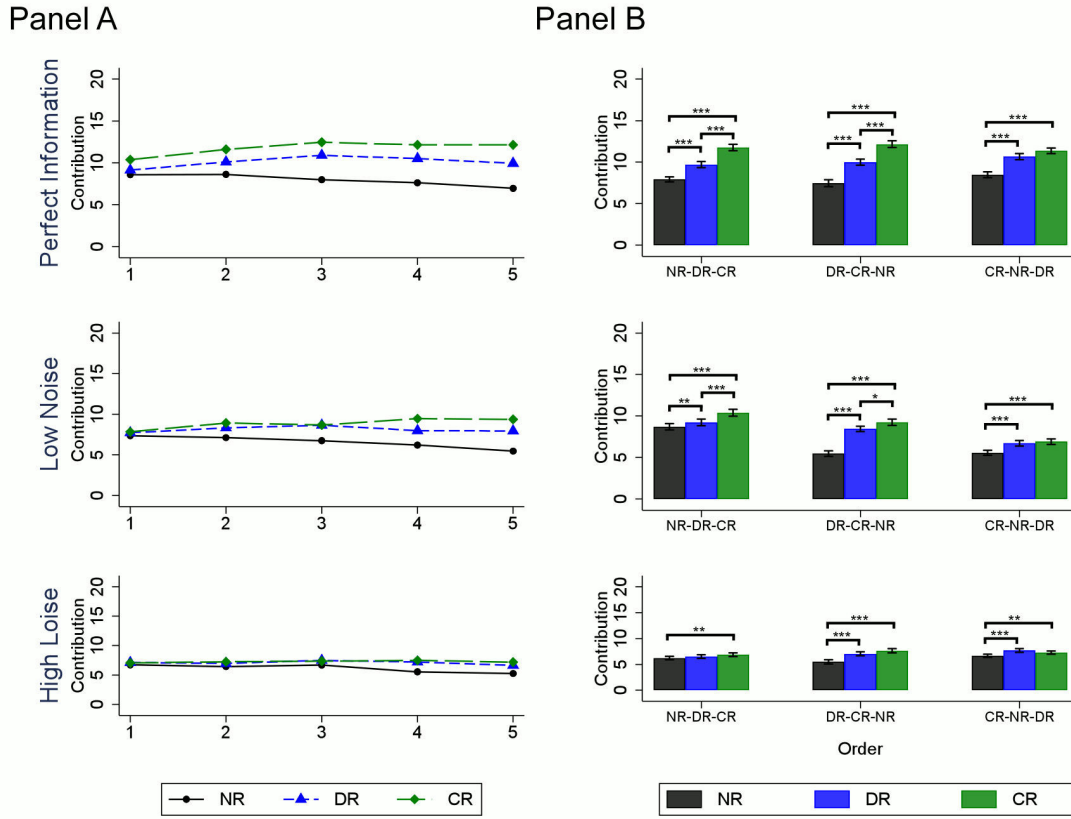


Figure 1. Average Contributions in Treatments and Institutions. Panel A: Average contributions over time in *NR*, *DR* and *CR* by treatment. Panel B: Average contributions over all periods in *NR*, *DR* and *CR* by treatment and the order of institutions. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

To test for institution difference nonparametrically, we apply Wilcoxon matched-pairs signed-rank tests to our data of the first part. Table 1 lists the average contribution, reward, and welfare observed in *NR*, *DR* and *CR* under our three treatments²³. Figure 1a and 2a display the evolution of public good contributions and welfare over time. Moreover, Figure 1b and 2b show the comparison on contributions and welfare of three institutions among three orders.

Table 1. Average Contributions, Reward, and Welfare in Institutions and Treatments

	<i>N</i> participants	Average contribution	Average reward	Average welfare
<i>Perfect-information</i>				
<i>NR</i>	156	7.96		27.77
<i>DR</i>	156	10.12	1.50	32.08
<i>CR</i>	156	11.75	2.22	34.49

Low-noise

²³ The average contribution, reward and welfare of each player in each institution are used for Wilcoxon matched-pairs signed-rank tests.

<i>NR</i>	156	6.58		26.95
<i>DR</i>	156	8.13	1.33	30.53
<i>CR</i>	156	8.86	1.85	32.01
<i>High-noise</i>				
<i>NR</i>	156	6.13		26.68
<i>DR</i>	156	7.08	1.26	29.77
<i>CR</i>	156	7.27	2.31	31.98

As Table 1 reveals, both contributions and welfare are highest in *CR*, followed by *DR* and *NR*. The effect of reward on the contribution is significant. Contributions increase from both *NR* to *DR* and *NR* to *CR*, under *perfect-information*, *low-noise*, and *high-noise* (all p -values < 0.001). These findings provide evidence for our proposition 3. Similarly, welfare also increases from both *NR* to *DR* and *NR* to *CR*, under *perfect-information*, *low-noise*, and *high-noise* (all p -values < 0.001). Figure 1a and 2a suggest that for each treatment, contributions and welfare have a tendency to decrease over time in *NR*, however tend to stabilize over time in *DR* and *CR*. The effect of reward on contributions and welfare is smallest in the first period and largest in the last period. This finding is in agreement with our proposition 4. Figure 1b and 2b reveal robust positive effect of reward institutions on contributions and welfare whatever the order of institutions participated by players in the first part is.

Comparing the effects of reward institutions, *CR* leads to higher contributions than *DR*. This is, however, only statistically significant under *perfect-information* and *low-noise* treatments (all p -values < 0.001), and not significant under *high-noise* treatment ($p = 0.387$). With respect to the average number of assigned reward points, there are significantly less reward points assigned in *DR* under all three information treatments (all p -values < 0.001). Therefore, *CR* always leads to significantly higher welfare than *DR* (all p -values < 0.001).

To complement the nonparametric analysis, we ran ordinary least square (OLS) regressions with robust standard errors clustered on the matching group²⁴. Table 2 lists the results from OLS estimates explaining contributions. In model (1), We find a moderate positive effect of reward on contributions to public good. However, this effect is weakened by noise. In models (2)-(4), we study the effect of reward on contributions for the three treatments separately. Under all three treatments, contributions are significantly higher in both *CR* and *DR* than *NR* ($p = 0.001$ for *DR* vs. *NR* under *perfect-information*, $p < 0.001$ for *CR* vs. *NR* under *perfect-information*, $p = 0.003$ for *DR* vs. *NR* under *low-noise*, $p < 0.001$ for *CR* vs. *NR* under *low-noise*, $p = 0.027$ for *DR* vs. *NR* under *high-noise*, $p = 0.010$ for *CR* vs. *NR* under *high-noise*). Furthermore, the coefficient for *DR* is smaller than for *CR*, and the difference is significant under *perfect-information* ($p < 0.001$), marginally significant under *low-noise* ($p = 0.084$), and not significant under *high-noise* ($p = 0.643$). It seems that higher levels of noise close the gap in contributions between *CR* and *DR*.

²⁴ We also use Tobit regressions to instead of all OLS regressions in this study and get similar estimates.

In the remaining models (5)-(7), we concentrate on the effect of noise on contributions in three institutions. The results confirm that *low-noise* results in lower contributions for all institutions. This is, however, only significant in *DR* ($p = 0.012$) and *CR* ($p = 0.001$), and not significant in *NR* ($p = 0.106$). *High-noise* results in significantly lower contributions in all three institutions ($p = 0.027$ for *NR*, $p < 0.001$ for *DR*, and $p < 0.001$ for *CR*). Furthermore, the coefficient for *low-noise* is smaller than for *high-noise*, and the difference is significant in *CR* ($p = 0.018$), marginally significant in *DR* ($p = 0.089$), and not significant in *NR* ($p = 0.613$). Interestingly, contributions significantly drop within a phase in *NR* ($p < 0.001$), but significantly rise in *CR* ($p = 0.001$).

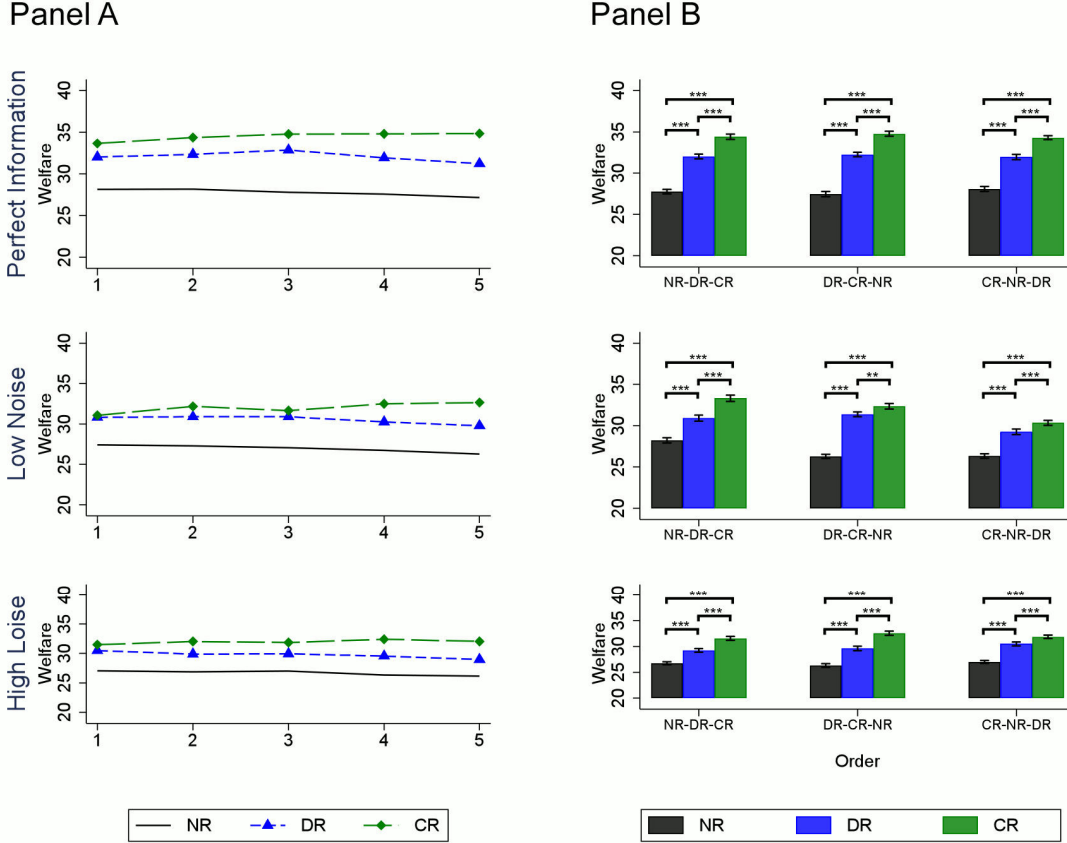


Figure 2. Average Welfare in Treatments and Institutions. Panel A: Average welfare over time in *NR*, *DR* and *CR* by treatment. Panel B: Average welfare over all periods in *NR*, *DR* and *CR* by treatment and the order of institutions. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

These results are similar when using welfare instead of contribution as the dependent variable (models not shown in the table): welfare is highest in *CR*, followed by *DR*, and *NR*. The differences are significant under all three treatments (all p -values < 0.001). Moreover, both *low-noise* and *high-noise* lead to lower welfare in all three institutions. This is at least marginally significant ($p = 0.093$ for *low-noise* vs. *perfect-information* in *NR*, $p = 0.040$ for *high-noise* vs. *perfect-information* in *NR*, $p = 0.008$ for *low-noise* vs. *perfect-information* in *DR*, $p < 0.001$ for *high-noise* vs. *perfect-information* in *DR*, $p = 0.001$ for *low-noise* vs. *perfect-information* in *CR*, $p < 0.001$ for *high-noise* vs. *perfect-information* in *CR*). Furthermore, welfare significantly drops within a phase in *NR* ($p < 0.001$) and *DR* ($p < 0.001$), but significantly rises in *CR* ($p = 0.001$).

Result 1. The two institutions with reward sustain higher contributions and welfare than *NR* even

if the information about other players' behavior becomes noisy. These differences become smaller as the noise level increases. Additionally, contributions and welfare in *CR* tend to be higher than *DR*.

Table 2. OLS Estimations of Contributions Based on Institution and Treatment Dummies

Dependent variable: Contribution							
	All observations	Perfect-Information	Low-Noise	High-Noise	NR	DR	CR
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
<i>DR</i>	2.17*** (0.59)	2.17*** (0.6)	1.54*** (0.48)	0.96** (0.41)			
<i>CR</i>	3.80*** (0.59)	3.80*** (0.59)	2.27*** (0.52)	1.14*** (0.42)			
<i>Low-noise</i>	-1.40* (0.82)				-1.34 (0.82)	-1.97** (0.77)	-3.03*** (0.86)
<i>High-noise</i>	-1.98** (0.80)				-1.73** (0.77)	-3.29*** (0.73)	-4.79*** (0.74)
<i>DR</i> × <i>Low-noise</i>	-0.62 (0.76)						
<i>DR</i> × <i>High-noise</i>	-1.21* (0.72)						
<i>CR</i> × <i>Low-noise</i>	-1.52* (0.78)						
<i>CR</i> × <i>High-noise</i>	-2.66*** (0.72)						
Period in phase	-0.04 (0.06)	0.06 (0.11)	-0.03 (0.08)	-0.14 (0.11)	-0.43*** (0.07)	0.04 (0.09)	0.27*** (0.08)
Constant	5.36* (2.84)	-3.04 (5.79)	2.52 (3.89)	10.03*** (2.97)	11.72*** (3.21)	7.78** (3.15)	2.54 (3.37)
Demography	Yes	Yes	Yes	Yes	Yes	Yes	Yes
N	6480	2160	2160	2160	2160	2160	2160
Cluster	108	36	36	36	108	108	108
Adjusted R ²	0.085	0.083	0.056	0.01	0.031	0.053	0.12

Notes: OLS estimates. *Low-noise* and *High-noise* are treatment dummies with *Perfect-information* as an omitted case, *DR* and *CR* are institution dummies with *NR* as an omitted case. Gender and age are demographic variables. Robust standard errors, clustered on matching group, in parentheses. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

4.2 Choice of institution

As Figure 3 reveals, for the choice of institution in part 2, *CR* attracts the majority of the population under all treatments. About 54% players choose this institution under *low-noise*, and even more under the other treatments (57.64% for *perfect-information* and 61.11% for *high-noise*). *DR* attracts about 40% players under *low-noise*, compared to 37.50% and 34.72% under *perfect-information* and *high-noise*, respectively. In sum, more than 95% of players favor the two reward

institutions ($p = 0.835$, Pearson chi-square test).

We then focus on the contributions and welfare of reward institutions after players' choices. We apply Wilcoxon rank-sum tests to our data of the second part. We find similar results compared to part 1. *CR* leads to significantly higher contributions than *DR* under *perfect-information* ($p = 0.004$). Under both *low-noise* and *high-noise* treatments, contributions are comparable in *CR* and *DR* ($p = 0.348$ for *low-noise* and $p = 0.653$ for *high-noise*). However, in *CR* there is more welfare than *DR* under all three treatments (all p -values < 0.001).

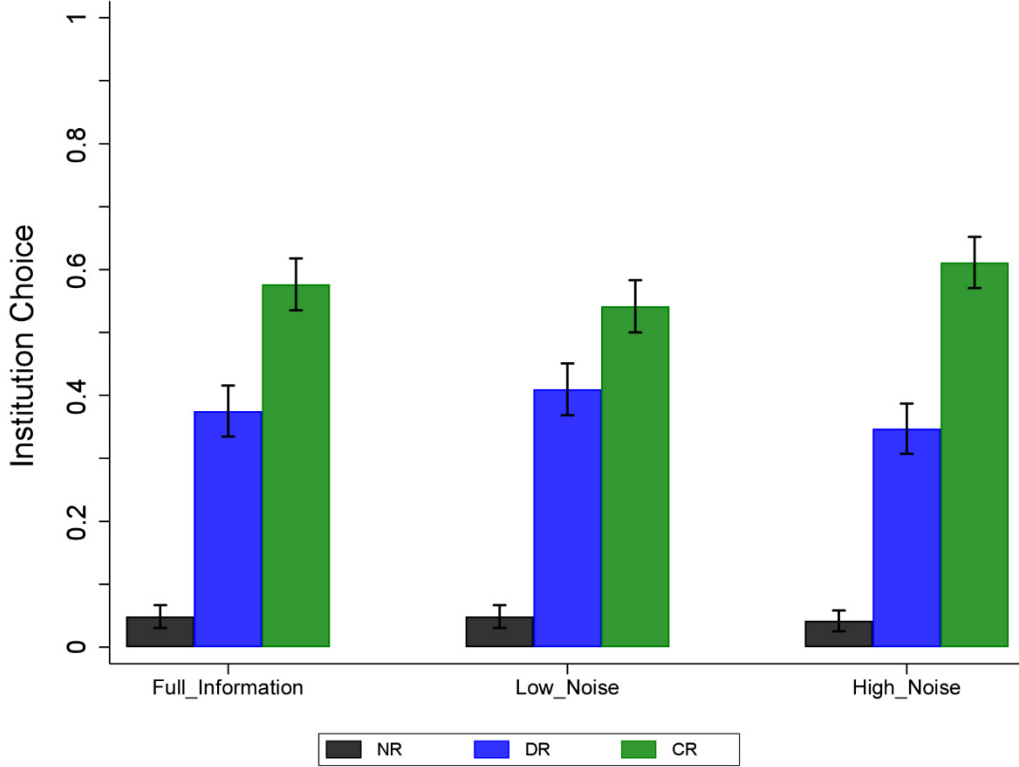


Figure 3. Average choice of institution in the second part.

Result 2. Players predominately choose two reward institutions than *NR* no matter whether the information is noisy or not. These differences are not affected by noise significantly. Additionally, share of choice for *CR* tends to be higher than *DR*.

4.3 Reward strategies

Tables 3 displays the results from OLS estimates explaining the number of reward points assigned to other players in *DR*. In model (1), we find reward points decrease with the increasing of the noise level. That is, in *DR*, *high-noise* leads to the lowest assigned reward points, followed by *low-noise* and *perfect-information*. However, only the difference between *high-noise* and *perfect-information* is significant ($p = 0.031$). Moreover, we find a significantly decaying trend of players' assignment of reward points in *DR* ($p < 0.001$).

In model (2), we find that players assign reward points based on the signals for others' contributions²⁵. More specifically, players assign more reward points to the target as they receive a

²⁵ We also use variable *AverageContr* which is the average of the rewarder's own contribution and other two players' contributions observed by the rewarder instead of *AverageSignal*. Both AIC and BIC statistics demonstrates that

larger signal for the target's contribution (*Signal*) and as they receive smaller signals for another two players' contributions (*AverageSignal*). In addition, noise leads to smaller reward points even if we control the received signals. The difference is just marginally significant between *high-noise* and *perfect-information*. This result suggests that players assign less reward points under *high-noise* than *perfect-information* even if they observe the same signals.

In model (3), we explore the effect of noise on the reward strategy in *DR*. We find that *high-noise* makes the players' assignment of rewards significantly less dependent on their received signals ($p = 0.014$ for the signal of the target's contribution and $p < 0.001$ for signals of other players' contributions). However, the dependence of the assigned reward on signal for target's contribution is not significantly influenced by *low-noise* ($p = 0.944$), and that for other players' contributions is significantly reduced by *low-noise* ($p = 0.007$). In the remaining models (4)-(6), we explore the effect of noise on players' reward strategies for the three treatments separately. The results confirm that players' assignment of reward points to target significantly depends on both the signal for the target's contribution and the signals for another two players' contributions even under *low-noise*, although this dependence is smaller than *perfect-information*.

Table 3. OLS Estimations of Reward Strategy in DR

Dependent variable: Reward point						
	All observations			Perfect- Information	Low- Noise	High- Noise
	(1)	(2)	(3)	(4)	(5)	(6)
<i>Signal</i>		0.040*** (0.003)	0.048*** (0.005)	0.048*** (0.005)	0.048*** (0.005)	0.032*** (0.003)
<i>Low-noise</i>	-0.056 (0.044)	-0.023 (0.041)	-0.209** (0.083)			
<i>High-noise</i>	-0.089** (0.041)	-0.067* (0.04)	-0.168** (0.079)			
<i>AverageSignal</i>		-0.023*** (0.005)	-0.040*** (0.005)	-0.040*** (0.005)	-0.020*** (0.004)	-0.014*** (0.003)
<i>Signal</i> × <i>Low-noise</i>			0.001 (0.007)			
<i>Signal</i> × <i>High-noise</i>			-0.016*** (0.006)			
<i>AverageSignal</i> × <i>Low-noise</i>			0.020*** (0.007)			
<i>AverageSignal</i> × <i>High-noise</i>			0.026*** (0.006)			
<i>Period in phase</i>	-0.050*** (0.005)	-0.050*** (0.005)	-0.050*** (0.005)	-0.055*** (0.007)	-0.051*** (0.008)	-0.045*** (0.009)
Constant	-0.089** (0.213)	0.556*** (0.203)	0.662*** (0.205)	0.506 (0.485)	0.549* (0.277)	0.516 (0.322)
Demography	Yes	Yes	Yes	Yes	Yes	Yes

AverageSignal is better to explain the variance of reward points than *AverageContr*.

N	6300	6300	6300	2160	1980	2160
Cluster	108	108	108	36	36	36
Adjusted R ²	0.016	0.176	0.186	0.197	0.208	0.144

Notes: OLS estimates. *Low-noise* and *High-noise* are treatment dummies with *Perfect-information* as an omitted case, *Signal* is a continuous variable representing the target's contribution level observed by the rewarder, *AverageSignal* is the average of rewarder's received signals except the target in *DR*. Gender and age are demographic variables. Robust standard errors, clustered on matching group, in parentheses. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Tables 4 displays the results from OLS estimates explaining the number of reward points assigned by authorities to other players in *CR*. In model (1), we find *low-noise* leads to lowest reward points, followed by *perfect-information* and *high-noise*. That is, in *CR*, *low-noise* decreases the reward points but *high-noise* increases the reward points compared to *perfect-information*. However, only the difference between *low-noise* and *perfect-information* is significant ($p = 0.013$). Unlike *DR*, we find a significantly growing trend of authorities' assignment of reward points in *CR* ($p = 0.011$).

In model (2), we find authorities assign reward points based on the signals for other players' contributions. Similar to players in *DR*, authorities assign more reward points to the target as they receive a larger signal for the target's contribution (*Signal*) and as they receive smaller signals for other three players' contributions (*AverageSignal*). Unlike *DR*, *high-noise* leads to significantly larger reward points even if we control the received signals. This result suggests that authorities assign more reward points under *high-noise* than *perfect-information* even if they observe the same signals.

In model (3), we explore the effect of noise on the reward strategy in *CR*. Merely a weak effect of noise on authorities' reward strategies is observed. To be specific, only the dependence of the assigned reward on the signal for the target's contribution is marginally significantly reduced by *high-noise* ($p = 0.059$). This result suggests that the reward strategy of authority in *CR* is less affected by noise than that of players in *DR*. In the remaining models (4)-(6), we explore the effect of noise on authorities' reward strategies for the three treatments separately. The results confirm that authorities' assignment of reward points to target significantly depends on both the signal for the target's contribution and the signals for another three players' contributions even under *high-noise*, although this dependence is slightly smaller than *perfect-information*.

Table 4 OLS Estimations of Reward Strategy in CR

Dependent variable: Reward point						
	All observations			Perfect- Information	Low- Noise	High- Noise
	(1)	(2)	(3)	(4)	(5)	(6)
<i>Signal</i>		0.17*** (0.01)	0.18*** (0.022)	0.18*** (0.022)	0.20*** (0.02)	0.16*** (0.014)
<i>Low-noise</i>	-0.44** (0.17)	-0.16 (0.16)	-0.79** (0.32)			
<i>High-noise</i>	0.19 (0.19)	0.44** (0.17)	0.50 (0.33)			

<i>AverageSignal</i>	-0.079***	-0.11***	-0.13***	-0.076***	-0.076***
	0.013	(0.023)	(0.022)	(0.023)	(0.018)
<i>Signal</i> × <i>Low-noise</i>		0.023			
		(0.029)			
<i>Signal</i> × <i>High-noise</i>		-0.046*			
		(0.026)			
<i>AverageSignal</i> × <i>Low-noise</i>		0.042			
		(0.032)			
<i>AverageSignal</i> × <i>High-noise</i>		0.035			
		(0.029)			
<i>Period in phase</i>	0.05**	0.028	0.025	-0.003	0.023
	(0.018)	(0.019)	(0.019)	0.039	0.03
Constant	0.88	0.77	1.27	-0.27	-1.29
	(1.44)	(1.25)	(1.24)	(1.70)	(1.69)
Demography	Yes	Yes	Yes	Yes	Yes
N	2160	2160	2160	720	720
Cluster	108	108	108	36	36
Adjusted R ²	0.023	0.40	0.41	0.45	0.49

Notes: OLS estimates. *Low-noise* and *High-noise* are treatment dummies with *Perfect-information* as an omitted case, *Signal* is a continuous variable representing the target's contribution level observed by the authority, *AverageSignal* is the average of authority's received signals except the target in *CR*. Gender and age are demographic variables. Robust standard errors, clustered on matching group, in parentheses. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Result 3. Noise has different effects on assigned reward points in *DR* and *CR*. In *DR*, assigned reward points are less under *high-noise* than *perfect-information* but comparable under *low-noise* to *perfect-information*. In *CR*, however, assigned reward points are less under *low-noise* than *perfect-information*, but comparable under *high-noise* to *perfect-information*.

As the assignment of both players and authorities is less dependent on signals, which are positively related to true contribution levels when the contribution information is quite inaccurate, we suspect that the linkage between contributions and assigned rewards is reduced by noise. We then examine this conjecture. Tables 5 displays the results from OLS estimates explaining the linkage between the number of reward points assigned and the contribution levels in *DR* and *CR*. In model (1), we find in *DR*, *high-noise* significantly reduces the positive correlation between contributions and assigned rewards ($p < 0.001$), while this effect is not observed for *low-noise* ($p = 0.399$). In models (2)-(4), we explore the linkage between contributions and assigned rewards for the three treatments separately. The results confirm that the more contribution levels, the more assigned reward points. This is significant under all three treatments ($p < 0.001$ for *perfect-information*, $p < 0.001$ for *low-noise*, and $p = 0.002$ for *high-noise*). In the remaining models (5)-(8), we find similar results for *CR*. That is, there exists a positive linkage between contributions and assigned rewards under all treatments in *CR*, and this linkage is reduced by *high-noise*.

Table 5. OLS Estimations of Assigned Reward Point Based Contribution and Treatment

Dummies

		Dependent variable: Assigned reward point							
		DR				CR			
		All	Perfect- Informati on	Low- Noise	High- Noise	All	Perfect- Informa tion	Low- Noise	High- Noise
		(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
<i>Contribution</i>		0.11*** (0.01)	0.11*** (0.01)	0.13*** (0.01)	0.04*** (0.01)	0.14*** (0.02)	0.13*** (0.02)	0.15*** (0.02)	0.06*** (0.02)
<i>Contribution</i>	×	0.02 (0.02)				0.02 (0.02)			
<i>Contribution</i>	×	-0.07*** (0.02)				-0.08*** (0.02)			
<i>Low-noise</i>									
<i>High-noise</i>									
<i>Low-noise</i>		-0.08 (0.15)				-0.10 (0.26)			
<i>High-noise</i>		0.59*** (0.14)				1.27*** (0.27)			
<i>Period in phase</i>		-0.16*** (0.02)	-0.18*** (0.02)	-0.14*** (0.02)	-0.14*** (0.03)	0.01 (0.02)	-0.04 (0.05)	0.01 (0.03)	0.06* (0.03)
Constant		1.50*** (0.43)	2.43** (1.15)	1.78** (0.72)	1.28** (0.48)	0.71 (0.85)	-1.11 (1.06)	-0.86 (1.11)	4.11*** (0.93)
Demography	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
N		2160	720	720	720	2160	720	720	720
Cluster		108	36	36	36	108	36	36	36
Adjusted R ²		0.23	0.28	0.32	0.065	0.23	0.31	0.32	0.064

Note: OLS estimates. *Low-noise* and *High-noise* are treatment dummies with *Perfect-information* as an omitted case. Gender and age are demographic variables. Robust standard errors, clustered on matching group, in parentheses. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Result 4. There exists a significantly positive linkage between contributions and assigned rewards in both *DR* and *CR* under all three treatments. However, if the contribution information is quite inaccurate, this linkage is significantly reduced.

4.4 Conditional cooperation and belief updating

Subject to the conditional cooperation norm, players contribute to the public good based on their beliefs about the average amount of other players' contributions. Table 6 reveals this phenomenon. In model (1), we find players increase their contributions with their beliefs about the average amount of other players' contributions (*Belief*). As their beliefs increase 1 token, their contributions increase about 0.86 token ($p < 0.001$). As shown in model (2), this effect is neither influenced by reward institutions nor noise. Moreover, given the same belief, *DR* and *CR* still lead the beliefs higher about 1 token than *NR* (all p -values < 0.001). These Differences are significant and provide evidence for our proposition 2.

Table 6. OLS Estimations of Contribution Based Belief and Treatment Dummies

	Dependent variable: Contribution	
	(1)	(2)
<i>Belief</i>	0.86*** (0.04)	0.90*** (0.06)
<i>Belief</i> × <i>DR</i>		-0.04 (0.04)
<i>Belief</i> × <i>CR</i>		0.00 (0.04)
<i>Belief</i> × <i>Low-noise</i>		-0.02 (0.09)
<i>Belief</i> × <i>High-noise</i>		-0.08 (0.10)
Constant	-0.36 (1.62)	-0.65 (1.78)
Control	Yes	Yes
Demography	Yes	Yes
N	6480	6480
Cluster	108	108
Adjusted R ²	0.44	0.44

Note: OLS estimates. Control variables include *Low-noise*, *High-noise*, *DR* and *CR*. *Low-noise* and *High-noise* are treatment dummies with *Perfect-information* as an omitted case. *Belief* is a continuous variable represents player's expectation of other group members' average contribution. Gender and age are demographic variables. Robust standard errors, clustered on matching group, in parentheses. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Table 7 reveals how reward institutions and noise change players' beliefs. In model (1), we find both *DR* and *CR* lead to significantly higher beliefs than *NR* ($p = 0.004$ for *DR* vs. *NR* and $p < 0.001$ for *CR* vs. *NR*). This result supports our proposition 3. Differences in beliefs between reward institutions and *NR* are reduced by noise. However, the adverse effect of *low-noise* on the difference between *DR* and *NR* is not significant ($p = 0.323$). We then explore the mechanism players update their beliefs.

In model (2), we estimate the current-period belief of players based on the positive beliefs (*Belief₀*) and average signals from other players in previous periods (*Presignal*, *Pre2signal*, *Pre3signal*, *Pre4signal*). We find both positive beliefs and signals from previous two periods have significantly positive effects on current beliefs. That is, players anchor their beliefs to positive beliefs and update them based on the observed signals from previous two periods. Furthermore, the weight for positive beliefs and signals from the previous period are twice as large as the weight for signals from two periods ago.

In the next step, we examine whether noise influences players' positive beliefs and their belief-updating procedures. In model (3), we find that players updating their current beliefs depend more on positive beliefs and less on signals from the previous period under noise. This is significant for both *low-noise* and *high-noise*. In the remaining models (4)-(6), we explore the effect of noise on the belief-updating procedure for the three treatments separately. The results confirm that not the belief-updating procedure per se, but the weight for positive beliefs and signals is changed by noise.

In model (7), we find players' positive beliefs decrease with the level of noise ($p = 0.046$ for *low-noise* vs. *perfect-information* and $p = 0.006$ for *high-noise* vs. *perfect-information*), although the coefficient for *low-noise* is not significantly different from that for *high-noise* ($p = 0.481$). Moreover, the positive beliefs in reward institutions are not significantly different from that in *NR*.

Table 7. OLS Estimations of Current Belief Based on Previous Signals

Dependent variable:	Belief						Belief ₀
	All Observations		Perfect-Information	Low-Noise	High-Noise	All Observations	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
<i>DR</i>	1.41*** (0.48)	0.62*** (0.23)	0.61*** (0.16)	0.42* (0.22)	0.91*** (0.26)	0.51 (0.32)	-0.1 -0.35
<i>CR</i>	2.58*** (0.51)	0.81*** (0.21)	0.77*** (0.14)	0.60*** (0.18)	1.14*** (0.21)	0.57* (0.31)	0.62 -0.39
<i>Low-noise</i>	-1.00 (0.7)	0.00 (0.28)	-0.17 (0.53)				-1.25** -0.62
<i>High-noise</i>	-1.13* (0.65)	-0.78** (0.31)	1.05 (0.76)				-1.67*** -0.6
<i>DR</i> × <i>Low-noise</i>	-0.61 (0.61)						-0.19 -0.46
<i>DR</i> × <i>High-noise</i>	-1.02* (0.55)						-0.05 -0.43
<i>CR</i> × <i>Low-noise</i>	-1.11* (0.66)						-0.19 -0.5
<i>CR</i> × <i>High-noise</i>	-2.06*** (0.59)						-0.6 -0.52
<i>Belief₀</i>		0.38*** (0.05)	0.24*** (0.06)	0.24*** (0.07)	0.52*** (0.06)	0.49*** (0.06)	
<i>Presignal</i>		0.36*** (0.04)	0.62*** (0.05)	0.62*** (0.05)	0.33*** (0.05)	0.22*** (0.03)	
<i>Pre2signal</i>		0.18*** (0.04)	0.11** (0.04)	0.11** (0.05)	0.14*** (0.04)	0.07** (0.03)	
<i>Pre3signal</i>		0.06 (0.05)					
<i>Pre4signal</i>		0.00 (0.03)					
<i>Belief₀</i> × <i>Low-noise</i>			0.29*** (0.09)				
<i>Belief₀</i> × <i>High-noise</i>			0.24*** (0.09)				
<i>Presignal</i> × <i>Low-noise</i>			-0.28*** (0.07)				
<i>Presignal</i> × <i>High-noise</i>			-0.40*** (0.06)				

<i>Pre2signal</i> × <i>Low-noise</i>			0.02				
			(0.06)				
<i>Pre2signal</i> × <i>High-noise</i>			-0.04				
			(0.05)				
Constant	6.87***	1.11	1.64	-1.27	2.48	3.70**	7.62***
	(1.96)	(1.56)	(1.27)	(2.09)	(2.23)	(1.77)	-1.77
Demography	Yes	Yes	Yes	Yes	Yes	Yes	Yes
N	6480	1296	3888	1296	1296	1296	1296
Cluster	108	108	108	36	36	36	108
Adjusted R ²	0.074	0.52	0.53	0.66	0.53	0.28	0.043

Note: OLS estimates. *Low-noise* and *High-noise* are treatment dummies with *Perfect-information* as an omitted case, *DR* and *CR* are institution dummies with *NR* as an omitted case. Variable *Belief₀* is the players' positive beliefs. Variable *Presignal* is the average signals from other players in the previous period. Similarly, variables *Pre2signal*, *Pre3signal*, and *Pre4signal* are the average signals from other players two periods ago, three periods ago and four periods ago, respectively. Gender and age are demographic variables. Robust standard errors, clustered on matching group, in parentheses. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Result 5. Both *DR* and *CR* increase players' beliefs about others' average contribution which results in higher contributions under *perfect-information* treatment. However, for *DR* this effect is marginally significantly impaired by *high-noise*. For *CR*, this effect is marginally significantly reduced by *low-noise* and significantly reduced by *high-noise*.

4.5 Reactions to received reward

We employ OLS regression analysis to analyze reactions to received reward and other previous experience of the players. In model (1) of Table 8, we estimate the current-period contribution of players based on the number of reward points she received in the previous period (*PreRew*, multiplied with *PreContr*). We control for the previous-period contribution (*PreContr*) of this player, the current period belief about the average amount of other players' contributions (*Belief*), number of periods in the current phase (*Period in phase*), demographic factors and other treatment factors (i.e., *DR*, *CR*, *high-noise* and *low-noise*). We find that the previous-period contribution has a large and significant positive effect ($p < 0.001$).

Our main interest, however, lies in the interaction term (*PreContr* × *PreRew*). For noncontributors, we find a marginally significant negative effect of the previous-period received reward points ($p = 0.093$)²⁶. This negative effect, however, is significantly decreased with the increasing previous-period contributions ($p < 0.001$). When the previous-period contribution is larger than 4, such effect changes to be positive, that is, the higher the received reward points, the larger contribution in the next period. In model (2) and (3), we explore the effect of reward institution on the reactions to received rewards. We find similar reaction patterns between *DR* and *CR*. In contrast to model (1), for noncontributors, we find the negative effect of the previous-period received reward points is insignificant ($p > 0.1$).

In the remaining models (4)-(6), we examine the effect of noise on the reactions to received

²⁶ Although previous-period noncontributors are not able to decrease their contribution level, the negative effect of the previous-period received reward points might represent a lower probability of free-ride in the current-period.

rewards. The results reveal an essential role of *high-noise* in changing the reaction patterns of players. Under *perfect-information* and *low-noise*, we find an insignificant effect of the previous-period received rewards on the current-period contributions ($p > 0.1$). This suggests that when the information is perfect or the noise is small, players are able to expect the rewards assigned to them accurately, and hence learning effect does not exist²⁷. Moreover, we can infer that received rewards per se hardly have an effect on players' subsequent contributions²⁸. Under *high-noise*, however, we find a significant interaction effect of the previous-period received reward points and contributions ($p < 0.001$). For noncontributors, we find a significant negative effect of the previous-period received reward points ($p = 0.001$). This negative effect, nevertheless, is significantly decreased with the increasing previous-period contributions ($p < 0.001$). When the previous-period contribution is larger than 6, such effect changes to be positive. This result reveals a learning procedure of players for the linkage between assigned rewards and contributions under *high-noise*. For players with lower contributions received rewards, they learn the correlation between their assigned rewards and contributions is small, and hence reduce their contributions in the next period. On the other hand, for players with higher contributions received rewards, they learn the correlation between their assigned rewards and contributions is large, and hence increase their contributions in the next period.

Table 8. OLS Estimations of Current Contribution Based on Previous Period Behavior

	Dependent variable: Contribution					
	All observations			Perfect- Information	Low- Noise	High- Noise
	All	<i>DR</i>	<i>CR</i>			
(1)	(2)	(3)	(4)	(5)	(6)	
<i>PreContr</i>	0.40*** (0.05)	0.38*** (0.06)	0.41*** (0.05)	0.35*** (0.1)	0.54*** (0.06)	0.28*** (0.07)
<i>PreRew</i>	-0.17* (0.1)	-0.15 (0.16)	-0.18 (0.12)	-0.03 (0.29)	0.04 (0.14)	-0.47*** (0.12)
<i>PreContr</i> × <i>PreRew</i>	0.04*** (0.01)	0.05*** (0.02)	0.03** (0.01)	0.02 (0.02)	0.02 (0.01)	0.07*** (0.02)
<i>Belief</i>	0.51*** (0.04)	0.50*** (0.06)	0.52*** (0.05)	0.59*** (0.08)	0.43*** (0.07)	0.52*** (0.07)
<i>CR</i>	0.13 (0.13)			0.29 (0.22)	0.19 (0.22)	0.04 (0.25)
<i>Period in phase</i>	-0.22*** (0.06)	-0.23** (0.09)	-0.20*** (0.07)	-0.38*** (0.11)	-0.20** (0.1)	-0.11 (0.11)
<i>Low-noise</i>	-0.44* (0.24)	-0.46 (0.31)	-0.42* (0.25)			
<i>High-noise</i>	-0.76*** (0.25)	-0.64** (0.29)	-0.84*** (0.31)			

²⁷ If players wrongly expect the reward points assigned to them, they will change their expectation of assigned reward based on the reward points they truly received and hence adjust their subsequent choice of contribution levels. If this was true, we should have found a significant effect of received reward points on subsequent contributions. This finding supports our assumption 4 in the theoretical model.

²⁸ This inference is consistent with our psychological utility function in contribution stage, which describes that received rewards per se have no effect on the players' choice of contribution.

Constant	1.02 (1.26)	1.01 (1.51)	1.04 (1.39)	-0.52 (3.56)	-0.25 (1.3)	1.88 (2.07)
Demography	Yes	Yes	Yes	Yes	Yes	Yes
N	3456	1728	1728	1152	1152	1152
Cluster	108	108	108	36	36	36
Adjusted R ²	0.56	0.51	0.6	0.55	0.61	0.43

Notes: OLS estimates. *Low-noise* and *High-noise* are treatment dummies with *Perfect-information* as an omitted case. *PreContr* and *PreRew* are continuous variables representing player's contribution and assigned reward points in the previous period, *Belief* indicates player's expectation of other group members' average contribution. Gender and age are demographic variables. Robust standard errors, clustered on matching group, in parentheses.

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Result 6. Under *perfect-information* and *low-noise*, players choose their current-period contributions independently on their previous-period received reward points. Under *high-noise*, however, players have a different reaction pattern. For players contributing lower than 6, the larger the received reward points, the lower contribution in the next period. On the other hand, for players contributing higher than 6, the direction is opposite.

4.6 Simulation results of contributions

According to our experimental setup, we assign values to some parameters of our theoretical model. We set $m = 20$, $\gamma = 0.4$, $\kappa = 3$, $\mu = 10$, $\lambda = 1$ for *perfect-information*, $\lambda = 0.895$ for *low-noise*, and $\lambda = 0.475$ for *high-noise*²⁹. For the parameters α and δ , we set a series of values from 0.5 to 5, spaced at intervals of 0.5^{30} . Based on these parameters, we simulate players' contributions in 50 periods. Figure 4 displays the simulation results of differences in contributions between reward institutions and *NR* in periods 1-5 and 45-50.

In figure 4a, we find that for all values of parameters α and δ , *DR* leads to contributions no less than *NR* under three treatments in the first 5 periods. Specifically, for $\delta \geq 0.1$ *DR* results in a higher contribution under *perfect-information*. For $\delta \geq 0.1$ and $\alpha \geq 0.1$ or $\delta \geq 0.15$, *DR* results in a higher contribution under *low-noise*. For $\delta \geq 0.2$ and $\alpha \geq 0.1$ or $\delta \geq 0.25$, *DR* results in a higher contribution under *high-noise*. Moreover, we find that for all values of parameters α and δ , *CR* leads to contributions higher than *NR* under all treatments in the first 5 periods.

In figure 4b, we find that for all values of parameters α and δ , the effects of reward institutions on contributions are bigger in the last 5 periods (periods 45-50) than in the first 5 periods (periods 1-5) under *perfect-information* and *low-noise*. However, this trend is undermined by *high-noise* because the contributions sustain a higher level in *NR* in the last 5 periods under a high noisy environment than under perfect information and low noisy environment. Even so, for any values of parameters α and δ making contribution levels for reward institutions higher than *NR* in the first 5 periods, result in higher contributions for reward institutions in the last 5 periods as well.

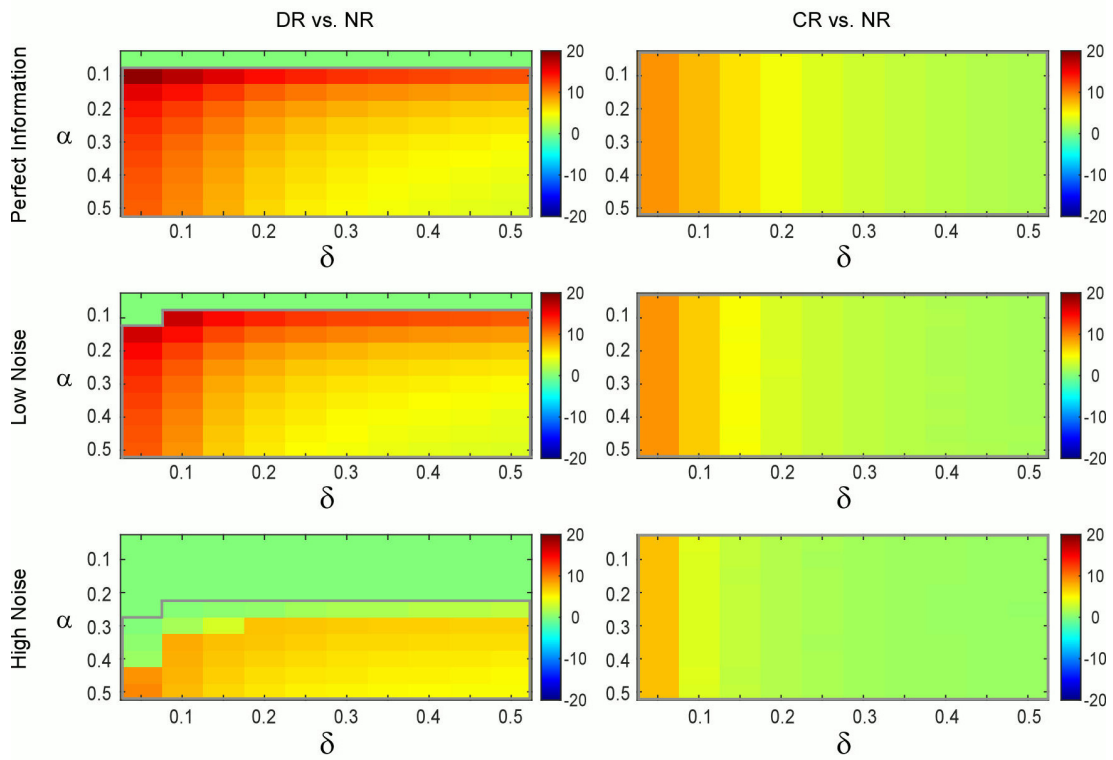
²⁹ When $\lambda = 0.895$, the probability of an accurate signal is 0.9, when $\lambda = 0.475$, the probability of an accurate signal is 0.5.

³⁰ We set the values of parameters α and δ no more than 0.5 because these parameters have a little effect on contributions when they are more than 0.5.

These results are similar when using welfare instead of contributions as the simulation object, and provide evidence to support our proposition 4 and 5.

Result 7. The two institutions with reward improve contributions and welfare in the long run even if the information about other players' behavior becomes noisy. This effect becomes smaller as the noise level increases.

Panel A. Average Contribution Difference over periods 1-5.



Panel B. Average Contribution Difference over periods 45-50.

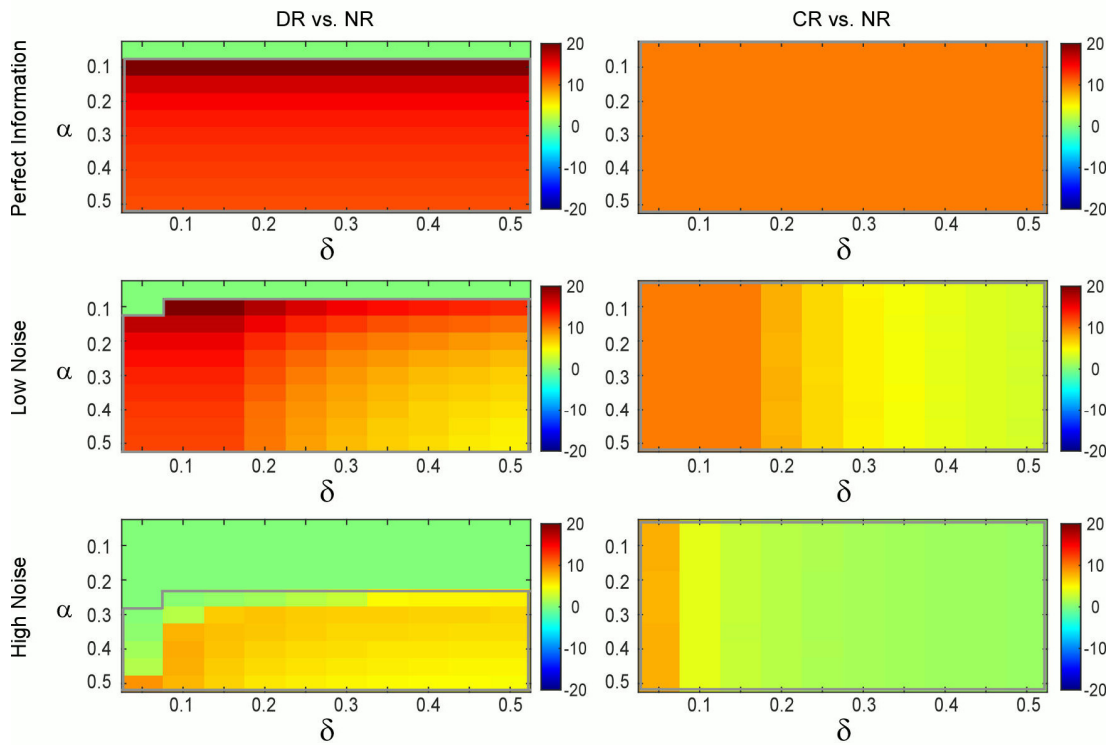


Figure 4. Simulation results of average contribution difference between reward institutions and institution without reward. The vertical axis represents the values of parameter α and the horizontal axis represents the values of parameter δ , the color in each cell represents the simulated contribution levels. The grey box marks the parameters that make average contributions higher in *DR* and *CR* than *NR* over periods 1-5.

5 Discussion and Conclusion

Our study examines whether the institution with reward option is an efficient and satisfactory way to solve social dilemma problem under different information conditions. We vary the accuracy of the information concerning the contributions of other participants in the group. We show that reward institutions sustain higher cooperation levels and let participants get more welfare under all information conditions. Furthermore, we find most participants tend to favor reward institutions no matter whether the information is perfect.

In line with previous literature on decentralized reward institution under perfect-information, we observe that contributions in the first period of *DR* are similar to *NR* (Walker,2004; Sefton,2007; Rand et al., 2009). Subjects, on average, allocate about 40%-50% of their endowments to the public good at the beginning of the repeated game in *NR* and these behaviors are consistent with findings in standard voluntary contribution mechanism (VCM) of other studies (Fehr & Gächter, 2000,2002; Walker,2004; Sefton,2007). As periods go on, contributions in *NR* decay quickly while that in *DR* maintains. Therefore, participants in *DR* tend to contribute more than *NR*. This finding replicates the results of Sefton (2007) and Rand et al. (2009).

Compared to the literature on punishment under perfect-information, the effects of reward institutions on cooperation in our study are relatively weaker. In classical studies of Fehr and Gächter, institutions with punishment opportunity maintain the average level of contributions at about 60%-80% of subjects' endowments in the final period (Fehr & Gächter, 2000,2002). In our study, participants contribute about 50% of their endowments to the public good in the final period of *DR* under perfect-information, and this contribution level increases to 60% in the final period of *CR*.

However, in a realistic environment with inaccurate information about others' contributions, the effects of reward institutions are close to punishment institutions. In a slightly noisy environment, contributions of decentralized punishment institutions continue to decay and decline to about 10% and 40% of subjects' endowments in the final period in Ambrus et al. (2019) and Ambrus & Greiner (2012) respectively. In our study, we find a steady contribution level in reward institutions and participants contribute about 40% and 45% of their endowments in the final period of *DR* and *CR*, respectively. As noise increases to a high level, the contribution levels in punishment institutions drop dramatically (about 20%-50% of endowment) while that in reward institutions only drop about 10% of endowment. This suggests that reward institutions might be more efficient than punishment institutions in a realistic world with noisy information.

The welfare benefits of reward institutions better than punishment institutions are self-evident. This is because the reward per se improves the welfare while punishment per se damages the welfare. Under perfect-information, the advantages of punishment institutions in cooperation compensate the flaw of punishment institutions on welfare. However, as information tends to be noisier, the advantages of punishment institutions in cooperation shrink while the flaw of punishment institutions in welfare is amplified. Therefore, punishment institutions lead to higher welfare in a perfect information environment while lead to similar or lower welfare in an imperfect information environment (Gächter et al., 2008; Ambrus & Greiner, 2012; Nicklisch et al., 2016; Ambrus et al., 2019). In contrast to punishment, we find that reward institutions lead to higher welfare even if the level of noise is quite large. Our findings highlight the superiority of reward institutions over punishment institutions in promoting cooperation and welfare in a realistic world with noisy information.

Of the two reward institutions, centralized reward is the most efficient and popular under all information conditions. This might account for the choices of reward institutions, especially the prevalence of centralized reward (i.e., commendation meeting) in reality. Nicklisch et al. (2016) analyze information conditions, which lead subjects to voluntarily subordinate themselves to punishment institutions. They demonstrate that an environment with perfect-information tends to favor the decentralized punishment institution, while a high level of noise favors an institution with no punishment. Under low levels of noise, they observe undifferentiated preference on institutions with punishment and without punishment. In our study, however, subjects consistently favor the centralized reward institution over the institution without reward under all information environments. So, we infer subjects are highly likely to favor centralized reward institutions over all other institutions (i.e., institutions with punishment, the institution without punishment or reward, and the institution with decentralized reward) in a realistic world.

Subjects contribute more in reward institutions under high noisy environment for three reasons. First, in reward institutions, one main aim of subjects is to get rewards to earn a higher payoff. If they believe that more contributions lead them to earn more rewards, they will choose higher amounts of contribution in reward institutions than the institution without reward. Under high noisy environment, we still find a significant positive correlation between contribution levels and assigned rewards. Subjects learned this positive correlation, and hence increase their contributions in the reward institutions. Second, unlike antisocial punishment which reduces players' subsequent contributions through the channel of anger (Herrmann et al., 2008; Hopfensitz & Reuben, 2009; Ambrus & Greiner, 2012), antisocial reward per se hardly has any effect on players' contributions. Therefore, even though players with lower contributions often receive reward points because the quite inaccurate information, the assigned rewards per se doesn't decrease their contributions in the next period. Third, through the reinforcement learning process, subjects expect others to contribute more in reward institutions than the institution without reward. Following the conditional cooperation norm that comes up in Fehr and Schurtenberger (2018), subjects increase their contributions with their beliefs about other group members' average contributions. This is corroborated by our data as well. Therefore, higher beliefs lead to higher contribution levels in reward institutions.

This study contributes to understand the evolution of social structure in humans. Early societies allowing for nearly perfect observation of others tend to apply decentralized punishment regimes. In maturing societies, production increases complexity. It becomes somewhat difficult to monitor others' contributions. These are the circumstances, in which people are willing to employ centralized punishment or reward regime instead. Under the condition under which high level of cooperation is essential to the survival of groups, centralized punishment is preferred. Otherwise, centralized reward is more desirable. In the past decades, innovation has grown up to be an important driver of economic growth. In the process of innovation, intellectual capital occupies an increasing proportion of input. Because of the high degree of uncertainty in the outcome of innovation and the extreme difficulty of monitoring intellectual capital input, centralized reward becomes fashionable in this circumstance.

Recently, the epidemic of COVID-19 has given rise to another interesting development, as it increases the space distance in people's work and decreases the transparency of actions among group members. As a consequence, we might expect that reward institutions are used more often to strengthen cooperation. A promising direction for future research is examining this guess using data

from real-world environments.

Data availability

The datasets generated during the current study are available upon request.

Reference

- Ambrus, Attila and Ben Greiner.** 2012. "Imperfect Public Monitoring with Costly Punishment: An Experimental Study." *American Economic Review*, 102(7), 3317-32.
- Ambrus, Attila and Ben Greiner.** 2019. "Individual, Dictator, and Democratic Punishment in Public Good Games with Perfect and Imperfect Observability." *Journal of Public Economics*, 178, 104053.
- Andreoni, James and Laura K. Gee.** 2012. "Gun for Hire: Delegated Enforcement and Peer Punishment in Public Goods Provision." *Journal of Public Economics*, 96(11-12), 1036-46.
- Battigalli, Pierpaolo and Martin Dufwenberg.** 2022. "Belief-Dependent Motivations and Psychological Game Theory." *Journal of Economic Literature*, 60(3), 833-82.
- Bereby-Meyer, Yoella and Alvin E Roth.** 2006. "The Speed of Learning in Noisy Games: Partial Reinforcement and the Sustainability of Cooperation." *American Economic Review*, 96(4), 1029-42.
- Boyd, R.; H. Gintis and S. Bowles.** 2010. "Coordinated Punishment of Defectors Sustains Cooperation and Can Proliferate When Rare." *Science*, 328(5978), 617-20.
- Dickinson, David L.; David Masclet and Marie Claire Villeval.** 2015. "Norm Enforcement in Social Dilemmas: An Experiment with Police Commissioners." *Journal of Public Economics*, 126, 74-85.
- Dong, Yali; Tatsuya Sasaki and Boyu Zhang.** 2019. "The Competitive Advantage of Institutional Reward." *Proceedings of the Royal Society B*, 286(1899), 20190001.
- Dufwenberg, Martin and Georg Kirchsteiger.** 2004. "A Theory of Sequential Reciprocity." *Games and economic behavior*, 47(2), 268-98.
- Faillo, Marco; Daniela Grieco and Luca Zarri.** 2013. "Legitimate Punishment, Feedback, and the Enforcement of Cooperation." *Games and Economic Behavior*, 77(1), 271-83.
- Fehr, Ernst and Simon Gächter.** 2002. "Altruistic Punishment in Humans." *Nature*, 415(6868), 137-40.
- Fehr, Ernst and Simon Gächter.** 2000. "Cooperation and Punishment in Public Goods Experiments." *American Economic Review*, 90(4), 980-94.
- Fehr, E. and I. Schurtenberger.** 2018. "Normative Foundations of Human Cooperation." *Nat Hum Behav*, 2(7), 458-68.
- Fischbacher, Urs.** 2007. "Z-Tree: Zurich Toolbox for Ready-Made Economic Experiments." *Experimental Economics*, 10(2), 171-78.
- Gächter, Simon; Elke Renner and Martin Sefton.** 2008. "The Long-Run Benefits of Punishment." *Science*, 322(5907), 1510-10.
- Grechenig, Kristoffel; Andreas Nicklisch and Christian Thöni.** 2010. "Punishment Despite Reasonable Doubt—a Public Goods Experiment with Sanctions under Uncertainty." *Journal of Empirical Legal Studies*, 7(4), 847-67.
- Herrmann, Benedikt; Christian Thoni and Simon Gächter.** 2008. "Antisocial Punishment across

Societies." *Science*, 319(5868), 1362-67.

Hopfensitz, Astrid and Ernesto Reuben. 2009. "The Importance of Emotions for the Effectiveness of Social Punishment." *The Economic Journal*, 119(540), 1534-59.

Kamijo, Y.; T. Nihonsugi; A. Takeuchi and Y. Funaki. 2014. "Sustaining Cooperation in Social Dilemmas: Comparison of Centralized Punishment Institutions." *Games and Economic Behavior*, 84, 180-95.

Markussen, Thomas; Louis Putterman and Jean-Robert Tyran. 2014. "Self-Organization for Collective Action: An Experimental Study of Voting on Sanction Regimes." *The Review of Economic Studies*, 301-24.

Nicklisch, Andreas; Kristoffel Grechenig and Christian Thöni. 2016. "Information-Sensitive Leviathans." *Journal of Public Economics*, 144, 1-13.

Ostrom, Elinor; James Walker and Roy Gardner. 1992. "Covenants with and without a Sword: Self-Governance Is Possible." *American political science Review*, 86(2), 404-17.

Rabin, Matthew. 1993. "Incorporating Fairness into Game Theory and Economics." *The American economic review*, 1281-302.

Rand, David G; Anna Dreber; Tore Ellingsen; Drew Fudenberg and Martin A Nowak. 2009. "Positive Interactions Promote Public Cooperation." *Science*, 325(5945), 1272-75.

Sefton, Martin; Robert Shupp and James M. Walker. 2007. "The Effect of Rewards and Sanctions in Provision of Public Goods." *Economic Inquiry*, 45(4), 671-90.

Sigmund, K.; H. De Silva; A. Traulsen and C. Hauert. 2010. "Social Learning Promotes Institutions for Governing the Commons." *Nature*, 466(7308), 861-3.

Sutter, Matthias; Stefan Haigner and Martin G. Kocher. 2010. "Choosing the Carrot or the Stick? Endogenous Institutional Choice in Social Dilemma Situations." *Review of Economic Studies*, 77(4), 1540-66.

Walker, James M and Matthew A Halloran. 2004. "Rewards and Sanctions and the Provision of Public Goods in One-Shot Settings." *Experimental Economics*, 7(3), 235-47.

Yamagishi, Toshio. 1986. "The Provision of a Sanctioning System as a Public Good." *Journal of Personality and social Psychology*, 51(1), 110.

Yang, Chun-Lei; Boyu Zhang; Gary Charness; Cong Li and Jaimie W Lien. 2018. "Endogenous Rewards Promote Cooperation." *Proceedings of the National Academy of Sciences*, 115(40), 9968-73.

Acknowledgments

We are grateful to the individuals who participated in this experiment. This work was supported by the National Social Science Foundation of China (Grant number: 20AZD044), the National Natural Science Fund of China (Grant numbers: 71673152; 71942002), Major Research Project of Humanities and Social Sciences of Shandong University (NO. 21RWZD15) and the Ministry of Education in China Project of Humanities and Social Sciences (Grant number: 18JJD630001).

Author Contributions

CZ, JP and JL designed the experiments. CZ and JP conducted the experiments, analyzed the data, and wrote the paper. CZ, JP, JL revised the paper. JL, JP contributed equally to this work.

Competing Interests Statement

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.